Improving galaxy group finders

Facundo Rodriguez

Instituto de Astronomía Teórica y Experimental (UNC-CONICET), Observatorio Astronómico de Córdoba
Laprida 854, Córdoba, X500BGR, Argentina
facundo.rodriguez@unc.edu.ar

Abstract

The identification of galaxy systems is a challenging task and, therefore, it is necessary to improve these techniques as much as possible to achieve a better understanding of the structure of the universe on a large scale and link the properties of galaxies with those of the inhabiting halos. Given the large present and forthcoming galaxy catalogues, we propose, implement and evaluate an algorithm that combines the two most popular techniques to identify galaxy systems that can be easily applied to any spectroscopic galaxy catalogue. Assuming that a group of galaxies is a gravitationally bounded system that has at least one bright galaxy, we begin by identifying groups with a Friend-Of-Friend algorithm adapted to fit this definition. Knowing the shortcomings of this method, we improve it by adding a halos-based procedure. To assess the performance of the algorithm, we construct a mock catalogue from a semi-analytical model to compare the groups identified using our method with those obtained from the simulation.

Introduction

We focus our attention on improving the redshift space identification of groups obtained by the two most used techniques nowadays: FOF and Halo-based.
The FOF method proposed by Huchra & Geller(1982) [3] is based on spatial criteria for assigning galaxies to groups. Beyond the demonstrated virtues of this algorithm and flexibility to adapt to different scientific objectives, one of the negative aspects that can be extracted from these analyses is that the poor groups obtained with this algorithm have low reliability and those which are numerous are likely to have many interlopers. Consequently, these catalogues are not suitable for certain studies, e.g., HOD analysis.
The halo-based group finder developed by Yang et al. (2005, 2007) [8, 9], has gained in importance in recent years because it has the virtue that finds groups with few members. In this method the luminosity of the galaxies is related to the mass of the dark matter halos, assigning masses using the abundance matching technique on luminosity. The group catalogues from these identifications allow the study of the HOD in a wide range of masses and luminosities. However, It starts with a FOF algorithm with an overdensity that is not representative of the galaxy groups and an imposed mass-luminosity ratio, that can bias the characteristics of the resulting groups.
In this work, we propose a method to identify galaxies groups that combine the two methods described above: a proper FOF identification is performed in the first step and, using the resulting groups, in the second step we implemented an algorithm that follows the halo-based method. Understanding each step as an improvement of the previous one, we asses the performance of this process applying this algorithm in a mock catalogue. This procedure allows us to learn about the reliability of the obtained groups.

Proposed group finder

Part I: FOF implementation

We began our group identification using the conventional FOF algorithm implemented by Merchán & Zandivarez (2002, 2005) [4, 5]. This method links galaxies that satisfy certain conditions in the projected distance and the line-of-sight velocity for a given definition of dark matter halos overdensity. Throughout this work, we will use an overdensity of $\delta\rho/\rho = 200$. It is worth to mention that this procedure takes into account the number density variation due to the apparent magnitude limit of the survey to evolve the transverse and radial linking lengths.
After the implementation of this algorithm, we have a sample of galaxy groups with two or more members. But, as it is well known, the groups determined following a FOF algorithm are more reliable as more galaxies they have. Thus, to improve the reliability of poor groups, we restrict the luminosity of group members: they have to have at least one bright galaxy $M_r < M_{gr_{lim}} = -19.5$ (absolute magnitude in r-band). With this in mind, we reject all groups that do not meet this criterion and, additionally, we add all bright galaxies as potential groups. With these restrictions, a group of galaxies is a system of galaxies gravitationally bounded that has at least one bright galaxy. Thus, our group definition is not limited to simple geometric considerations, also takes into account the galaxies astrophysical properties that comprise it.

Part II: Halo-based group finder

This part of the method is performed following the halo-based group finder proposed by Yang et al. (2005, 2007) [8, 9], which is iterative and based on an adaptive filter modelled after the general properties of dark matter halos. Since this method requires a catalogue of potential groups and our intention is to improve the reliability of the systems previously identified, we use the resultant galaxy groups of Part I and, based on its properties, implement a similar iterative procedure.

