

DOI 10.18524/1810-4215.2021.34.244253

DETERMINATION OF GALAXY DISTRIBUTION WITH STATISTICAL MOMENTS

A.V. Tugay¹, S.Yu. Shevchenko², L.V.Zadorozhna¹¹ Taras Shevchenko National University of Kyiv,
Kyiv, Ukraine, *e-mail: tugay.anatoliy@gmail.com*² Schmalhausen Institute of Zoology, NASc of Ukraine,
Kyiv, Ukraine, *e-mail: astromott@gmail.com*

ABSTRACT. In this report we discuss topological studies of large scale structure of the Universe (LSS) from XMM-Newton, Sloan Digital Sky Survey and simulated data of galaxy distribution. Early works in this mentioned field were based on genus statistics, which is averaged curvature of isosurface of smoothed density field. Later, significant number of other methods was developed. This comprise Euler characteristics, Minkowski functionals, Voronoi clustering, alpha shapes, Delanuy tessellation, Morse theory, Hessian matrix and Soneira-Peebles models.

In practice, modern topology methods are reduced to calculation of the three Betti numbers which shall be interpreted as a number of galaxy clusters, filaments and voids. Such an approach was applied by different authors both for simulated and observed LSS data.

Topology methods are generally verified using LSS simulations. Observational data normally includes SDSS, CFHTLS and other surveys. These data have many systematical and statistical errors and gaps. Furthermore, there is also a problem of underlying dark matter distribution. The situation is not better in relation to calculations of the power spectrum and its power law index which does not provide a clear picture as well. In this work we propose some tools to solve above problems. First, we performed topology description of simple LSS models such as cubic, graphite-like and random Gaussian distribution of matter. Our next idea is to set a task for LSS topology assessment using X-ray observations of the galaxies. Although, here could be a major complication due to current lack of detected high energy emitting galaxies. Nevertheless, we are expecting to get sufficient results in the future encouraging comprehensive X-ray data. Here we present analysis of statistical moments for four galaxy samples and compare them with the behavior of Betti numbers. Finally, we consider the options of applying artificial neural networks to observed galaxies and fill the data deficiency. This shall enable to define topology at least for superimposed superclusters and other LSS elements.

Keywords: Galaxies, large-scale structure, XMM-Newton, X-ray.

АНОТАЦІЯ. У цій роботі ми розглянули топологічні дослідження великомасштабної структури всесвіту (LSS) з використанням даних обсерваторій XMM-Newton, SDSS та змодельованих розподілів галактик. Ранні роботи у зазначеній галузі базувалися на типовій статистиці, що є усередненою кривизною ізоповерхні згладженого поля щільності. Пізніше було розроблено значну кількість інших методів. Сюди входять характеристики Ейлера, функціонали Мінковського, кластеризація Вороного, альфа-форми, теселяція Делано, теорія Морзе, матриці Геса та моделі Сонейра-Піблза. На практиці сучасні методи топології зводяться до обчислення трьох чисел Бетті, які слід інтерпретувати як ряд скупчень галактик, філаментів та порожнин. Такий підхід був застосований різними авторами як до змодельованих, так і до спостережуваних даних LSS. Методи топології зазвичай перевіряються на моделюванні LSS, а дані спостережень зазвичай включають SDSS, CFHTLS та інші огляди неба. Ці дані мають багато систематичних та статистичних помилок та недостатньо кількість інформації. Більш того, існує також проблема розподілу темної матерії. Не кращий стан і з розрахунками спектру потужності та степеневого коефіцієнту потужності, які також не дають чіткої картини. Ми пропонуємо деякі інструменти для вирішення вищевказаних проблем. По-перше, нас цікавить опис топології простих моделей LSS, таких як кубічна, графітова та випадкова гаусівська модель розподілу речовини. Наша наступна ідея це поставити задачу оцінити топологію LSS за допомогою рентгенівських спостережень галактик. Хоча тут може виникнути серйозне ускладнення через поточну відсутність виявлених галактик, що випромінюють високі енергії. Тим не менш, ми очікуємо в майбутньому отримати достатні результати, заховаючи вичерпні рентгенівські дані. У даній роботі представлено аналіз статистичних моментів для чотирьох вибірок

галактик і їх порівняння з поведінкою чисел Бетті. Нарешті, ми розглядаємо можливість застосування штучної нейронної мережі до спостережуваних галактик і заповнити дефіцит даних. Це дозволить визначити топологію принаймні для накладених суперкластерів та інших елементів великомасштабної структури всесвіту.

