Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 22 January 2013 (MN LATEX style file v2.2)

A method for testing the cosmic homogeneity with Shannon entropy

Biswajit Pandey[1,2⋆]
[1] *Max-Planck Institute for Astrophysics, Karl-Schwarszchild Str. 1, D85748, Garching, Germany*
[2] *Department of Physics, Visva-Bharati University, Santiniketan, Birbhum, 731235, India*

22 January 2013

ABSTRACT

We propose a method for testing Cosmic homogeneity based on the Shannon entropy in Information theory and test the potentials and limitations of the method on Monte Carlo simulations of some homogeneous and inhomogeneous 3D point process in a finite region of space. We analyze a set of N-body simulations to investigate the prospect of determining the scale of homogeneity with the proposed method and show that the method could serve as an efficient tool for the study of homogeneity.

Key words: methods: numerical - galaxies: statistics - cosmology: theory - large scale structure of the Universe.

1 INTRODUCTION

The cosmological principle which assumes that the Universe is statistically homogeneous and isotropic on very large scales is one of the fundamental pillars of modern cosmology. This can not be proved in a mathematical sense and can be only verified from observations and various predictions of the physical theories based on it. The cosmic microwave background is by far the best conclusive evidence in favour of isotropy (Penzias & Wilson 1965; Smoot et al. 1992; Fixsen et al. 1996) which also strongly supports large scale homogeneity in the early Universe. Various other observations like the isotropy in angular distributions of radio sources (Blake & Wall 2002) and the isotropy of the X-ray background (Peebles 1993; Wu et al. 1999; Scharf et al. 2000) support the assumption of cosmic homogeneity on large scales. The isotropy does not by itself guarantee homogeneity and it implies homogeneity only when there is isotropy around every points. The present Universe is known to be highly inhomogeneous on small scales and there are important consequences if the inhomogeneities persist on large scales. The most important implication of inhomogeneities comes from the averaging problem in General Relativity through their effect on the large scale dynamics known as backreaction mechanism. The backreaction mechanism can cause a global cosmic acceleration without any additional dark energy component (Buchert & Ehlers 1997; Schwarz 2002; Kolb et al. 2006; Buchert 2008) although it seems unlikely that this can explain all of it (Paranjape & Singh 2008). The implications of inhomogeneities and backreaction for Cosmology are still considered to be important even if it does not provide an alternate explanation of dark energy (Paranjape 2009; Kolb et al. 2010; Ellis 2011).

The principle of cosmic homogeneity demands that the statistical properties of the observed galaxy distribution in a given finite volume does not depend on the location of that volume in the Universe. The statistical properties of galaxy distributions are characterized by the correlation functions (Peebles 1980). The two point correlation function on small scales $0.1\,h^{-1}$Mpc $\leq$ r $\leq 10\,$h^{-1}Mpc, is well described by a power law of the form $\xi(r) = (\frac{r}{r_0})^{-\gamma}$, with correlation length $r_0 \sim 5\,h^{-1}$Mpc and slope $\gamma \sim 1.8$. $\xi(r)$ vanishes at scales $> 20\,h^{-1}$Mpc which is consistent with large scale homogeneity. However the problem with correlation function analysis is that it assumes a mean density on the scale of survey which is not a defined quantity below the scale of homogeneity. Most of the statistical tools for homogeneity analysis are based on the simple number counts $n(< r)$ in spheres of radius r which is expected to scale as $\sim r^3$ for a homogeneous distribution. The conditional density (Hogg et al. 2005; Sylos Labini 2011a) measures the average density in these spheres which is expected to flatten out beyond the scale of homogeneity. The fractal analysis (Martinez & Jones 1990; Coleman & Pietronero 1992; Borgani 1995) uses the scaling of different moments of $n(< r)$ to characterize the scale of homogeneity. Some of the studies carried out with these methods on different galaxy surveys claim to have found a transition to homogeneity on sufficiently large scales $70 - 150\,h^{-1}$Mpc (Martinez & Coles 1994; Guzzo 1997; Martinez et al. 1998;

⋆ E-mail: biswa@mpa-garching.mpg.de

Bharadwaj et al. 1999; Pan & Coles 2000; Kurokawa et al. 2001; Hogg et al. 2005; Yadav et al. 2005; Sarkar et al. 2009; Scrimgeour et al. 2012) whereas some studies claim the absence of any such transition out to scale of the survey (Coleman & Pietronero 1992; Amendola & Palladino 1999; Sylos Labini et al. 2007, 2009a,b; Sylos Labini 2011b). The disagreements between various studies indicate the need for some alternative measures of homogeneity which could capture interesting information on different aspects of a homogeneous distribution and serve as complimentary and alternative tool to the existing methods in the literature. In the present work we introduce a method to asses homogeneity which is based on the evenness, a more general and robust aspect of any homogeneous distribution. We employ the Shannon entropy (Shannon 1948) to measure the unevenness characterizing inhomogeneities in a distribution and explore the possibility of employing the proposed method for exploring the scale of homogeneity.

