On the Fractal Distribution of HII Regions in Disk Galaxies

Néstor Sánchez and Emilio J. Alfaro

Abstract In this work we quantify the degree to which star-forming events are clumped. We apply a precise and accurate technique to calculate the correlation dimension D_c of the distribution of HII regions in a sample of disk galaxies. Our reliable results are distributed in the range $1.5 \lesssim D_c \lesssim 2.0$. We get significant variations in the fractal dimension among galaxies, contrary to a universal picture sometimes claimed in literature. The faintest galaxies tend to distribute their HII regions in more clustered (less uniform) patterns. Moreover, the fractal dimension for the brightest HII regions within the same galaxy seems to be smaller than for the faintest ones suggesting some kind of evolutionary effect.

1 Introduction

There is clear evidence that gas observed in external galaxies (both irregulars and spirals) follows hierarchical and self-similar patterns [20, 12, 21, 3, 5]. This fractal picture is consistent with the distribution of star fields and star-forming sites on galaxy-wide scales [7, 15, 6, 4, 13, 2]. However, it is not clear whether this kind of fractal distributions are connected/related or not to some properties of the host galaxies, such as radius, rotation, brightness, morphology, etc. In spite of the great variety of fractal dimension values reported in the literature for different galaxies (for both the gas and the distribution of star forming sites) most of the authors argue in favor of a more or less universal picture. In this universal description, the constancy of the fractal dimension is a natural consequence of the fact that the same physical processes are structuring these systems. However, there are some indications that the situation could be more complicated. The fractal dimension of the distribution of HII regions could be different in grand design and flocculent galaxies [11], and the brightest galaxies could have fractal dimensions higher than faintest

Néstor Sánchez & Emilio J. Alfaro
Instituto de Astrofísica de Andalucía, CSIC, Granada, Spain. e-mail: nestor@iaa.es, emilio@iaa.es

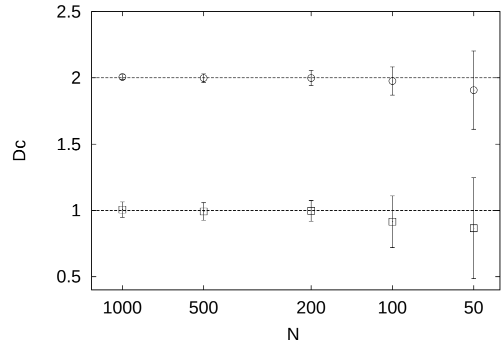

Fig. 1 Calculated dimension D_c as a function of the sample size N for fractals with expected dimensions $D_c = 2$ (open circles) and $D_c = 1$ (open squares). Each point is the average of 50 random realizations and the bars show the standard deviations.

ones [15, 13]. On the contrary, [7] do not find any correlation between the fractal dimension and the galactic properties, but their uncertainties are so large that the robustness of this conclusion is questionable.

Part of the problem that prevents achieving unequivocal conclusions lies in the great diversity of analysis techniques used in the literature and/or the application to not large enough samples of galaxies. Here we use a carefully designed method that has been tested previously on simulated data and that clearly establishs its accuracy and applicability depending on the sample itself. We apply this method to the most complete sample of galaxies that we have found in literature expecting to draw significant conclusions regarding this matter.

2 Fractal dimension calculation

We use the correlation dimension [10] which gives robust results when dealing with distributions of points in space. First we have to calculate correlation integral $C(r)$ which represents the probability of finding a point within a sphere of radius r centered on another point. For a fractal set $C(r)$ scales at small r as

$$C(r) \sim r^{D_c} \quad , \tag{1}$$

being D_c the correlation dimension. When evaluating $C(r)$ for real data this power-law behavior is valid only within a limited range of r values. For relatively small or high r values $C(r)$ deviates from the expected behavior and D_c tends to be underestimated. We have developed an algorithm (see details in [18] and [19]) to calculate D_c which implements objective and suitable criteria to avoid both boundary effects and finite-data problems at small scales. Moreover, the algorithm estimates an uncertainty associated to D_c by using bootstrap techniques (σ_{boot}).

