Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **435,** 3444–3461 (2013) doi:10.1093/mnras/stt1538
Advance Access publication 2013 September 13

The ALHAMBRA survey: reliable morphological catalogue of 22 051 early- and late-type galaxies

M. Pović,[1⋆] M. Huertas-Company,[2,3] J. A. L. Aguerri,[4] I. Márquez,[1] J. Masegosa,[1] C. Husillos,[1] A. Molino,[1] D. Cristóbal-Hornillos,[5] J. Perea,[1] N. Benítez,[1] A. del Olmo,[1] A. Fernández-Soto,[6,7] Y. Jiménez-Teja,[1] M. Moles,[1,5] E. Alfaro,[1] T. Aparicio-Villegas,[1,8] B. Ascaso,[1] T. Broadhurst,[9,10] J. Cabrera-Caño,[11] F. J. Castander,[12] J. Cepa,[4,13] M. Fernandez Lorenzo,[1] M. Cerviño,[1,4] R. M. González Delgado,[1] L. Infante,[14] C. López-Sanjuan,[5] V. J. Martínez,[7,15] I. Matute,[1] I. Oteo,[4,11] A. M. Pérez-García,[4,13,16] F. Prada[1] and J. M. Quintana[1]

[1]*Instituto de Astrofísica de Andalucía (IAA-CSIC), E-18008, Granada, Spain*
[2]*GEPI, Paris-Meudon Observatory, F-75014 Meudon, France*
[3]*University of Paris, F-75205 Paris, France*
[4]*Instituto de Astrofísica de Canarias (IAC), E-38200 La Laguna, Tenerife, Spain*
[5]*Centro de Estudios de Física del Cosmos de Aragón (CEFCA), E-44001 Teruel, Spain*
[6]*Instituto de Física de Cantabria (CSIC-UC), E-39005 Santander, Spain*
[7]*Unidad Asociada Observatori Astronòmic (IFCA – UV), Valencia, Spain*
[8]*Observatório Nacional-MCT, Rua Jos Cristino, 77. CEP 20921-400 Rio de Janeiro-RJ, Brazil*
[9]*Department of Theoretical Physics, University of Basque Country, UPV/EHU, PO Box 644, E-48080 Bilbao, Spain*
[10]*IKERBASQUE, Basque Foundation for Science, Alameda Urquijo 36-5, E-48008 Bilbao, Spain*
[11]*Facultad de Física, Departamento de Física Atómica, Molecular y Nuclear, Universidad de Sevilla, E-41012 Sevilla, Spain*
[12]*Institut de Ciències de l'Espai, IEEC/CSIC, E-08193 Barcelona, Spain*
[13]*Departamento de Astrofísica, Facultad de Física, Universidad de la Laguna, E-38200 La Laguna, Spain*
[14]*Departamento de Astronomía, Pontificia Universidad Católica, Santiago de Chile, Chile*
[15]*Observatori Astronòmic de la Universitat de València, Valencia, Spain*
[16]*Asociacíon ASPID, Apartado de Correos 412, E-38200 La Laguna, Tenerife, Spain*

Accepted 2013 August 13. Received 2013 August 5; in original form 2013 April 9

ABSTRACT

Advanced Large Homogeneous Area Medium Band Redshift Astronomical (ALHAMBRA) is photometric survey designed to trace the cosmic evolution and cosmic variance. It covers a large area of ∼4 deg^2 in eight fields, where seven fields overlap with other surveys, allowing us to have complementary data in other wavelengths. All observations were carried out in 20 continuous, medium band (30 nm width) optical and 3 near-infrared (*JHK*) bands, providing the precise measurements of photometric redshifts. In addition, morphological classification of galaxies is crucial for any kind of galaxy formation and cosmic evolution studies, providing the information about star formation histories, their environment and interactions, internal perturbations, etc. We present a morphological classification of >40 000 galaxies in the ALHAMBRA survey. We associate to every galaxy a probability to be early type using the automated Bayesian code GALSVM. Despite of the spatial resolution of the ALHAMBRA images (∼1 arcsec), for 22 051 galaxies, we obtained the contamination by other type of less than 10 per cent. Of those, 1640 and 10 322 galaxies are classified as early- (down to redshifts ∼0.5) and late-type (down to redshifts ∼1.0), respectively, with magnitudes $F613W \leq 22.0$. In addition, for magnitude range $22.0 < F613W \leq 23.0$, we classified other 10 089 late-type galaxies with redshifts ≤1.3. We show that the classified objects populate the expected regions

