Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **458,** 226–232 (2016)
Advance Access publication 2016 January 29

doi:10.1093/mnras/stw215

The MeSsI (merging systems identification) algorithm and catalogue

Martín de los Rios,[1⋆] Mariano J. Domínguez R.,[1⋆] Dante Paz[1] and Manuel Merchán[1,2,3]

[1]*Instituto de Astronomía Teórica y Experimental (CCT Córdoba, CON ICET, UNC), Laprida 854, X5000BGR, Córdoba, Argentina*
[2]*Observatorio Astronómico de Córdoba, Universidad Nacional de Córdoba, Laprida 854, X5000BGR, Córdoba, Argentina*
[3]*Consejo Nacional de Investigaciones Científicas y Técnicas, Rivadavia 1917, C1033AAJ Buenos Aires, Argentina*

Accepted 2016 January 25. Received 2016 January 25; in original form 2015 May 28

ABSTRACT

Merging galaxy systems provide observational evidence of the existence of dark matter and constraints on its properties. Therefore, statistically uniform samples of merging systems would be a powerful tool for several studies. In this paper, we present a new methodology for the identification of merging systems and the results of its application to galaxy redshift surveys. We use as a starting point a mock catalogue of galaxy systems, identified using friends-of-friends algorithms, that have experienced a major merger, as indicated by its merger tree. By applying machine learning techniques in this training sample, and using several features computed from the observable properties of galaxy members, it is possible to select galaxy groups that have a high probability of having experienced a major merger. Next, we apply a mixture of Gaussian techniques on galaxy members in order to reconstruct the properties of the haloes involved in such mergers. This methodology provides a highly reliable sample of merging systems with low contamination and precisely recovered properties. We apply our techniques to samples of galaxy systems obtained from the Sloan Digital Sky Survey Data Release 7, the Wide-Field Nearby Galaxy-Cluster Survey (WINGS) and the Hectospec Cluster Survey (HeCS). Our results recover previously known merging systems and provide several new candidates. We present their measured properties and discuss future analysis on current and forthcoming samples.

Key words: galaxies: clusters: general – galaxies: kinematics and dynamics – dark matter.

1 INTRODUCTION

Merging galaxy systems, such as the Bullet Cluster (Clowe et al. 2006), Abell 520 (Mahdavi et al. 2007; Jee et al. 2012; Clowe et al. 2012; Jee et al. 2014), Baby Bullet (Bradac et al. 2008), Pandora (Merten et al. 2011), Musket Ball (Dawson et al. 2012), El Gordo (Menanteau et al. 2012; Ng et al. 2015; Molnar & Broadhurst 2015), Abell 1758 (Durret, Lima Neto & Forman 2005) and Abell 3716 (Andrade-Santos et al. 2015), have provided observational evidence for the existence of dark matter. Most of them have been used to test the cold dark matter (CDM) paradigm itself (Markevitch et al. 2004; Hayashi & White 2006; Farrar & Rosen 2007; Milosavljevic et al. 2007; Springel & Farrar 2007; Randall et al. 2008; Mastropietro & Burkert 2008; Lee & Komatsu 2010; Forero-Romero, Gottlober & Yepes 2010; Thompson & Nagamine 2012; Watson et al. 2014; Thompson, Davé & Nagamine 2015). Several statistical techniques have been proposed to measure dark matter properties using merging systems (such as the self-interaction cross-section; Massey, Kitching & Nagai 2011; Harvey et al. 2014; Kahlhoefer et al. 2014; Harvey et al. 2015). However, the lack of a complete and uniformly selected sample of merging systems prevents efforts to derive robust constraints. In order to overcome this limitation, different approaches have been proposed: the Merging Cluster Collaboration uses radio emission due to induced shocks in the intracluster medium (ICM; Feretti et al. 2012) to obtain high-redshift merging system candidates. These systems have been studied using pan-chromatic observations and detailed merging kinematic Bayesian reconstructions (Dawson 2013). X-ray imaging, spectra and Sunyaev–Zeldovich (SZ) effect observations have been used to identify unrelaxed clusters of galaxies (Mann & Ebeling 2012), cluster mergers (Harvey et al. 2014), substructures and any departures from hydrostatic equilibrium, mainly for the most massive galaxy clusters.

Galaxy redshift surveys are very useful to trace the dynamical state of galaxy systems and to search for substructures (Dressler & Shectman 1988). To this end, some methods look for departures in the global Gaussian redshift distribution of system members (Solanes, Salvador-Solé & González-Casado 1999; Hou, Parker & Harris 2014; Yu et al. 2015). Even though all these methods

⋆ E-mail: martin@oac.unc.edu.ar (MdlR); mardom@oac.uncor.edu (MJDR)

aim to identify the substructures and recover their properties, they suffer from false identifications and incompleteness, at least to some extent.

