Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **485,** 498–512 (2019)
Advance Access publication 2019 January 28

doi:10.1093/mnras/stz294

AMICO galaxy clusters in KiDS-DR3: sample properties and selection function

Matteo Maturi,[1]★ Fabio Bellagamba,[2,3] Mario Radovich,[4] Mauro Roncarelli,[2,3] Mauro Sereno,[2,3] Lauro Moscardini,[2,3,5] Sandro Bardelli[2] and Emanuella Puddu[6]

[1]*Zentrum für Astronomie, Universität Heidelberg, Philosophenweg 12, D-69120 Heidelberg, Germany*
[2]*Dipartimento di Fisica e Astronomia, Alma Mater Studiorum – Università di Bologna, via Piero Gobetti 93/2, I-40129 Bologna, Italy*
[3]*INAF - Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Gobetti 93/3, I-40129 Bologna, Italy*
[4]*INAF - Osservatorio Astronomico di Padova, vicolo dell'Osservatorio 5, I-35122 Padova, Italy*
[5]*INFN - Sezione di Bologna, Viale Berti Pichat 6/2, I-40127 Bologna, Italy*
[6]*INAF - Osservatorio Astronomico di Capodimonte, Salita Moiariello 16, I-80131 Napoli, Italy*

Accepted 2019 January 18. Received 2019 January 16; in original form 2018 October 5

ABSTRACT

We present the first catalogue of galaxy cluster candidates derived from the third data release of the Kilo-Degree Survey. The sample of clusters has been produced using the Adaptive Matched Identifier of Clustered Objects (AMICO) algorithm. In this analysis, AMICO considers the luminosity, spatial distribution, and photo-z of galaxies, without performing any selection based on their colours. In this way, we minimize the dependence of the selection function on the detectability or even absence of the red sequence in the clusters. The catalogue comprises 7988 candidate galaxy clusters in the redshift range $0.1 < z < 0.8$ down to signal-to-noise ratio > 3.5 with a purity approaching 95 per cent over the entire redshift range. In addition to the catalogue of galaxy clusters, we also provide a catalogue of galaxies with their probabilistic association to the detected clusters. We quantify the sample purity, completeness, and the uncertainties of the detections properties, such as richness, redshift, and position, by means of mock galaxy catalogues. The simulations are derived directly from the data to fully reproduce their statistical properties including photo-z uncertainties, unknown absorption across the survey, missing data, spatial correlation of galaxies, and galaxy clusters. Being based on real data, such mock catalogues do not have to rely on the assumptions on which numerical simulations and semi-analytic models are based on. This paper is the first of a series of papers in which we discuss the details and physical properties of the sample presented in this work.

Key words: galaxies: clusters: general – cosmology: observations – large-scale structure of Universe.

1 INTRODUCTION

Clusters of galaxies are one of the fundamental probes to study the nature of dark matter, dark energy (Umetsu et al. 2014; de Haan et al. 2016; Planck Collaboration XXIV 2016b; Sartoris et al. 2016; Smith et al. 2016; Wang et al. 2016; Schellenberger & Reiprich 2017; Corasaniti et al. 2018; Giocoli et al. 2018), gravity itself (Llinares & Mota 2013; L'Huillier et al. 2017), neutrinos (Costanzi et al. 2013; Roncarelli, Carbone & Moscardini 2015) as well as the far universe, and the early stages of star and galaxy formation when used as gravitational lensing telescopes (Zheng et al. 2012; Coe et al. 2013; Bradley et al. 2014; Kelly et al. 2015; Rydberg et al. 2018). There are many ways to identify galaxy clusters: through the X-ray emission (Böhringer et al. 2004; Pace et al. 2008; Piffaretti et al. 2011; Merloni et al. 2012; Clerc et al. 2014), the Comptonization of the cosmic microwave background photons by the hot plasma they contain (Reichardt et al. 2013; Planck Collaboration XIII 2016a; Hilton et al. 2018), the gravitational lensing distortion they induce on background galaxies (Maturi et al. 2005; Pace et al. 2007; Bellagamba et al. 2011), and the optical emission of their population of galaxies. In this paper, we discuss the latter. Various methods have been proposed and used for their detection in photometric catalogue of galaxies. For instance, wavelength filters (Benoist 2014; Gonzalez 2014), friend-of-friends (Farrens et al. 2011), methods based on Voronoi tessellation (Iovino 2014), minimal spanning trees (Adami & Mazure 1999), red-sequenced finders (Rykoff et al. 2014; Licitra et al. 2016), and matched optimal filters (Bellagamba et al. 2011, 2018b). A comparison of the performance of the different algorithms in the context of the

★ E-mail: maturi@uni-heidelberg.de

Euclid survey (Laureijs et al. 2011) can be found in Adam et al. (in preparation).

In this work, we searched for galaxy clusters in the third data release of the Kilo-Degree Survey (KiDS-DR3, de Jong et al. 2017). With respect to our previous studies on the second data release (Radovich et al. 2016, 2017), the analysis presented here benefits of a larger survey area, better data quality, and significant improvements in the cluster detection algorithm. For this task, we use the Adaptive Matched Identifier of Clustered Objects algorithm (AMICO, Bellagamba et al. 2018b), an optimal matched filter that takes advantage of the known statistical properties of the field galaxies and of galaxy clusters. Even if AMICO can deal with an arbitrary number of quantities describing galaxies, in this specific application we consider their spatial coordinates, magnitude, and photo-z only. We avoided using explicitly the information on galaxy colours: This aims at achieving a selection function minimally sensitive to the detectability, presence, or even absence of the red sequence of clusters.

We derived the sample purity, the sample completeness, and the uncertainty of the properties of the detections through a series of realistic mock catalogues of galaxies based on the real KiDS data. In doing so, we took care to preserve the actual masked areas in the data, the statistics of the photometry, and photo-zs as well as the galaxies large-scale correlation, the correlation of clusters among themselves, and of clusters with the large-scale structures (LSS).

The structure of the paper is as follows. In Section 2, we describe the data set. In Section 3, we summarize the characteristics of the detection algorithm and the new features used specifically for this work. The catalogue of galaxy clusters and the comparison with existing catalogues are presented in Sections 4 and 5, respectively. The uncertainties on the detection properties, the completeness, and purity of the sample are quantified in Section 6. Finally, the conclusions are summarized in Section 7.

2 THE DATA SETS

We analysed the galaxy catalogue coming from the KiDS Data Release 3 (de Jong et al. 2017) obtained with the OmegaCAM wide-field imager (Kuijken 2011) mounted at the VLT Survey Telescope, a 2.6 m telescope sited at the Paranal Observatory (VST; Capaccioli & Schipani 2011). OmegaCAM contains a mosaic of 32 science CCDs offering a field of view of 1 deg^2 with a resolution of 0.21 arcsec pixel^{-1}. The data cover an area of about 438 deg^2 split into two main stripes, one equatorial (KiDS-N) and one centred around the South Galactic Pole (KiDS-S).

The galaxy catalogue provides the coordinates, the photometry in four bands (u, g, r, i), and the photometric redshifts for all galaxies down to the 5σ limiting magnitudes of 24.3, 25.1, 24.9, and 23.8 in the four bands, respectively. As described in detail in de Jong et al. (2017), the photometry includes aperture magnitudes, Kron magnitudes (*MAG_AUTO*), and the so-called Gaussian Aperture and point spread function (*GAap*) magnitudes. The *GAap* magnitudes provide an optimal measurement of galaxy colours, as they correct the effect of seeing variations across the image and in different bands. For the photometry, we used the Kron magnitudes and selected all galaxies with magnitude $r < 24$ for a total of about 32 million objects. We used the r band as a reference band because it is deeper (by 1.3 magnitudes on average) and much more uniform, from tile to tile, than the i band. This choice also allows us to take advantage of the larger number of faint blue cluster galaxies that statistically enhance the detected signal (see next). The photometric redshifts of the galaxies have been obtained with the *GAap* magnitude and Bayesian photo-z (BPZ), an estimator based on a template-fitting method (Benítez 2000; de Jong et al. 2017). BPZ returns a photo-z posterior probability distribution function that is fully exploited by AMICO (see next). Other two sets of photometric redshifts obtained via machine learning techniques, MLPQNA and ANNz2, are available in KiDS-DR3 (de Jong et al. 2017; Bilicki et al. 2018). An extensive analysis of the probability distribution functions derived for these two sets will be presented in Amaro et al. (2018): since it was still in progress when our cluster catalogue has been derived, we opted to use the well-tested BPZ photometric redshifts. In future releases, we will also investigate the usage of machine learning photometric redshifts.

3 AMICO: THE DETECTION ALGORITHM

For the detection of the galaxy clusters, we used the AMICO code (Bellagamba et al. 2018b). In this section, we briefly summarize its main concepts and the new features recently implemented and adopted in the following analysis.

3.1 Linear optimal matched filtering

AMICO is based on a linear optimal matched filter approach (Maturi et al. 2005, 2007; Viola, Maturi & Bartelmann 2010; Bellagamba et al. 2011). Within this framework, the data, $\mathrm{d}(\vec{x}) = s(\vec{x}) + n(\vec{x})$, are modelled as the superimposition of the signal we are interested in, i.e. the signal relate to galaxy clusters $s(\vec{x}) = Ac(\vec{x})$, and a noise component, $n(\vec{x})$, describing the contamination of the field galaxies. The filter itself is a kernel used to convolve the data, and it is derived through a constrained minimization procedure. Such a procedure aims at estimating the signal amplitude, A, is unbiased, and with minimum variance. Despite the fact that AMICO can deal with an arbitrary number of galaxy properties, we restrict the analysis to the simple case in which the properties of the ith galaxy, $\vec{x}_i = \left(\vec{\theta}_i, m_i, p_i(\vec{z})\right)$, are limited to the sky coordinates $\vec{\theta}_i$, an r-band magnitude m_i and a photometric redshift distribution $p_i(z)$. The filter returning the estimate, A, is evaluated on a 3D grid $(\vec{\theta}_c, z_c)$, with a resolution of 0.3 arcmin across the sky and a sampling of 0.01 in redshift. Furthermore, it is discretized to deal with counts of galaxies:

$$A(\vec{\theta}_c, z_c) = \alpha^{-1}(z_c) \sum_{i=1}^{N} \frac{C(z_c; \vec{\theta}_i - \vec{\theta}_c, m_i) p_i(z_c)}{N(m_i, z_c)} - B(z_c). \quad (1)$$

Here, N and C describe the properties of the field and cluster galaxies at redshift z_c, respectively, (see Section 3.4), the factor α takes care of the filter normalization, and B is a background subtraction term quantifying the average contribution of the field galaxies to the total signal amplitude. The expected r.m.s. of the amplitude is given by

$$\sigma_A(\vec{\theta}_c, z_c) = \alpha(z_c)^{-1} + A(\vec{\theta}_c, z_c) \frac{\gamma(z_c)}{\alpha(z_c)^2}, \quad (2)$$

where the first term refers to the stochastic fluctuations of the background and the second one is related to the Poissonian fluctuations given by the galaxies of a cluster with amplitude A. The factors B, α, and γ are properties of the filter and solely depend on the cluster[1] and field models. The definition of B is given in Bellagamba et al. (2018b), while the new definitions of α and γ, implementing the new features of the algorithm, are given in Section 3.2.