A brief outline of the paper follows. We describe our method and various effects in Section 2, describe the tests and the data in Section 3 and presents the results and conclusions in Section 4.

2 METHOD OF ANALYSIS

Our method is based on the Shannon entropy in Information theory originally proposed by Claude Shannon to quantify the information content in strings of text.

In Information theory entropy is a measure of the amount of information required to describe the random variable. The Shannon entropy for a discrete random variable X with n outcomes $\{x_i : i = 1, ...n\}$ is a measure of uncertainty denoted by $H(X)$ defined as,

$$H(X) = -\sum_{i=1}^{n} p(x_i)\, log\, p(x_i) \quad (1)$$

where $p(x)$ is the probability distribution of the random variable X. Increase in Shannon entropy increases the uncertainty and decreases the information about the knowledge of the random variable. Another interesting aspect of Shannon entropy is its entropy-maximizing property for an uniform distribution.

We propose a method to study inhomogeneities in a 3D distribution of points. Given a set of N points distributed in 3D we consider each of the i^{th} points as center and determine $n_i(< r)$ the number of other points within a sphere of radius r as,

$$n_i(< r) = \sum_{j=1}^{N} \Theta(r - \mid \mathbf{x_i} - \mathbf{x_j} \mid) \quad (2)$$

where Θ is the Heaviside step function and $\mathbf{x_i}$ and $\mathbf{x_j}$ are the radius vector of i^{th} and j^{th} points respectively. To avoid any edge effects we discard all the points as centers which lie within a distance r from the survey boundary. Clearly the number of valid centers will decrease with increasing r for any finite volume sample. We define a separate random variable X_r for each radius r which has $M(r)$ possible outcomes each given by, $f_{i,r} = \frac{\rho_{i,r}}{\sum_{i=1}^{M(r)} \rho_{i,r}}$ with the constraint $\sum_{i=1}^{M(r)} f_{i,r} = 1$. Here $\rho_{i,r} = \frac{n_i(<r)}{\frac{4}{3}\pi r^3}$ is the density at the i^{th} center. Note that for a given sample with a finite volume, $M(r)$ is the maximum numbers of valid centers available at a radius r i.e. there are no provision for projecting further another sphere of radius r within the given volume.

The Shannon entropy associated with the random variable X_r can be written as,

$$\begin{aligned} H_r &= -\sum_{i=1}^{M(r)} f_{i,r}\, log\, f_{i,r} \\ &= log(\sum_{i=1}^{M(r)} n_i(< r)) - \frac{\sum_{i=1}^{M(r)} n_i(< r)\, log(n_i(< r))}{\sum_{i=1}^{M(r)} n_i(< r)} \end{aligned} \quad (3)$$

where the base of the logarithm is arbitrary and we choose it to be 10. Note that in X_r, f_r and H_r, r is a just a label for a number and not an argument.

In an ideal situation when all the spheres around the $M(r)$ valid centers are equally populated then one gets an uniform value of $f_{i,r} = \frac{1}{M(r)}$ for all the centers maximizing the uncertainty. Then the Shannon entropy H_r (equation 3) has its maximum value $(H_r)_{max} = log\, M(r)$ for radius r. The relative Shannon entropy $\frac{H_r}{(H_r)_{max}}$ at any r quantifies the degree of uncertainty in the knowledge of the random variable X_r. When $\frac{H_r}{(H_r)_{max}} = 1$ the knowledge about the random variable X_r becomes most uncertain. The distribution of f_r also become completely uniform when $\frac{H_r}{(H_r)_{max}} = 1$ is reached. Equivalently one can use $1 - \frac{H_r}{(H_r)_{max}}$ to quantify the information available in X_r at any r.

The joint Shannon entropy of a set of independent random variables is simply the sum of the individual entropies associated with each of the random variables. But if they are dependent then the total entropy is a sum of the conditional entropies. Entropy measures uncertainty in the random variable whereas the information is a difference in uncertainty that is a difference in entropies. The mutual information characterizes the reduction in the uncertainty in the one random variable due to the knowledge of the other.

In our case the mutual information between the random variables X_r are always positive as the random variables are not independent due to the fact that the density measurements around $M(r)$ centers at each radius r are taken from the same finite volume sample. Given that the mutual information are always positive for correlated random variables, an increase in information in X_r at one r would lead to decrease in information in X_rs at other r values. Apart from the correlations from finite sample, the random variables could have extra correlations if the points are clustered or if the points are distributed in a preferred way. These extra correlations increase the mutual information in the random variables X_r. The decrease of information in X_r with increasing r is evident irrespective of a homogeneous/inhomogeneous distribution but it would diminish differently depending on the nature and the degree of inhomogeneity present in the distribution. Maximum uncertainty or the complete loss of information in the knowledge of the random variable X_r at a radius r suggests that beyond r the random variables X_r would be independent and completely uninformative about each other. In an infinite perfectly homogeneous system if the set of spheres used for density measurements are completely independent then the