⋆E-mail: mpovic@iaa.es

in the colour–mass and colour–magnitude planes. The presented data set is especially attractive given the homogeneous multiwavelength coverage available in the ALHAMBRA fields, and is intended to be used in a variety of scientific applications. The low-contamination catalogue (<10 per cent) is made publicly available with this paper.

Key words: surveys – galaxies: fundamental parameters – galaxies: statistics.

1 INTRODUCTION

One of the first steps in any research work is to group objects with common properties (e.g. shape, weight, colour, etc.). This taxonomy is a powerful tool in order to understand the physics behind the formation and evolution of the studied objects. Probably the most popular classification of galaxies is based on the shapes or morphologies (started with Reynolds 1920; Hubble 1926, 1936). This first order classification has survived over time since the different morphological classes of galaxies have also different physical properties and probably different evolutionary tracks. In general, galaxies can be divided into two main classes: early-types (hereafter ETs) and late-types (hereafter LTs). ETs include elliptical and lenticular galaxies, while LTs include spirals and irregular galaxies. ETs appear to be a family of objects showing old stellar populations, spheroidal-like dynamical properties and a small fraction of cold gas whereas LTs are more gas-rich objects, present younger stellar populations, and are mainly rotation supported.

Visual inspection is the traditional way to classify galaxies. By definition, it is subjective and not reproducible, but for bright and extended objects there is a general good agreement between different observers. They are however time consuming. In the past, galaxy samples contained from dozens up to hundreds of galaxies while present galaxy surveys have up to millions of galaxies. This makes impossible to give detailed morphological classifications unless a large amount of classifiers are involved (see the Galaxy Zoo project; Lintott et al. 2008).

Over the past few years, different automated methods of morphological classification of galaxies have been developed. Automated classifications resolve the two main problems raised above. They provide indeed reproducible information, and the errors can be fully understood by using extensive simulations (e.g. Trujillo et al. 2001; Simard et al. 2002, 2011). In addition, modern galaxy classification algorithms are able to classify large samples of galaxies in a reasonable amount time.

We usually distinguish between three broad groups of automated galaxy classifications: parametric based on galaxy physical parameters, non-parametric and parametric based on mathematical parameters. Parametric classifications use parametric models in order to reproduce some galaxy measurements. One of the most popular parametric methods based on galaxy physical parameters classifies galaxies according to some properties of their structural parameters obtained by fitting the surface brightness (e.g. de Vaucouleurs 1948; Sérsic 1963; Prieto et al. 1997, 2001; Peng et al. 2002, 2010; Simard et al. 2002, 2011; Aguerri et al. 2004, 2005; de Souza, Gadotti & dos Anjos 2004; Méndez-Abreu et al. 2008). On the other hand, non-parametric galaxy classifications are based on the measurements of a set of galaxy parameters that correlate with the morphological types. These methods have the advantage that they assume non-parametric models and can hence classify regular and irregular objects. Several galaxy parameters have been used to discriminate between different morphological types, i.e. colours (e.g. Strateva et al. 2001), spectral properties (e.g. Humason 1931; Morgan & Mayall 1957; Baldwin, Phillips & Terlevich 1981; Folkes, Lahav & Maddox 1996; Sánchez-Almeida et al. 2010), or light distribution (e.g. Abraham et al. 1994; Conselice et al. 2000; Abraham et al. 1996; Abraham, van den Bergh & Nair 2003; Lotz, Primack & Madau 2004; Scarlata et al. 2007). Current works are led to more complex fitting models, using orthonormal mathematical bases to decompose the galaxies, and then correlating the physical properties of the objects with the coefficients of the decomposition by means of a principal component analysis. This is the case of the shapelets (Kelly & McKay 2004, 2005; Andrae, Jahnke & Melchior 2011a), or future works with the sersiclets (Ngan et al. 2009; Andrae, Melchior & Jahnke 2011b), and the CHEFs (Jiménez-Teja & Benítez 2012). The main advantage of these latter methods with regard to the former is their flexibility and reliability to reproduce every feature in the galaxies, and then to efficiently model every kind of morphology, including irregular objects. Moreover, they do not need any a priori knowledge and they do not impose any profile to fit the galaxies. However, these mathematical models are computationally more expensive and the parameters of the final decomposition do not provide any physical information. ALH data will be used for further development of the CHEFs method (Jiménez et al., in preparation).