[1]We base the redshift distribution of the model on the average $p(z)$ of the input galaxies as detailed in Bellagamba et al. (2018b).

Once the amplitude A has been evaluated for all angular and redshift positions, the first cluster candidate is identified as the location with positive amplitude and the largest likelihood:

$$\mathcal{L}(\vec{\theta}_c, z_c) = \mathcal{L}_0 + A^2(\vec{\theta}_c, z_c)\alpha(z_c)\ ,\ \Delta\mathcal{L}(\vec{\theta}_c, z_c) = A^2(\vec{\theta}_c, z_c)\alpha(z_c). \tag{3}$$

Here, $\mathcal{L}_0$ is a constant of no relevance that shall not be discussed further. Subsequently, we evaluate the probability of all galaxies in the surroundings, labelled with i, to be members of this detection, now labelled with j:

$$P_i(j) = \tilde{P}_{f,i}\frac{A_j C_j(\vec{\theta}_i - \vec{\theta}_j, \vec{m}_i)p_i(z_j)}{A_j C(\vec{\theta}_i - \vec{\theta}_j, \vec{m}_i)p_i(z_j) + N(\vec{m}_i, z_j)}. \tag{4}$$

Here, $\tilde{P}_{f,i} = 1 - \sum_k^{j-1} P_i(k)$ accounts for the previous memberships assigned to the ith galaxy. Note that $\tilde{P}$ is not the probability of the ith galaxy to belong to the field,[2] which instead is

$$P_{f,i} \equiv 1 - \sum_k P_i(k). \tag{5}$$

To proceed with the search for further clusters, we use the probabilities $P_i(j)$ to remove the contribution of the last detection from the amplitude map:

$$A_{new}(\vec{\theta}_j, z_k) = A(\vec{\theta}_j, z_k) - \sum_{i=1}^{N} P_i(j)\frac{C_j(\vec{\theta}_i - \vec{\theta}_j, \vec{m}_i)p_i(z_k)}{N(\vec{m}_i, z_k)}, \tag{6}$$

and to re-evaluate the likelihood as well as the variance. This signal subtraction, which we refer to as '*cleaning*', facilitates the identification of objects that might have been blended. This iterative process proceeds down to a desired minimum signal-to-noise ratio $(S/N) = A/\sigma_A$, which in this work is set to $(S/N)_{\rm min} = 3.0$.

3.2 New features of the algorithm

In order to correctly normalize the amplitude A and to estimate its uncertainty σ_A, AMICO calculates the quantities $\alpha(z_c)$ and $\gamma(z_c)$, which depend on the properties of the redshift probability distributions of the galaxy sample. In Bellagamba et al. (2018b), this was done by introducing $q(z_c, z)$, the typical redshift probability distribution for a galaxy, that lies at redshift, z_c, computed as

$$q(z_c, z) = \left(\sum_{i=1}^{N_{\rm gal}} p_i(z_c)\right)^{-1} \sum_{i=1}^{N_{\rm gal}} p_i(z - z_c + z_{{\rm peak,i}})\ p_i(z_c), \tag{7}$$

where $z_{{\rm peak},i}$ is the most probable redshift for the ith galaxy. In this analysis, we refined this treatment in two ways. First of all, we expressed the photo-zs as a function of the r-band magnitude to capture the accuracy of photo-zs depending on the quality of the galaxy photometry. Here, we did not want to include an explicit dependence on the colours of galaxies to avoid any direct link to the cluster red sequence. Then, we replaced q with two different statistics q_1 and q_2 defined as

$$q_1(m, z_p, z_c) = \left(\sum_{z_{{\rm peak},i}=z_p} p_i(z_p)\right)^{-1} \sum_{z_{{\rm peak},i}=z_p} p_i(z_p)\ p_i(z_c), \tag{8}$$

[2]In general, a galaxy can be associated with more detections because the latter can overlap one with each other. We store the information of all cluster members with $P_i(j) > 0$.

and

$$q_2(m, z_c, z_p) = \left(\sum_{i=1}^{N_{\rm gal}} p_i(z_c)\right)^{-1} \sum_{z_{{\rm peak},i}=z_p} p_i(z_c)\ p_i(z_p), \tag{9}$$

where $z_{{\rm peak},i} = z_p$ stands for a sum running only on the galaxies whose probability redshift distribution peaks at z_p. In practice, q_1 describes the typical $p(z)$ that peaks at z_p, while q_2 describes the probability distribution for the peak, z_p, of a galaxy that is located at redshift z_c. Together, they allow to measure the typical precision of the redshift probability distribution as a function of z and the small-scale features of the $p(z)$-peaks distribution to avoid the smoothing that is implicit in equation (7). With these two new quantities, the constants $\alpha(z_c)$ and $\gamma(z_c)$ are now defined as

$$\alpha(z_c) = \int \frac{M_c^2(\vec{\theta} - \vec{\theta}_c, m)\ q_1(m, z_p, z_c)q_2(m, z_c, z_p)}{N(m, z_c)}\ {\rm d}^2\theta\ {\rm d}m\ {\rm d}z_p, \tag{10}$$

and

$$\gamma(z_c) = \int \frac{M_c^3(\vec{\theta} - \vec{\theta}_c, m)\ q_1^2(m, z_p, z_c)q_2(m, z_c, z_p)}{N^2(m, z_c)}\ {\rm d}^2\theta\ {\rm d}m\ {\rm d}z_p. \tag{11}$$

3.3 Mass proxies and cluster richness

As discussed in Section 3.1, the natural output of the linear optimal matched filter is the amplitude, A, expressed by equation (6). In this section, we derive two other mass proxies based on the probabilistic membership association of galaxies to detections (equation 4). The first one is the *apparent richness* that is defined as the sum of the probabilities of all galaxies associated with the jth detection:

$$\lambda_j = \sum_{i=1}^{N_{gal}} P_i(j). \tag{12}$$

This quantity represents the number of visible galaxies belonging to a detection. Clearly, this number depends on the cosmic distance at which a cluster is located. Consequently, λ is a redshift-dependent quantity: the further the cluster, the smaller the number of visible members is. The advantage of this definition with respect to the amplitude A is that it is related to a direct observable, namely the number of visible galaxies.

The second mass proxy is the *intrinsic richness* that is defined in a similar fashion but by summing over the galaxies brighter than $m_* + 1.5$ and within the virial radius, R_{200}:

$$\lambda_{*j} = \sum_{i=1}^{N_{\rm gal}} P_i(j) \quad \text{with} \quad \begin{cases} m_i < m_*(z_j) + 1.5 \\ r_i(j) < R_{200}(z_j) \end{cases}. \tag{13}$$

The radial cut R_{200} and m_* are parameters of the model used for the construction of the filter, see Section 3.4. We adopt these values for internal consistency, but they can be arbitrarily choosen. In this case, our goal is to set a magnitude cut-off within the survey limit to keep the proxy redshift independent. The choice of the radial cut-off is less crucial because the probability $P_i(j)$ already weights the galaxies according to their angular separation from the detection. Obviously, each detection has its own R_{200} and we could scale the radial cut-off according to the amplitude, but this would lead to an additional stochastic scatter given by the uncertainty of A.

Note that the definitions of λ and λ_* depend on the conversion from angular to physical distances and therefore imply the

Table 1. The three mass proxy delivered by AMICO.

Name	Description
A	Amplitude, natural output of the filter
λ	Apparent richness, number of visible galaxies
$\lambda*$	Intrinsic richness, as λ but for $r < R_{200}$ and $m < m_* + 1.5$

assumption of a specific cosmological model.[3] Despite that such dependence is very mild being a simple cut-off. In Table 1, we list the mass proxies provided by AMICO and delivered with the catalogue of galaxy clusters.

3.4 Model description: cluster and field galaxies

The model, $C(z_c;\ r,\ m)$, describes the expected distribution of galaxies as a function of distance from the centre, $r = |\vec{\theta}_i - \vec{\theta}_c|$, and of r-band magnitude, m, for a cluster at redshift z_c. In this work, the cluster model is constructed from a luminosity function $\Phi(m)$ and a radial profile $\Psi(r)$ as

$$C(r, m) = \Phi(m)\Psi(r), \tag{14}$$

where we made implicit the dependence on the redshift z_c. The parameters for these distributions are taken from the analysis of a sample of Sunyaev–Zel'dovich (SZ) detected clusters observed by the Dark Energy Survey (DES; Zenteno et al. 2016; Hennig et al. 2017). These clusters cover a redshift range $0 < z < 1.1$, which is broadly comparable to ours. Their detection via the SZ effect avoids any selection bias related to the optical properties of the galaxies in clusters, which could introduce systematics in the detection process.

In particular, the luminosity function $\Phi(m)$ follows a Schechter function (Schechter 1976):

$$\Phi(m) = 10^{-0.4(m-m_\star)(\beta+1)} \exp\left[-10^{-0.4(m-m_\star)}\right], \tag{15}$$

where the normalization is not relevant for this work because it can be absorbed by the constants during the filter construction. The typical magnitude $m_\star$ as a function of redshift is derived from a stellar population evolutionary model with a decaying starburst at redshift $z = 3$ (decay time $= 0.4$ Gyr) and a Chabrier initial mass function (Bruzual & Charlot 2003). This model has been described and confirmed in DES data by Zenteno et al. (2016), where they also derived a mean faint-end slope $\beta = -1.06$, which we adopt. Note that the depth of the r band allows to reach the m$*$+1.5 limit over the entire redshift range covered by our search. This threshold is used only to define the λ_* mass proxy and does not enter in the cluster detection.

For the radial profile $\Psi(r)$, we assume a Navarro–Frenk–White profile (NFW; Navarro, Frenk & White 1997):

$$\Psi(r) = \frac{C_0}{\frac{r}{r_s}\left(1 + \frac{r}{r_s}\right)^2}, \tag{16}$$

where the scale radius depends on the concentration c via $r_s \equiv R_{200}/c_{200}$. Hennig et al. (2017) found that the NFW distribution is a good description of the observed radial density distribution of galaxies in clusters, with a mean $c = 3.59$. For our cluster model, we used this value and an R_{200} corresponding to a mass $M_{200} =$ $10^{14}\,\mathrm{M}_\odot\,h^{-1}$, typical for the cluster sample we want to target. The normalization parameter C_0 is such that the total number of galaxies N_{200} inside R_{200} and below $m_\star + 2$ is coherent with the relation with M_{200} found by Hennig et al. (2017). For a mass $M_{200} = 10^{14}\,\mathrm{M}_\odot\,h^{-1}$, this corresponds to $N_{200} = 22.9$. The choice of a fixed R_{200} lowers the slope of the mass proxies scaling relations. This is not an issue because the scatter around the scaling laws is small as shown in Bellagamba et al. (2018a).