random variables X_r at each r has no knowledge about each other making them maximally uncertain and devoid of any information. When no inhomogeneities are present due to other sources, the departure of $\frac{H_r}{(H_r)_{max}}$ from 1 would be solely due to the correlations among X_r caused by the finite volume of the sample. Given the other sources are present one would expect a larger departure of this ratio from 1 and a more uneven distribution of f_r. The scale where this departure levels up with 1 indicates absence of any correlations among X_r beyond that scale. But this ideal situation may never occur in an exact sense in a finite sample as the correlations between X_rs introduced by the finite volume and confinement bias persists over the whole range of length scales. This would always give some residual information in X_r even at the largest value of r. But the mutual information content due to clustering or any other source of inhomogeneity are expected to diminish with increasing length scale provided a scale of homogeneity exist for the distribution. As the exact transition marked by $\frac{H_r}{(H_r)_{max}} = 1$ may never happen in a finite volume sample so one could set a very small limiting value for $1 - \frac{H_r}{(H_r)_{max}}$ to identify the scale of homogeneity. In our analysis we set this limit to 10^{-4}.

2.1 Effects of clustering, intrinsic inhomogeneity and finite volume

Clustering of the points is the most important source of correlations between the random variables X_r. Clustering produces fluctuations in f_r which is directly related to the fluctuations in $n(< r)$. When the points are clustered we expect on average $\bar{n}(< r)$ number of points in a volume V where $\bar{n}(< r)$ is given by,

$$\bar{n}(< r) = \lambda \int_V (1 + \xi(x))\, d^3x = \lambda\, V + 4\pi\lambda \int_0^r x^2\, \xi(x)\, dx \quad (4)$$

Here λ is the mean density of the distribution and $V = \frac{4}{3}\pi r^3$ is the volume of the sphere used. The first term has fluctuations from the Poisson noise of the same order. The Poisson noise rapidly decreases with increasing r and is only important on small scales. If the distribution itself is intrinsically inhomogeneous then Poisson noise will be modulated according to the spatially dependent intensity parameter $\lambda(x)$. The second term takes into account of clustering of the points. The variance of $n(< r)$ is,

$$\sigma^2_{n(<r)} = \overline{n^2(< r)} - (\bar{n}(< r))^2 \quad (5)$$

Going back to our definition of $f_{i,r} = \frac{\rho_{i,r}}{\sum_{i=1}^{M(r)} \rho_{i,r}}$, the fluctuations in this quantity is closely related to the fluctuations in $n_i(< r)$. The variance in f_r is,

$$\sigma^2_{f_r} = \overline{f_r^2} - (\bar{f}_r)^2 = \frac{\sigma^2_{n(<r)}}{[M(r)\,\bar{n}(< r)]^2} \quad (6)$$

which is $\frac{1}{M(r)^2}$ times the normalized variance of $n(< r)$.

One can quantify the correlations of any two random variables X_r at two different r by estimating their covariance. The correlation coefficient is,

$$C_{X_{r_i},X_{r_j}} = \frac{Cov\left(f_{r_i}, f_{r_j}\right)}{\sqrt{\sigma^2_{f_{r_i}}\sigma^2_{f_{r_j}}}} \quad (7)$$

where the indices i and j takes value between 1 to n. Here n is the total number of r values used in the analysis. In this case we have a positive correlation $0 < C_{X_{r_i},X_{r_j}} < 1$ between X_{r_i} and X_{r_j} and the correlations would be higher when the differences between r_i and r_j are smaller. The random vector $X = (X_{r_1}, X_{r_2}, X_{r_3}....X_{r_n})$ has $M(r_1)M(r_2)M(r_3).....M(r_n)$ equally likely outcomes and in principle one can estimate the full covariance matrix of all the random variables X_r at different rs to compute their correlations.

Even in the absence of any clustering and inhomogeneity one would expect a positive correlations between the random variables as their probability distribution is derived from the same finite volume sample. The set of centers at each r is a subset of the centers at the preceding r. More specifically all the set of centers at each r is a subset of the set of the centers at the smallest r implying the spheres at different r share some common regions. The fractional amount of share is larger when the difference between r is smaller thereby introducing larger correlations between the X_rs from neighboring r values. So the finite volume introduces correlations between the random variables X_r at different r which increases the mutual information in the random variables.

2.2 Effects of overlapping

One should also keep in mind that the spheres used for density measurements are not independent and could share large overlapping regions. At each r we use a finite set of spheres $S_r = \{s_{1,r}, s_{2,r}, s_{3,r}....., s_{M(r),r}\}$ for density measurements. In general these spheres overlap with each other. The probability that a random point drawn out of the distribution would lie in a particular sphere $s_{i,r}$ is $\frac{n_i(<r)}{N}$ where N is the total number of points in the distribution. The probability that the randomly drawn point would appear somewhere in the sample is 1. If the spheres are disjoint then at any r there are $M(r) + 1$ outcome of this experiment i.e the point would lie either in any one of the $M(r)$ spheres or somewhere in the sample outside the spheres. But given the fact that the spheres overlap the point could also appear at the intersections of multiple spheres. Given a finite sample A the total probability can be written as,

