PASJ: Publ. Astron. Soc. Japan **61**, 85–95, 2009 February 25

Analysis of the Spatial Distribution of Galaxies in the Numerical Galaxy Catalog

Haruhiko UEDA
Faculty of Education and Human Studies, Akita University, 1-1 Tegata-Gakuen, Akita, Akita 010-8502
ueda@ipc.akita-u.ac.jp
Masahiro NAGASHIMA
Faculty of Education, Nagasaki University, 1-14 Bunkyo-machi, Nagasaki, Nagasaki 852-8521
masahiro@nagasaki-u.ac.jp
and
Hideki YAHAGI
Research Institute for Information Technology, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581
yahagi@cc.kyushu-u.ac.jp

(Received 2006 August 15; accepted 2008 September 27)

Abstract

We examine the spatial distribution of galaxies in the Numerical Galaxy Catalog, which is based on a hierarchical clustering framework. This catalog is constructed from a semianalytic model of galaxy formation combined with high-resolution N-body simulations in a Λ-dominated flat cold dark-matter cosmological model. To quantify the spatial distribution of galaxies, we use a graph-theoretical approach, because it is one of the most powerful statistical ways of estimating spatial data. In particular, three types of graphs are adopted: the Delaunay graph, the minimal spanning tree, and the constellation graph. To quantify the galaxy distributions, we apply statistical measures suitable for each type. The mass distributions in the cold dark-matter universe are also examined to clarify differences in the spatial distribution of dark matter and galaxies. The spatial distribution of galaxies in the two-degree field (2dF) galaxy redshift survey is finally compared with that in the Numerical Galaxy Catalog. From our analysis, we definitely show that galaxy distributions in the Numerical Galaxy Catalog are different from the dark-matter distributions. We also find that the Numerical Galaxy Catalog considerably improves the theoretical prediction of spatial galaxy distributions, although it does not do enough to reproduce the 2dF galaxy redshift survey. Finally, we show that the constellation graph and the minimal spanning tree are convenient for quantifing the galaxy distributions in an objective manner.

Key words: galaxies: formation — galaxies: statistics — methods: numerical — methods: statistical

1. Introduction

Large-scale structure of the universe has been eagerly studied to test hypotheses of dark-matter identification, the nature of initial-density perturbations, and the mechanism of structure growth. In particular, the values of fundamental cosmological parameters are estimated to quantify the galaxy distributions. Because considerable progress has been made in cosmology, the consensus is that cold dark matter is the most likely candidate for dark matter, and the cosmic structure grows due to the gravitational amplification of random-Gaussian initial-density fluctuations. In addition, results of measuring anisotropies of the cosmic microwave background with the Wilkinson Microwave Anisotropy Probe (Spergel et al. 2003) and the type Ia supernova rate (Riess et al. 1998; Perlmutter et al. 1999) reveal that the fundamental cosmological parameters have the following values: the density parameter, $\Omega_0 \simeq 0.3$, the cosmological constant term, $\Omega_\Lambda \simeq 0.7$, and the Hubble constant (in units of $100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$), $h \simeq 0.7$. The fundamental frame of a cosmological model is now established, and we will advance form the present stage to the next stage; namely, studies of the galaxy formation and evolution.

To approach this subject, new galaxy surveys of high qualitiy and correct modeling of galaxy formation are critical. Fortunately, a new generation of galaxy surveys that provides the key to solving the mechanism of galaxy formation and evolution has become available. To construct correct models of galaxy formation was once known as a very difficult work for cosmologists (Peacock 1999). Due to progress in computer capacity and the theory of galaxies, however, the idea of realistic calculations of galaxy formation and evolution has gradually been put into practice. There exist two strategies to provide theoretical galaxy formation and evolution. One is a hydrodynamical simulation that incorporates gas dynamics into cosmological structure models (Monaghan 1992). Because the hydrodynamical simulation directly calculates dynamics related to galaxy formation processes, the computer complexity is very vast. The resolution is therefore limited, and the mass scale that can be studied is restricted. For this reason, a numerical hydrodynamical simulation of galaxy formation still has large uncertainties.

Recently, another strategy has attracted attention. This is a semianalytic model of galaxy formation combined with the high-resolution N-body simulations (Roukema et al. 1997; Kauffmann et al. 1999a, 1999b; Diaferio et al. 1999, 2001; Somerville et al. 2001; Helly et al. 2003a, 2003b; Hatton et al.

2003). Because this method does not follow the dynamics related to galaxy-formation processes, the computation is much faster. This means that semianalytic modeling is flexible for studying the effects of varying the assumptions and parameter values. Therefore, semianalytic modeling is an important approach, complementary to hydrodynamical simulations for understanding the physics of galaxy formation. Lately, theoretical data that are based on semianalytic models combined with N-body simulations have been constructed by some groups (Kauffmann et al. 1999a; Hatton et al. 2003). Among them, the Numerical Galaxy Catalog (νGC) constructed by Nagashima et al. (2005) is one of the most reliable models, because the high-resolution N-body simulations in a Λ-dominated flat cold dark matter (ΛCDM) cosmological model are used.

In this paper, we use this catalog and examine the spatial distribution of galaxies. Only a few attempts have so far been made at this point, and studies of the spatial distribution of galaxies in νGC provide useful information about galaxy formation and evolution. In particular, we can check the reliability of theoretical predictions by comparing our results with recent observations of high quality.

Two-point correlation function is a famous measure to quantify spatial distributions of galaxies. However, spatial properties are fully specified by the two-point correlation function, only if the distributions have a Gaussian form. To recognize nonlinear properties, it is necessary to calculate the higher-order correlation functions (Peebles 1980; Suto & Matsubara 1994). However, it is very difficult to calculate the higher-order correlation functions from real galaxy surveys. Many other statistical methods are therefore proposed for estimating the large-scale structure (Saslaw 2000; Martínez & Saar 2002). Among them, Minkowski functionals (Mecke & Wagner 1991; Mecke et al. 1994; Schmalzing et al. 1996; Sahni et al. 1998; Kerscher et al. 2000) and genus statistics (Gott et al. 1986; Weinberg et al. 1987; Melott 1990; Coles 1992; Matsubara & Suto 1996; Hikage et al. 2001; 2002) are famous measures. In addition to these two well-known measures, the graph-theoretical approach is also useful for quantifying the galaxy distributions (Ueda & Itoh 1999).

In this paper, we use the graph-theoretical approach for quantifying the galaxy distributions. Because the distribution of galaxies is discrete, it is natural to apply discrete mathematics to the statistical descriptions of galaxies. Graph theory is one of the most typical topics of discrete mathematics, and we therefore examine the galaxy distributions by means of a graph-theoretical approach. This approach has been examined by many people, and is known as a useful procedure for describing spatial patterns of galaxies (Barrow et al. 1985; Bhavsar & Nigel Ling 1988; Gurzadyan & Kocharyan 1994; Graham et al. 1995; Pearson & Coles 1995; Krzewina & Saslaw 1996). In this paper, we adopt three types of graph (the Delaunay graph, the minimal spanning tree, and the constellation graph) to quantify the galaxy distributions.

The Delaunay graph (Delone 1934) is a dual graph of Voronoi tessellations (Voronoi 1908). This graph is well known to computer technicians who are interested in computational geometry (Preparata & Shamos 1985); e.g., the Delaunay triangulations are utilized for hight interpolations of the terrain (de Berg et al. 1997). In addition, this graph is used for finite elements (Livesley 1983), computational fluid dynamics (Chung 2002), two-dimensional interpolation (Späth 1995), image analysis (Seul et al. 2000), and computer vision (Forsyth & Ponce 2002). In cosmology, however, the Voronoi tessellation is more famous, because it was regarded as being a toy model of large-scale structures. For example, van de Weygaert (1994) presented a detailed study as to how several quantities (area, number of vertices, edge length, etc.) relate to the cells of a Voronoi tessellation. It was also used as a description of the geometric skeleton of galaxy distributions by Icke and van de Weygaert (1987) and van de Weygaert and Icke (1989). In addition, the dynamical Voronoi tessellation was used for examining the observed three-dimensional structure of galaxies (Zaninetti 1991). The Voronoi tessellation was also applied to find clusters of galaxies by Ramella et al. (2001). On the other hand, the Delaunay graph was used for spatial interpolation by Schaap and van de Weygraert (2000). Although applications in cosmology have been in decreasing numbers recently, we revive this attractive graph to quantify the galaxy distributions in νGC.