[3]We assume a ΛCDM cosmological model with $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, and $h = 0.7$ throughout the paper.

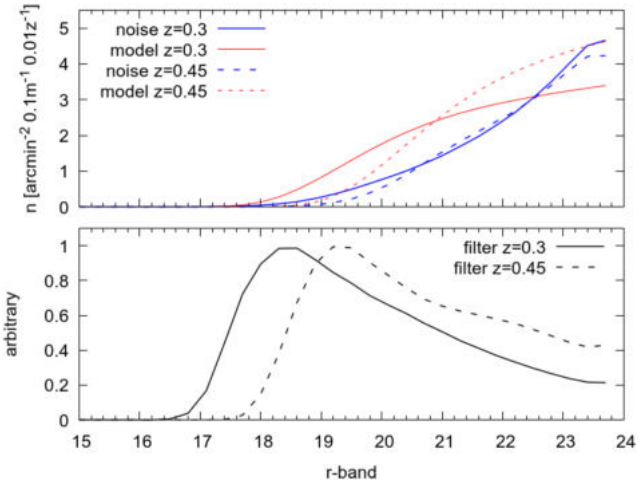

Figure 1. Top panel: the luminosity function of the model of the field galaxies in blue (the noise component) and the one of the cluster members in red (the cluster model) at redshift $z = 0.3$ (solid lines) and at redshift $z = 0.45$ (dashed lines). Bottom panel: the magnitude distribution of the resulting filters for the same two redshifts.

The field galaxies distribution $N(m,z_c)$ can be approximated by the total distribution in the galaxy sample, as the cluster component is small. For each redshift z_c, we build $N(m,z_c)$ weighing each galaxy by its redshift probability distribution $p(z_c)$.

For illustration purposes, we show in the top panel of Fig. 1 the luminosity function of both field galaxies and cluster members at redshift $z = 0.3$ (solid lines) and $z = 0.45$ (dashed lines). The magnitude dependence of the filter resulting from the use of these luminosity functions is shown in the bottom panel of the same figure. The filter turns out to be a band-pass filter that gives more weight to the galaxies on the bright-end side: the higher the redshift, the larger the magnitude where the filter peaks at.

4 CLUSTERS DETECTIONS AND GALAXY MEMBERS

In this section, we describe the catalogue of galaxy clusters detected in the KiDS-DR3 with the AMICO code. We discuss the main statistical properties of the sample and present few examples of detections.

4.1 The catalogue of galaxy clusters

The survey covers a total area of 438 deg^2, but we rejected all galaxies falling in those regions severely affected by the haloes produced by bright stars ('Primary halo' masks, see de Jong et al. 2015) leaving us with 414 deg^2. The remaining galaxies have been used to produce an initial set of detections down to $S/N \geq 3.0$ as explained in Section 3. In a second step, this first sample of detections has been filtered out by rejecting all objects falling in the more restrictive masks used by the KiDS collaboration for

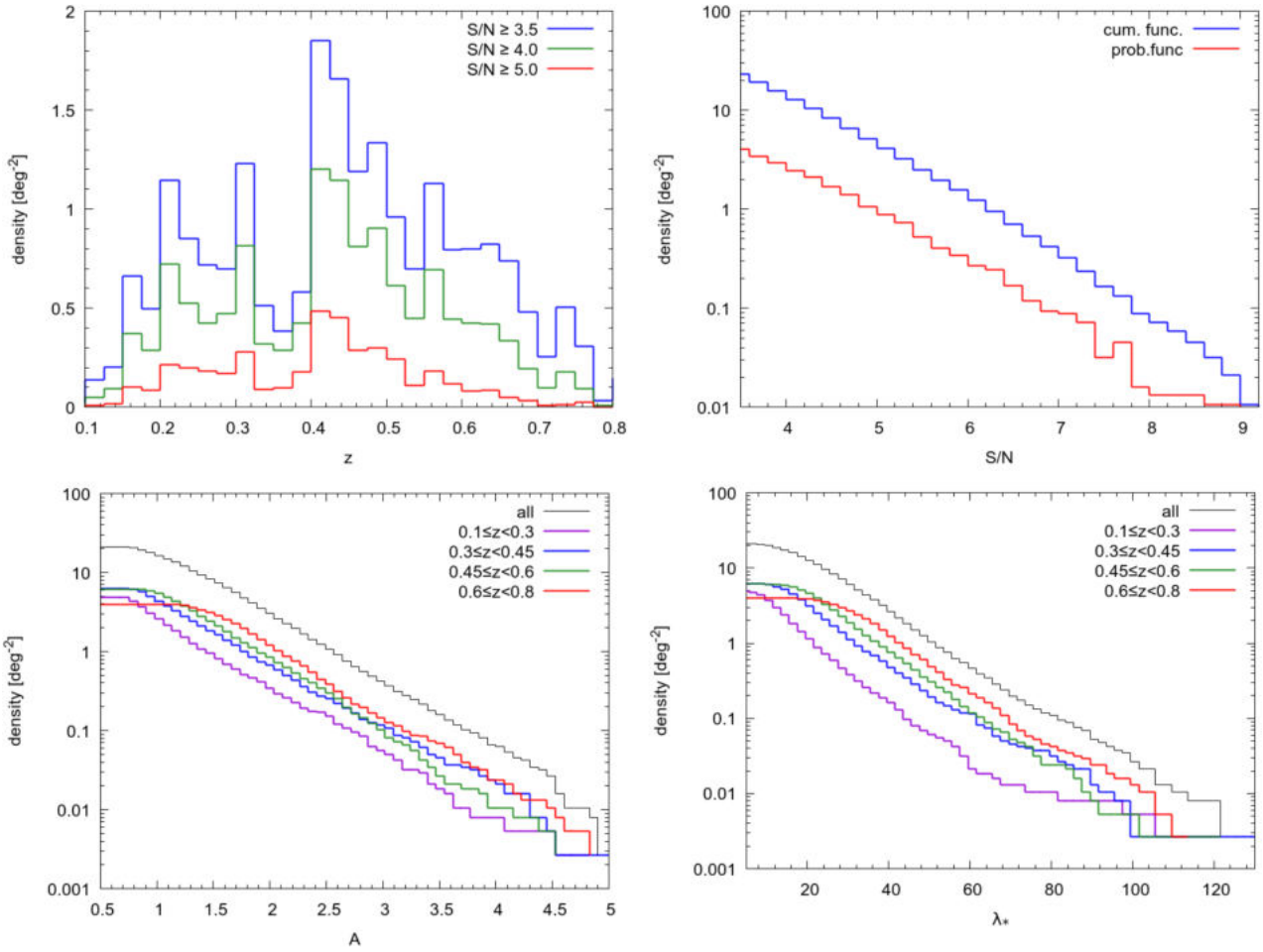

Figure 2. The cumulative distributions of the number density of detections as a function of various properties. In the top left-hand panel, as a function of redshift and for three different *S*/*N*s; in the top right-hand panel, the number density distribution (red curve) and its cumulative (blue curve) as a function of *S*/*N*; in the bottom panels, as a function of amplitude *A* (left) and intrinsic richness λ_* (right).

the weak lensing analysis ('Secondary/tertiary halo' masks, see de Jong et al. 2015). The final effective area is of 377 deg^2, i.e. 86 per cent of the total area of the survey. All detections with $S/N > 3.0$ have been used for the construction of the mock simulations discussed in Section 6.1, but for the final catalogue of clusters we kept only those with $S/N > 3.5$ obtaining a final sample of 7988 cluster candidates.[4] This buffer in *S*/*N* is necessary for constructing reliable mock catalogues and derive solid statistical properties for our cluster sample as it will be detailed in Section 6.1.2. The entries of the catalogue are specified in Table 3. For all 6972 objects falling in the redshift range $0.1 < z < 0.6$, we also provide mass estimates using the scaling relations obtained via stacked weak gravitational lensing (Bellagamba et al. 2018a).

In Fig. 2, we summarize the main statistical properties of the detections listed in the catalogue by showing their number density as a function of redshift (top left-hand panel), the *S*/*N* (top right-hand panel), and amplitude *A* and intrinsic richness λ_* (left and right bottom panels, respectively). The drop in density at $z \sim 0.38$ is not caused by a degradation of the detection efficiency but rather by a redshift misplacement, i.e. clusters located at $z \sim 0.38$ are detected at a slightly higher redshift. In fact, the galaxies photo-*z*s delivered by BPZ suffer from the non-optimal covering of the 4000 Å break at that redshift caused by the shape of the *g* and *r* filters (see e.g. Padmanabhan et al. 2005). This also explains the sharp peak at z 0.41.

Finally, Fig. 3 shows the amplitude *A* against the intrinsic richness λ_* for the galaxy clusters falling in four different redshift bins. The mass reported on the secondary axes is derived from the scaling relations of λ_* and *A* based on weak lensing measurements, at a fiducial value of $z = 0.35$[5] (Bellagamba et al. 2018a).

As an example of detections, we show in the left-hand panels of Fig. 4 the image cut-outs of a rich cluster, with $\lambda_* = 121$, at relatively low redshift, $z = 0.28$ (top row), and an intermediate one, with $\lambda_* = 49$, at a higher redshift, $z = 0.69$ (bottom row). The first cluster has been detected with $S/N = 10.4$, while the second one has $S/N = 4.6$. Both cut-outs are of 400 arcsec in size. In the corresponding right-hand panels, we show how the AMICO code 'sees' the same two clusters. The circles mark the position of the galaxies in the image cut-outs on their left, and their size and colour represents the

[4]The catalogue is available on request.

[5]We verified that the redshift evolution of the scaling relation is small and can be neglected for this purpose.

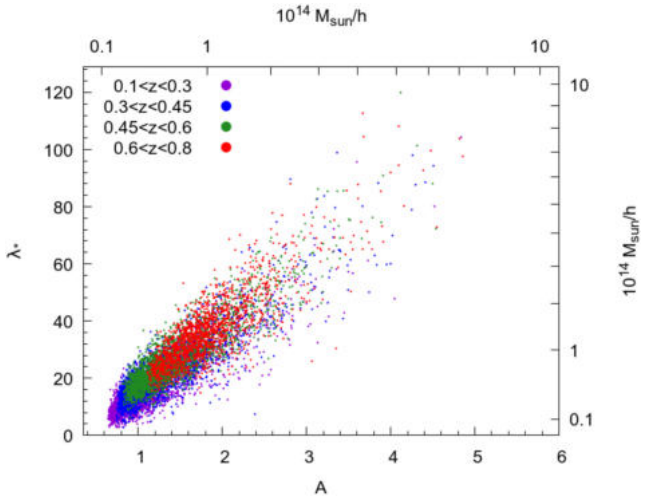

Figure 3. Correlation between the amplitude A as returned by the matched filter and the intrinsic richness λ_* defined as the sum of the membership probabililistic association of the galaxies to clusters (within $r < r_{200}$ from the detection center and with $m < m_* + 1.5$). Each cluster is colour-coded according to its redshift, as labelled in the figure. The masses indicated in the secondary axes come from the scaling relations derived in Bellagamba et al. (2018a).

galaxies probabilistic association to the detection. The plot shows clearly the dependence of the membership probabilistic association on the angular distance from the detection center. However, the value of $P_i(j)$ has also a dependence on magnitude and on the $p(z)$ of the galaxy (see equation 4). Other examples are shown in Appendix (A).