$$\begin{aligned} P(A) &= P(\cup_{i=1}^{M(r)} s_{i,r}) + P((\cup_{i=1}^{M(r)} s_{i,r})^c) \\ &= \sum_{i=1}^{M(r)} P(s_{i,r}) - \sum_{i\neq j,1}^{M(r)} P(s_{i,r} \cap s_{j,r}) \\ &+ \sum_{i\neq j\neq k,1}^{M(r)} P(s_{i,r} \cap s_{j,r} \cap s_{k,r}) - ... \\ &+ (-1)^{M_r - 1}\, P(\cap_{i=1}^{M(r)} s_{i,r}) + P((\cup_{i=1}^{M(r)} s_{i,r})^c) \\ &= \sum_{i=1}^{M(r)} P(s_{i,r}) - P_{overlap} + P((\cup_{i=1}^{M(r)} s_{i,r})^c) \\ &= 1 \end{aligned} \quad (8)$$

The first sum in above equation gives the sum of probabilities for individual spheres provided they are disjoint. As the spheres overlap one needs to take into account the probabilities for all the possibilities. The second sum is over all distinct pairs of spheres, third sum is over all distinct triples of spheres and so forth. The last term which gives the probability that the point would come up in a region $(s_{1,r} \cup s_{2,r} \cup s_{3,r}... \cup s_{M(r),r})^c$ outside the union of spheres is important only at small r and becomes insignificant afterwards. It may be noted that in an overlapping scenario the different terms in equation 8 can be much larger than 1 but when summed together they will always give 1. In equation 8, $P_{overlap}$ contains all the terms from overlapping spheres and successive terms in $P_{overlap}$ has alternate sign . At smaller r the most dominant contribution to $P_{overlap}$ comes from all pairs of overlapping spheres. The contribution from all triples of overlapping spheres would be the next dominant term and so on. Thus at smaller r all the individual sums in $P_{overlap}$ are different and the magnitude of each sum is lesser than its preceding sum resulting into a positive value of $P_{overlap}$. New sums appear in $P_{overlap}$ with increasing r thereby increasing its value. But at the same time the differences between the successive sums starts decreasing due to larger overlap thereby decreasing the value of $P_{overlap}$. There is a competition between this two effect and initially the first effect is more dominant than the second leading to an overall increase in $P_{overlap}$ with increasing r. As we increase the radius r the numbers of valid centers $M(r)$ decreases and they preferentially get more confined near the center of the sample. Due to the finite volume of the sample ultimately the second effect will dominate on some scale depending on the size of the finite sample and also depending on the nature of inhomogeneity to some extent. For example the confinement bias is expected to be even higher for an inhomogeneous distribution which preferentially has more particles residing near the center of the sample. When the differences between the successive sums become smaller then there would be a larger net cancellation leading to smaller values of $P_{overlap}$. As the second effect starts dominating the first one, the value of $P_{overlap}$ would start decreasing and finally on the scale of the largest sphere that would fit inside the finite sample, all the individual terms in all the individual sums in $P_{overlapp}$ would be of the same order i.e. $P(s_i) \sim P(s_i \cap s_j) \sim P(s_i \cap s_j \cap s_k) \sim P(s_1 \cap s_2 \cap s_3 ..... \cap s_{M(r)}) \sim 1$. This would lead to a very large mutual cancellation of all the sums in $P_{overlapp}$ decreasing its value to $P(\cup_{i=1}^{M(r)} s_{i,r}) - 1$ or $M(r) - 1$. This finite volume effect eventually introduces an artificial evenness in the distribution of f_r at some r depending on the sample size. We see that the effect of overlapping starts dominating much before the scale of the largest possible sphere (Figure 1) making an interpretation difficult on large scales. We note that all the other methods of testing homogeneity based on count in spheres $n(< r)$ are also affected by the same problem making the interpretations on large scales equally difficult.

In order to avoid these complicacy of correlations introduced by the overlapping of the spheres we also consider non-overlapping spheres of different radii. But the statistics becomes too noisy due to very small number of independent spheres at progressively larger radii which consequently prohibits us to address the issue of homogeneity on large scales.

A point of caution in this context is that in our method $\frac{H_r}{H_{r_{max}}} \sim 1$ at very large length scales does not necessarily indicate a real transition to homogeneity which is an obvious outcome forced by the confinement bias resulting from the finite volume of the sample. But with a large enough sample which ensures spheres upto a sufficiently large r without significant confinement of the centers could detect the scale of homogeneity if the transition happens before the scale where the confinement bias completely dominates the statistics.

3 TESTS ON HOMOGENEOUS, INHOMOGENEOUS AND CLUSTERED DISTRIBUTIONS

In order to study the prospects and limitations of the proposed method we carry out some preliminary tests by applying it to some simple distributions. We consider (i) homogeneous distribution without any clustering, (ii) inhomogeneous distribution without any clustering and (iii) strongly clustered distribution.