Ключові слова: Галактики, великомасштабна структура Всесвіту, ХММ-Newton, рентгенівське випромінювання.

1. Introduction

Observable galaxy distribution is affected by large selection effects and only minor part of galaxies detected. Hence for cosmological studies only general parameters of galaxy distribution can be evaluated. Correlation functions and power spectrum are the most popular approaches to it (Elyiv et al., 2012). It is also important to characterize the shape of large scale galaxy structures, which is significant to reveal the galaxies origin. Topological methods are the most straightforward to it.

Genus was the first topological index applied to LSS (Weinberg et al., 1987). It can be defined as the difference between number of voids and isolated regions.

Since 2000 much more topological indexes for LSS study were developed and applied. Use of Minkowski functionals is one of the main modern topological methods in this realm (van de Weygaert et al., 2013).

While Genus (G) is global topological index, Minkowski functionals describes local curvature of isodensity surface. For example, Euler characteristics are compositions of number of blobs (b0), tunnels (b1) and voids (b2):

$$E = b_0 - b_1 + b_2; G = b_0 - E.$$

Here b0, b1, b2 – Betti numbers, ranks of homology groups of isodensity surface or number of d-dimensional holes (d=0,1,2). Idea of persistent homology study is to find stable correspondence between abel group and topological space (isodensity surface). Minkowski functionals are calculated by the method of alpha shapes (Pranav et al., 2016). Alpha shape is the most intuitive case of Delanauy triangulation. It can be calculated by Voronoi clustering tool.

It should be stressed that all LSS topology studies since 1980th are based on the analysis of isosurfaces of smoothed galaxy density field. The problems of recovering such surfaces are discussed below in this work. Minkowski functionals are calculated by eigenvectors of Hessian matrix of that isosurfaces. The Hessian matrix of a function f is the Jacobian matrix of the gradient of the function f. If the gradient (the

vector of the partial derivatives) of a function f is zero at some point x, then f has a critical point (or stationary point) at x. The determinant of the Hessian at x is called, in some contexts, a discriminant. If this determinant is zero then x is called a degenerate critical point of f, or a non-Morse critical point of f. Otherwise it is non-degenerate, and called a Morse critical point of f. The determinant of the Hessian matrix, when evaluated at a critical point of a function, is equal to the Gaussian curvature of the function considered as a manifold. The eigenvalues of the Hessian at that point are the principle curvatures of the function, and the eigenvectors are the principle directions of curvature.

Morse theory was applied for topological analysis of SDSS in (Sousbie et al., 2010; Sousbie et al., 2010). Filaments were detected based on Hessian critical points although they are affected by significant statistical errors.

By the time, Betti numbers were only theoretically discussed or applied to simulated galaxy distributions. In (Pranav et al., 2016) Betti numbers were calculated for three main LSS models where galaxy clusters, filaments and walls respectively dominating. The dependencies of Betti numbers from isosurface density threshold were determined. These relations have maximum that indicate approximate boundaries of superclusters. Since such result looks rather simple, we supposed that it could be reproduced applying more simple methods without recovering undetermined isodensity surface and Hessian calculation.