In this paper, we develop a uniform identification algorithm of merging systems based on galaxy redshift catalogues. These methods can be applied to low-mass systems and should increase the number of merging systems, such as the bullet group recently identified by Gastaldello et al. (2014). This paper is organized as follows. In Section 2, we apply machine learning techniques to a number of observable features and present its calibration based on the result of simulations. We also introduce techniques to recover properties of merging dark matter haloes. In Section 3, we apply our techniques to samples of galaxy systems identified from low-redshift galaxy surveys such as the Sloan Digital Sky Survey Data Release 7 (SDSS-DR7), the Wide-Field Nearby Galaxy-Cluster Survey (WINGS) and the Hectospec Cluster Survey (HeCS).

Finally, in Section 4, we summarize the main results of this work and discuss the uses of this new sample of merging systems. We adopt the standard cosmological model used in the Millennium Simulation (Springel et al. 2005) when necessary ($H_0 = 73$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.25$ and $\Omega_\Lambda = 0.75$).

2 METHODOLOGY

2.1 Mock galaxy and halo catalogues

From the point of view of the current theory of galaxy formation, the most direct route for defining galaxy systems is via its host dark matter halo (Mo, Van den Bosch & White 2010).

The Millennium Simulation (Springel et al. 2005), used in this work, provides a catalogue of dark matter (sub)haloes constructed using a traditional three-dimensional friends-of-friends algorithm (FoF) that percolates nearby particles. The German Astrophysical Virtual Observatory (GAVO) Millennium data base also provides merger trees for each halo (Roukema, Quinn & Peterson 1993). Therefore, given all the haloes belonging to a FoF group, it is possible to infer its merger tree. We compute for each FoF group a parent list of FoF progenitors (identified in previous snapshots) that have contributed with at least one subhalo to the current FoF group. We define a major merger of FoF groups as the merger between two groups where the total mass of the involved haloes represents at least 20 per cent of the mass after the merger. It is worth noting that this condition imposes a minimal value of 0.25 for the mass ratio of the interacting systems. We construct mock catalogues of the SDSS-DR7 redshift survey based on the results of a semi-analytical model (Guo et al. 2011) and we use it to calibrate our identification method for merging systems. This process has been extensively described in previous works (Lares, Lambas & Domínguez 2011; Domínguez Romero, García Lambas & Muriel 2012).

We define a recent major merger as the major merger of two FoF groups, where its principal haloes are still present as different haloes in the final FoF group. With these selection criteria, the mean of the look-back merger time is around 3 Gyr, consistent with other work (Pinkney et al. 1996). The FoF groups identified as recent major mergers and their member galaxies identified in the mock catalogues are used to train different machine learning methods.

2.2 Identification technique of merging systems

Galaxy systems were identified in the mock catalogue reproducing the FoF process as applied to real redshift catalogues (Merchán & Zandivarez 2002, 2005). Using the properties of the galaxies in these FoF systems, we compute several features relevant to the problem, as follows.

(i) The DS test developed by Dressler & Shectman (1988) uses the deviation of the local radial velocity, defined as the mean radial velocity of the closest n galaxies to each galaxy, from the global radial velocity in order to find substructures in clusters of galaxies. A global cluster value of the DS test is then obtained by summing up individual galaxy values. Following Pinkney et al. (1996), we select those systems of galaxies with an occupancy $Ngal > 30$ galaxy members in order to have a better identification. We performed the DS test for $n = 10$ and $n = \sqrt{Ngal}$, and used global and individual DS values as features.

(ii) Well-known tests measuring the departures from a normal Gaussian distribution include the Anderson–Darling test, the Cramer–von Mises test, the Kolmogorov–Smirnov test, the Pearson chi-square test, the Shapiro–Francia test (provided by the package NORTEST; Gross & Ligges 2015) and the Shapiro–Wilk test (provided by the STATS package).

(iii) Astrophysical properties of galaxies and clusters include SDSS magnitudes, $g - r$ colour index and occupancy of clusters.

With this set of features, we test different machine learning algorithms such as the logistic regression (Davison & Hinkley 1997; Canty & Ripley 2015), support vector machines (Cortes & Vapnik 1995; Meyer et al. 2014) and random forest (hereafter RF; Breiman 2001; Liaw & Wiener 2002), provided by the R statistical programming language, with the aim of finding merging systems in the complete sample of galaxy clusters in our simulated catalogues.

In order to measure its performance, we run a standard cross-validation test in eight folds; that is, we divide the total sample into eight individual and independent subsets and train each machine learning algorithm with seven of them in order to predict the dynamical status of the clusters of the remaining fold. As we know both the underlying and predicted dynamical status, we are able to compute the true positive rate (TPR), defined as the ratio between the number of merging clusters found in the final sample and the total number of merging clusters that were in the test fold. In the same way, we are able to compute the false positive rate (FPR), defined as the number of relaxed clusters classified as merging clusters divided by the number of relaxed clusters that the studied fold has. This information allows us to construct the receiver operating characteristic (ROC) curve shown in Fig. 1(a) in which it can be seen that the best performance is obtained by the RF algorithm.