In above three cases the unevenness in type (i) is only due to Poisson noise. The unevenness in type (ii) distribution is controlled by Poisson noise too but it has the additional complexity that the contribution to the unevenness is governed by the distribution of the spatially dependent intensity parameter $\lambda(x)$. In type (iii) distribution both Poisson noise and the clustering together contributes to the unevenness.

For the first two types we generate a set of Monte Carlo realizations of some simple homogeneous and inhomogeneous point process. For the third type we use a set of N-body simulations where the points are strongly clustered.

3.1 Monte Carlo simulations of homogeneous and inhomogeneous point processes

We generate a set of Monte Carlo realizations of different types of homogeneous and inhomogeneous distributions.

For the sake of simplicity we consider some simple radial density distributions $\rho(r) = K\,\lambda(r)$ where (i) $\lambda(r) = \frac{1}{r}$, (ii) $\lambda(r) = \frac{1}{r^2}$ and (iii) $\lambda(r) = 1$. K is a normalization constant. The distributions in (i) and (ii) are inhomogeneous Poisson distributions which are isotropic about only one point and the inhomogeneities in these distributions persist on all scales. The distribution in (iii) is a homogeneous Poisson point process which has a constant density everywhere.

Enforcing the desired number of particles N within radius R one can turn the radial density function into a probability function within $r = 0$ to $r = R$ which is normalized to one when integrated over that interval. So the probability of finding a particle at a given radius r is $P(r) = \frac{r^2\lambda(r)}{\int_0^R r^2\lambda(r)\,dr}$ which is proportional to the density at that radius implying more particles in high density regions.

We generate the Monte Carlo realizations of these distributions using a Monte Carlo dartboard technique. The maxima of the function $r^2\lambda(r)$ in $P(r)$ is at $r = R$ in (i) and (ii) whereas in (iii) it is same and constant everywhere. We label the maximum value of $P(r)$ as P_{max}. We randomly choose a radius r in the range $0 \leq r \leq R$ and a probability value is randomly chosen in the range $0 \leq P(x) \leq P_{max}$. The actual probability of finding a particle at the selected

radius is then calculated using expression for $P(r)$ and compared to the randomly selected probability value. If the random probability is less than the calculated value, the radius is accepted and assigned isotropically selected angular coordinates θ and ϕ, otherwise the radius is discarded. In this way, radii at which particle is more likely to be found will be selected more often because the random probability will be more frequently less than the calculated actual probability. We choose $R = 200$ in h^{-1} Mpc unit and $N = 10^5$. We generate 10 realizations of each of the above density distributions and analyze them separately with the method described in section 2.

3.2 N-body simulations

We simulate the dark matter distribution using a Particle-Mesh (PM) N-body code. The simulations use 256^3 particles on a 512^3 mesh and cover a comoving volume of $[921.6h^{-1}\text{Mpc}]^3$. We use $(\Omega_{m0}, \Omega_{\Lambda 0}, h) = (0.27, 0.73, 0.71)$ for the cosmological parameters along with a ΛCDM power spectrum with spectral index $n_s = 0.96$ and normalization $\sigma_8 = 0.812$ (Komatsu et al. 2009). A simple "sharp cutoff" biasing, scheme (Cole, Hatton & Weinberg 1998) was used to extract particles from the simulations that are biased relative to the dark matter and are labelled as galaxies. The bias parameter b of each simulated biased sample was estimated using the ratio

$$b = \sqrt{\frac{\xi_g(r)}{\xi_{dm}(r)}} \tag{9}$$

where $\xi_g(r)$ and $\xi_{dm}(r)$ are the galaxy and dark matter two-point correlation functions respectively. This ratio is found to be constant at length-scales $r \geq 5h^{-1}$Mpc and we use the average value over $5 - 40h^{-1}$Mpc. We use this method to generate galaxy samples with bias values $b = 1.5, 2$ and 2.5. We run the simulation for three different realizations of the initial density fluctuations and extract biased samples from each of them. We extract randomly $N = 10^5$ particles from three non overlapping spherical regions of radius $R = 200\,h^{-1}$Mpc from each of the simulation boxes giving us total nine samples for each of the bias values. The numbers N and R and the specific spherical geometry are chosen just to maintain the uniformity in all the analysis presented here. Generally one can choose any number of particles and any geometry for the samples.