2. Method

Instead of Betti numbers derivation from Hessian eigenventors and critical points we calculated statistical moments of spatial galaxy distribution. These moments include:

1. Mean value. $\bar{N} = \sum N_i / M$
2. Standard deviation. $\sigma = \frac{\sqrt{M \sum N_i^2 - (\sum N_i)^2}}{M}$
3. Skewness. $g_1 = \frac{m_3}{m_2^{3/2}} = \frac{\frac{1}{M} \sum_{i=1}^M (N_i - \bar{N})^3}{\left[\frac{1}{N} \sum_{i=1}^N (N_i - \bar{N})^2 \right]^{3/2}}$
4. Kurtosis. $g_2 = \frac{m_4}{m_2^2} - 3 = \frac{\frac{1}{M} \sum_{i=1}^M (N_i - \bar{N})^4}{\left[\frac{1}{N} \sum_{i=1}^M (N_i - \bar{N})^2 \right]^2} - 3$

We divided given volume containing N galaxies by $M = K^3$ equal cubic cells and calculated the number of galaxies in each cell – N_i . We also build probability distribution function (PDF) as dependence of $\delta = N_i / \bar{N} - 1$ from K. We counted numbers of cubic cells (P) in 8 bins with values of δ step equal to 0.25. Finally we plotted the distributions of P and a function of K and δ step number (J).

3. Samples

We considered two toy models of LSS and two observed galaxy samples.

1. Cubic cells of round clusters with random background of field galaxies. This sample contains 99936 points.

2. 3D version of smoothed network of straight filaments connecting randomly distributed clusters. 2D version of this model was presented in (Voycehovskiy & Tugay, 2018). This second sample contains 260500 simulated galaxies.

3. Based on SDSS (Abazajian et al., 2009) observed galaxies positions we selected a cube with the following spatial boundaries: lower cube side was oriented along celestial equator; bottom vertex was placed in the Solar System; two sides have RA equal to 45 and 135 degrees respectively; cube facet length defined as 420 Mpc ($z < 0.1$). The third sample contains 246003 optical galaxies.

4. The last sample is Xgal catalog of 3933 X-ray galaxies (Zadorozhna et al., 2021).

4. Results

Points distribution, statistical moments and PDF for Sample 1 of the cubic cluster grid are plotted on Figs. 1, 2 and 3 respectively. Significant maximum on dispersion plot indicates transition of the cell size from cluster interior to exterior. Similar plots were obtained for Betti number in (Pranav et al., 2016). In case of correct generalisation of such behavior we could consider LSS topology description applying our method instead of Betti numbers.

Unfortunately, statistical moments distributions for other samples (Figs 5, 10, 11) do not demonstrate any specific extremes. This could be explained by weak manifestation of LSS and large content of the field galaxies. Maximum on dispersion plot for Xgal sample (Fig. 11) the most probably is the result of computational artifact due to low number of galaxies in the sample.

PDF for two simulated samples (Figs 3 and 6) demonstrate significant difference of their spatial structure. Unfortunately, PDF for observed samples suffer from lack of galaxies for some ranges of averaging parameters. We suppose that deeper analysis of real galaxy distribution will help to determine LSS topology.

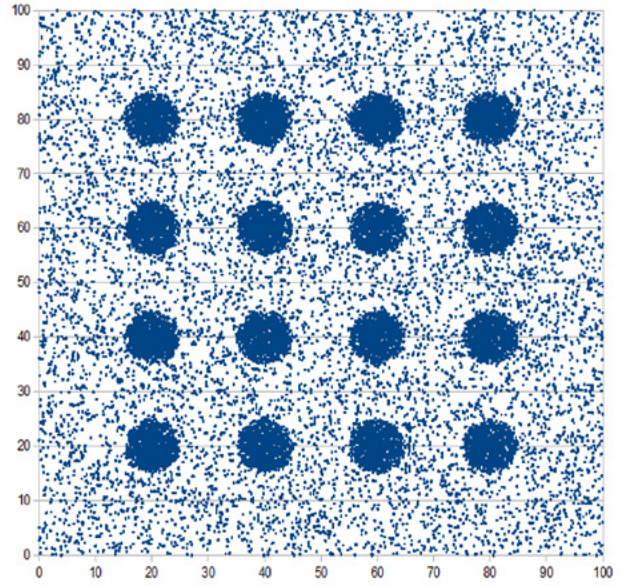


Fig. 1: Points distribution in Sample 1. X and Y axis are the facets of the simulated cube.