For each cluster, the RF computes a statistic that is related to the probability that the cluster is undergoing a merger, by building many decision trees from bootstrap training data, where the final classification is based on the average assignation of the ensemble of decision trees. Each tree is grown using randomly selected features from the training data set previously described. We impose a threshold to the RF statistic of each cluster in order to classify the sample of merging clusters. In order to select an optimal threshold, we study in Fig. 1(b) the impact of different values on several statistics of the sample of merging systems, namely the TPR, the FPR, the effectiveness (number of identified true mergers divided by the total number of identified mergers) and the normalized length (number of identified mergers divided by the maximum number of identified mergers of the different thresholds). As can be seen, the classification threshold affects the performance of the RF classifier. Consequently, we select a threshold value of 0.3. This selection criterion guarantees a low false positive detection (high effectiveness) in the selected merging systems, but it should be recalled that we are only detecting just a fraction of the overall merging systems in the

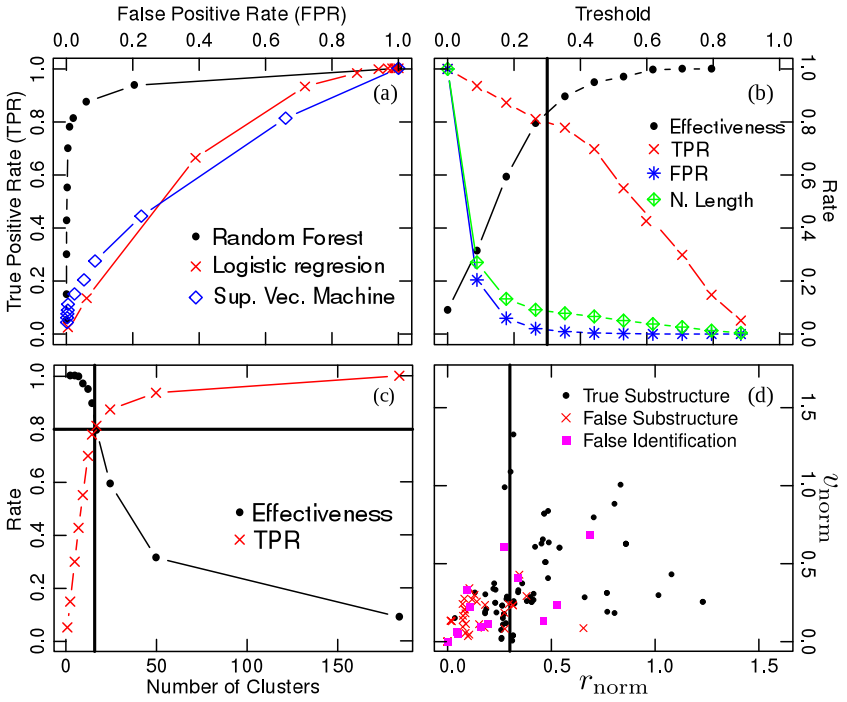

Figure 1. (a) ROC curve for logistic regression (red crosses), support vector machines (blue diamonds) and RF (black dots) classifiers evaluated using the eight-fold cross-validation test on the SDSS mock training catalogues. (b) Effectiveness (black dots), true positive rate (red crosses), false positive rate (blue asterisks) and normalized length (green diamonds) as a function of the threshold imposed to the RF statistic (see text). The vertical line indicates the selected threshold value. (c) Effectiveness (black dots) and true positive rate (red crosses) as a function of the number of galaxy systems for the RF model, with the values at the selected threshold indicated by continuous lines. (d) Normalized projected distances $r_{\rm norm}$ and velocity difference $v_{\rm norm}$ of the merging cluster haloes recovered by our methodology (black points). False positive merging clusters are indicated by magenta squares. Merging clusters with misidentified substructures are indicated by red crosses. The vertical line indicates the selection cut introduced in order to avoid LOS contamination.

simulated catalogue, as can be seen in Fig. 1(c). Note that Fig. 1(d) is analysed later in Section 2.3.

The RF implementation also allows us to assess the relative importance of the features as described in Ehrlinger (2015). In our case, the most important features are the number of galaxy members, the p-value of the Shapiro–Wilk test and the Dressler–Shectman test. Nevertheless, it is worth noting that all the previously described features are used in our RF implementation.

2.3 Measured properties of the merging haloes

The RF algorithm gives us a list of galaxies with high probability of belonging to the merging system. However, this method does not provide information about galaxy membership to the individual substructures. Using a mixture of Gaussian algorithms (R package MCLUST; Fraley et al. 2012), it is possible to cluster these member galaxies into the two merging substructures. Assigning galaxies, we are able to compute the centre position (angular and redshift) of the substructures, their velocity dispersion (Gapper estimator) and virial radius, and therefore to measure a dynamical mass. In order to estimate the associated errors, we implemented a bootstrap technique. As shown in Figs 2(a) and (b) the individual masses and the mass ratios are well recovered.