4 RESULTS AND CONCLUSIONS

In the Figure 2 we show the variations of $\frac{H_r}{(H_r)_{max}}$ with distance r for some homogeneous and inhomogeneous distributions described in subsection 3.1. For the Poisson distribution the value of $\frac{H_r}{(H_r)_{max}}$ shows a small departure from 1 at smaller values of r which decreases with increasing r. The ratio reaches a value ~ 1 at $r \sim 15$. As the points are uncorrelated, the Void Probability function (White 1979) for a homogeneous Poisson distribution is $e^{-\lambda V}$. It is expected that such a distribution would become homogeneous on a scale $r \sim \lambda^{-\frac{1}{3}}$, which is a measure of the average size of voids in the distribution. Since the correlations between the random variables X_r in the homogeneous Poisson distribution mainly comes from the finite volume effect one

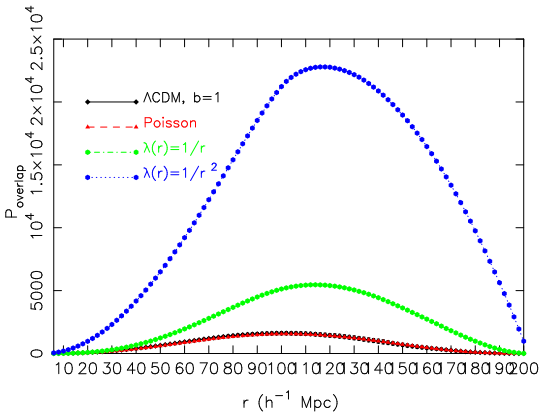

Figure 1. This shows the $P_{overlap}$ (equation 8) as a function of r for different distributions as indicated in the figure. The error bars overplotted on the data points here are not visible due to their very small sizes.

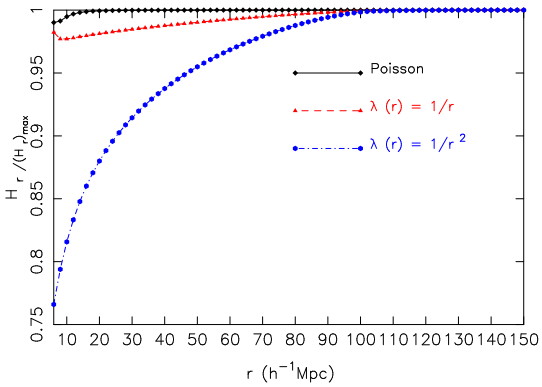

Figure 2. This shows the ratio $\frac{H_r}{(H_r)_{max}}$ as a function of r for homogeneous Poisson distribution and two different anisotropic distributions. The tiny error bars overplotted on the data points are invisible here.

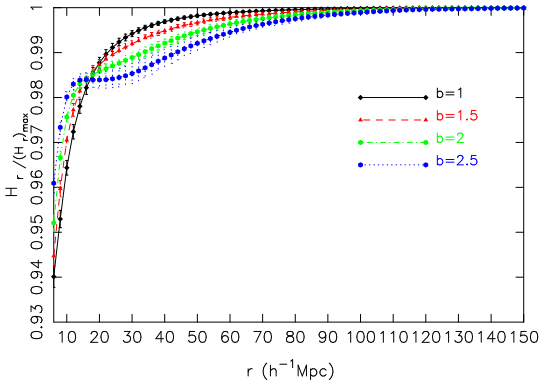

Figure 3. This shows the ratio $\frac{H_r}{(H_r)_{max}}$ as a function of r for the unbiased ΛCDM model and its different biased variants. $1-\sigma$ error bars obtained from 9 different samples are overplotted at each points.

would expect a lesser amount of mutual information in the set of random variables and consequently a faster information leakage but the information in X_r i.e. $1-\frac{H_r}{(H_r)_{max}}$ would never be exactly zero due to persistence of the finite volume correlations at all r.

The $\frac{H_r}{(H_r)_{max}}$ values at smaller r values shows relatively larger departure from 1 for both the inhomogeneous distributions considered here. The ratio has a larger departure for the distribution with $\lambda(r) = \frac{1}{r^2}$ than the distribution with $\lambda(r) = \frac{1}{r}$ on all scales. This shows that the ratio $\frac{H_r}{(H_r)_{max}}$ describes the degree of inhomogeneity present in two different anisotropic distributions in correct order. Both these distributions are only isotropic about one point and have a radial density distribution which does not allow them to be homogeneous on any scales. The ratio $\frac{H_r}{(H_r)_{max}} \sim 1$ at $r \sim 120$ for both of them simply shows the importance of confinement bias which ultimately is expected to take over the situation at some r depending on the sample size and type of inhomogeneity present. The confinement bias gets a larger boost in both of these distributions (Figure 1) as there are already more particles preferentially located near the center of the samples. This effect produces larger overlap and hence larger mutual information are stored in the random variables from neighbouring rs. Consequently this accelerates the leakage of information in X_r with increasing r finally forcing the ratio $\frac{H_r}{(H_r)_{max}} \sim 1$. It may be noted that for the homogeneous Poisson distribution there are no such additional boost coming from preferential deposition and further there are no extra correlations due to clustering. The transition $\frac{H_r}{(H_r)_{max}} \sim 1$ is reached there at a r much before the scale where the confinement bias play a dominant role indicating a real transition to homogeneity.