4.2 The catalogue of cluster members

As discussed in Section 3.1, AMICO returns a probabilistic association of galaxies to cluster candidates, see equatiton (4). Using this information, we present a catalogue of cluster members as well as their probability to be field galaxies, $P_{f,i}$. We note that in the catalogue we do not apply any cut in the membership probability but for each galaxy we store a maximum of 20 associations to clusters.[6] This catalogue can be used to study the properties of clusters, and their connection with the process of galaxy formation. Moreover, it can help in the removal of the foreground for weak lensing studies, and in improving the strong lensing estimates derived with photometric-based algorithms (Stapelberg, Carrasco & Maturi 2017; Carrasco et al. 2018), among many other applications.

5 CORRELATION WITH OTHER DATA SETS

In this section, we compare our sample with other catalogues of galaxy clusters published in the literature. The match of our detections with those in external catalogues is based on the following conditions on the difference in redshift and projected distance: $\Delta z = 0.1$ and $\Delta R = 1\mathrm{Mpc\,h^{-1}}$.

[6]This limit is reached in very rare cases.

5.1 RedMaPPer optically selected clusters

We compared our detections with the cluster candidates listed in Rykoff et al. (2014). This catalogue has been obtained by running the redMaPPer algorithm on the Sloan Digital Sky Survey Data Release 8 (SDSS-DR8) and contains entries within the redshift range $0.08 < z < 0.55$. Of the 681 detections falling in the unmasked regions of KiDS-DR3 data, 624, i.e. 92 per cent, find a match. Of the remaining 57 detections, 20 have been detected by AMICO with an $S/N > 3$ but below our restrictive threshold of $S/N = 3.5$, while other 13 satisfy the matching criteria but lay inside a KiDS 'Secondary/tertiary halo' mask. The remaining 24, i.e. less then 4 per cent of the redMaPPer sample, do not have a counterpart with $S/N > 3$. These redMaPPer cluster candidates have a richness measured by redMaPPer, λ_{RM}, close to the detection limit of this algorithm as illustrated in Fig. 5. Since the KiDS and the SDSS data sets are very different in terms of depth and image quality, no further comparison would be meaningful.

5.2 Planck SZ-selected clusters

Of all clusters listed in the second Planck catalogue of SZ sources (Planck Collaboration XIII 2016a), 19 fall in the area we processed, and all of them have been detected by AMICO. The redshift of 10 of these objects has not been reported in the literature to our knowledge and for this we list in Table 2 our estimates as measured with the AMICO code. Note that the cluster with $\mathrm{ID_p} = 1606$ has not been detected with redMaPPer because located at a redshift, $z = 0.77$, which exceeds their maximum limit.

5.3 MCXC X-ray selected clusters

We then compared our mass proxies with the X-rays mass estimates listed in the Meta-Catalog of X-ray detected Clusters of galaxies (MCXC; Piffaretti et al. 2011). The MCXC catalogue comprises X-rays selected clusters collected in archival data and includes the *ROSAT* All Sky Survey-based (NORAS, REFLEX, BCS, SGP, NEP, MACS, and CIZA) and serendipitous cluster catalogues (160SD, 400SD, SHARC, WARPS, and EMSS) for a total of 1743 objects. Since the data have been taken with different instruments and exposure times, they have been homogenized in order to provide a coherent picture for the whole sample. All 13 clusters of the MCXC catalogue falling in the KiDS-DR3 foot-print have been identified by AMICO. In Fig. 6, we compare our two mass proxies, A and λ_*, with the value of M_{500} derived with the X-Rays observations. A well-defined correlation is evident.

5.4 GAMA spectroscopy

We finally used the GAMA-I galaxy group catalogue ($\mathrm{G^3C}$; Driver et al. 2009, 2011; Liske et al. 2015) to verify the redshift estimate of the clusters provided by the AMICO code. GAMA is a highly complete spectroscopic survey up to a Petrosian r-band magnitude of 19.8 and comprises 110 192 galaxies, 40 per cent of which belong to 14 388 galaxy groups identified with a friends-of-friends algorithm in the redshift range $0 \leq z \leq 0.5$ (Robotham et al. 2011). In Fig. 7, we plot the relative scatter between the redshift estimates of the groups as measured by AMICO and those listed in the $\mathrm{G^3C}$ catalogue. A clear bias, well described by $\Delta z/(1 + z) \sim 0.02$, emerges from this comparison. This bias corresponds to what was found by de Jong et al. (2017) when comparing the KiDS photo-zs to the GAMA spec-z (see their table 8). Since the sample of the $\mathrm{G^3C}$

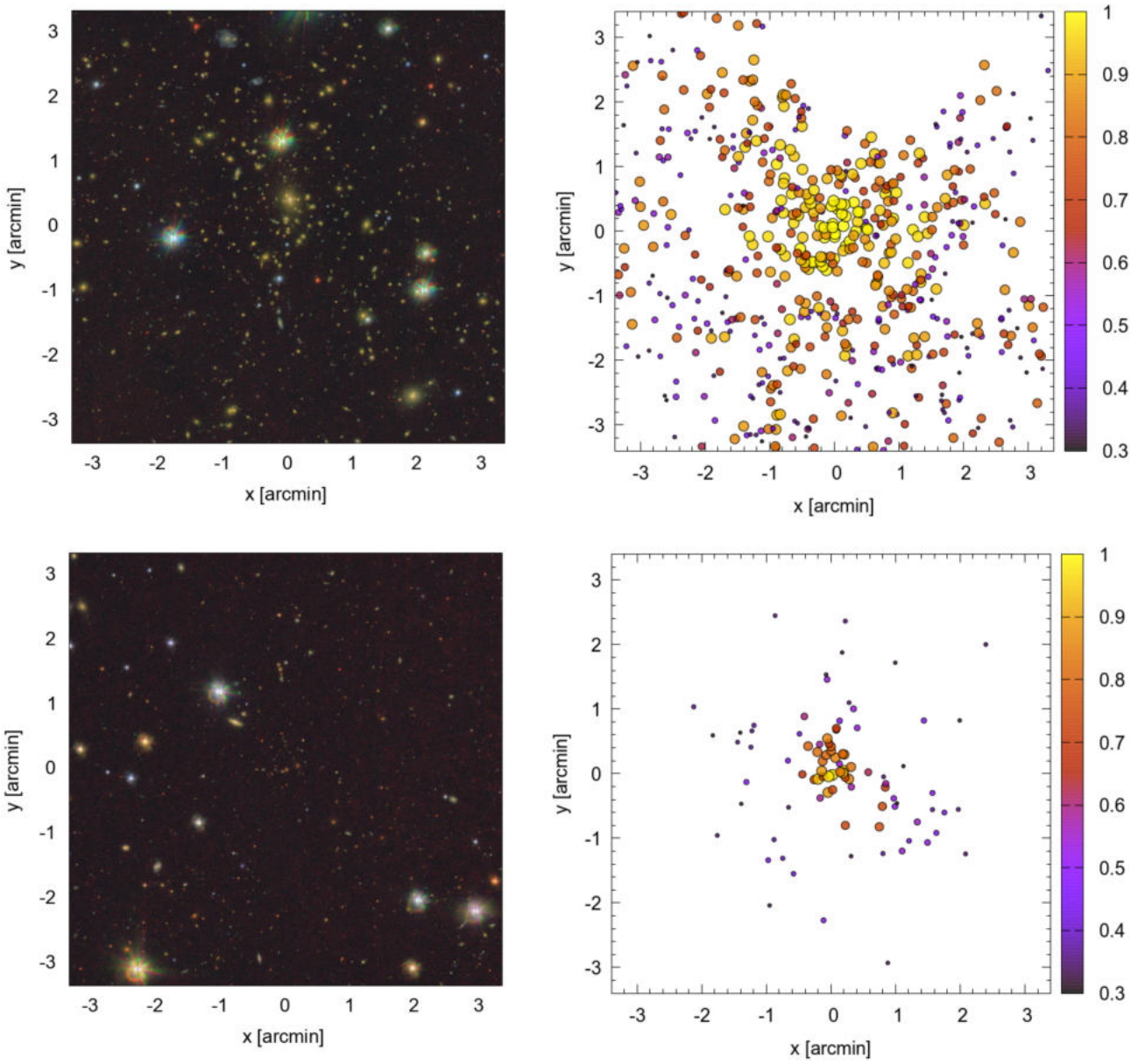

Figure 4. In the left-hand panels, we show the colour composite (*g, r, i*) image of a rich cluster, with $\lambda_* = 121$, at $z = 0.28$, detected with $S/N = 10.4$ (top panel) and a smaller one, with $\lambda_* = 49$, at a higher redshift, $z = 0.69$, detected with $S/N = 4.6$ (bottom panel). Both stamps are centred at the position identified by AMICO and have a size of 400 arcsec on a side. In the corresponding right-hand panels, we show the galaxies in these fields as circles, the size of which is proportional to the probabilistic association to the detection. The numerical values of the probability are colour-coded as in the bar on the right.

is limited to $z < 0.5$, we cannot draw any conclusions for clusters at higher redshifts. More details regarding this bias and how we deal with it will be given in Section 6.2 together with an extensive discussion of all other uncertainties.

6 ASSESSING THE QUALITY OF THE DETECTIONS

In this section, we describe a method to produce realistic mock catalogues of galaxies that we use to estimate the uncertainties of the quantities characterizing the detections as well as the purity and completeness of the entire cluster sample. These mocks are completely phenomenological, i.e. they do not rely on any (astro)physical assumption and are solely based on the original KiDS data, whose properties they aim at reproducing.