1. *Estimation of the characteristic luminosity.*
2. *Properties of the dark matter halo derived from its luminosity.*
3. *Membership Assignment.*
4. *Iteration.*

To start the iterative process, we take the membership of FOF groups described in Part I to estimate the first luminosities. In the following loops, we will use the members determined using the procedure described in Part II. As a result of this process, we obtain two group samples. The first is obtained from Part I, which consists of a suitable FOF algorithm and adapted to our group definition, hereinafter FOF modified. And the second that results from the whole process, which introduces a halo-based technique described in Part II, called the Final Sample.

Algorithm Performance

To test the effectiveness of the proposed group finder, we use a mock catalogue that reproduces the features of the SDSS DR12 spectroscopic catalogue [1] built up from a semi-analytical model. This allows us to exactly know how galaxies populate dark matter halos and, with this information, we can evaluate the assignment made by our identifier and estimate the errors that will be committed in the real catalogues.

Mock catalogue

Our mock catalogue is built using the semi-analytical galaxy formation model developed by Guo et al. (2010)[2] applied on the Millennium Simulation [7], which offers high spatial and time resolution within a large cosmological volume. This simulation evolves more than 10 billion dark matter particles in a 500 h^{-1} Mpc periodic box, using a comoving softening length of 5 h^{-1} kpc. The procedure followed to build the mock is the same as described in detail in Rodriguez et al. (2015) [6].

Purity and completeness

A traditional way to assess the quality of group finding algorithms is through the purity and completeness concepts. Groups purity, P, quantifies if galaxies that are identified as members of given group reside actually in the same dark matter halo:

$$P = \frac{n_{gh}}{N_g} \qquad (1)$$

where N_g is the total number of galaxies of the identified group and n_{gh} is the number of galaxies in the identified group g that reside in the halo h. With this definition, the value of $1-P$ indicates the percentage of interlopers. If the group shares galaxies with more than one halo, we will choose the one with the highest number n_{gh}.
The next quantity to be defined is the halo completeness, C. Conceptually, C quantifies the fraction of visible galaxies in a dark matter halo that is included in some identified galaxy group:

$$C = \frac{n_{hg}}{N_h} \qquad (2)$$

where N_h is the total number of visible galaxies in the halo h (i.e. the number of survivor galaxies of the halo once the apparent magnitude cut off has been applied) and n_{hg} is the number of visible galaxies in the halo h that were also identified as members of the group g. As in the case of purity, if the halo shares galaxies with more than one identified group, we will choose the one with the largest n_{hg}.