In the Figure 3 the variations of $\frac{H_r}{(H_r)_{max}}$ with distance r are shown for ΛCDM N-body simulations with different linear bias values. The assumption of linear bias holds reasonably well on large scales. Different types of galaxies are biased differently with respect to the dark matter and the inhomogeneities in the distributions of different types of galaxies could be modulated by their bias values. We would also like to investigate here how well the ratio $\frac{H_r}{(H_r)_{max}}$ can track the variation in inhomogeneities in a distribution on different scales and give some useful information about the characteristics of the inhomogeneities present in them. We see that the ratio initially depart from 1 at small r for all the distributions with higher bias values showing systematically lesser inhomogeneity (lesser information in X_r) than lower bias values. In the unbiased ΛCDM model the centers used for calculating $n(< r)$ are residing in all types of environments (clusters, sheets, filaments and voids). The fact that the centers are distributed across different types of nonlinear structures increases the unevenness in the distribution of f_r and the mutual information in X_r. Whereas with increasing bias values the particles in a biased distribution preferentially represent progressively higher density peaks in the density field homogenizing the distribution on the corresponding length scales. This is primarily due to the fact that in a biased distribution the centers are located in less diverse environments. A complete reversal of this behaviour is seen at $15-20\,h^{-1}$Mpc after which with increasing bias values the distributions systematically show larger inhomogeneity and more information in X_r at all scales r until they eventually merge to $\frac{H_r}{(H_r)_{max}} = 1$. This behaviour is related to the fact that increasing the radius r beyond the typical scale of the nonlinear structures in the biased distributions would lead to larger disparity in $n(< r)$ and f_r values and hence the larger inhomogeneities. The scales corresponding to the reversal in the behaviour of ratio corresponds to the typical scale of the nonlinear structures present in the biased distributions. Further on large scales the correlation functions $\xi(r)$ in biased distributions are b^2 times larger than the $\xi(r)$ in unbiased distributions introducing larger correlations among all the random variables X_r at larger r. Eventually all the curves merge to $\frac{H_r}{(H_r)_{max}} = 1$ in the range $100 - 150\,h^{-1}$Mpc with the difference that the transition appears to happen at a relatively larger r values for larger bias. Although this can not be emphasized due to the size of the overlapping error bars at those length scales. It may be noted that the inhomogeneities in the ΛCDM model with different biases are much smaller as compared to the inhomogeneous radial density distributions given by $\lambda(r) = \frac{1}{r^2}$. The information in X_r for a Poisson distribution mostly comes from the finite volume effect whereas the very large mutual information in the random variables in two inhomogeneous distributions is the combined outcome of finite volume correlations and large degree of confinement bias. In the unbiased and biased ΛCDM models the mutual information are generated by correlations from confinement bias and clustering together. It may be noted that the confinement biases are much smaller and very similar in the ΛCDM models and the homogeneous Poisson distributions (Figure 1). The confinement biases become very large in the two anisotropic distributions considered here which eventually force the ratio $\frac{H_r}{(H_r)_{max}}$ to 1 at some scales.

If a scale of homogeneity exists then ideally one would expect a naturally emerging evenness in f_r when the spheres of that radius around each centers statistically include similar numbers of different types of nonlinear structures. So provided that the correlations introduced by the finite volume of the sample and the confinement bias are much smaller compared to the correlations induced by clustering then the leakage of information in X_r would ideally track the variation of inhomogeneity in the distribution with scale r. But unfortunately in a real situation one has to also deal with the correlations due to finite volume and confinement bias modulating the information in X_r or the evenness of f_r. We find that the confinement bias could start to dominate even much before the scale defined by the largest sphere included in any finite sample (Figure 1). The degree of correlations introduced among X_rs would also depend on the type of inhomogeneity present in the distribution specially if it is overpopulated near the center of the sample with respect to the rest of the volume resulting into larger overlap between the spheres. When dealing with galaxy redshift surveys one has to also properly take into account the different selection effects involved, redshift space distortions and the specific geometry of the samples. We plan to carry out analysis in the publicly available catalogues from large galaxy redshift surveys (e.g. 2dFGRS, Colles et al. 2001; SDSS, Stoughton et al. 2002; 2MASS, Huchra et al. 2012) in future works and investigate these issues further. It may be noted here that a caveat of our method is we assume that we always have access to the data samples on a spatial hypersurface of constant time. Though this assumption is approximately valid in case of low redshift galaxy samples

in the nearby Universe or data samples from a snapshot of N-body simulations, it is not strictly true as the entire observational samples do not consist of objects on a constant time hypersurface but rather on a light cone. Taking this into account it is hard to distinguish radial inhomogeneity from time evolution without assuming a cosmological model. Consequently this could make an inhomogeneous but isotropic model (e.g. anti-Copernican void models) to look like a homogeneous one on large scales. Fortunately these models can be constrained with other observations such as SNe, CMB, BAO and measurements of Hubble parameter (Zibin et al. 2008; Clifton et al. 2008; Biswas et al. 2010; Clarkson 2012).

Finally we note that the method presented here has the desired ability to characterize inhomogeneities and their variations on different length scales in any 3D distribution of points. The method has the potential to successfully identify any existing scale of homogeneity given a sufficiently large volume which can ensure negligible confinement bias and less effective overlap between the spheres upto large length scales. Given the large survey volumes that are currently available from the modern redshift surveys the method could provide an efficient tool for exploring the issue of Cosmic homogeneity.