The recovered geometries of the merging systems can be seen in Fig. 1(d) in terms of the separation of the components in the line of sight (LOS) and in the plane of the sky. Specifically, on the x-axis, we compute the normalized projected distance $r_{\rm norm} = d_{1,2}/(r_{\rm vir\,1} + r_{\rm vir\,2})$, where $d_{1,2}$ is the angular separation between both components of the merging system and $r_{\rm vir\,1}$ and $r_{\rm vir\,2}$ are the corresponding virial radii. On the y-axis, we show the velocity difference $v_{\rm norm} = |v_{1,2}|/(\sigma_1 + \sigma_2)$, where $v_{1,2}$ is the velocity distance between components and σ_1 and σ_2 are the corresponding velocity dispersions of the substructures. Using this identification method, we found three different cases: relaxed clusters, which we classify as merging clusters (indicated by magenta squares); merging clusters in which we are unable to recover the real substructures (indicated by red crosses); and merging clusters in which we do recover the true substructures that are undergoing a merger (indicated by black dots). As can be seen, the FPR cases (FPR $\sim$15 per cent), indicated by magenta squares, are evenly distributed as a function of the r_{norm} parameter. The merging systems where we are not able to recover the real substructures are concentrated below a value of $r_{norm} = 0.22$. Such a selection cut will be used in order to report the properties of the substructures. Although we are able to find spatially coincident merging systems ($r_{norm} < 0.22$; i.e. systems merging along the LOS or systems with a small projected angular separation), we are not able to recover well the intervening substructures. Using this methodology, we are able to define samples of merging systems with high levels of purity, low contamination and almost exact computation of the centre position of each component, as can be seen in Figs 2(c) and (d). In Fig. 2(c), we show the angular separation θ between the actual and the recovered positions of each component, normalized to the actual virial radius. Similarly, in Fig. 2(d), we show the differences Δv between actual and measured radial velocities, divided by the real velocity dispersion. It should be recalled that our machine learning method is able to recover the correct substructures that belong to the major merger, as identified using the merger tree.

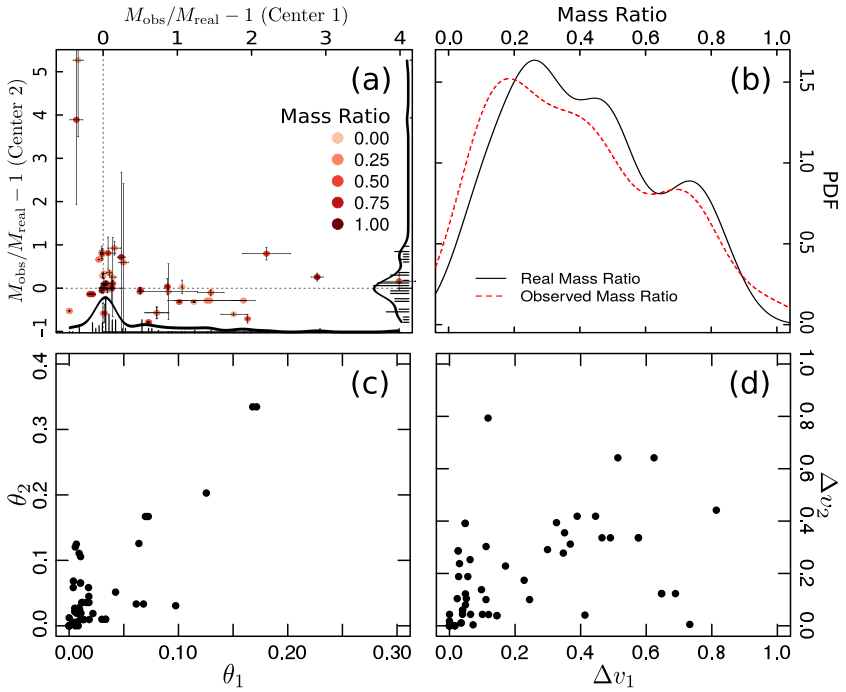

Figure 2. Recovered properties of merging galaxy systems identified in the simulated mock catalogue. The subscripts 1 and 2 refer to the main and the merging subhaloes, respectively. Panel (a) shows the estimated mass bias ($M_{obs}/M_{real} - 1$) for the main and merging substructure; the colour scale represents different mass ratios. Panel (b) shows the probability distribution function for the real and observed mass ratio (M_2/M_1). Panel (c) displays the cluster centre separation on the sky plane (real and recovered) normalized to the real virial radius. Panel (d) shows the absolute value of the velocity separation (real and recovered) normalized to the real velocity dispersion. See text for further explanations.

Table 1. Here we present the low-redshift merging cluster sample, including the properties of the two main substructures identified by our algorithm. In column (1), we present the name of the cluster. Columns (2)–(5) give the estimated mass and the position of the main substructures. Columns (6)–(9) give the estimated mass and the position of the other substructure. Finally, Column (10) lists previous work on each cluster. Clusters that have been previously reported as merging systems are indicated with √. References: 2, Wen & Han (2013), 3, Einasto et al. (2012), 4, Cohen et al. (2014), 6, Abdullah et al. (2011), 7, Rines & Diaferio (2006), 8, Rhee, van Haarlem & Katgert (1991), 9, Parekh et al. (2015), 16, Ramella et al. (2007), 21, Wang et al. (2010), 22, Johnston-Hollitt et al. (2008), 23, Boschin et al. (2012), 26, Ragozzine et al. (2012b), 27, Durret, Laganá & Haider (2011), 29, Smith & Taylor (2008), 30, Korngut et al. (2011). A full version of the table is available as supporting information in the online version.