GALSVM (Huertas-Company et al. 2008) provides a generalization of the non-parametric classifications by using an unlimited number of dimensions and providing a probabilistic output following a Bayesian approach (see also Fasano et al. 2012 for a similar approach). The algorithm has been shown to be especially useful when dealing with low-resolution and high-redshift data (Huertas-Company et al. 2009) and has been successfully applied to several large samples at low and high redshift (Huertas-Company et al. 2009, 2011), including the Cosmological Evolution Survey (COSMOS;[1] Scoville et al. 2007) and Sloan Digital Sky Survey (SDSS;[2] Castander 1998) samples.

In this paper, we present the morphological classification of a large sample of galaxies from the Advanced Large Homogeneous Area Medium Band Redshift Astronomical, ALHAMBRA survey (hereafter ALH; see Moles et al. 2008), located at different redshifts. Taking into account the resolution of ALH data ($\sim$1 arcsec) and all the advantages of GALSVM code mentioned above, we used this method for our morphological classification. The ALH survey imaged $\sim$4 deg^2 of the sky through 23 optical and near-infrared (NIR) filters. The large number of filters provide accurate photometric redshifts by fitting the spectral energy distributions (hereafter SEDs) of the galaxies for about one million sources. This galaxy sample is ideal in order to study evolutionary properties of galaxies in the last 8 Gyr (Moles et al. 2008; Cristóbal-Hornillos et al. 2009; Matute et al. 2012; Oteo et al. 2013a,b).

The ALH survey, data and sample selection are introduced in Section 2. The methodology used for morphological classification

[1] http://cosmos.astro.caltech.edu/
[2] http://www.sdss.org/

Table 1. ALH observations used in this paper: ALH fields and their central coordinates, number of objects and covered areas in the *F*613*W* band down to magnitudes 23.0, range of seeing of individual observations used to create the final ones, and the averaged seeing.

Field	RA (J2000) (h m s)	Dec. (J2000) (° ′ ″)	Num. of obj. in *F*613*W* band (at mag ≤ 23.0)	Area (deg^2)	Min–max seeing (arcsec)	Averaged seeing (arcsec)
ALH-2/DEEP2	02 28 32.0	+00 47 00	14 322	0.5	0.86–1.40	1.04
ALH-3/SDSS	09 16 20.0	+46 02 20	12 508	0.5	0.70–1.18	0.89
ALH-4/COSMOS	10 00 28.6	+02 12 21	7104	0.25	1.06–1.32	1.17
ALH-5/HDF-N	12 35 00.0	+61 57 00	6274	0.25	0.95–1.40	1.23
ALH-6/GROTH	14 16 38.0	+52 25 05	13 614	0.5	0.81–1.30	1.11
ALH-7/ELAIS-N1	16 12 10.0	+54 30 00	15 887	0.5	0.84–1.40	1.04
ALH-8/SDSS	23 45 50.0	+15 34 50	14 128	0.5	0.72–1.40	0.91