The minimal spanning tree (MST) has been applied in many fields. For example, this is used in computer science (Aho et al. 1983; Eiselt & Sandblom 2000), and geographic information systems (Wise 2002). MST has also been applied rigorously by many cosmologists. Barrow et al.(1985) constructed MST from the Zwicky catalog (Zwicky et al. 1961–68) and the Harvard-Smithsonian Center for Astrophysics (CfA) catalog (Huchra et al. 1983), and showed that the distribution of edge length of the MST in Zwicky and CfA catalogs are very different from that in a Poisson model. Bhavsar and Nigel Ling (1988) showed that MST is useful for seeing whether the filaments are real or not. The nature of MST was examined by van de Weygaert (1991) and Krzewina and Saslaw (1996). Doroshkevich et al. (1999) used the MST edge length to analyze the galaxy distribution, and Adami and Mazure (1999) applied MST to cluster galaxies. Martínez et al. (1990) used MST for obtaining the multifractal dimension. Bhavsar and Splinter (1996), Doroshkevich et al. (2001), Demianski and Doeoshkevich (2004), and Colberg (2007) also used MST. Although MST may be regarded as an old-fashioned statistical measure, this is still a good measure to quantify galaxy distributions.

The constellation graph has been examined by Ueda and Itoh (1997) in the context of cosmology. From previous analyses, the constellation graph is one of the most useful methods for estimating galaxy distributions; this is useful for discriminating differences in the galaxy distributions among power-law models (Ueda & Itoh 1997) and the cold dark-matter models (Ueda & Itoh 1999; Ueda et al. 2003). Moreover, this approach is also useful when applied to two-dimensional galaxy distributions (Ueda et al. 2001; Ueda & Takeuchi 2006).

The remainder of this paper is organized as follows. Data that we use are described in section 2. An overview of the graph-theoretical approach is given in section 3, and in section 4 the results of our analyses are presented. Finally, we provide a summary in section 5.

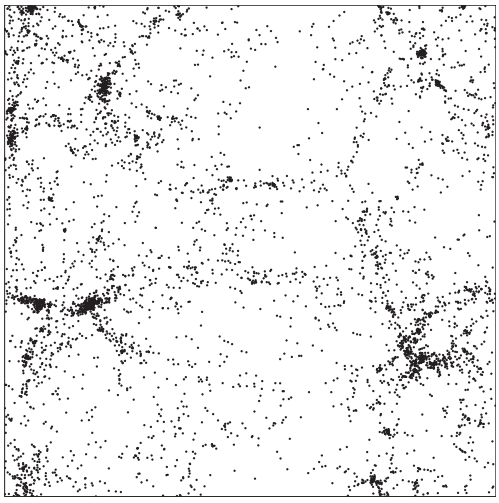

Fig. 1. Example of a thin slice of the spatial galaxy distributions of νGC in real space with a B-band absolute magnitude of $M_B \leq -16.5$.

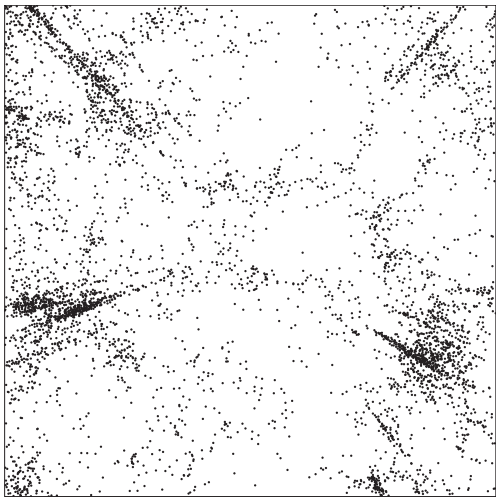

Fig. 2. Example of a thin slice of the spatial galaxy distributions of νGC in redshift space with a B-band absolute magnitude of $M_B \leq -16.5$.

2. Data

In this paper, we use three kinds of data: galaxies in νGC, dark matter in an N-body simulation, and galaxies in 2dFGRS. We will begin by explaining these data in order.

2.1. Numerical Galaxy Catalog (νGC)

The catalog νGC is constructed by a semianalytic model of galaxy formation combined with high-resolution cosmological N-body simulations. This catalog was created by Nagashima et al. (2005) as follows: the distribution of dark matter is constructed by means of high-resolution N-body simulations carried out by using the parallel version (Yahagi 2005) of the adaptive mesh refinement N-body code (Yahagi & Yoshii 2001). In this simulation, cosmological parameters are fixed as $(\Omega_0, \Omega_\Lambda, h, \sigma_8) = (0.3, 0.7, 0.7, 0.9)$ with a particle number of $N = 512^3$, and a box size of $L = 70\,h^{-1}$ Mpc. Dark halos are identified in this simulation using a friends-of-friends algorithm with a linking length of $b = 0.2$. The minimum number of particles identifying a dark halo is 10, which corresponds to $3.04 \times 10^9\,M_\odot$. As pointed out by Nagashima et al. (2005), it is important to notice that the fully resolved mass in νGC might be larger than $5 \times 10^{11}\,M_\odot$. Because the mass of dark-matter particles in νGC is smaller than other theoretical data in the public domain, νGC is a better and more reliable model of galaxy formation and evolution.

The galaxy is then assigned to a position within the halo, by accounting for the tidal stripping of subhalos, hot halo gas cooling, star formation, and supernova feedback. In these processes, the semianalytic model proposed by Nagashima and Yoshii (2004) is used to predict the properties of the galaxy. Finally, three-dimensional galaxy distributions are obtained. Notice that there exists uncertainty as to how to estimate the supernova feedback. Here, we adopt the strong supernova feedback model, because this is in much better agreement with near-infrared faint galaxy number counts and redshift distribution than the weak feedback model (Nagashima et al. 2005).

In this paper, volume-limited samples as well as apparent magnitude-limited samples are used. In volume-limited samples, we construct two types whose B-band absolute magnitudes are $M_B \leq -16.5$ and $M_B \leq -18.0$. The galaxy numbers in each type are $N = 29065$ and $N = 9436$, respectively. Figure 1 is a thin slice of νGC, in which galaxies with $M_B \leq -16.5$ are indicated by points. In addition, the volume-limited galaxy distributions in redshift space are also displayed in figure 2. The apparent magnitude-limited samples are used to compare the galaxy distributions in νGC with observations (see subsection 2.4).

2.2. N-Body Simulation

In order to estimate the difference between the galaxy and dark-matter distributions, we performed a low-resolution cosmological N-body simulation in ΛCDM cosmology. Various codes are available to perform N-body simulations, and we used Hydra, obtained from the Hydra Consortium Web page. Hydra is an adaptive particle-particle, particle-mesh N-body simulation program. It can be used with periodic boundary conditions (Couchman et al. 1995),

Here, we adopted the same cosmological parameters, $(\Omega_0, \Omega_\Lambda, h, \sigma_8) = (0.3, 0.7, 0.7, 0.9)$, with a particle number of $N = 64^3$, a box size of $L = 70\,h^{-1}$ Mpc, and particle mass of $2.1 \times 10^{11}\,M_\odot$ that is almost the same as typical galaxy mass or fully resolved mass in νGC. We purposely used this low-resolution simulation (resolution length $r_c \sim 1.56$ Mpc) because it has the same resolution that we used to examine the properties of the large-scale structure. To compare the distribution of dark matter with galaxies in νGC, we selected the same number of particles ($N = 9436$, 29065) randomly. For a statistical treatment, we produced three samples from this simulation. Figure 3 is a slice of the dark-matter distributions in ΛCDM cosmology. In addition, the dark-matter distributions in redshift space are also displayed in figure 4.