Name	M_1 (10^{14} $M_\odot$)	RA_1 (°)	$Dec._1$ (°)	z_1	M_2 (10^{14} $M_\odot$)	RA_2 (°)	$Dec._2$ (°)	z_2	References
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Abell 1424	4.9	179.38	5.08	0.0760	5.1	179.19	5.01	0.0746	3,6.7,8
	±2.3	±0.09	±0.02	±0.0004	±1.4	±0.1	±0.04	±0.0005	
Abell 2142	18.3	239.61	27.23	0.0901	11.3	239.33	27.5	0.0893	2
√	±0.6	±0.005	±0.005	±0.0004	±1.8	±0.005	±0.005	±0.0001	
Abell 3158	37.24	55.75	−53.63	0.0633	4.6	55.37	−53.48	0.0622	4,9,16,21,22
√	±1.5	±0.07	±0.004	±0.0001	±0.2	±0.007	±0.001	±0.0001	
Abell 2382	77.7	327.90	−15.66	0.0676	6.12	328.167	−15.62	0.0642	
	±10.2	±0.006	±0.006	±0.0003	±1.1	±0.003	±0.01	±0.0002	
Abell 1758N	59.3	203.07	50.59	0.2768	29.1	203.25	50.57	0.2783	2,23,26,27
√	±9	±0.02	±0.01	±0.0002	±15.8	±0.02	±0.03	±0.0007	
Abell 1835	105	210.25	2.87	0.2516	17	210.29	2.75	0.2479	2,29,30
	±14	±0.01	±0.01	±0.0006	±25	±0.09	±0.04	±0.002	

3 APPLICATION TO LOW-REDSHIFT CLUSTERS

After testing our algorithm on simulated data, we applied our identification algorithms to galaxy systems with more than 30 galaxy members identified in the updated Merchán & Zandivarez (2002) catalogue, based on SDSS-DR7 (Abazajian et al. 2009) data and two samples of galaxy cluster measurements: WINGS (Cava et al. 2009) and HeCS (Rines et al. 2013). For the WINGS clusters, we compute the $g - r$ colour of the individual galaxies based on the observed $b - v$ colour, applying the formulae presented by the 2df Collaboration based on the results of Fukugita, Shimasaku & Ichikawa (1995). We report the following clusters as spatially coincident merging systems candidates: A2593, A2199√, A2048√, A3266√, A3497, A667, A1201√, A267√, Zw8197, A697√, A750, Zw2701, Zw3146, A1246, A1302, A1413, A1682√, A1763√, A1902, A1918, A1930, A2009, A2034√, A2069√, A2111√, A2219√, A2050, A2259 and RXC 1504. Note that we do not provide the substructure properties due to the LOS projection effect (as discussed