6.1 Mock simulations

We base our mocks on the original KiDS-DR3 data rather than on numerical simulations; this is done to minimize the number of assumptions and to account for all expected and unexpected properties of the survey across the sky, such as photometric and photo-z uncertainties, absorption, masks, variation in depth as well as the clustering of galaxies etc. The central idea of the mocks relies on a Monte Carlo extraction of the galaxies based on their probabilistic association to the entries in our cluster sample,

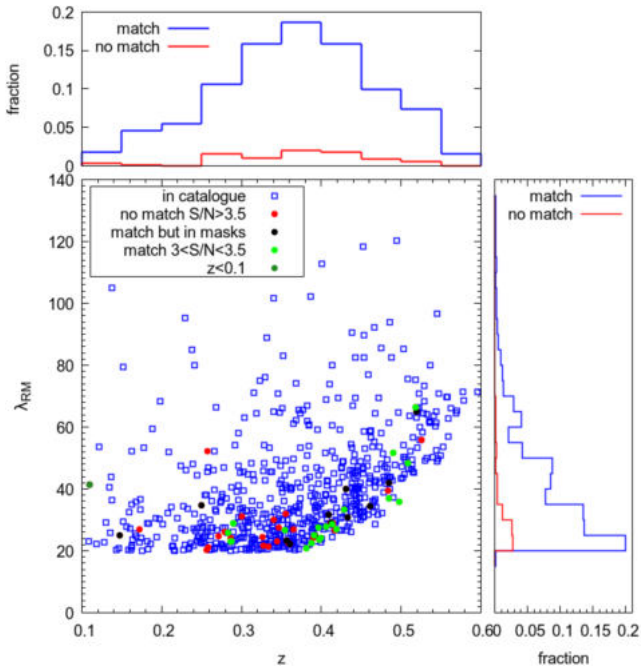

Figure 5. The correlation of richness λ_{RM} and redshift z, as measured with redMaPPer, of detections matched (blue curves) and non-matched (red curves) to our sample. Most of the non-matched detections are close to the detection limit of redMaPPer. Some of the detections are matched but they fall at the very border of masked areas (black circles), above $S/N > 3$ but below the minimum $S/N = 3.5$ we considered (green circles) or because outside the redshift range we imposed on our sample (the only dark green circle on the left). The distribution of the detections with respect to z and λ_{RM} are plotted in the top and left-hand panels, respectively.

Table 2. The identification number ID, the sky coordinates (RA and Dec.) and the redshift z for the clusters of galaxies listed in the Planck PSZ2 catalogue for which the redshift information was previously missing. The subscript 'p' indicates the values listed in the Planck catalogue, while 'a' those listed in our catalogue.

ID_p	RA_p	$Dec._p$	ID_a	RA_a	$Dec._a$	z_a
11	358.351	−33.2932	7485	358.385	−33.2837	0.67
19	350.538	−34.5752	6948	350.468	−34.6173	0.23
39	354.054	−32.1343	7226	354.053	−32.1320	0.41
41	342.976	−33.3942	6473	342.963	−33.4027	0.24
44	356.853	−31.1509	7361	356.884	−31.1470	0.45
48	341.633	−32.2011	6295	341.717	−32.2280	0.50
50	351.169	−30.6511	7002	351.156	−30.6723	0.31
59	340.637	−30.3150	6241	340.618	−30.4084	0.24
1033	39.6692	−30.8391	7601	39.6215	−30.8968	0.55
1606	216.108	−2.73976	4120	216.089	−2.83463	0.77

equation (4), and to the field, equation (5). A scheme with all steps involved in the mock generation is shown in Fig. 8.

In the following, we describe the procedure to create the mock field population (Section 6.1.1) and the mock clusters (Section 6.1.2). We will discuss several other aspects and considerations related to the methodology in Section 6.1.3.

6.1.1 *Mock field galaxies*

The field galaxies are extracted from the KiDS data catalogue via a Monte Carlo sampling based on the probability, $P_{f,i}$. In detail, for each galaxy we extract a uniform random number r_i between 0 and 1, and assign the galaxy to the field if $r_i < P_{f,i}$. For example, a galaxy with $P_f = 0.32$ has the 32 per cent chance to be extracted and associated with the field. All observed properties of these

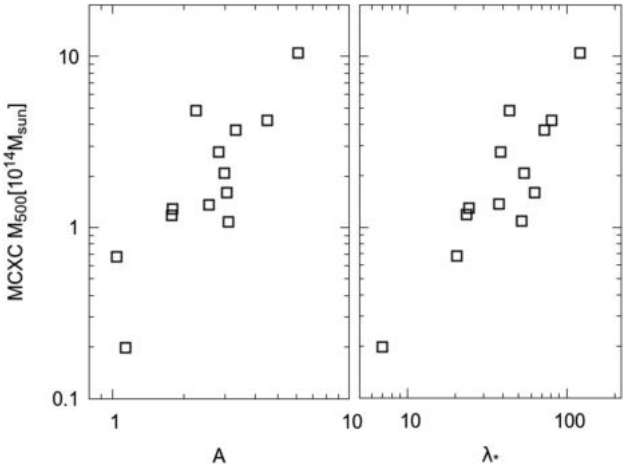

Figure 6. Correlation between the mass estimates reported in the MCXC catalogue against the amplitude A (left-hand panel) and against the intrinsic richness λ_* (right-hand panel).

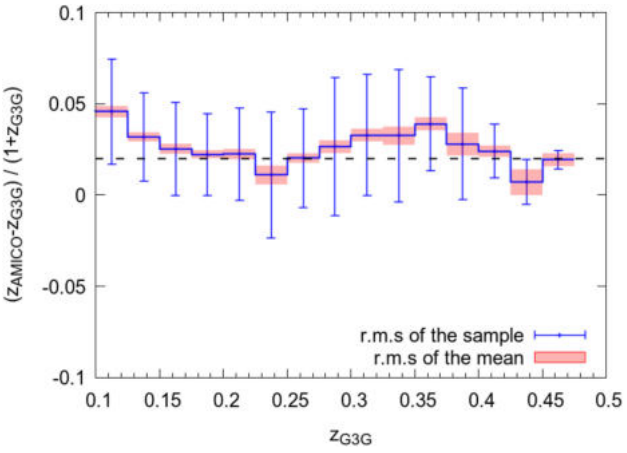

Figure 7. The difference between the clusters redshifts measured with AMICO and the spectroscopic ones taken from the GAMA G^3C catalogue. The bias affecting the KiDS photometric redshifts around is evident.

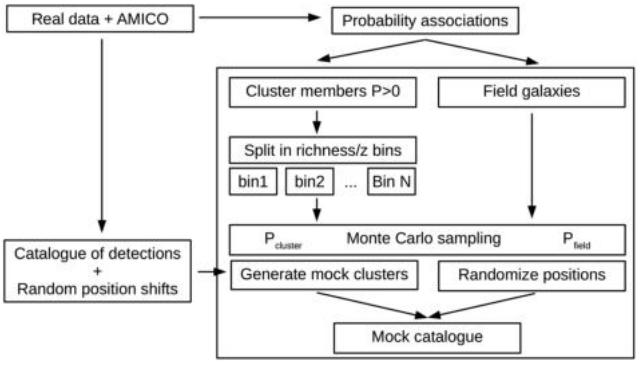

Figure 8. Flow chart showing the process used to create the mock simulations.

galaxies are preserved except for their position in the sky, which is slightly perturbed by introducing a random angular displacement. The maximum random displacement is a free parameter that we set to $r_{\rm rnd} \leq 1\,{\rm Mpc\,h^{-1}}$. This scale is large enough to dump the presence of clusters, which might have not been detected by the algorithm, but it is small enough to preserve the correlation of the LSS on larger scales and, more importantly, the local number density of the field galaxies, which varies across the survey.

6.1.2 Mock galaxy clusters

To generate the mock clusters, we started by defining bins of apparent richness λ and redshift z in which we collect all galaxies associated with clusters. All galaxies with $P_i(j) > 0$ have been considered and those with more than one cluster association have been attributed to more than one bin accordingly. In this way, each bin contains all galaxies potentially belonging to clusters with the richness and redshift defining the bin itself. Each mock cluster is then generated by randomly extracting galaxies out of the corresponding bin via a Monte Carlo sampling based on the cluster membership probability (see equation 4). In doing so, we account for the presence of the masked areas in the actual survey. The number of visible members assigned to each mock cluster is λ by definition, i.e. $\lambda_{\rm expected} \equiv \lambda \in \mathbb{N}$. The other mass proxies are univocally related to the cluster members through their own definition:

$$A_{\rm expected}(\vec{\theta}_{\mathcal{C}}, z_{\mathcal{C}}) = \alpha^{-1}(z_{\mathcal{C}}) \sum_{k}^{\lambda} \frac{C(z_{\mathcal{C}}; r_k, m_k)}{N(m_k, z_{\mathcal{C}})}\,, \tag{17}$$

and

$$\lambda_{*\rm expected} = \sum_{k}^{\lambda} \delta(m_k, r_k)\,, \tag{18}$$

where $z_{\mathcal{C}}$ is the redshift of the mock cluster, r_k is the distance of the member from the centre of the mock cluster, $\delta(m_k,r_k) = 1$ if $m_k < m_*(z_{\mathcal{C}}) + 1.5 \ \wedge \ r_k < R_{200}(z_{\mathcal{C}})$ and $\delta(m_k, r_k) = 0$ otherwise. In short, the resulting mock cluster is a random realization based on the overall statistical properties of all original detections with similar λ and z.

The mock clusters are then injected into the field to maintain the angular position, apparent richness, and redshift of all cluster detections with $S/N \geq 3$ found in the original KiDS data. In this way, we avoid any assumption, and we rely solely on the statistics of the data in terms of the correlation of clusters with the LSS and of clusters with clusters, as well as blending, missing data, non-uniform absorption across the survey, photometric, and photo-z uncertainties. Clusters that are far and/or small enough to have $S/N < 3$ are not generated in our mocks. This does not have a substantial impact on the results because the final catalogue is limited to the detections with $S/N \geq 3.5$. In fact, even if objects with lower S/N would be generated, their probability to exceed the $S/N = 3.5$ threshold, when measured in the mocks, would be very small so that their contribution to the final sample would be anyway negligible.

In total, the simulations contain 9018 mock-clusters over 200 fields covering a total area of 189 deg^2.

6.1.3 Further considerations

This mock generation has the advantage to fully preserve all statistical properties of the original data by construction. The overall process boils down to a data bootstrapping based on the probabilities $P_i(j)$ and $P_{i,f}$, which obviously sum up to unity. In other words, the mock catalogue is a random realization based on the original data where only the galaxies are labelled as cluster members or field galaxies. The only assumption within this procedure is hidden in the membership probabilistic association that directly depends on the cluster model used to define the filter, see equation (4). In spite of that, such dependence of the mock clusters on the assumed model is softened by the fact that the mock generation goes through the original catalogue once more after the detection process has been completed. This reiteration through the data helps reproduce the properties of the actual clusters present in the data. For example, let us consider an extreme case in which the cluster model has a flat density radial profile. In this scenario, the membership probabilistic association of galaxies to clusters has no dependence on the position of galaxies. Galaxies with different distances from the cluster centre are equally likely to be extracted during the Monte Carlo sampling. Nevertheless, the pool of galaxies out of which we extract the members has a population that follows the actual density profile of clusters, for the simple reason that they are the actual galaxies in the data. Therefore, there are more galaxies to be sampled close to the centre rather than in the outskirts, and hence the mock clusters will have a radial density profile closer to reality than to the initial assumption. The mock clusters are not a mere representation of the model.