2.3. Two-Degree Field Galaxy Redshift Survey (2dFGRS)

The 2dFGRS is an optical spectroscopic survey of objects brighter than $b_J = 19.45$ selected from the APM (Automatic

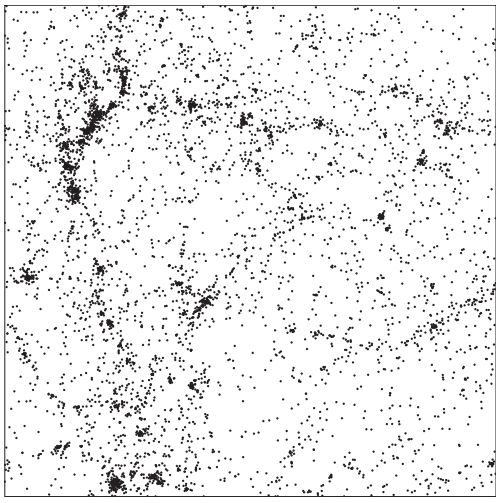

Fig. 3. Example of a thin slice of the spatial dark-matter distributions of N-body simulations in real space.

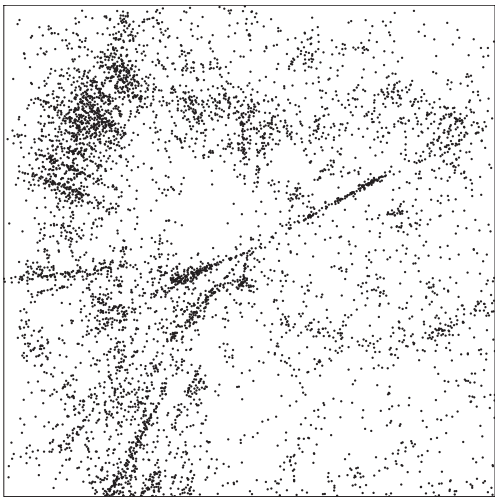

Fig. 4. Example of a thin slice of the spatial dark-matter distributions of N-body simulations in redshift space.

Plate Measuring) Galaxy Survey (Maddox et al. 1990a, b). Therefore, 2dFGRS is a magnitude-limited sample. The 2dFGRS was divided into two main regions: the region of the North Galactic Pole (NGP), which covers $147.^\circ5 < \alpha < 222.^\circ5$ and $-7.^\circ5 < \delta < 2.^\circ5$, and the region of the South Galactic Pole (SGP), which covers $325^\circ < \alpha < 52.^\circ5$ and $-37.^\circ5 < \delta < -22.^\circ5$. In addition to these two main regions, random fields were observed, but we did not utilize these fields.

The data that we analyzed were from a public release of the final version (Colless et al. 2001 and references therein); 102426 galaxy redshifts were contained in the data release. All redshift identifications were assigned to a quality parameter Q in the range of 1–5; $Q \geq 3$ redshifts were 98.4% reliable, and we used galaxies with $Q \geq 3$. In addition, we only selected nearby galaxy distributions within $15 \leq r \leq 300\,\mathrm{Mpc}$ to avoid any effects of galaxy evolution. [We supposed cosmological parameters $(\Omega_0, \Omega_\Lambda, h) = (0.3,0.7,0.7)$]. Although there exists a huge number of galaxies in the original 2dFGRS, the galaxies that are available are 18377 in NGP, and 25909 in SGP. Although the galaxy number that we utilized in our analysis was rather smaller than that of the original 2dFGRS, it was sufficient to examine the galaxy distributions in a statistical way.

2.4. *Mock Sample*

To compare the galaxy distributions in 2dFGRS with those in theoretical models, we constructed artificial mock surveys from νGC and N-body simulations. To make mock surveys, a periodic boundary condition was used. Putting νGC side by side, we constructed a large-volume sample. Based on the peculiar velocity of each galaxy, we transferred the galaxy distributions from real space to redshift space. By cutting off excess regions, we obtained a map with the same geometry as the 2dFGRS. To select galaxies whose apparent magnitude is brighter than $b_J \leq 19.45$, we finally constructed a mock sample of νGC. In this procedure, we used the following color relation:

$$b_J = B - 0.28(B - V). \tag{1}$$

For a statistical treatment, we produced three mock samples. In fact, the galaxy numbers in the mock samples were slightly larger than 2dFGRS, and we chose galaxies randomly to set the same galaxy number. For the sake of completeness, we directly calculated the luminosity functions of our mock samples and 2dFGRS, and confirmed that these were consistent with each other.

The mock sample of dark matter was constructed by the use of same procedures. However, particles do not have any information about the absolute luminosity. We therefore randomly assigned the luminosity to each particle under the condition that it corresponds to a luminosity function determined by the Schechter function (Schechter 1976),

$$\phi(L)\mathrm{d}L = \phi_* \left(\frac{L}{L_*}\right)^\alpha \exp\left(-\frac{L}{L_*}\right) \mathrm{d}\left(\frac{L}{L_*}\right). \tag{2}$$

The characteristic galaxy luminosity, L_*, relates to the characteristic magnitude, M_*, as

$$\frac{L_\odot}{L_*} = 10^{-0.4(M_\odot - M_*)}. \tag{3}$$

We adopted parameter values of $M_* = -19.66 + 5\log h$ and $\alpha = -1.21$ (Norberg et al. 2002). The parameter ϕ_* was set suitably since the galaxy number in 2dFGRS and the particle number in the N-body simulation were the same.

3. Overview of the Graph-Theoretical Approach

3.1. *Graph*

Although the graph theory has long been studied in mathematics and computer science, technical terms for graphs are not unified (Harary 1969). To avoid confusion, we define some terms that are needed to understand the graph-theoretical approach. A graph (or undirected graph) is a set of vertices connected by edges. Strictly speaking, a graph G is a pair $G = (V, E)$ with a set of vertices V and a set of unordered pairs of vertices called edges E. A subgraph of graph G is a graph whose vertices and edges are subsets of G. In particular,

a spanning subgraph of a graph G is a subgraph if it has the same vertex set as G.

A graph is labeled when the vertices are distinguished from one another by name. As is well known, a labeled graph is expressed by an adjacency matrix. The adjacency matrix $A = (a_{pq})$ for a graph to n vertices is an $n \times n$ matrix in which $a_{pq} = 1$ if vertex p is adjacent to vertex q, and $a_{pq} = 0$ otherwise. Since the adjacency matrix is symmetric, eigenvalues of this matrix are always real numbers. A graph structure can be extended by assigning a weight to each edge of the graph. In particular, we weight each edge by assigning a factor of edge length l^j (j is a real number). In this case, the element of the adjacency matrix is $a_{pq} = l^j$ if vertex p is adjacent to vertex q, and $a_{pq} = 0$ otherwise.

Since there exists no absolute criterion as to how to construct graphs from the galaxy distributions, we here adopted three types of graphs: the Delaunay graph, the MST, and the constellation graph.

3.2. *Delaunay Graph (Delaunay Tetrahedron)*

A Delaunay triangulation is a two-dimensional Delaunay graph, and to see the Delaunay graph intuitively, we first explain this case. The Delaunay triangulation of a point set is a collection of edges satisfying an empty circle property, i.e., for each edge one can find a circle containing the edge's endpoint, but not containing any other points. Figure 5 is an example of a Delaunay triangulation constructed from two-dimensional point distributions. A Delaunay tetrahedron is a three-dimensional Delaunay graph. The Delaunay tetrahedron is a collection of four points of a discrete point distribution whose circumsphere does not contain any other points. More detailed discussions and figures of three-dimensional Delaunay tetrahedrons can be found in van de Weygaert (1994). In this paper, we used the Delaunay tetrahedron to quantify the galaxy distributions.

Although the Delaunay graph has not been examined in the context of cosmology, it has many attractive features. For example, the Delaunay graph is convenient, because for a given set of vertices it is unique, and its existence is guaranteed. We here highlight the homogeneousness of the galaxy distributions using volumes of the tetrahedron. This is achieved by counting the number of Delaunay tetrahedrons as a function of volumes. If the galaxy distributions are rather homogeneous, almost all tetrahedrons have the same volumes. As clustering becomes strong, the number of tetrahedrons that have small or large volumes increases. We therefore used Delaunay graph to estimate homogeneity of the galaxy distributions.