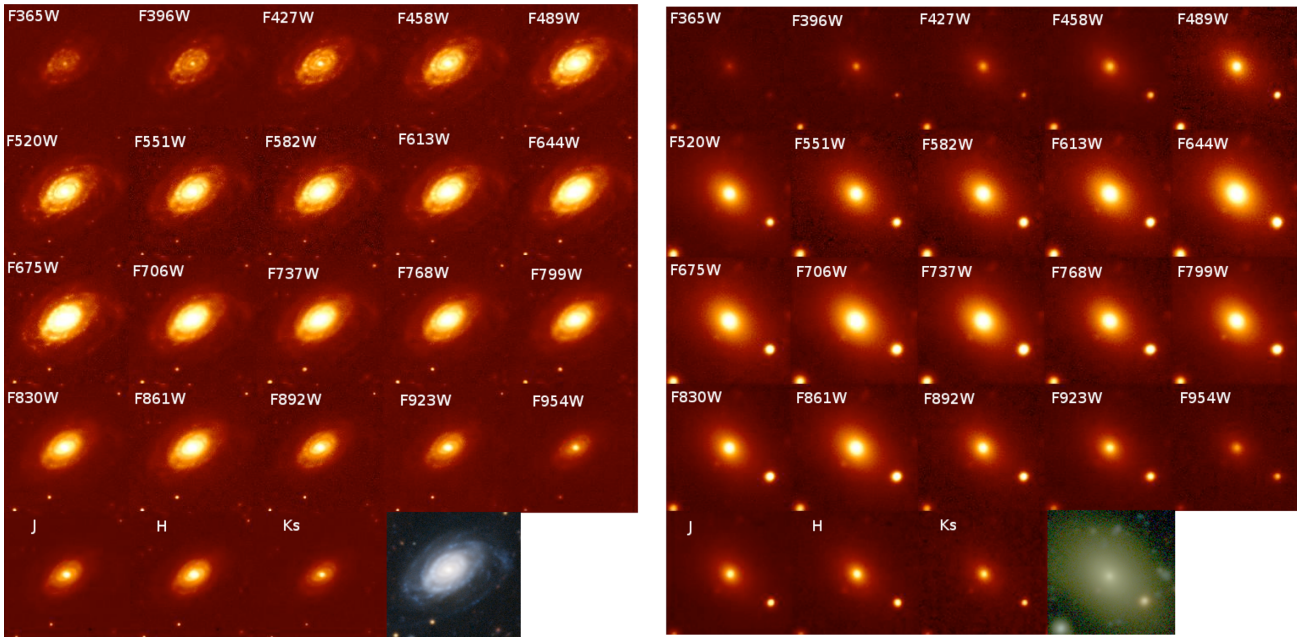

Figure 1. Example of two galaxies observed with 20 optical and 3 NIR ALH filters.

is described in Section 3, and the main results are discussed in Sections 4–6. Finally, Appendix A describes the content of the first published morphological catalogue of galaxies in the ALH survey.

We assumed the following cosmological parameters throughout the paper: $\Omega_\Lambda = 0.7$, $\Omega_M = 0.3$ and $H_0 = 70$ km s^{-1} Mpc^{-1}. Unless otherwise specified, all magnitudes are given in the AB system (Oke & Gunn 1983).

2 DATA

The ALH observations were carried out at the Calar Alto German-Spanish Astronomical Center (CAHA[3]), under the Spanish guaranteed time of 110 nights. Eight fields were observed in the Northern hemisphere sky, having a seeing lower than 1.6 arcsec (ranging mainly between 0.8 and 1.2 arcsec). All fields (except ALH-1) have the multiwavelength information available from other (deep) extragalactic surveys (see Moles et al. 2008 and Table 1). Each ALH detection was observed in 23 bands, with 20 optical and three standard *JHK* NIR filters. Optical range is covered in a continuous way from 340 to 970 nm, with non-overlapping and equal 30 nm width medium bands (Aparicio-Villegas et al. 2010). This set of optical filters was specially designed for the ALH survey to achieve a good accuracy of photometric redshifts of $\sim\delta z/(1+z) = 0.015$ for galaxies brighter than $F814W \leq 24.5$, three times better than the one achieved when using the 4–5 filter systems (for more information, see Benítez et al. 2009). Fig. 1 shows the example of two galaxies observed in all ALH bands.

Optical data were observed with Large Area Imager for Calar Alto (LAICA[4]), with the total exposure time of 100 000 s per pointing. On the other hand, OMEGA2000[5] instrument was used in the NIR, with the total exposure time of 60 000 s per pointing. Data reduction was carried out using the standard set of IRAF[6] packages. Table 1 gives the summary of ALH observations used in this work.