We show in the top left-hand panel of Fig. 9 the amplitude map for a slice at redshift $z = 0.35$ resulting from the analysis of about 1 deg^2 per side of the real KiDS data (top left-hand panel). The green contour lines represent the areas masking the artefacts caused by bright stars. In the top right-hand panel, we show the same analysis but performed on the corresponding mock catalogue of galaxies. All main features are clearly preserved. The differences between the left and right maps are both due to the displacement of the galaxy positions and to the Monte Carlo sampling process. In practice, our mock map is a realization of the population of galaxies, statistically independent of the original one. In the two bottom panels of the same figure, we plot with two different colour scales the contribution to the amplitude of the mock field galaxies alone. The two top panels and bottom left one have the same colour scale, while in the right bottom panel we stretched the colour contrast to better highlight the details such as the LSS pattern.

6.2 Uncertainties on the detections properties

The uncertainties on the properties of the detections (position, redshift, amplitude, richness, etc.) are evaluated by analysing the mock catalogues with the AMICO code as done with the real data and by comparing the measured values with the expected ones. One detection is assigned to one mock cluster present in the simulations if it lays within a cylinder of $\Delta r \leq 1\,{\rm Mpc}\,h^{-1}$ in radius and of $\Delta z = 0.1$ in length. The detections without a match are considered spurious and allow us to derive the purity of the sample. The results are presented in Fig. 10 and are discussed next.

Angular position: The scatter along the declination of the detections, $\Delta Dec.$, is larger for clusters at lower redshift. This is because of their larger angular extension, which does not allow for a well-defined localization of their centre. At higher redshift, $z > 0.45$, the angular resolution is dictated by the pixel size we have chosen for our maps, which is of $\sim$0.3 arcmin. Since the scatter along RA is identical to the one along Dec., we do not display it.

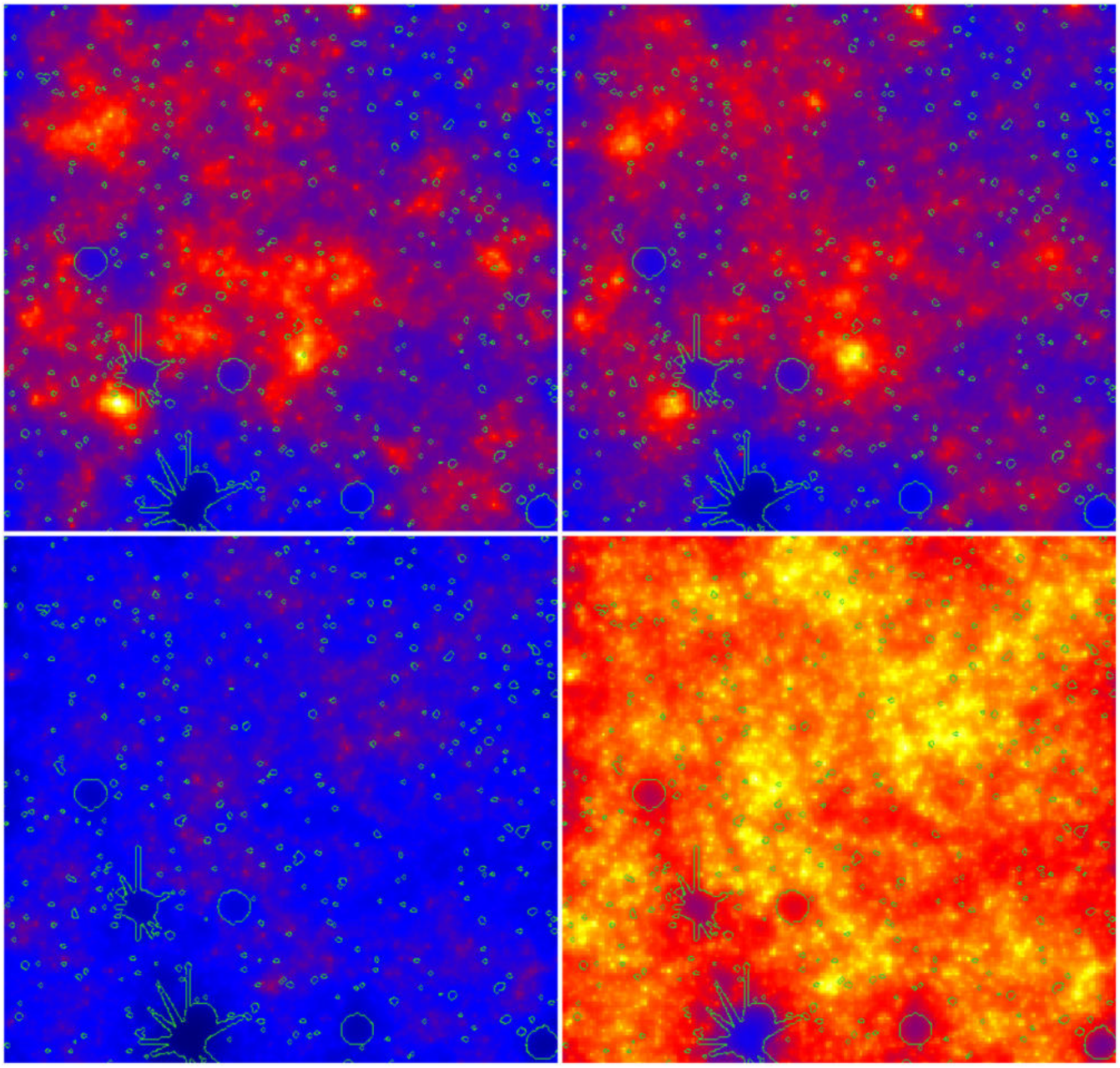

Figure 9. Map of 1 deg^2 per side of the amplitude A at redshift $z = 0.35$ for the original KiDS data (top left), the mock catalogue with both field and cluster galaxies (top right), and the mock catalogue with the field galaxies only, with the same colour scale of the top panels (bottom left) and with a colour scale stretched to better highlight the details (bottom right). The green contour lines outline the areas masking the artefacts caused by bright stars.

Redshift: The relative scatter in redshift is constant over the whole redshift interval, and it amounts to $\Delta z/(1 + z) \sim 0.02$, which is smaller than one of the galaxy population $\Delta z_{\rm gal}/(1 + z_{\rm gal}) \sim 0.044$, as expected (de Jong et al. 2017). The error estimate obtained with the mocks agrees very well with the scatter of our measures around the spectroscopic redshifts of the G^{3}C sample, see Section 5.4. The known bias of the galaxy photo-z already discussed in de Jong et al. (2017) surely affects the redshift estimates of the detections, but this does not appear in the top left-hand panel of Fig. 10. This is because the reference redshift of the mock clusters has been taken from the data and is therefore dragged by the very same bias. Nevertheless, this test serves to estimate the redshift uncertainty in a reliable way and shows that AMICO returns unbiased results with respect to the input photo-z catalogue. We decided not to apply any bias correction to the redshifts of our catalogue because more data against which to calibrate will be available and the accuracy of the photo-z of the galaxies might improve with time. The final redshift correction is thus left to the user: Based on the analysis discussed in Section 5.4, we suggest to apply the following relation to our sample $z_{\rm corrected} = z - 0.02\,(1 + z)$, in agreement with de Jong et al. (2017). Our approach differs from the one adopted in redMaPPer, where such calibration

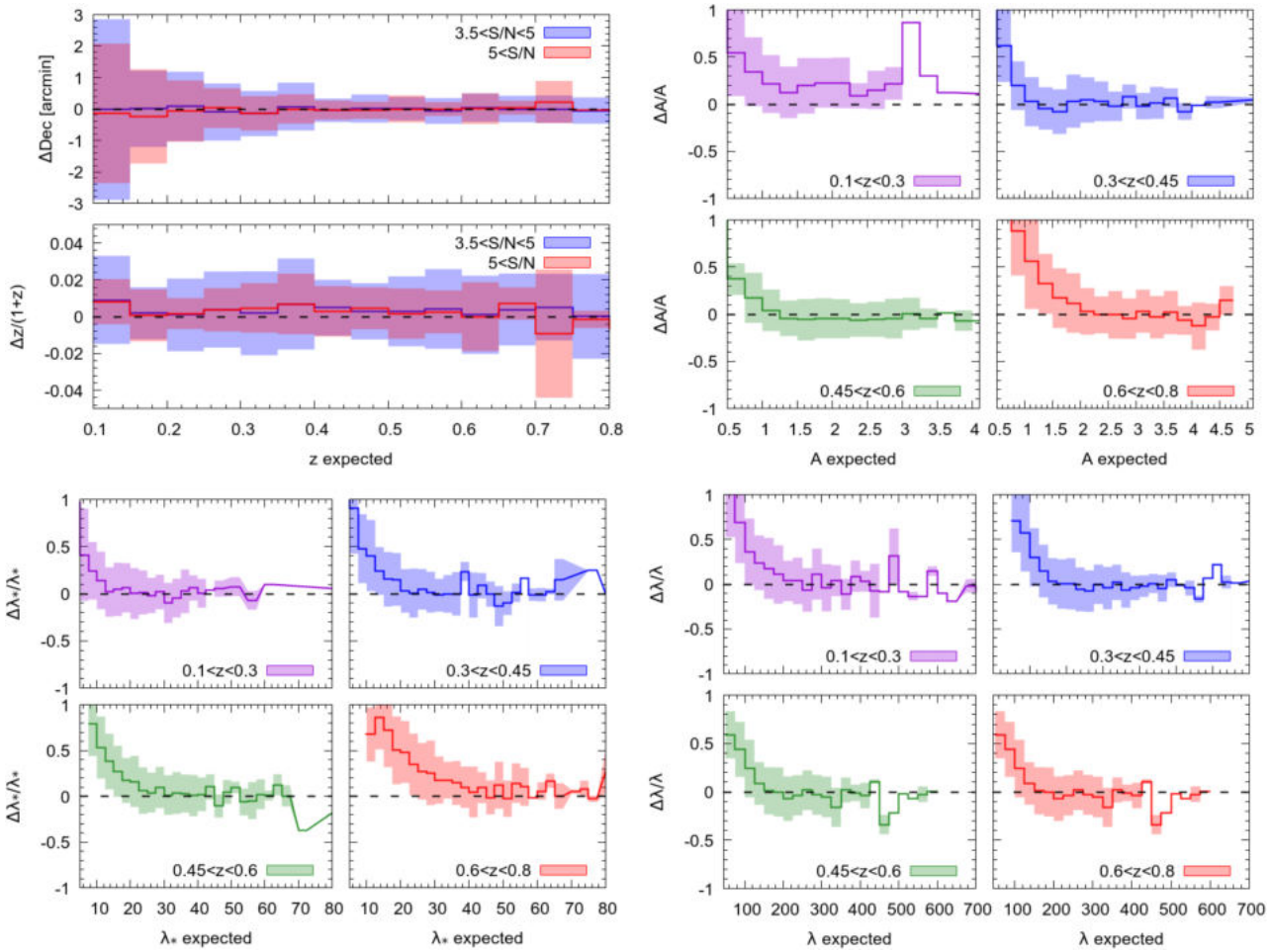

Figure 10. Top left-hand panels: expected statistical 1σ error (shaded areas) on the redshift and angular position (declination) as a function of redshift for two different *S/N* intervals. In the other panels, we show that the relative error of the amplitude A (top right), the intrinsic richness λ_* (bottom left), and the apparent richness λ (bottom right) for four different intervals of redshift.

against external spectroscopic data is done internally (Rykoff et al. 2014).