An important point is the dependence of the measured dimension on the sample size because many times the number of available data points is rather small. Figure 1 shows what happens when the calculation is done on the *same set of simulated*

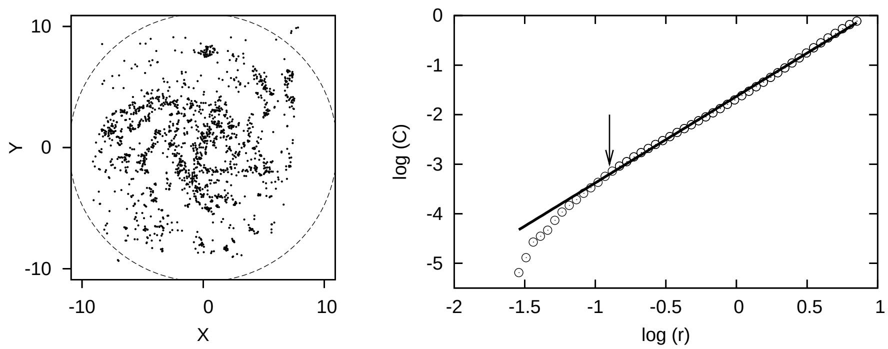

Fig. 2 Left panel: positions (units are in kpc) of the HII regions in NGC 6946. The dashed line circle indicates the radius R_{25}. Right panel: the calculated correlation integral $C(r)$ for NGC 6946. The lower limit of the best linear fit (solid line) is indicated by the vertical arrow.

fractals but randomly removing points to reach smaller sample sizes N. Since we are dealing exactly with the same original data, the observed decreasing in the average D_c value with decreasing N have to be attributed exclusively to sample size itself. Thus, our algorithm is able to estimate the fractal dimension in a reliable way when applied to random subsamples of fractal distributions of points, but if the sample size is too small ($N \lesssim 100$) a bias tending to underestimate the dimension is produced. The existence of this bias is related to the random variations in the simulated (or observed) data set and not related to the uncertainties in the determination of D_c (σ_{boot}). The latter ones are always smaller than the error bars in Figure 1.

3 Results

We have used VizieR and ADS databases, in conjunction with the papers [8] and [9], to search for catalogs of external galaxies containing positions of HII regions available in machine readable format. We found a total of 93 spiral galaxies with positions for at least 50 HII regions. We have also included data for 8 irregular galaxies with HII positions provided by D. Hunter [17]. The properties of the selected galaxies are listed in Table 2 of [19] available through VizieR at CDS. This table contains the galaxy name, Hubble type, position ϕ and inclination i angles, morphological de Vaucouleurs type T, arm class A, distance D, B-band absolute magnitude M_B, radius of the isophotal 25 mag/arcsec^2 R_{25}, and the maximum rotation velocity V_{rot}. The deprojected positions of the HII regions for the sample of 101 disk galaxies, which can be seen in Figure 3 of [19], are also available through VizieR.

Figure 2 shows the result for one example galaxy. The left side panel in this figure shows the distribution of HII regions in NGC 6946. The right side panel shows the

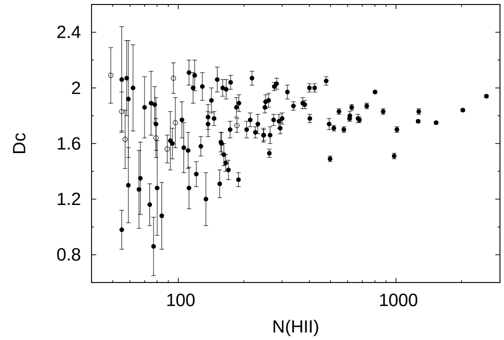

Fig. 3 Calculated dimension D_c as a function of the number of available data $N(\mathrm{HII})$ for spiral galaxies (solid circles) and irregular galaxies (open circles) in the sample. The error bars indicate the uncertainties obtained from bootstrapping.

corresponding correlation integral $C(r)$. The vertical arrow indicates the value of r below which $C(r)$ is poorly estimated [18, 19]. The rest of the data exhibits a characteristic fractal behavior. The slope of the best linear fit gives D_c, in this case $D_c = 1.75$.