3.3. *Minimal Spanning Tree (MST)*

A tree is a graph in which any two of the vertices are connected by exactly one path, and a spanning tree for G is a tree which contains all of the vertices of G. Given a weighted, undirected graph G, an MST is one for which the sum of the weights on its edges is less than, or equal to, the sum of the weights on the edges of any other spanning tree for G. The MST is also known to be unique. Figure 6 is an example of the MST that is constructed from the same two-dimensional point distributions in figure 5. (We construct MST in figure 6 for weighting each edge by assigning the edge length.) Notice that

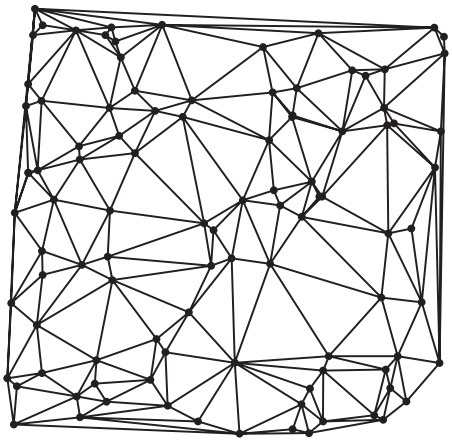

Fig. 5. Example of the Delaunay graph constructed from two-dimensional point distributions.

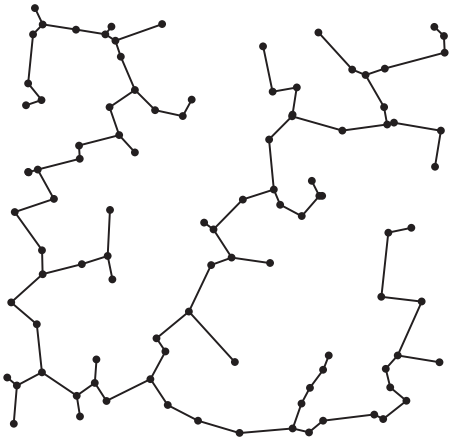

Fig. 6. Example of the MST constructed from the same two-dimensional point distributions in figure 5. Notice that the MST is a spanning subgraph of the Delaunay graph.

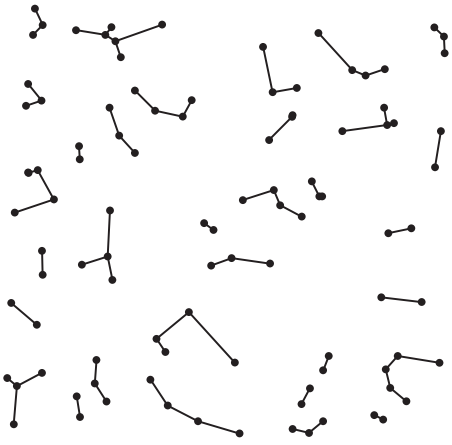

Fig. 7. Example of the constellation graph constructed from the same two-dimensional point distributions in figure 5 or figure 6. Notice that the constellation graph is a spanning subgraph of the Delaunay graph or the MST.

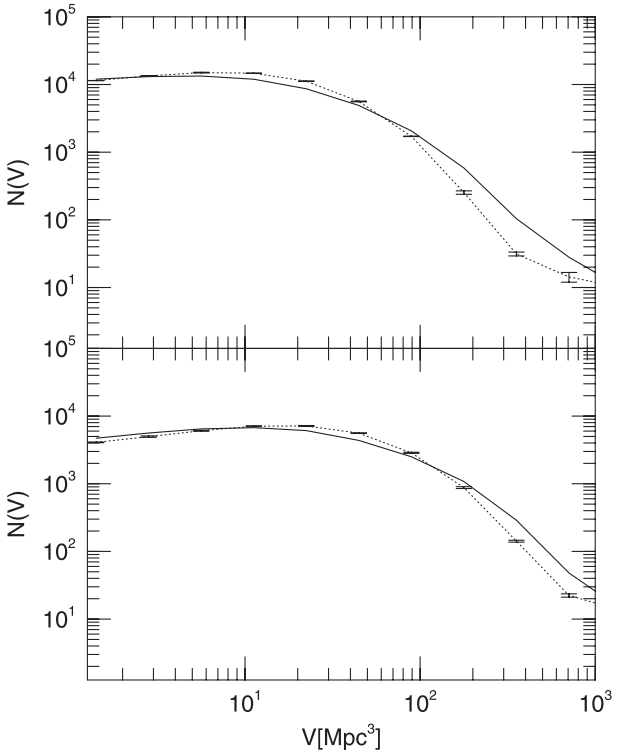

Fig. 8. Number counts of Delaunay tetrahedrons in the Delaunay graph as a function of volume V. Solid and short-dashed curves represent $N_{\nu\mathrm{GC}}(V)$ and $N_{\mathrm{NB}}(V)$ in real space. The upper panel (lower panel) is the $M_B \leq -16.5$ ($M_B \leq -18.0$) case.

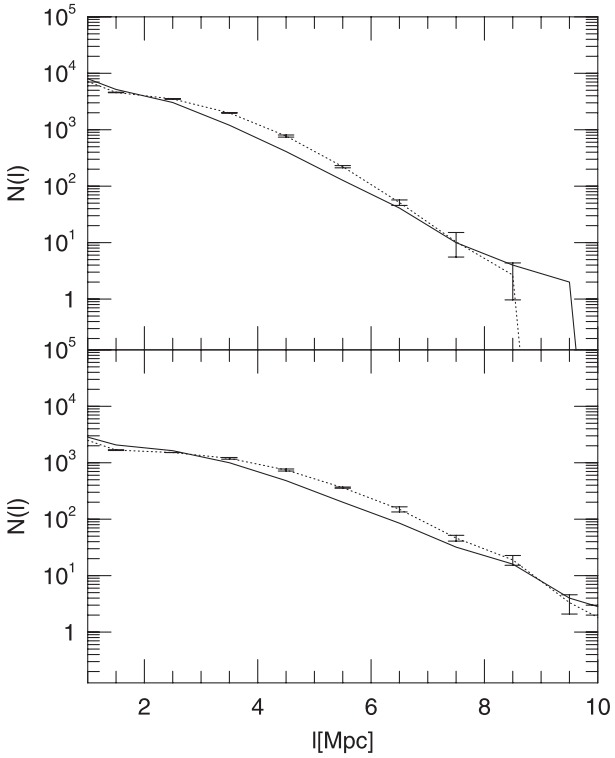

Fig. 9. Number counts of edge lengths in the MST as a function of the edge length l. Solid and short-dashed curves represent $N_{\nu\mathrm{GC}}(l)$ and $N_{\mathrm{NB}}(l)$ in real space. The upper panel (lower panel) is the $M_B \leq -16.5$ ($M_B \leq -18.0$) case.

the MST is a spanning subgraph of the Delaunay graph.

The MST has many attractive features to quantify the galaxy distributions; in particular, it highlights the filamentary nature. In this paper, we also utilized the MST to quantify the galaxy distributions counting the number of edges as a function of the edge length.

3.4. Constellation Graph

The constellation graph is constructed by connecting the nearest neighbor pair of galaxies. This graph is also uniquely determined, and details of this graph can be found in Ueda and Itoh (1997). The constellation graph is not frequently used as compared with the Delaunay graph and the MST, although it is useful to cosmologists. Figure 7 is an example of the constellation graph that is constructed from the same two-dimensional point distributions in figure 5 or figure 6. Notice that the constellation graph is a spanning subgraph of the Delaunay graph or the MST.

The constellation graph connects with pattern recognition of human eyes. Humans have regarded the arrangement of stars in the night sky as being animals or heroes. The constellation graph is related to the constellation as its name indicates, although it resembles not the constellations of Greece, but the constellations of China. In particular, it highlights the pattern in a small area. In this paper, we utilized the constellation graph to quantify the galaxy distributions, estimating the eigenvalues of adjacency matrix.

4. Analysis

The main purpose of this section is to quantify the galaxy or dark-matter distributions in νGC, an N-body simulation, and 2dFGRS. Analyses of this subject fall into three categories: In the first place, we compared the galaxy distributions in νGC with the dark-matter distributions in real space. In the second place, a comparison in redshift space was performed. Lastly, the galaxy (or dark matter) distributions in mock samples were compared with observations.

In addition to the above main subject, we also reexamined the usefulness of the graph-theoretical approach. This is because the graph-theoretical approach has not been compared with other famous statistics, although this has the advantage of dealing with discrete spatial data directly. We shall discuss these two subjects in detail.