Amplitude: The relative scatter of the amplitude A is almost constant over the whole range of values and in all four redshift intervals displayed in the top right four panels of Fig. 10. For low values, the sample is incomplete and the amplitude is affected by the Malmquist bias, as expected. Where the sample is complete, the amplitude is unbiased for all redshift bins except for the lower redshift one, where it is biased high by a constant factor. We verified that this is not due to the masks that are more relevant for the lower redshift clusters because of their larger angular extension. The origin of this amplitude bias is not completely understood.

Richness: Similar considerations hold for the intrinsic richness λ_* and apparent richness λ except for the fact that these quantities are unbiased at all redshifts.

Membership probabilistic association: In Fig. 11, we show the fraction of real members as a function of the measured probabilistic association of the galaxies to clusters (see equation 4). Two different intervals of *S/N*, which are $3.5 < S/N < 5$ and $S/N > 5$, are shown. The correlation well satisfies the identity proving that the association of galaxies to clusters is reliably estimated.

6.3 Purity and completeness

The completeness and the purity of the final sample are shown in Fig. 12. The purity is a measure of the contamination level of the cluster sample and is defined as the fraction of detections successfully matched with the clusters in the simulations over the total number of detections. This is shown in the right-hand panels. The sample purity is extremely high. The completeness as a function of amplitude or richness, shown in the left-hand panels, is defined as the number of correctly identified detections over the total number of mock clusters in the simulations. The minimum amplitude A and intrinsic richness λ_* for which the sample is complete grows with redshift. This is because these two quantities are direct mass proxies. Clearly, the larger is the redshift, the larger is the mass necessary for a cluster to reach the minimum *S/N*, $S/N > 3.5$, for its detection. This is not the case for the apparent richness, λ, because it quantifies the number of visible cluster members, which is mostly determined by the depth of the data. This missing or weak redshift dependence makes λ a better probe for cosmological studies.

In Fig. 13, we show the completeness as a function of redshift for different levels of amplitude A, intrinsic richness λ_*, and apparent richness λ. Note that the completeness is measured not by setting

Table 3. Meaning of the different entries listed in the catalogue of galaxy clusters.

SURVEY	Sample identifier: AK3 (AMICO Kids dr3)
NAME	Unique identification name: JHHMMSS.d+DDMMSS
ID	Unique identification number
FIELD	Identification number of the tile in which the detection has been found
XPIX, YPIX, ZPIX	The indexes of the pixel of the amplitude map in which the detection falls
XSKY, YSKY, ZSKY	Sky coordinates RA, Dec., and the redshift corresponding to XPIX, YPIX, and ZPIX
A	Amplitude, A, as defined in equation (6)
LAMBDA	Apparent richness, λ, as defined in equation (12)
LAMBDASTAR	Intrinsic richness, λ_*, as defined in equation (13)
XPIX_ERR, YPIX_ERR, ZPIX_ERR	1σ error of the position in pixel units
XSKY_ERR, YSKY_ERR, ZSKY_ERR	1σ error of the position in RA, Dec., and the redshift
A_ERR	1σ error of the amplitude defined in equation (2)
LAMBDA_ERR	1σ error of λ based on the mock catalogues
LAMBDASTAR_ERR	1σ error of λ_* based on the mock catalogues
SN	S/N based on the amplitude, AMPLITUDE, and its r.m.s, A_ERR
LIKELIHOOD	Likelihood derived in equation (3)
MASKFRAC	Fraction of the detection that is masked
ID_LITERATURE	Identification number for those clusters already listed in the literature

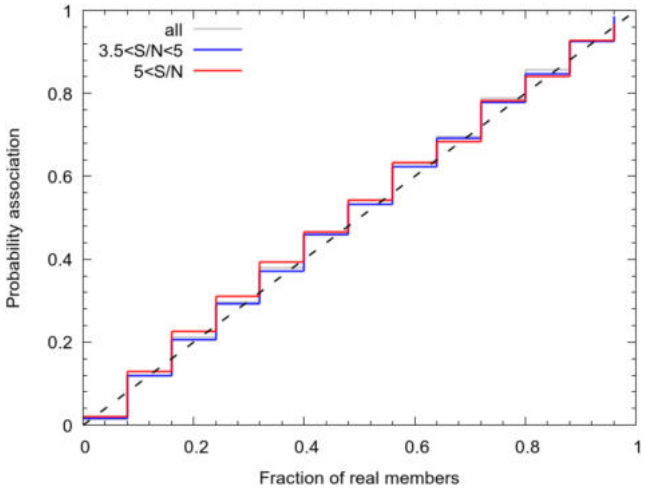

Figure 11. Membership probabilistic association of the galaxies against the fraction of actual members.

a fixed mass threshold for all redshifts but instead it refers to a richness limit which, on average, grows with redshift. This definition is a consequence on how the mocks have been constructed. The population of clusters below such threshold is observationally unaccessible and could only be introduced by assuming, for example, a mass function and a model relating dark matter haloes to visible galaxies, but such a study is not of our interest because it is fully model dependent. What this method is aiming at is a model-independent selection function based on observables only, which can then be used to study, for instance, cosmological models or star formation history. In fact, the mass proxies discussed in this paper can be related to actual masses for a direct comparison with theoretical models because of the scaling relations based on weak lensing mass measurements derived in Bellagamba et al. (2018a).

7 CONCLUSIONS

We detected galaxy clusters in the KiDS-DR3 data with the AMICO code (Bellagamba et al. 2018b). In the analysis, we avoided those regions of the sky affected by the presence of image defects produced, for example, by bright stars, thus covering an effective area of 377 deg^2. With respect to our previous study of the KiDS-DR2 (Radovich et al. 2017), the work presented here takes advantage of the new data and of new features introduced in the AMICO detection algorithm described in this work, which are cleaning procedure, a probability redshift distribution of the filter that now depends on the individual magnitude of the galaxies, and a more robust approach to deal with possible biases in the galaxies photo-zs. We detected 7988 galaxy clusters over a redshift range of $0.1 < z < 0.8$ with a minimum $S/N = 3.5$. The catalogue lists for each detection its unique identification number, sky position, redshift, amplitude A, intrinsic richness λ_*, apparent richness λ, S/N, likelihood $\mathcal{L}$, masked fraction, and its name in the literature if present. In the process, we also derived the probabilistic association of galaxies to each cluster detection. This is a useful information to study the population of galaxies in clusters, galaxy formation in general, or help in the removal of the foreground for weak lensing studies, just to mention some possible applications.

We compared our sample to other catalogues of galaxy clusters overlapping our fields, in particular: (1) We matched the cluster candidates detected by the redMaPPer algorithm on the Sloan Digital Sky Survey Data Release 8 (SDSS-DR8): of the 681 redMaPPer detections, 97 per cent match our sample; (2) we detected all 19 Planck SZ-selected clusters present in our sky area and provide, for the first time, a redshift estimate for 10 of them; (3) we used the X-rays derived masses listed in the MCXC sample of clusters to test our mass proxies: even if the set of common objects is not large enough to achieve conclusive results, the clear correlation we find is extremely encouraging; and (4) finally, we used the GAMA-I galaxy group catalogue (G^{3}C) to verify our redshift estimates: This allowed us to confirm the already known bias in the photo-z that affects the KiDS data and to derive the required calibration for its correction: $\Delta z/(1+z) \sim 0.02$.

Finally, we proposed a new methodology, based on the galaxy membership probabilistic association provided by AMICO, to create realistic mock catalogues. We use them to evaluate the uncertainties of all the properties of the detections, such as their angular position in the sky, redshift, and mass proxies. Most importantly, we use this method to derive the selection function of the sample in a fully model-independent way. As it turned out, the sample has an extremely high purity that approaches 95 per cent over the whole redshift range. This method provides the first step towards the mea-