The result of calculating the correlation dimension of the distribution of HII regions in the full sample of galaxies is shown in Table 3 of [19] (also available via VizieR). Figure 3 shows the calculated correlation dimension for all the galaxies in the sample as a function of the number of available data points. We see that as $N(\mathrm{HII})$ decreases, the uncertainties increase and the obtained dimensions are substantially more spread out toward lower values. This trend is very similar to the behavior shown in Figure 1. It is clear that at least part of this trend is due to a bias in the estimated value of D_c for galaxies with small number of HII regions. To overcome this bias, we focus our analysis on galaxies having more than 200 HII regions (46 spiral galaxies). The average fractal dimension in this case is $\langle D_c \rangle = 1.81$ with a standard deviations of $\sigma = 0.14$, and the average σ_{boot} is ~ 0.03. By considering the associated uncertainties we can say that the D_c values do differ among themselves.

3.1 Dependence on galactic properties

To examine possible correlations we propose, in a first approximation, a linear model linking all variables and we evaluate its "goodness" via the Akaike's Information Criterion [1]. This criterion not only rewards goodness of fit, but also includes a penalty that is an increasing function of the number of free parameters to be estimated (discouraging overfitting). Then, starting from a linear regression model based on n variables we apply a stepwise regression process to select the subset of variables providing the best model. We begin with the variables T, A, M_B, R_{25}, V_{rot}, and another variable that we have called the average surface density of star forming regions ($N(\mathrm{HII})/R_{25}^2$). The application of this procedure using the **R** environment for statistical computing [16] yields that M_B and R_{25} represent the best

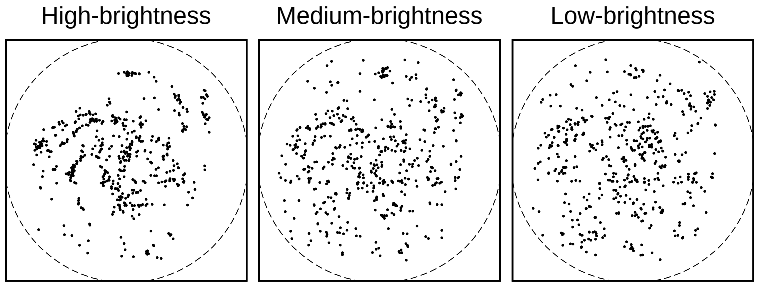

Fig. 4 Positions of the HII regions in NGC 6946 for high, medium, and low brightness regions, for which the resulting fractal dimensions were $D_c = 1.64$, $D_c = 1.82$, and $D_c = 1.79$, respectively.

linear model fitting the data:

$$D_c = -(0.069 \pm 0.025) M_B - (0.006 \pm 0.003) R_{25} + 0.456 \quad . \tag{2}$$

The overall fit is significant at the confidence level of 95% and the regression coefficients of the variables M_B and R_{25} are significant at the confidence levels of 99% and 95%, respectively.

According to our results D_c correlates with M_B and, to a lesser extent, with R_{25}. Brightest galaxies have fractal dimensions higher than faintest ones, while [7] find a pure scatter diagram when comparing D_c and M_B which, together with the lack of correlation with other properties is used as argument favoring a universal D_c value in spiral galaxies. In contrast, [15] and [13] report the same trend in their samples of dwarf galaxies, i.e. the highest fractal dimensions for the brightest galaxies.