4.1. νGC Galaxy versus Dark Matter in Real Space

In this subsection, we compare the galaxy distributions in νGC with the dark-matter distributions found by an N-body simulation. Figure 8 is the number counts of the volume of Delaunay tetrahedrons in the Delaunay graph. As stated in subsection 2.1, two different cases, $M_B \leq -16.5$ (upper panel) and $M_B \leq -18.0$ (lower panel), are represented. In each panel, solid and short-dashed curves are the number counts of volume in νGC, $N_{\nu\mathrm{GC}}(V)$, and the dark matter in N-body simulations, $N_{\mathrm{NB}}(V)$, respectively. For a statistical treatment, we produced three samples from N-body simulations, and the errors in

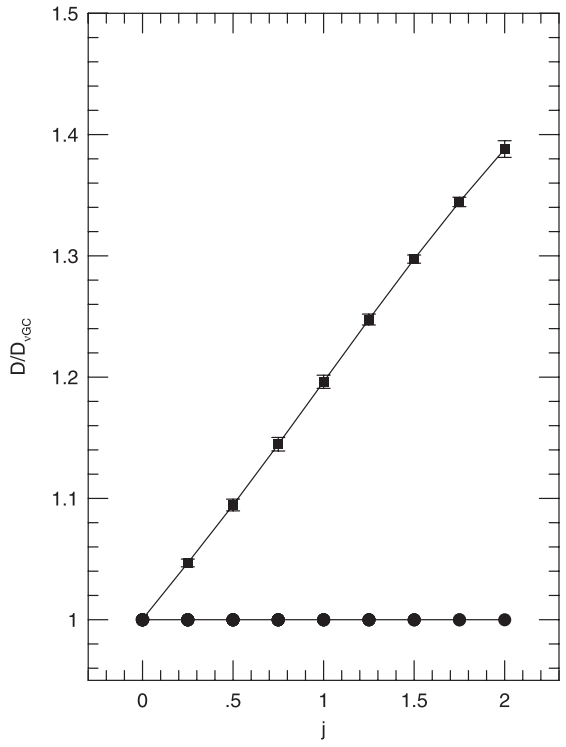

Fig. 10. Relative mean absolute deviation of galaxy and dark matter in real space with $M_B \leq -18.0$, as a function of index j. Filled circles and squares represent the relative mean absolute deviations of νGC, and an N-body simulation, respectively.

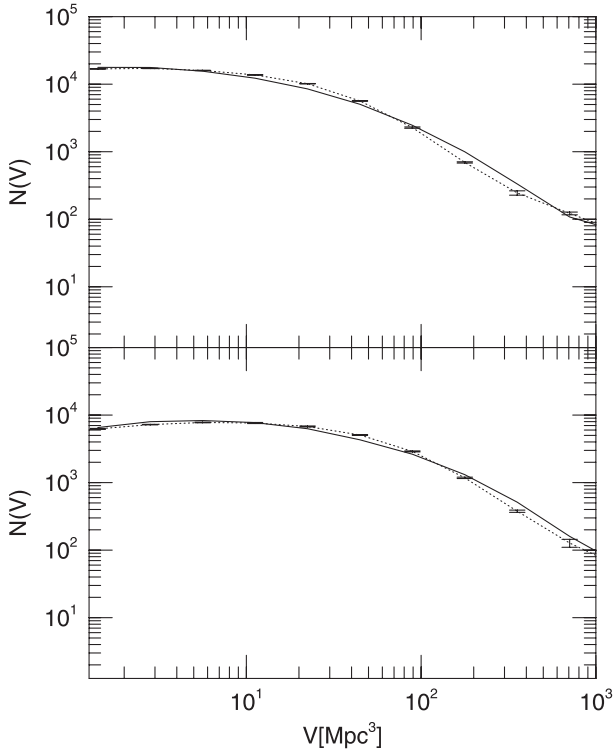

Fig. 11. Number counts of Delaunay tetrahedrons in the Delaunay graph as a function of volume V. Solid and short-dashed curves represent $N_{\nu\mathrm{GC}}(V)$, and $N_{\mathrm{NB}}(V)$ in redshift space, respectively. The upper panel (lower panel) is the $M_B \leq -16.5$ ($M_B \leq -18.0$) case.

$N_{\mathrm{NB}}(V)$ were estimated by treating three samples differently, i.e., we treated $N_{\mathrm{NB}}(V)$ of each sample independently, and calculated the average value and the standard deviation. This standard deviation is shown as an error bar. From these panels, we found that clustering of the galaxies in νGC is stronger than that of the dark matter, because small or large volume tetrahedrons emerge as the clustering increases. This fact seems to harmonize with the spatial distributions in figure 1 and figure 3. Although it is not clear, the Delaunay graph succeeds in highlighting the homogeneousness of the galaxy distributions. We therefore conclude that this graph is useful to some degree for describing features of the spatial galaxy distributions in real space.

Figure 9 is the number counts of the edge length in the MST. In this figure, solid and short-dashed curves represent the number counts of the edge length in νGC, $N_{\nu\mathrm{GC}}(l)$, and N-body simulations, $N_{\mathrm{NB}}(l)$, respectively. Two different magnitudes, $M_B \leq -16.5$ and $M_B \leq -18.0$, are also represented. From these panels, we also infer that galaxy clustering is stronger than the dark matter. This is because $N_{\nu\mathrm{GC}}(l)$ is larger than $N_{\mathrm{NB}}(l)$ around in the small or large scale, and $N_{\mathrm{NB}}(l)$ is larger than $N_{\nu\mathrm{GC}}(l)$ in the middle scale. Although the difference is not clear, the MST that highlights the filamentary structure succeeds in describing the spatial galaxy distributions in real space.

We finally quantify the galaxy or dark-matter distributions by means of a constellation graph. Analysis of the constellation graph is not simple; we first construct an adjacency matrix of each constellation graph. In this procedure, the weighted constellation graphs that are assigned to the edge length l^j (j = 0–2.0) are used. The eigenvalues of the adjacency matrices are calculated, and the distribution functions of the eigenvalues are obtained. We finally estimate a mean absolute deviation $D(j)$ of the distribution function of the eigenvalue as a function of the weight index, j (Ueda & Itoh 1997).

Figure 10 is the relative mean absolute deviation $D/D_{\nu\mathrm{GC}}$ as a function of index j. Filled circles and squares represent the relative mean absolute deviations of the νGC galaxy and the dark matter, respectively. Since the constellation graph was constructed by connecting the nearest neighbor pair of galaxies, the edge length in the constellation graph is rather small. In order to decrease resolution effects in N-body simulation, we only analyzed the $M_B \leq -18.0$ case, because the average edge length of the constellation graph in $M_B \leq -18.0$ was larger than r_{c}, and because that was smaller than r_{c} in $M_B \leq -16.5$. Notice that the long edge was weighed for our analysis, which also decreased the resolution effects.

From the analysis of the constellation graph, we found that D_{NB} is larger than $D_{\nu\mathrm{GC}}$. This means that the galaxy clustering in νGC is stronger than the dark-matter clustering in N-body simulations, because the mean absolute deviation becomes smaller as the degree of clustering increases (Ueda & Itoh 1999). This result agrees with our intuitive perception, although it is expected from the Delaunay graph or MST analysis (see lower panels in figures 8 and 9). Since the constellation graph is constructed by the nearest-neighbor pair, the global structure does not reflect this analysis. The

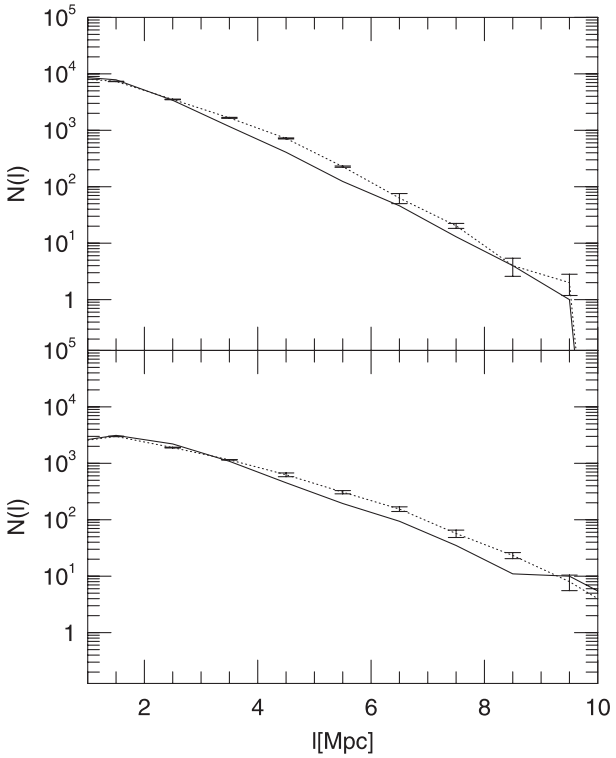

Fig. 12. Number counts of edge in the MST as a function of the edge length, l. Solid and short-dashed curves represent $N_{\nu\mathrm{GC}}(l)$ and $N_{\mathrm{NB}}(l)$ in redshift space. The upper panel (lower panel) is the $M_B \leq -16.5$ ($M_B \leq -18.0$) case.