ACKNOWLEDGMENT

The author thanks Somnath Bharadwaj for useful comments and discussions. The author would also like to thank an anonymous refree for insightful comments about the paper. The author acknowledges the Alexander von Humboldt Foundation for financial support and the Computing Center of the Max Planck Society in Garching (RZG) for the computing facilities provided for the work.

REFERENCES

Amendola, L., & Palladino, E. 1999, ApJ Letters, 514, L1
Biswas, T., Notari, A., & Valkenburg, W. 2010, JCAP, 11, 30
Bharadwaj, S., Gupta, A. K., & Seshadri, T. R. 1999, A&A, 351, 405
Borgani, S. 1995, Physics Reports, 251, 1
Blake, C., & Wall, J. 2002, Nature, 416, 150
Buchert, T. 2008, General Relativity and Gravitation, 40, 467
Buchert, T., & Ehlers, J. 1997, A&A, 320, 1
Clarkson, C. 2012, Comptes Rendus Physique, 13, 682
Clifton, T., Ferreira, P. G., & Land, K. 2008, Physical Review Letters, 101, 131302
Cole, S., Hatton, S., Weinberg, D. H., & Frenk, C. S. 1998, MNRAS, 300, 945
Coleman, P. H., Pietronero, L. 1992, Physics Reports, 213, 311
Colles, M. et al.(for 2dFGRS team) 2001,MNRAS,328,1039
Ellis, G. F. R. 2011, Classical and Quantum Gravity, 28, 164001
Fixsen, D. J., Cheng, E. S., Gales, J. M., et al. 1996, ApJ, 473, 576
Guzzo, L. 1997, New Astronomy, 2, 517
Hogg, D. W., Eisenstein, D. J., Blanton, M. R., Bahcall, N. A., Brinkmann, J., Gunn, J. E., & Schneider, D. P. 2005, ApJ, 624, 54
Huchra, J. P., Macri, L. M., Masters, K. L., et al. 2012, ApJS, 199, 26
Kolb, E. W., Matarrese, S., & Riotto, A. 2006, New Journal of Physics, 8, 322
Kolb, E. W., Marra, V., & Matarrese, S. 2010, General Relativity and Gravitation, 42, 1399
Komatsu, E., et al. 2009, ApJS, 180, 330
Kurokawa, T., Morikawa, M., & Mouri, H. 2001, A&A, 370, 358
Martinez, V. J., & Jones, B. J. T. 1990, MNRAS, 242, 517
Martinez, V. J., & Coles, P. 1994, ApJ, 437, 550
Martinez, V. J., Pons-Borderia, M.-J., Moyeed, R. A., & Graham, M. J. 1998, MNRAS, 298, 1212
Pan, J., & Coles, P. 2000, MNRAS, 318, L51
Paranjape, A., & Singh, T. P. 2008, Physical Review Letters, 101, 181101
Paranjape, A. 2009, arXiv:0906.3165
Penzias, A. A., & Wilson, R. W. 1965, ApJ, 142, 419
Peebles, P. J. E. 1993, Principles of Physical Cosmology. Princeton, N.J., Princeton University Press, 1993
Peebles, P. J. E. 1980, The large scale structure of the Universe. Princeton, N.J., Princeton University Press, 1980, 435 p.,
Sarkar, P., Yadav, J., Pandey, B., & Bharadwaj, S. 2009, MNRAS, 399, L128
Scharf, C. A., Jahoda, K., Treyer, M., et al. 2000, ApJ, 544, 49
Schwarz, D. J. 2002, arXiv:astro-ph/0209584
Scrimgeour, M. I., Davis, T., Blake, C., et al. 2012, MNRAS, 3412
Shannon, C. E. 1948, Bell System Technical Journal, 27, 379-423, 623-656
Smoot, G. F., Bennett, C. L., Kogut, A., et al. 1992, ApJ Letters, 396, L1
Stoughton, C., et al. 2002, AJ, 123, 485
Sylos Labini, F., Vasilyev, N. L., & Baryshev, Y. V. 2007, A&A, 465, 23
Sylos Labini, F., Vasilyev, N. L., & Baryshev, Y. V. 2009a, Europhysics Letters, 85, 29002
Sylos Labini, F., Vasilyev, N. L., Pietronero, L., & Baryshev, Y. V. 2009b, Europhysics Letters, 86, 49001
Sylos Labini, F. 2011a, Classical and Quantum Gravity, 28, 164003
Sylos Labini, F. 2011b, Europhysics Letters, 96, 59001
White, S. D. M. 1979, MNRAS, 186, 145
Wu, K. K. S., Lahav, O., & Rees, M. J. 1999, Nature, 397, 225
Yadav, J., Bharadwaj, S., Pandey, B., & Seshadri, T. R. 2005, MNRAS, 364, 601
Zibin, J. P., Moss, A., & Scott, D. 2008, Physical Review Letters, 101, 251303