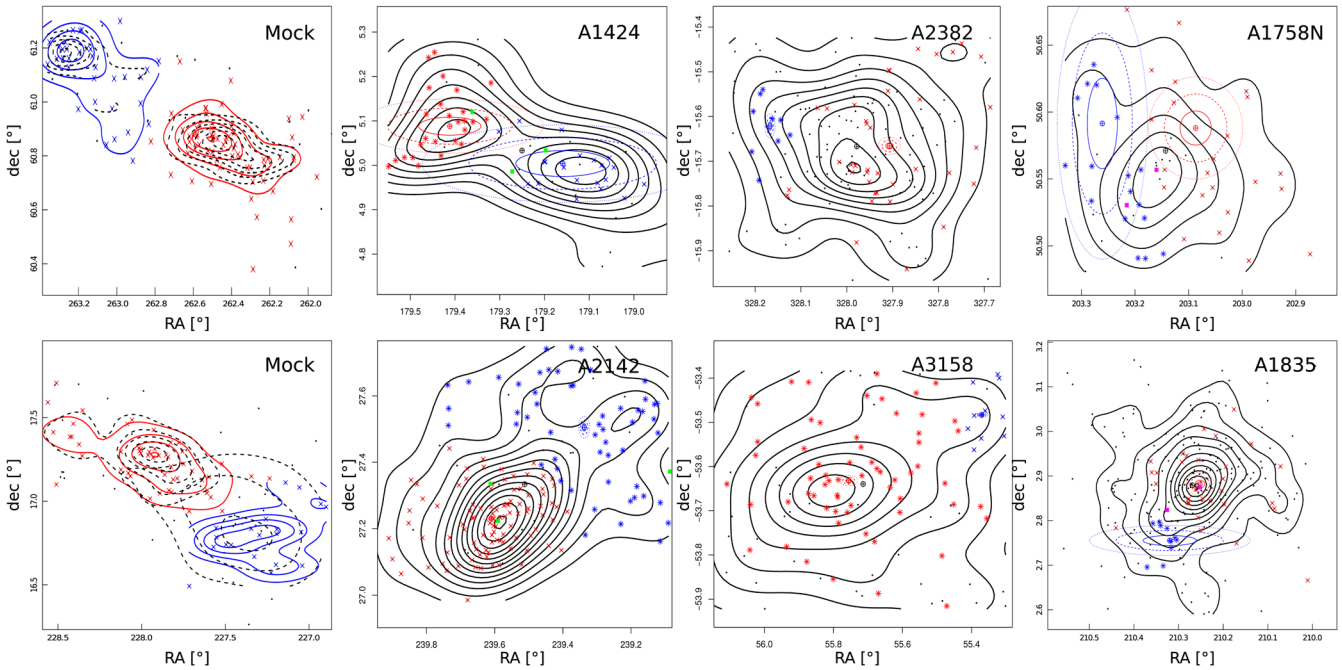

Figure 3. Galaxy angular distribution of some merging systems. From left to right we show in columns: two simulated systems, two clusters from the SDSS, two WINGS clusters and two HeCs clusters. Member galaxies of both substructures are shown as red and blue dots while the black lines plot the iso-density contours. The ellipses indicate the 1σ, 2σ and 3σ errors of the identified substructure (on blue or red colours depending on the component), with solid, dashed and dotted lines, respectively. For comparison, green dots show the angular positions of the X-ray sources, and magenta dots show the angular positions of the substructures identified by other authors. For the simulated merging systems, we also show as a dashed line the iso-density of the true substructures that are colliding and as red and blue lines the iso-density contours for the identified substructures.

in Section 2.3), and therefore minor mergers could be included in this sample (Ma et al. 2012). We display in Table 1 the properties of the two interacting structures of the merging clusters identified using our machine learning (RF) classification techniques. The errors for each property are the standard deviation computed on a hundred measurements obtained from the RF and the clustering realizations. Many of them are well-known merging systems (indicated by $\surd$in the table); however, it is important to emphasize that our method was able to find several new candidates and also to measure their properties. It should be noted that the decomposition of these structures is only indicative, because galaxy velocities are strongly affected by the gravitational attraction of the two haloes. Therefore, a tomographic reconstruction is necessary (including lensing and X-ray/SZ data) to recover the substructures accurately.

3.1 Case of multiple major mergers

It is well known that there are some clusters that are the result of the merging of more than two systems, although these represent only a small fraction of the total sample (27 of 132 in eight SDSS mock catalogues). In order to recover all the merging substructures, we performed a mixture of more than two Gaussians. We found that our algorithm is only able to recover, with reliable properties, the two more important substructures. Hence, the remaining structures may appear as contamination, or not appear at all. We discuss an individual analysis of Abell 1758 as an example of the multiple major merger case. This cluster is known as a merger of four substructures, two in the north and two in the south (Ragozzine et al. 2012a). At a first iteration, our algorithm classified correctly Abell 1758 as a merging cluster, but failed to properly reconstruct the merging substructures. Considering this extra information, we separate the cluster into north and south components and perform a new analysis of each separately. We find that our algorithm is able to classify both as merging clusters and well recovers the merging substructure properties of both components (north and south).

4 CONCLUSIONS AND FUTURE WORK

In this paper, we have introduced a method aimed at detecting merging systems of galaxies in redshift surveys. We select a RF algorithm between other machine learning algorithms, and use as features the quantities derived from the galaxy redshift space information and from photometry (e.g. colours). Our detection method was trained and calibrated using a sample of merging systems extracted from mock catalogues. By studying the merger trees, we check that we do find the two substructures that experienced a major merger and we recover their fundamental properties (positions and masses). We apply our techniques to a sample of systems of galaxies identified in SDSS-DR7, WINGS and HeCS.

The resulting merging system sample, in which we are able to recover the merging substructures, comprises 12, 4 and 16 systems, respectively. Additionally, we report 29 spatially coincident merging system candidates. Several of these systems were previously reported by other authors as interacting systems of galaxies. We also report for the first time 40 new nearby candidates as merging systems, which were previously overlooked.

We emphasize that our method detects in a reliable way the merging system candidates and substructure properties, but we also wish to note that, in the case of multiple mergers, some merging substructures may be joined by our algorithm and hence are reported as one component, causing a possible bias in some measured properties. The kinematical reconstructions (see Fig. 3 and Table 1) will be

corroborated using tomographic techniques including X-ray and weak lensing information (Gonzalez et al. 2015) in forthcoming papers. We also plan in further works to train our algorithm with light cone mocks in order to apply our technique to high-redshift catalogues, such as CLASH-VLT (Biviano et al. 2013), FRONTIERs (Ebeling, Ma & Barrett 2014), EDiCs (Milvang-Jensen et al. 2008), DESI, etc., and to study any environmental dependence of galaxy properties (e.g. star formation rate, stellar mass, morphology) at different stages in the merging process.