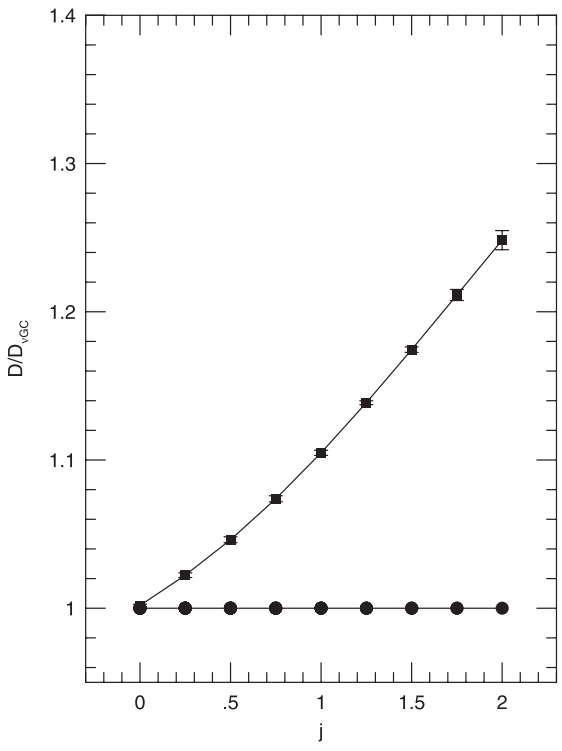

Fig. 13. Relative mean absolute deviations of galaxy and dark-matter distributions in redshift space with $M_B \leq -18.0$, as a function of index j. Filled circles and squares represent the relative mean absolute deviations of νGC and an N-body simulation, respectively.

constellation graph is, however, very useful for quantifying galaxy distributions, and is an important approach that serves as complementary to the other graph-theoretical approaches.

4.2. νGC Galaxy versus Dark Matter in Redshift Space

In this subsection, we compare the distribution of galaxies with that of dark-matter in redshift space. Although the galaxy distributions are different from the dark-matter distributions in real space (see subsection 4.1), we examine how this difference changes in redshift space. Figure 11 is the number counts of the volume of Delaunay tetrahedrons in redshift space. The symbols in this figure are the same as in figure 8. From these panels in figure 11, we find that the difference between $N_{\nu\mathrm{GC}}(V)$ and $N_{\mathrm{NB}}(V)$ becomes unclear in redshift space. As is well known, it is not easy to construct the Delaunay graph from given points (Okabe et al. 2000); we thus conclude that the volume analysis of the Delaunay tetrahedron is not attractive.

Figure 12 is the number counts of the edge length in the MST. The symbols in this figure are the same as in figure 9. From these panels, we find that the trends of the number count of the edge length in redshift space are also the same as in real space. Notice that the difference between the galaxy and dark matter in figure 12 is clearer than that in figure 11. We therefore conclude that MST is a more useful method for the description of the spatial galaxy distributions, in general.

If one can clearly see the difference between a galaxy and dark matter, we have to use the constellation graph. Figure 13 is the relative mean absolute deviations of galaxy and dark-matter distributions as a function of index j. The symbols in this figure are the same as in figure 10. Because of the effect of finger-of-god, the average edge length of the constellation graph in redshift space is larger than that in real space, and the analysis in the $M_B \leq -18.0$ case is still meaningful. From this analysis, we find that the difference between spatial distributions of galaxies and dark matter in redshift space is still clear. We also find that the analysis of the constellation graph still harmonizes with our intuitive perception. It is therefore concluded that the constellation graph is very useful for quantifying the nature of galaxy clustering.

4.3. νGC Galaxy, Dark Matter versus 2dFGRS

We now compare the galaxy distributions in the 2dFGRS with those in mock samples that are constructed from νGC or N-body simulations. Figure 14 is the number counts of the Delaunay tetrahedrons as a function of volume V. In this figure, the solid, short-dashed, and long-dashed curves represent the number counts of tetrahedrons that are constructed from 2dFGRS [$N_{\mathrm{2dF}}(V)$], νGC [$N_{\nu\mathrm{GC}}(V)$], and N-body simulations [$N_{\mathrm{NB}}(V)$], respectively. Two different regions are analyzed independently; the upper panel is the result in NGP and the lower panel is the result in SGP. From these panels, we find that the mock samples derived from νGC and dark matter do not reproduce the galaxy distributions in 2dFGRS; in a small volume, $N_{\nu\mathrm{GC}}(V)$ and $N_{\mathrm{NB}}(V)$ are considerably larger than $N_{\mathrm{2dF}}(V)$. Although this means that the clustering in the νGC galaxy and dark matter is stronger than the galaxy clustering in 2dFGRS, this fact does not harmonize with intuition

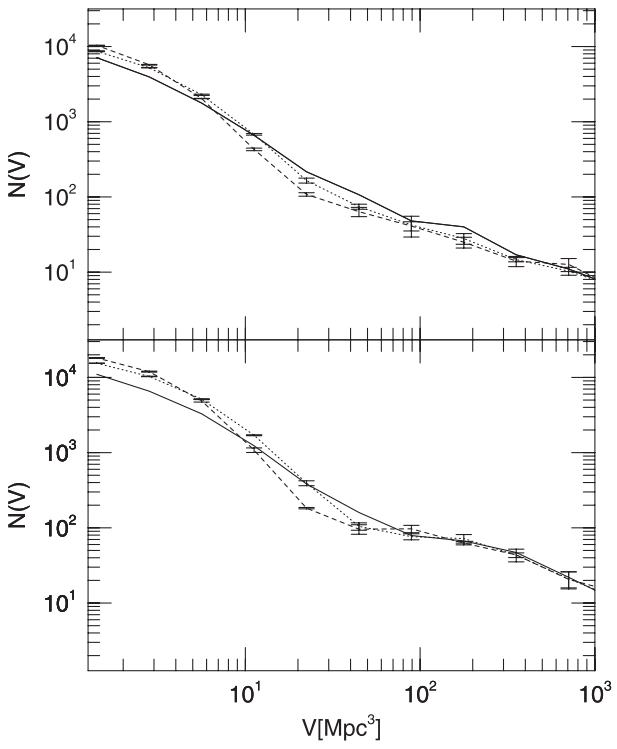

Fig. 14. Number counts of Delaunay tetrahedrons in a Delaunay graph as a function of volume V. Solid, short-dashed, and long-dashed curves represent $N_{\rm 2dF}(V)$, $N_{\nu\rm GC}(V)$, and $N_{\rm NB}(V)$, respectively. The upper panel (lower panel) is the result of NGP (SGP).