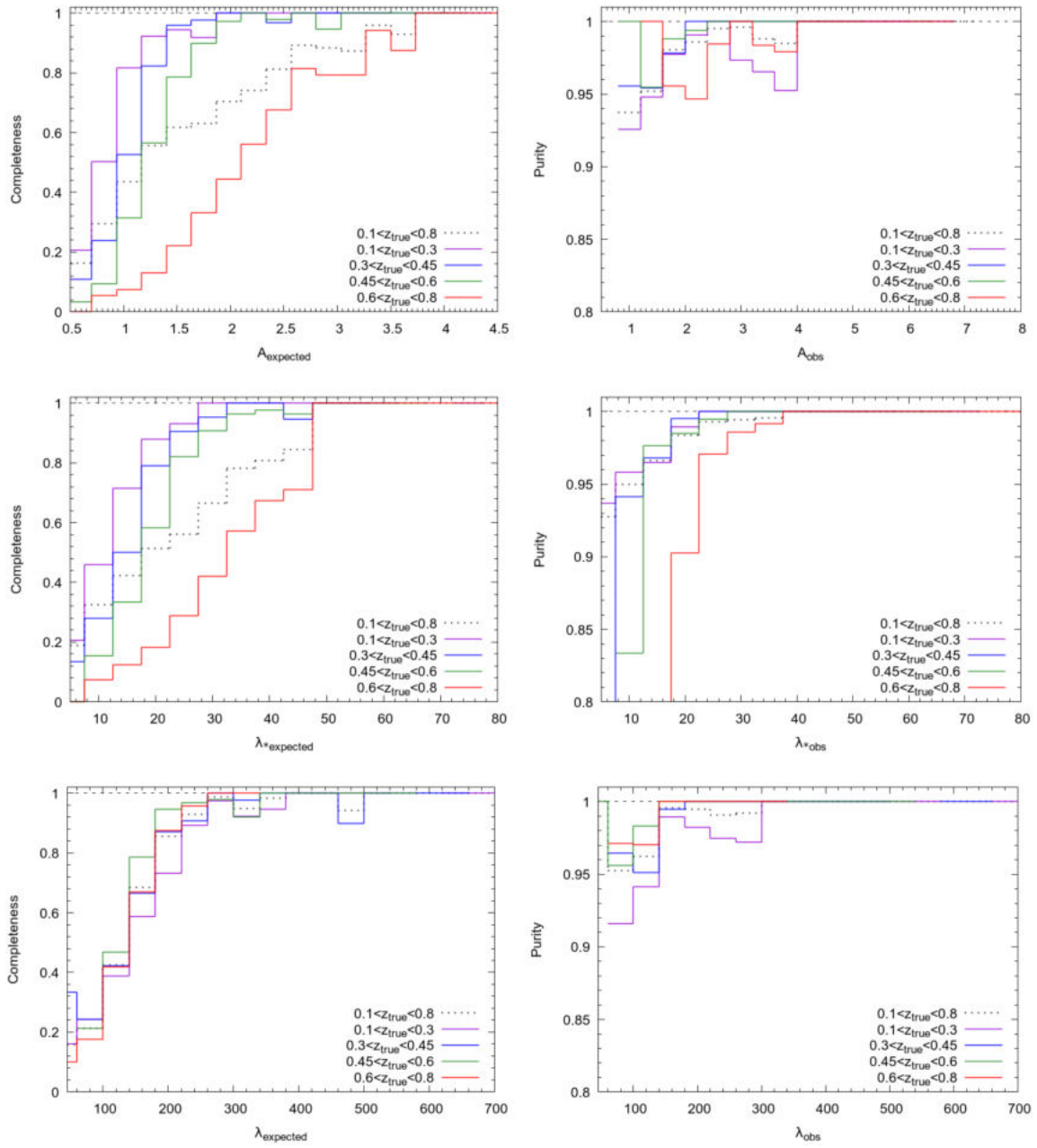

Figure 12. Completeness (left-hand panels) and purity (right-hand panels) for four different redshift intervals (coloured as in the label) as a function of amplitude A, intrinsic richness λ_*, and apparent richness λ (rows from top to bottom, respectively). The completeness of the sample, with respect to the amplitude A and intrinsic richness λ_*, changes with redshift. This is not the case when the apparent richness λ is adopted.

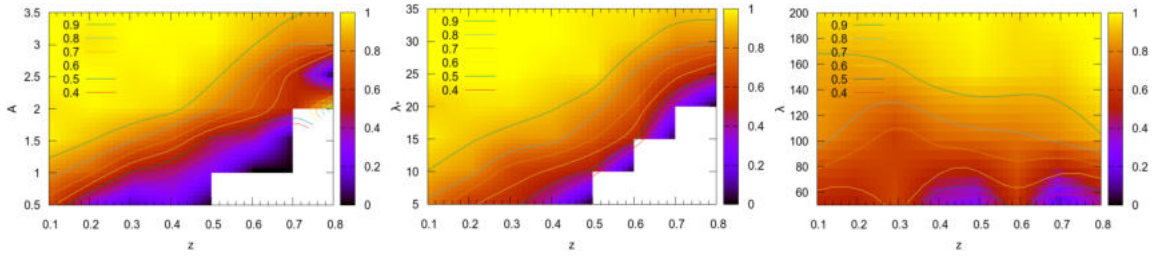

Figure 13. Left-hand panel: completeness of the cluster catalogue as a function of redshift z and amplitude A. Centre and right-hand panels: same as in the left-hand panel but for the intrinsic richness λ_* and the apparent richness λ. The iso-contours trace the completeness from 0.4 to 0.9 as indicated in the figure. Where the completeness is extremely low, there are no sufficient data (white areas).

sure of cosmological parameters through the use of photometrically selected galaxy clusters.

The catalogue of clusters and of the cluster members will be made publicly available but can already be obtained on request, after evaluation.

ACKNOWLEDGEMENTS

This work was supported by the Collaborative Research Center TR33 'The Dark Universe'. Fabio Bellagamba, Mauro Roncarelli, and Lauro Moscardini thank the support from the grants Agenzia Spaziale Italiana n.I/023/12/0 'Attività relative alla fase B2/C per la missione Euclid' and Progetti di Ricerca di Rilevanza Nazionale-Ministero dell'Istruzione dell'Università e della Ricerca 2015 'Cosmology and Fundamental Physics: Illuminating the Dark Universe with Euclid'. Mauro Sereno acknowledges financial support from the contracts Agenzia Spaziale Italiana-Istituto Nazionale di Astrofisica I/009/10/0, NARO15 Agenzia Spaziale Italiana-Istituto Nazionale di Astrofisica I/037/12/0, Agenzia Spaziale Italiana 2015-046-R.0, and Agenzia Spaziale Italiana-Istituto Nazionale di Astrofisica n.2017-14-H.0.

REFERENCES

Adami C., Mazure A., 1999, A&AS, 134, 393
Amaro V. et al., 2018, MNRAS, submitted
Bellagamba F., Maturi M., Hamana T., Meneghetti M., Miyazaki S., Moscardini L., 2011, MNRAS, 413, 1145
Bellagamba F. et al., 2018a, MNRAS, 484, 1598
Bellagamba F., Roncarelli M., Maturi M., Moscardini L., 2018b, MNRAS, 473, 5221
Benítez N., 2000, ApJ, 536, 571
Benoist C., 2014, in Proceedings of a Conference, Building the Euclid Cluster Survey – Scientific Program. p. 8
Bilicki M. et al., 2018, A&A, 616, A69
Böhringer H. et al., 2004, A&A, 425, 367
Bradley L. D. et al., 2014, ApJ, 792, 76
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Capaccioli M., Schipani P., 2011, The Messenger, 146, 2
Carrasco M., Stapelberg S., Maturi M., Bartelmann M., Seidel G., Erben T., 2018, preprint (arXiv:1807.03793)
Clerc N. et al., 2014, MNRAS, 444, 2723
Coe D. et al., 2013, ApJ, 762, 32
Corasaniti P. S., Ettori S., Rasera Y., Sereno M., Amodeo S., Breton M.-A., Ghirardini V., Eckert D., 2018, ApJ, 862, 40
Costanzi M., Villaescusa-Navarro F., Viel M., Xia J.-Q., Borgani S., Castorina E., Sefusatti E., 2013, J. Cosmol. Astropart. Phys., 2013, 012
de Haan T. et al., 2016, ApJ, 832, 95
de Jong J. T. A. et al., 2015, A&A, 582, A62
de Jong J. T. A. et al., 2017, A&A, 604, A134
Driver S. P. et al., 2009, Astron. Geophys., 50, 5.12
Driver S. P. et al., 2011, MNRAS, 413, 971
Farrens S., Abdalla F. B., Cypriano E. S., Sabiu C., Blake C., 2011, MNRAS, 417, 1402
Giocoli C., Moscardini L., Baldi M., Meneghetti M., Metcalf R. B., 2018, MNRAS, 478, 5436
Gonzalez A., 2014, in Proceedings of a Conference, Building the Euclid Cluster Survey – Scientific Program. p. 7
Hennig C. et al., 2017, MNRAS, 467, 4015
Hilton M. et al., 2018, ApJS, 235, 20
Iovino A., 2014, in Proceedings of a Conference, Building the Euclid Cluster Survey – Scientific Program. p. 5
Kelly P. L. et al., 2015, Science, 347, 1123
Kuijken K., 2011, The Messenger, 146, 8
L'Huillier B., Winther H. A., Mota D. F., Park C., Kim J., 2017, MNRAS, 468, 3174
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Licitra R., Mei S., Raichoor A., Erben T., Hildebrandt H., 2016, MNRAS, 455, 3020
Liske J. et al., 2015, MNRAS, 452, 2087
Llinares C., Mota D. F., 2013, Phys. Rev. Lett., 110, 151104
Maturi M., Meneghetti M., Bartelmann M., Dolag K., Moscardini L., 2005, A&A, 442, 851
Maturi M., Schirmer M., Meneghetti M., Bartelmann M., Moscardini L., 2007, A&A, 462, 473
Merloni A. et al., 2012, preprint (arXiv:1209.3114)
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Pace F., Maturi M., Meneghetti M., Bartelmann M., Moscardini L., Dolag K., 2007, A&A, 471, 731
Pace F., Maturi M., Bartelmann M., Cappelluti N., Dolag K., Meneghetti M., Moscardini L., 2008, A&A, 483, 389
Padmanabhan N. et al., 2005, MNRAS, 359, 237
Piffaretti R., Arnaud M., Pratt G. W., Pointecouteau E., Melin J.-B., 2011, A&A, 534, A109
Planck Collaboration XIII, 2016a, A&A, 594, A13
Planck Collaboration XXIV, 2016b, A&A, 594, A24
Radovich M., Puddu E., Bellagamba F., Moscardini L., Roncarelli M., Getman F., Grado A., 2016, Astrophysics and Space Science Proc., Vol. 42, The Universe of Digital Sky Surveys. Springer International Publishing, Switzerland, p. 189
Radovich M. et al., 2017, A&A, 598, A107
Reichardt C. L. et al., 2013, ApJ, 763, 127
Robotham A. S. G. et al., 2011, MNRAS, 416, 2640
Roncarelli M., Carbone C., Moscardini L., 2015, MNRAS, 447, 1761
Rydberg C.-E., Whalen D. J., Maturi M., Collett T., Carrasco M., Magg M., Klessen R. S., 2018, preprint (arXiv:1805.02662)
Rykoff E. S. et al., 2014, ApJ, 785, 104
Sartoris B. et al., 2016, MNRAS, 459, 1764
Schechter P., 1976, ApJ, 203, 297
Schellenberger G., Reiprich T. H., 2017, MNRAS, 471, 1370
Smith G. P. et al., 2016, MNRAS, 456, L74
Stapelberg S., Carrasco M., Maturi M., 2019, MNRAS, 482, 1824
Umetsu K. et al., 2014, ApJ, 795, 163
Viola M., Maturi M., Bartelmann M., 2010, MNRAS, 403, 859
Wang B., Abdalla E., Atrio-Barandela F., Pavón D., 2016, Rep. Prog. Phys., 79, 096901
Zenteno A. et al., 2016, MNRAS, 462, 830
Zheng W. et al., 2012, Nature, 489, 406

APPENDIX A: SOME EXAMPLES OF DETECTIONS

Here, we show a sample of 12 detections located at four different redshifts ($z = 0.2$, 0.3, 0.5, and 0.7, rows from top to bottom), and three different intrinsic richnesses ($\lambda_* \sim 100$, 50, and 5, columns from left to right). All cut-outs are of 400 arcsec in size.

Figure A1. A sample of 12 detections located at redshifts $z = 0.2$, 0.3, 0.5, and 0.7 (rows from top to bottom), and with an intrinsic richnesses of $\lambda_* \sim 100$, 50, and 5 (columns from left to right). All postage stamps have a size of 400 arcsec.

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