Diverse studies could be performed with a sample of merging systems such as the ones presented in this paper. Using the Bayesian reconstruction techniques presented by Dawson (2013), it is possible to recover the three-dimensional information of the merger. In a forthcoming paper, we will use such information in order to explore the implications for the properties of the DM particle. A web interface implementing these methods (the MeSsI Algorithm) is freely available at http://200.16.29.98/martin/merclust.

ACKNOWLEDGEMENTS

MdlR warmly thanks Monica Jauregui and Jorge de los Rios for their unconditional support. We also thank Dario Graña for technical support and Sebastían Gurovich for useful comments. This work was partially supported by the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, Argentina) and the Secretaría de Ciencia y Tecnología de la Universidad Nacional de Córdoba (SeCyT-UNC, Argentina). This research was developed while MJdLDR was a postdoctoral researcher at the Center for Gravitational Wave Astronomy (http://cgwa.phys.utb.edu/). This research has made use of the SAO/NASA Astrophysics Data System (ADS) (http://adsabs.harvard.edu/), the Cornell University xxx.arxiv.org repository, the SIMBAD data base, operated at CDS, Strasbourg, France, the R project for statistical computing and the following CRAN libraries: `mclust,rgl,randomForest,nortest,e1071,shiny`.

The Millennium Simulation data bases (Lemson & Virgo Consortium 2006) used in this paper and the web application providing online access to them were constructed as part of the activities of the GAVO.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society and the Higher Education Funding Council for England. The SDSS web site is http://www.sdss.org/.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Abdullah M. H., Ali G. B., Ismail H. A., Rassem M. A., 2011, MNRAS, 416, 2027
Andrade-Santos F. et al., 2015, ApJ, 803, 108
Biviano A., Rosati P., Balestra I., Mercurio A., 2013, A&A, 558, A1
Boschin W., Girardi M., Barrena R., Nonino M., 2012, A&A, 540, A43
Bradac M., Allen S. W., Treu T., Ebeling H., Massey R., Morris R. G., von der Linden A., Applegate D., 2008, ApJ, 687, 959
Breiman L., 2001, Machine Learning, 45, 5
Canty A., Ripley B. D., 2015, boot: Bootstrap R Functions
Cava A. et al., 2009, A&A, 495, 707
Clowe D., Bradac M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109
Clowe D., Markevitch M., Bradac M., Gonzalez A. H., Chung S. M., Massey R., Zaritsky D., 2012, ApJ, 758, 128
Cohen S. A., Hickox R. C., Wegner G. A., Einasto M., Vennik J., 2014, ApJ, 783, 136
Cortes C., Vapnik V., 1995, Machine Learning, 20, 273
Davison A. C., Hinkley D. V., 1997, Bootstrap Methods and Their Applications. CUP, Cambridge
Dawson W. A., 2013, ApJ, 772, 131
Dawson W. A. et al., 2012, ApJ, 747, L42
Domínguez Romero M., García Lambas D., Muriel H., 2012, MNRAS, 427, L6
Dressler A., Shectman S. A., 1988, AJ, 95, 985
Durret F., Lima Neto G., Forman W., 2005, A&A, 432, 809
Durret F., Laganá T. F., Haider M., 2011, A&A, 529, A38
Ebeling H., Ma C.-J., Barrett E., 2014, ApJS, 211, 21
Ehrlinger J., 2015, ggRandomForests: Graphical Exploration of Random Forests
Einasto M. et al., 2012, A&A, 540, A123
Farrar G. R., Rosen R. A., 2007, Phys. Rev. Lett., 98, 171302
Feretti L., Giovannini G., Govoni F., Murgia M., 2012, A&ARv, 20, 54
Forero-Romero J. E., Gottlober S., Yepes G., 2010, ApJ, 725, 598
Fraley C., Raftery A. E., Murphy T. B., Scrucca L., 2012, mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation, Technical Report No. 597. Department of Statistics, University of Washington
Fukugita M., Shimasaku K., Ichikawa T., 1995, pasp, 107, 945
Gastaldello F. et al., 2014, MNRAS, 442, L76
Gonzalez E. J., Foëx G., Nilo Castellón J. L., Domínguez Romero M. J., Alonso M. V., García Lambas D., Moreschi O., Gallo E., 2015, MNRAS, 452, 2225
Gross J., Ligges U., 2015, nortest: Tests for Normality
Guo Q. et al., 2011, MNRAS, 413, 101
Harvey D. et al., 2014, MNRAS, 441, 404
Harvey D., Massey R., Kitching T., Taylor A., Tittley E., 2015, Sci, 347, 1462
Hayashi E., White S. D. M., 2006, MNRAS, 370, L38
Hou A., Parker L., Harris W. E., 2014, MNRAS, 442, 406
Jee M. J., Mahdavi A., Hoekstra H., Babul A., Dalcanton J. J., Carroll P., Capak P., 2012, ApJ, 747, 96
Jee M. J., Hoekstra H., Mahdavi A., Babul A., 2014, ApJ, 783, 78
Johnston-Hollitt M., Sato M., Gill J. A., Fleenor M. C., Brick A.-M., 2008, MNRAS, 390, 289
Kahlhoefer F., Schmidt-Hoberg K., Frandsen M. T., Sarkar S., 2014, MNRAS, 437, 2865
Korngut P. M. et al., 2011, ApJ, 734, 10
Lares M., Lambas D., Domínguez M., 2011, AJ, 142, 13
Lee J., Komatsu E., 2010, ApJ, 718, 60
Virgo ConsortiumLemson G. 2006, preprint (arXiv:astro-ph/0608019)
Liaw A., Wiener M., 2002, R News, 2, 18
Mahdavi A., Hoekstra H., Babul A., Balam D. D., Capak P. L., 2007, ApJ, 668, 806
Ma C.-J., Owers M., Nulsen P. E. J., McNamara B. R., Murray S. S., Couch W. J., 2012, ApJ, 752, 139
Mann A. W., Ebeling H., 2012, MNRAS, 420, 2120
Markevitch M., Gonzalez A. H., Clowe D., Vikhlinin A., Forman W., Jones C., Murray S., Tucker W., 2004, ApJ, 606, 819
Massey R., Kitching T., Nagai D., 2011, MNRAS, 413, 1709
Mastropietro C., Burkert A., 2008, MNRAS, 389, 967
Menanteau F. et al., 2012, ApJ, 748, 7
Merchán M., Zandivarez A., 2002, MNRAS, 335, 216
Merchán M. E., Zandivarez A., 2005, ApJ, 630, 759
Merten J. et al., 2011, MNRAS, 417, 333
Meyer D., Dimitriadou E., Hornik K., Weingessel A., Leisch F., 2014, e1071, TU Wien
Milosavljevic M., Koda J., Nagai D., Nakar E., Shapiro P. R., 2007, ApJ, 661, L131
Milvang-Jensen B., Noll S., Halliday C., Poggianti B. M., 2008, A&A, 482, 419
Mo H., Van den Bosch F., White S., 2010, Galaxy Formation and Evolution. CUP, Cambridge