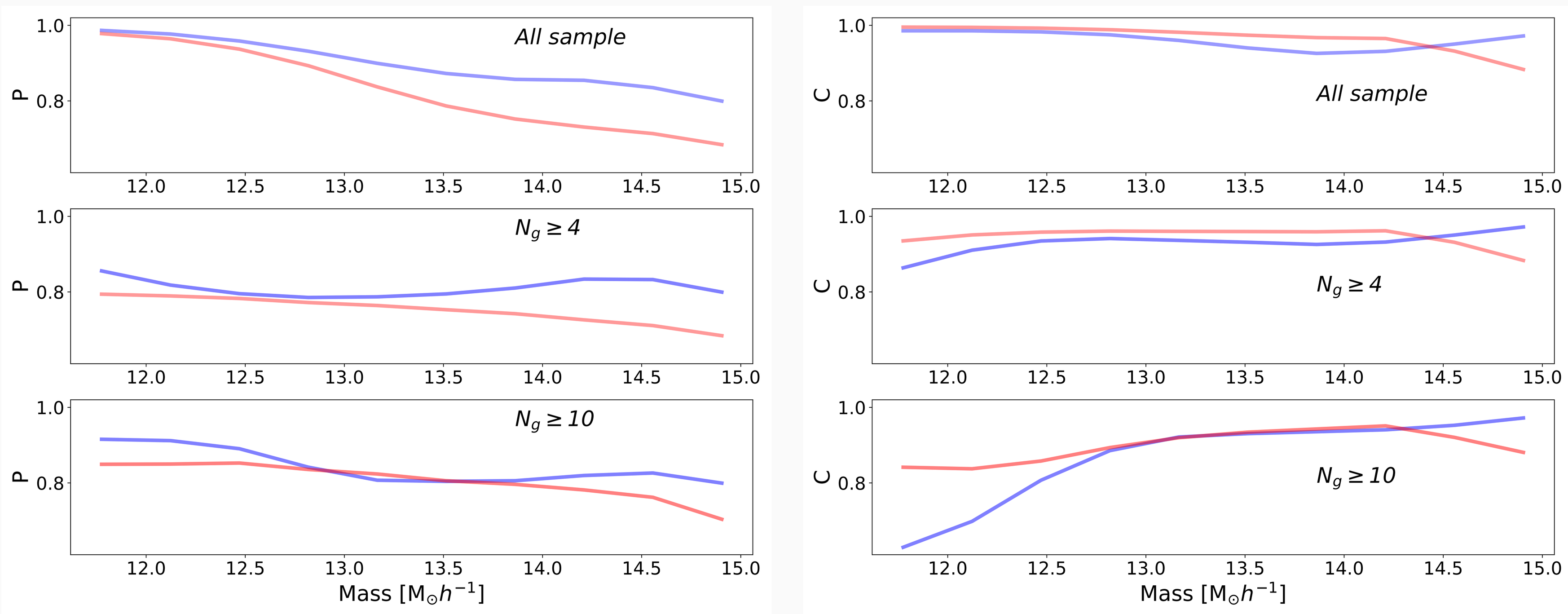

Figure 1: *Group purity and halo completeness as a function of halo mass for FOF modified group catalogue (red) and Final Sample (blue). Left panels show the purity results for all sample, groups with 4 or more members and groups with 10 or more (on top, middle and bottom panel, respectively) while rights panels display those for halo completeness.*

Halo Occupation Distribution (HOD)

One of the main reasons for spending time and effort on developing an identifier capable of including groups of galaxies that traditional FOF algorithms cannot reliably detect ($N_g < 4$) is studying the HOD. Now with the two samples of groups that we obtained we can study the HOD in a wide range of masses and, also, compare the results with the mock halos.