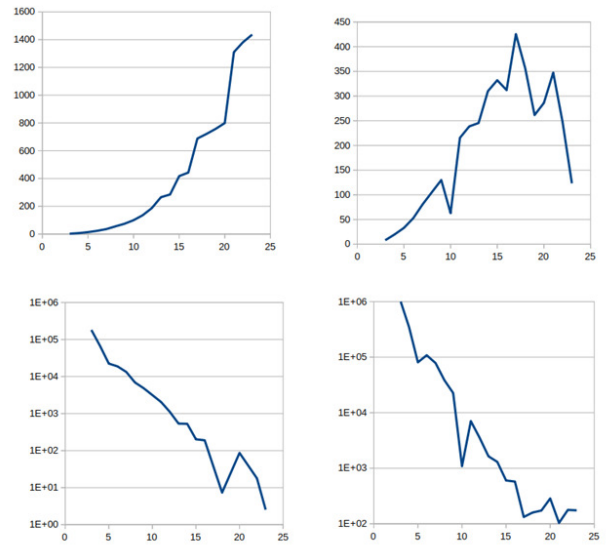


Fig. 2: X-axis – cells number along the modelled cube side (K), Y-axis – statistical moments values; Upper left – Average number; Upper right – Dispersion; Bottom left – Asymmetry; Bottom right – kurtosis.

5. Conclusion

Presented method based on statistical moments can not replace current LSS topology analysis by Betti numbers. Nevertheless, we suppose that further developments will be useful for topological description of observed LSS in optical and X-ray bands. In

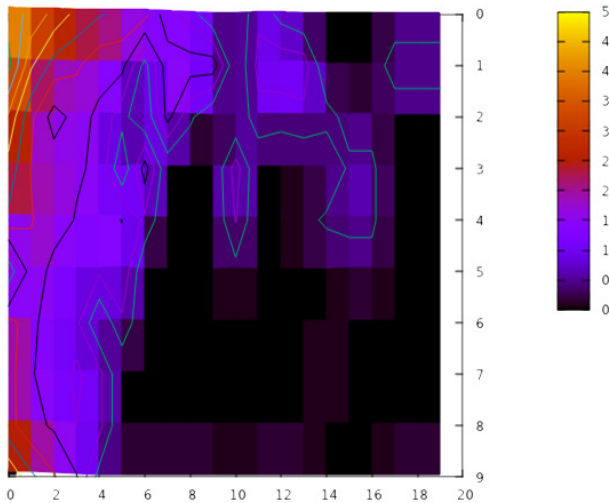


Fig. 3: PDF for Sample 1. X-axis – cells number on the simulated cube side (K). Y-axis – normalised density step number (J). Colour indicates $\lg P(J,K)$.

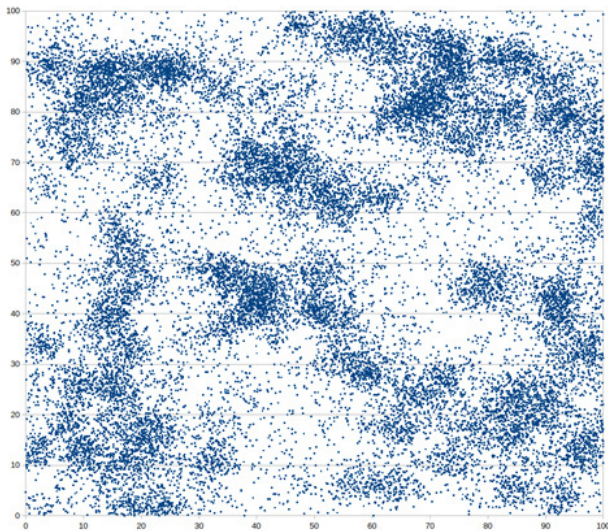


Fig. 4: Spatial distribution of simulated galaxies from Sample 2. A layer of 3D random filament distribution plotting.

particular, we propose to use neural networks in later studies. Neural networks are capable to generate unseen galaxies, clusters and superclusters. With these superimposed superclusters it will be possible to continue the analysis of topology of large-scale structure of the Universe at new level.