Molnar S. M., Broadhurst T., 2015, ApJ, 800, 37
Ng K. Y., Dawson W. A., Wittman D., Jee M. J., Hughes J. P., Menanteau F., Sifón C., 2015, MNRAS, 453, 1531
Parekh V., van der Heyden K., Ferrari C., Angus G., Holwerda B., 2015, A&A, 575, A127
Pinkney J., Roettiger K., Burns J. O., Bird C. M., 1996, ApJS, 104, 1
Ragozzine B., Clowe D., Markevitch M., Gonzalez A. H., Bradač M., 2012a, ApJ, 744, 94
Ragozzine B., Clowe D., Markevitch M., Gonzalez A. H., Bradač M., 2012b, ApJ, 744, 94
Ramella M. et al., 2007, A&A, 470, 39
Randall S. W., Markevitch M., Clowe D., Gonzalez A. H., Bradac M., 2008, ApJ, 679, 1173
Rhee G. F. R. N., van Haarlem M. P., Katgert P., 1991, A&A, 246, 301
Rines K., Diaferio A., 2006, AJ, 132, 1275
Rines K., Geller M. J., Diaferio A., Kurtz M. J., 2013, ApJ, 767, 15
Roukema B. F., Quinn P. J., Peterson B. A., 1993, in Chincarini G. L., Iovino A., Maccacaro T., Maccagni D., eds, ASP Conf. Ser. Vol. 51, Observational Cosmology. Astron. Soc. Pac., San Francisco, p. 298
Smith G. P., Taylor J. E., 2008, ApJ, 682, L73
Solanes J. M., Salvador-Solé E., González-Casado G., 1999, A&A, 343, 733
Springel V., Farrar G. R., 2007, MNRAS, 380, 911
Springel V. et al., 2005, Nat, 435, 629
Thompson R., Nagamine K., 2012, MNRAS, 419, 3560
Thompson R., Davé R., Nagamine K., 2015, MNRAS, 452, 3030
Wang Y., Xu H., Gu L., Gu J., Qin Z., Wang J., Zhang Z., Wu X.-P., 2010, MNRAS, 403, 1909
Watson W. A., Iliev I., Diego J., Gottlober S., Knebe A., Martinez-Gonzalez E., Yepes G., 2014, MNRAS, 437, 3776
Wen Z. L., Han J. L., 2013, MNRAS, 436, 275
Yu H., Serra A., Diaferio A., Baldi M., 2015, ApJ, 810, 37

SUPPORTING INFORMATION

Table 1. Here we present the low-redshift merging cluster sample, including the properties of the two main substructures identified by our algorithm.
(http://mnras.oxfordjournals.org/lookup/suppl/doi:10.1093/mnras/stw215/-/DC1).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.