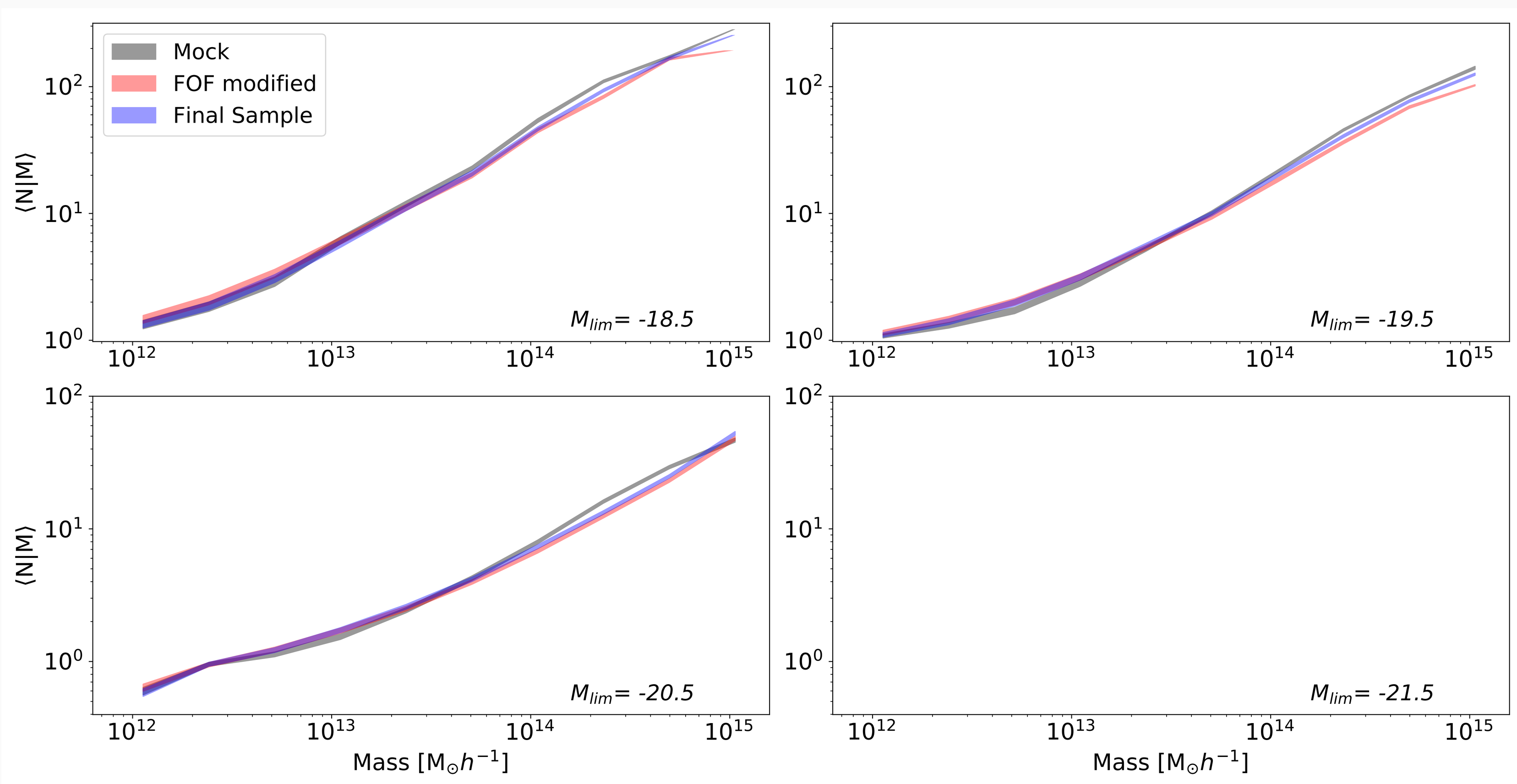

Figure 2: *HOD for halos in the mock (grey), FOF modified (red) and Final Sample (blue) for four absolute magnitude threshold (M_r =-18.5,-19.5,-20.5 and -21.5) as each panel indicate.*

Conclusions

In this work, we proposed a method of identifying galaxy systems, assuming a definition of galaxy groups and combining the two most popular methods: FOF and Halo-based.
We asses our algorithm using a mock catalogue of galaxies obtaining excellent results in terms of purity and completeness. Given that our method brings two output catalogues (FOF modified and Final Sample) we were able to compare the results corresponding to each of them. It is worth to mention the good performance of the FOF modified algorithm, from which high levels of reliability were obtained with an average purity greater than 0.7 together with a completeness greater than 0.9. Even so, the entire process succeeds to improve purity at the price of a small loss of completeness.
To conclude with the evaluation of the method, we compare the HODs obtained from our catalogues with those corresponding to the mock, succeeding to reproduce their behaviour in a wide range of masses and magnitudes for both catalogues. As an additional result, it is observed that HOD does not depend strongly on the identification method. This result suggests that a suitable group definition and a reliable mass estimation is more important than the singularities of each of the parts of our algorithm.
The next step is to apply this procedure to the SDSS DR12 to extract the galaxy groups and compare our result with that of other authors.

References

[1] Alam, S., Albareti, F. D., Allende Prieto, C., Anders, F., Anderson, S. F., Anderton, T., Andrews, B. H., and Armengaud, E. The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. *219*, 1 (Jul 2015), 12.

[2] Guo, Q., White, S., Li, C., and Boylan-Kolchin, M. How do galaxies populate dark matter haloes? *404*, 3 (May 2010), 1111–1120.

[3] Huchra, J. P., and Geller, M. J. Groups of Galaxies. I. Nearby groups. *257* (Jun 1982), 423–437.

[4] Merchán, M., and Zandivarez, A. Galaxy groups in the 2dF Galaxy Redshift Survey: the catalogue. *335* (Sept. 2002), 216–222.

[5] Merchán, M. E., and Zandivarez, A. Galaxy Groups in the Third Data Release of the Sloan Digital Sky Survey. *630* (Sept. 2005), 759–763.

[6] Rodriguez, F., Merchán, M., and Sgró, M. A. Taking advantage of photometric galaxy catalogues to determine the halo occupation distribution. *580* (Aug 2015), A86.

[7] Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., and Pearce, F. Simulations of the formation, evolution and clustering of galaxies and quasars. *435* (June 2005), 629–636.

[8] Yang, X., Mo, H. J., van den Bosch, F. C., and Jing, Y. P. A halo-based galaxy group finder: calibration and application to the 2dFGRS. *356* (Feb. 2005), 1293–1307.

[9] Yang, X., Mo, H. J., van den Bosch, F. C., Pasquali, A., Li, C., and Barden, M. Galaxy Groups in the SDSS DR4. I. The Catalog and Basic Properties. *671* (Dec. 2007), 153–170.