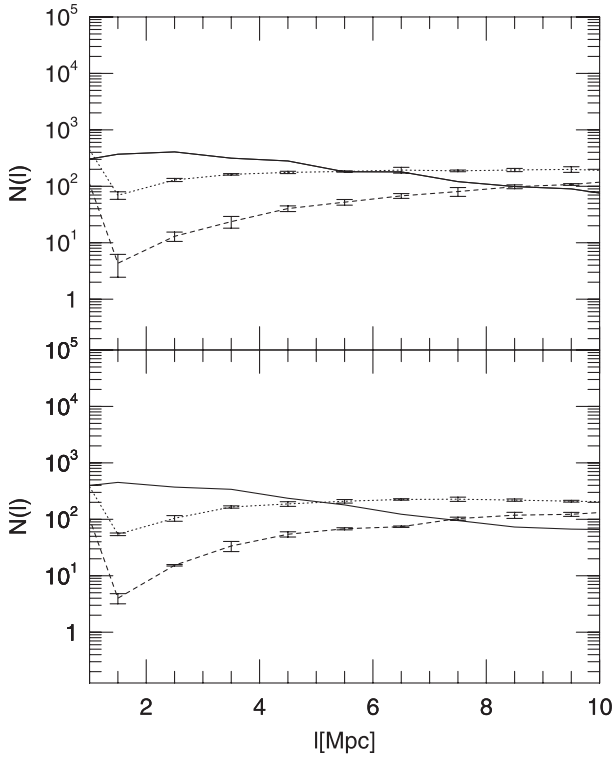

Fig. 15. Number counts of edges in the MST as a function of the edge length, l. Solid, short-dashed, and long-dashed curves represent $N_{\rm 2dF}(l)$, $N_{\nu\rm GC}(l)$, and $N_{\rm NB}(l)$, respectively. The upper panel (lower panel) is the result of NGP (SGP).

from human eyes. As we describe below, the galaxy distributions in 2dFGRS and mock samples are not the same, and that $N_{\nu\rm GC}(V)$ or $N_{\rm NB}(V)$ agrees with $N_{\rm 2dF}(V)$ in a large volume is not desirable. In addition, the difference between $N_{\nu\rm GC}(V)$ and $N_{\rm NB}(V)$ is unclear. After all, a volume analysis of the Delaunay tetrahedron becomes meaningless if the shape of a sample is far from a cube.

In order to understand the correct features of the galaxy spatial distribution, we performed a MST analysis. The results of the number counts of the edge length in the MST are given in figure 15. The symbols in this figure are the same as in figure 14. From these panels, we find that the number of small edges in 2dFGRS is larger than those in νGC and dark matter, and that the number of long edges is smaller. This means that the clustering is the strongest in 2dFGRS, and normal in νGC and weak in N-body simulations. We therefore confirm that the νGC considerably improves the theoretical prediction of the spatial galaxy distributions, although this do not sufficiently reproduce the 2dFGRS. We also conclude that the MST is useful if the shape of a sample is far from cubic.

For the sake of completeness, we finally quantify the galaxy distributions by means of a constellation graph. Figure 16 shows the relative mean absolute deviations of the galaxy in the 2dFGRS, νGC, and dark-matter distributions in N-body simulations as a function of index j. The filled circles, squares, and triangles represent the relative mean absolute deviations of the 2dFGRS, νGC, and N-body simulation, respectively. It is not surprising that the behaviors of $D_{\nu\rm GC}$ and $D_{\rm NB}$ are complex in NGP with a small j; this is due to the abundance of the very small edge. In the large-j region, or an analysis in SGP, we also see that the clustering in 2dFGRS is the strongest, and normal in νGC and weak in N-body simulations. In any case, we confirmed that the spatial galaxy distributions in νGC do not coincide with those in 2dFGRS.

5. Summary

In this paper, we have examined the galaxy distributions in νGC using the graph-theoretical approach. In order to understand the nature of νGC, the dark-matter distributions in N-body simulations and the galaxy distributions in 2dFGRS were also quantified. From our analysis, two important results were derived, and summarized as follows.

The most important result is that although νGC does not reproduce the observational galaxy distributions, this considerably improves the theoretical prediction of the spatial galaxy distributions. From graph-theoretical analyses, we confirm that the differences between νGC and 2dFGRS are always smaller than those between an N-body simulation and 2dFGRS. By using the high-resolution N-body simulations, νGC is found to be one of the most reliable models at the present stage. Comparisons of the galaxy spatial distributions between reliable theoretical modeling and high-quality observations are not fully understood. Our analysis clearly shows that νGC improves the theoretical prediction of the spatial galaxy distributions. It may therefore be considered that the analysis of semianalytic modeling is still important in examining the nature of galaxy formation.

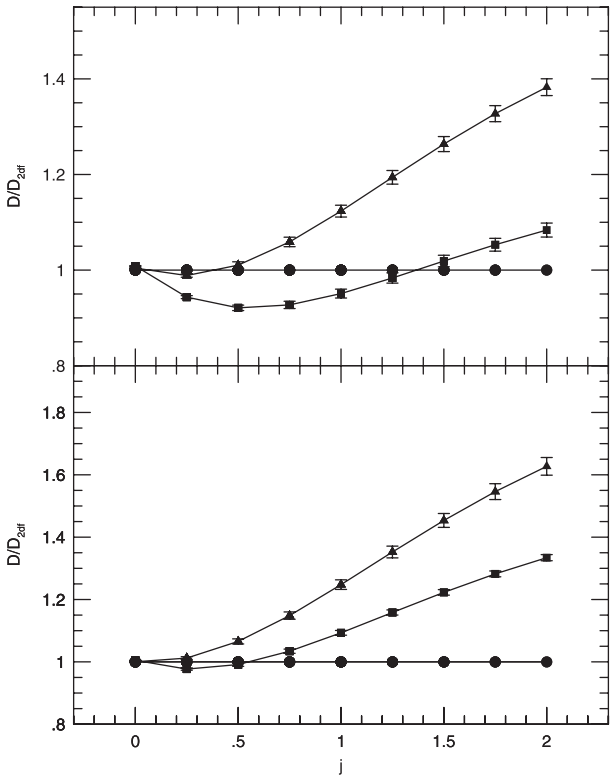

Fig. 16. Relative mean absolute deviations of the galaxy and dark matter, as a function of index j. Filled circles, squares, and triangles represent the relative mean absolute deviations of 2dFGRS, νGC, and N-body simulations, respectively. The upper panel (lower panel) is the result of NGP (SGP).

In addition to the above result, we prove the usefulness of the statistical measures used in the graph-theoretical approach. Three types of graphs are adopted for this purpose. Local clustering properties are described in the constellation graph. On the other hand, global galaxy distributions are examined in the Delaunay graph, and the MST highlights the filamentary structure. From our analysis, we conclude that the graph-theoretical approach is convenient for quantifying the galaxy distributions in an objective manner. In particular, the constellation graph is useful for this purpose. Because the MST contains information about the global structure, this is also an attractive graph. On the other hand, the Delaunay graph is less useful than the constellation graph and the MST. It is, therefore, important to improve the techniques concerning the constellation graph and the MST.

Finally, we come to the remaining problems of our analysis. Recent progress in computer capacity will allow researchers to perform more high-resolution simulations, and it will therefore be practicable to improve νGC. Comparing νGC with the hydrodynamical simulation to estimate the correctness of the semianalytic modeling is also interesting. We only examined the galaxy distributions in 2dFGRS, but more complete galaxy redshift surveys become available. It will therefore be attractive to compare the galaxy distributions in νGC with the next-generation surveys. We look forward to having the opportunity of improving our analysis, and to understanding the physics of galaxy formation and evolution clearly.

We would like to thank the 2dF team who supplied us with the galaxy catalog, and the Hydra consortium who supplied us with the software for the cosmological N-body simulation. We are also grateful to Koukichi Sugihara for providing codes for the Delaunay graph. The present analyses were carried out on computer systems in the Faculty of Education and Human Studies at Akita University, and in Yukawa Institute for Theoretical Physics in Kyoto University; νGC was carried out by using Fujitsu-made vector parallel processors VPP5000 installed at the Astronomical Data Analysis Center, National Astronomical Observatory of Japan (ADAC/NAOJ), under the ADAC/NAOJ large-scale simulation projects (group-ID: myy26a, yhy35b). This research was partially supported by a Grant-in-Aid for Scientific Research (C) (No. 18500648, 2006–2008) from the Ministry of Education, Culture, Sports, Science and Technology.