Acknowledgements. This work has been supported by Ministry of Education and Science of Ukraine: Grant of the Ministry of Education and Science of Ukraine for perspective development of a scientific direction

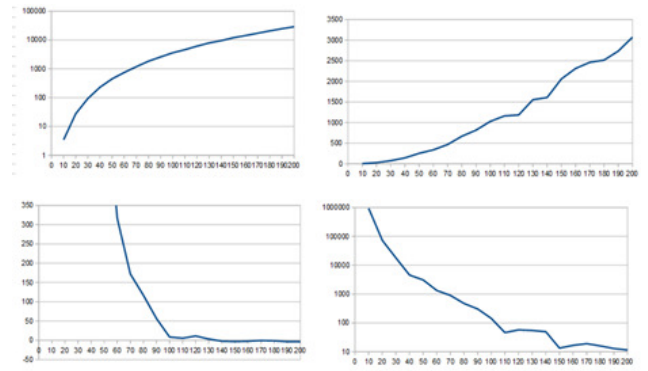


Fig. 5: Statistical moments for random distribution of the large scale structures.

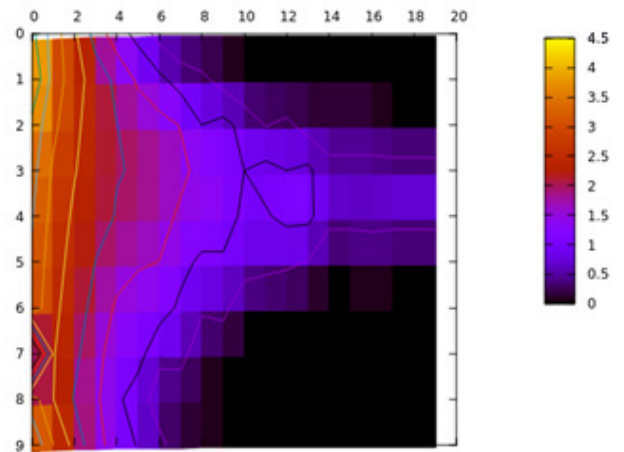


Fig. 6: Probability distribution function for randomly distributed large scale structures.

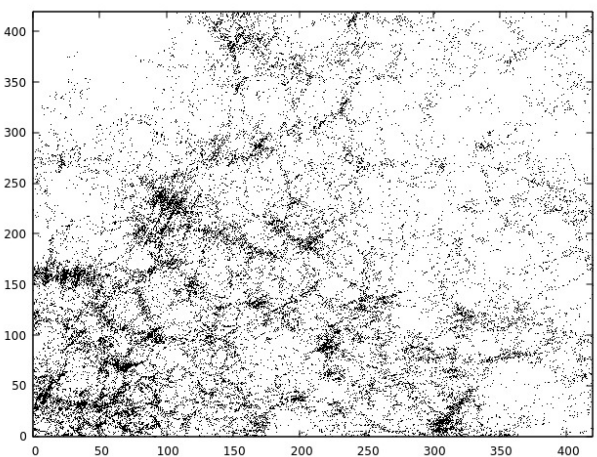


Fig. 7: Sample 3. A layer of SDSS galaxies distribution.

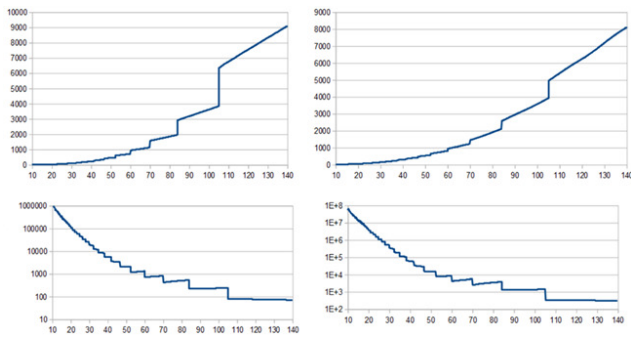


Fig. 8: Statistical moments for SDSS galaxies.