3.2 Possible evolutionary effects

We have analyzed the dependence of D_c on the brightness of the HII regions. We first selected galaxies from the sample for which we have data on HII region brightness and a sample size of $N(\mathrm{HII}) > 600$. There are 9 galaxies in our sample fulfilling these requiriments. We divided them in three equal subsamples ordered in descending brightness: the brightest HII regions, the medium bright ones, and the faintest ones. Figure 4 shows the resulting distributions in the galaxy NGC 6946. It can be seen that the obtained D_c values are consistent with the appearance of the HII region distributions. The high brightness map has a relatively small D_c value ($\simeq 1.6$) that corresponds to a more irregular distribution having clumps and/or filaments separated by low density (or empty) regions. The medium and low brightness maps have higher D_c values ($\simeq 1.8$) corresponding to more homogeneous distributions. We recalculated D_c for each case and found that for 7 of the 9 galaxies D_c is smaller for the brightest HII regions than for the rest of the data. The brightest regions should reflect, in a first approximation, the initial distribution of star-forming regions in each galaxy. It seems likely from our results that some kind of evolutionary effect tends to randomize (homogenize) in some degree the initial distributions of HII regions.

4 Final remarks

There are many other galactic properties from which D_c could depend, such as star formation activity, mass, age, metallicity, or a combination of them. It has been suggested that the fractal dimension of the distribution of star-forming sites could be increased during the star formation process [4]. Recently, a transition scale from a lower to a higher correlation dimension has been suggested as a possible explanation for the wide range of observed D_c values [14]. Obviously, a more complete analysis including more galactic variables and a wider and more diverse sample of galaxies would be necessary in order to obtain a clearer picture.

Acknowledgements We acknowledge financial support from MEC of Spain through grant AYA2007-64052 and from Consejería de Educación y Ciencia (Junta de Andalucía) through TIC-101.

References

1. Akaike, H. 1974, IEEE Transactions on Automatic Control, 19, 716
2. Bastian, N., Ercolano, B., Gieles, M., Rosolowsky, E., Scheepmaker, R. A., Gutermuth, R., & Efremov, Y. 2007, MNRAS, 379, 1302
3. Begun, A., Chengalur, J. N., & Bhardwaj, S. 2006, MNRAS, 372, L33
4. de la Fuente Marcos, R., & de la Fuente Marcos, C. 2006, MNRAS, 372, 279
5. Dutta, P., Begum, A., Bharadwaj, S., & Chengalur, J. N. 2008, MNRAS, 384, L34
6. Elmegreen, B. G., Elmegreen, D. M., Chandar, R., Whitmore, B., & Regan, M. 2006, ApJ, 644, 879
7. Feitzinger, J. V., & Galinski, T. 1987, A&A, 179, 249
8. Garcia-Gomez, C., & Athanassoula, E. 1991, A&AS, 89, 159
9. Garcia-Gomez, C., Athanassoula, E., & Barbera, C. 2002, A&A, 389, 68
10. Grassberger, P., & Procaccia, I. 1983, Phys. Rev. Lett., 50, 346
11. Hodge, P. 1985, PASP, 97, 688
12. Kim, S., Staveley-Smith, L., Dopita, M. A., Sault, R. J., Freeman, K. C., Lee, Y., & Chu, Y.-H. 2003, ApJS, 148, 473
13. Odekon, M. C. 2006, AJ, 132, 1834
14. Odekon, M. C. 2008, ApJ, 681, 1248
15. Parodi, B. R., & Binggeli, B. 2003, A&A, 398, 501
16. R Development Core Team 2008, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org, ISBN 3-900051-07-0
17. Roye, E. W., & Hunter, D. A. 2000, AJ, 119, 1145
18. Sánchez, N., Alfaro, E. J., Elias, F., Delgado, A. J., & Cabrera-Caño, J. 2007, ApJ, 667, 213
19. Sánchez, N. & Alfaro, E. J. 2008, ApJS, 178, 1
20. Westpfahl, D. J., Coleman, P. H., Alexander, J., & Tongue, T. 1999, AJ, 117, 868
21. Willett, K. W., Elmegreen, B. G., & Hunter, D. A. 2005, AJ, 129, 2186