References

Adami, C., & Mazure, A. 1999, A&AS, 134, 393
Aho, A. V., Ullman, J. D., & Hopcroft, J. E. 1983, Data Structures and Algorithms (Reading, Massachusetts: Addison-Wesley)
Barrow, J. D., Bhavsar, S. P., & Sonoda, D. H. 1985, MNRAS, 216, 17
Bhavsar, S. P., & Nigel Ling, E. 1988, ApJ, 331, L63
Bhavsar, S. P., & Splinter, R. J. 1996, MNRAS, 282, 1461
Chung, T. J. 2002, Computational Fluid Dynamics (Cambridge: Cambridge University Press)
Colberg, J. M. 2007, MNRAS, 375, 337
Coles, P. 1992, Statistical Challenges in Modern Astronomy I, ed. E. D. Feigelson & G. J. Babu (New York: Springer-Verlag) 57
Colless, M., et al. 2001, MNRAS, 328, 1039
Couchman, H. M. P., Thomas, P. A., & Pearce, F. R. 1995, ApJ, 452, 797
de Berg, M., Cheong, O., van Kreveld, M., & Overmars, M. 1997, Computational Geometry, Algorithms and Applications (Berlin: Springer-Verlag)
Delone, B. N. 1934, Bull. Acad. Science USSR (VII) Classe Sci. Mat., 793
Demiański, M., & Doroshkevich, A. G. 2004, A&A, 422, 423
Diaferio, A., Kauffmann, G., Balogh, M. L., White, S. D. M., Schade, D., & Ellingson, E. 2001, MNRAS, 323, 999
Diaferio, A., Kauffmann, G., Colberg, J. M., & White, S. D. M. 1999, MNRAS, 307, 537
Doroshkevich, A. G., Müller, V., Retzlaff, J., & Turchaninov, V. 1999, MNRAS, 306, 575
Doroshkevich, A. G., Tucker, D. L., Fong, R., Turchaninov, V., & Lin, H. 2001, MNRAS, 322, 369
Eiselt, H. A., & Sandblom, C.-L. 2000, Integer Programming and Network Models (Berlin: Springer-Verlag)
Forsyth D. A., & Ponce, J. 2002, Computer Vision: A Modern Approach (Upper Saddle River, New York: Prentice Hall)
Gott, J. R., III., Melott, A. L., & Dickinson, M. 1986, ApJ, 306, 341
Graham, M. J., Clowes, R. G., & Campusano, L. E. 1995, MNRAS, 275, 790

Gurzadyan, V. G., & Kocharyan, A. A. 1994, Paradigms of the Large-Scale Universe (Lausanne: Gordon and Breach)
Harary, F. 1969, Graph Theory (Massachusetts: Addison-Wesley Publishing Company)
Hatton, S., Devriendt, J. E. G., Ninin, S., Bouchet, F. R., Guiderdoni, B., & Vibert, D. 2003, MNRAS, 343, 75
Helly, J. C., Cole, S., Frenk, C. S., Baugh, C. M., Benson, A., & Lacey, C. 2003a, MNRAS, 338, 903
Helly, J. C., Cole, S., Frenk, C. S., Baugh, C. M., Benson, A., Lacey, C., & Pearce, F. R. 2003b, MNRAS, 338, 913
Hikage, C., et al. 2002, PASJ, 54, 707
Hikage, C., Taruya, A., & Suto, Y. 2001, ApJ, 556, 641
Huchra, J., Davis, M., Latham, D. W., & Tonry, J. 1983, ApJS, 52, 89
Icke, V., & van de Weygaert, R. 1987, A&A 184, 16
Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999a, MNRAS, 303, 188
Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999b, MNRAS, 307, 529
Kerscher, M., Szapudi, I., & Szalay, A. S. 2000, ApJ, 535, L13
Krzewina, L. G., & Saslaw, W. C. 1996, MNRAS, 278, 869
Livesley, R. K. 1983, Finite Elements: An Introduction for Engineers (Cambridge: Cambridge University Press)
Maddox, S. J., Efstathiou, G., Sutherland, W. J., & Loveday, J. 1990b, MNRAS, 242, 43P
Maddox, S. J., Sutherland, W. J., Efstathiou, G., & Loveday, J. 1990a, MNRAS, 243, 692
Martínez, V. J., Jones, B. J. T., Domínguez-Tenreiro, R., & van de Weygaert, R. 1990, ApJ, 357, 50
Martínez, V. J., & Saar, E. 2002, Statistics of the Galaxy Distribution (Boca Raton: Chapman & Hall/CRC)
Matsubara, T., & Suto, Y. 1996, ApJ, 460, 51
Mecke, K. R., Buchert, T., & Wagner, H. 1994, A&A, 288, 697
Mecke, K. R., & Wagner, H. 1991, J. Stat. Phys., 64, 843
Melott, A. L. 1990, Phys. Rep., 193, 1
Monaghan, J. J. 1992, ARA&A, 30, 543
Nagashima, M., Yahagi, H., Enoki, M., Yoshii, Y., & Gouda, N. 2005, ApJ, 634, 26
Nagashima, M., & Yoshii, Y. 2004, ApJ, 610, 23
Norberg, P., et al. 2002, MNRAS, 336, 907
Okabe, A., Boots, B., Sugihara, K., & Chiu, S. N. 2000, Spatial Tessellations — Concepts and Applications of Voronoi Diagrams, Second Ed. (Chichester: John Wiley and Sons)
Peacock, J. A. 1999, Cosmological Physics (Princeton: Cambridge University Press)
Pearson, R. C., & Coles, P. 1995, MNRAS, 272, 231
Peebles, P. J. E. 1980, The Large Scale Structure of the Universe (Princeton: Princeton University Press)
Perlmutter, S., et al. 1999, ApJ, 517, 565
Preparata, F. P., & Shamos, M. I. 1985, Computational Geometry — An Introduction (New York: Springer-Verlag)
Ramella, M., Boschin, W., Fadda, D., & Nonino, M. 2001, A&A, 368, 776
Riess, A. G., et al. 1998, AJ, 116, 1009
Roukema, B. F., Peterson, B. A., Quinn, P. J., & Rocca-Volmerange, B. 1997, MNRAS, 292, 835
Sahni, V., Sathyaprakash B. S., & Shandarin, S. F. 1998, ApJ 495, L5
Saslaw, W. C. 2000, The Distribution of the Galaxies — Gravitational Clustering in Cosmology (Cambridge: Cambridge University Press)
Schaap, W. E., & van de Weygaert, R. 2000, A&A 363, L29
Schechter, P. 1976, ApJ, 203, 297
Schmalzing, J., Kerscher, M., & Buchert, T. 1996, Dark Matter in the Universe, ed. S. Bonometto, J. R. Primack, & A. Provenzale (Amsterdam: IOS Press), 281
Seul, M., O'Gorman, L., & Sammon, M. J. 2000, Practical Algorithms for Image Analysis: Description, Examples, and Code (Cambridge: Cambridge University Press)
Somerville, R. S., Lemson, G., Sigad, Y., Dekel, A., Kauffmann, G., & White, S. D. M. 2001, MNRAS, 320, 289
Späth, H. 1995, Two Dimensional Spline Interpolation Algorithms (Wellesley: A.K. Peters, Ltd.)
Spergel, D. N., et al. 2003, ApJS, 148, 175
Suto, Y., & Matsubara, T. 1994, ApJ, 420, 504
Ueda, H., & Itoh, M. 1997, PASJ, 49, 131
Ueda, H., & Itoh, M. 1999, ApJ, 526, 560
Ueda, H., & Takeuchi, T. T. 2006, PASJ, 58, 283
Ueda, H., Takeuchi, T. T., & Itoh, M. 2001, PASJ, 53, 381
Ueda, H., Takeuchi, T. T., & Itoh, M. 2003, A&A, 399, 1
van de Weygaert, R. 1994, A&A, 283, 361
van de Weygaert, R., & Icke, V. 1989, A&A, 213, 1
Voronoi, G. 1908, J. reine angew. Math., 134, 198
Weinberg, D. H., Gott, J. R., III., & Melott, A. L. 1987 ApJ, 321, 2
Wise, S. 2002, GIS Basics (New York: CRC Press)
Yahagi, H. 2005, PASJ, 57, 779
Yahagi, H., & Yoshii, Y. 2001, ApJ, 558, 463
Zaninetti, L. 1991, A&A, 246, 291
Zwicky, F., Herzog, E., Wild, P., Karpowicz, M., & Kowal, C. T. 1961–68, Catalogue of Galaxies and of Clusters of Galaxies (Pasadena: California Institute of Technology)