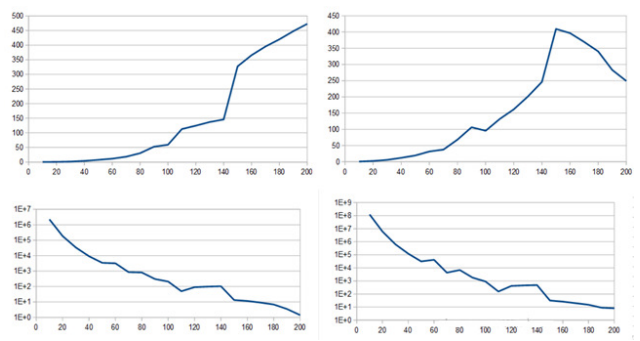


Fig. 11: Statistical moments of X-ray galaxies.

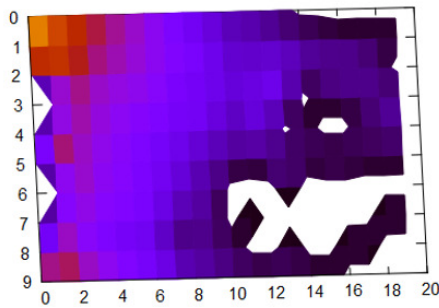


Fig. 9: PDF for SDSS galaxies.

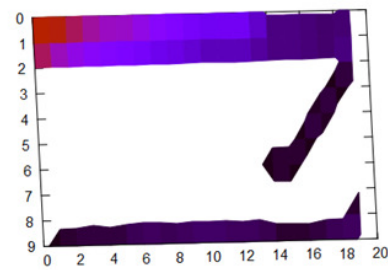


Fig. 12: Probability distribution function for X-ray galaxies.

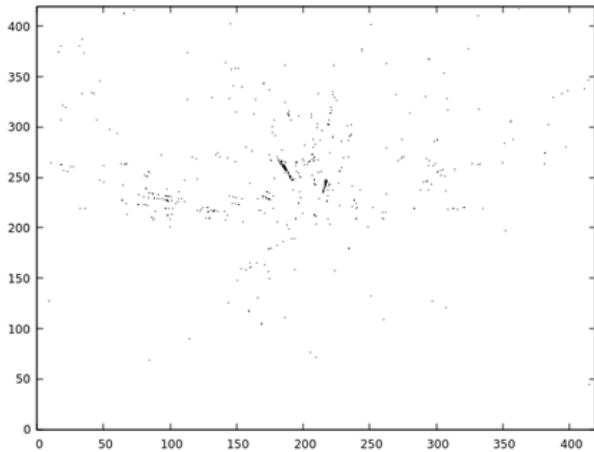


Fig. 10: X-ray galaxies distribution. Cube side equals to 420 Mpc.

References

Abazajian K.N., Adelman-McCarthy J.K., Agueros M.A.: 2009, *ApJS*, **182**, 543.
 Elyiv A., Clerc N., Plionis M.: 2012, *A&A*, **537A**, 131.
 Pranav P., Edelsbrunner H., van de Weygaert R.: 2016, *MNRAS*, **465**, 4281.
 Sousbie T., Colombi S., Pichon C.: 2009, *MNRAS*, **393**, 457.
 Sousbie T., Pichon C., Kawahara H.: 2011, *MNRAS*, **414**, 384.
 van de Weygaert R., Vegter G., Edelsbrunner H., et al.: 2011, *Trans. Comput. Sci.*, **XIV**, 60.
 Voitsekhovskiy V.V., Tugay A.V.: 2018, *Odessa Astron. Publ.*, **31**, 52.
 Weinberg D.H., Gott J.R., Melott A.L.: 1987, *Ap. J.*, **321**, 2.
 Zadorozhna L.V., Tugay A.V., Shevchenko S.Yu., et al.: 2021, *KPCB*, **37**, 149.

"Mathematical sciences and natural sciences" at Taras Shevchenko National University of Kyiv. This research has made with support of the Center for the Collective Use of Scientific Equipment "Laboratory of high energy physics and astrophysics".