[3] http://www.caha.es/

[4] http://www.caha.es/CAHA/Instruments/LAICA/index.html

[5] http://w3.caha.es/CAHA/Instruments/O2000/index2.html

[6] IRAF is distributed by the National Optical Astronomy Observatory (NOAO), which is operated by the Association of Universities for Research in Astronomy (AURA) under cooperative agreement with the National Science Foundation (NSF).

Source detection was performed by means of the SExtractor (v. 2.8.6) code (Bertin & Arnouts 1996). Both, detection of sources and creation of photometric catalogues, were carried out in all 23 bands, where each catalogue contains more than 180 photometric, astrometric and basic morphological parameters. More than 670 000 sources were detected in all ALH surveys, with a photometric completeness of $r \sim 25.0$ (corresponds to SDSS r band constructed from five individual ALH bands). All additional information respect to the source detection and creation of photometric catalogues can be found in Husillos et al. (in preparation), while the catalogues are publicly available via the survey webpage.[7]

For photometric redshift estimations (hereafter photo-z), the Bayesian Photometric Redshift code was used (BPZ; Benítez 2000; Benítez et al. 2004). BPZ was run on a separated point spread function (PSF) corrected photometry and a new library of templates (Benítez et al., in preparation) was implemented, composed by 11 SEDs (4 elliptical, 1 lenticular, 1 Sbc, 1 Scd and 4 starburst galaxies), originally drawn from PEGASE (Fioc & Rocca-Volmerange 1997), but then optimized using the FIREWORKS photometry and spectroscopic redshifts from Wuyts et al. (2008). Based on a sample of ∼7000 galaxies with known spectroscopic redshifts, drawn from different surveys (see Table 1), the final performance of resulting photo-zs was evaluated. The expected accuracy for galaxies with magnitudes <23.0 in the constructed $F814W$ nm band is $\sim\delta z/(1+z) = 0.011$ with a small fraction of catastrophic outliers no higher than 3 per cent. In this work, we used the BPZ measured photo-zs, and the same photometry that was used in photo-z estimations. All information about the photo-z measurements and catalogues can be found in Molino et al. (2013).

2.1 Sample selection

We carried out the morphological classification on a well-defined sample, free of stars, having reliable photo-z and photometric measurements. In this section, we describe the different selection criteria applied to obtain the final working sample.

(i) *Extended–point-like source separation*. To discriminate between the extended and point-like sources, we used the classification provided by Husillos et al. (in preparation). In all observations, a sample of more than 2000 real, bright, unsaturated and geometrically circular point-like ALH sources was used, scaled to lower random fluxes, and injected into the empty regions of optical image. After this, the sources were recovered, using the same parametrization as applied in source detection. Point-like sources are considered as successfully recovered when detected at distances lower than 1 arcsec from previous injection coordinates. The same procedure was implemented for extended sources. Finally, the locus of point-like and extended sources was defined in the apparent magnitude and surface brightness space (MAG_AUTO[8] and MU_MAX[9] parameters, respectively, obtained by SExtractor). By plotting each locus, it is possible to estimate the point-like source contamination in a quantitative way. GEOM_CLASS_STAR parameter was defined in this way, having value equal to 1 for point like, and equal to 0 for extended sources. Unclassified sources have GEOM_CLASS_STAR = 99. Taking into account the ALH resolution, it was confirmed that this method works well down to magnitudes $r = 23.0$, which is in good agreement with our magnitude-selection criteria (see below).

To select the extended sources, we used the GEOM_CLASS_STAR = 0. With this criteria, using the comparisons with the *Hubble Space Telescope* (*HST*) data from the COSMOS survey, we estimated to have in the selected sample contamination of point-like sources lower than 1 per cent down to magnitudes 22.0 in the $F613W$ band, and of 5–7 per cent between magnitudes 22.0 and 23.0.

(ii) *Photo-z selection*. We used two criteria to select the sources with good photo-z measurements. First, we only selected those sources detected in all ALH filters. And secondly, we selected sources with BPZ_ODDS[10] parameter above 0.2. With these two criteria, we expect to have less than 3 per cent of outliers (Molino et al. 2013). We checked that requiring the objects to be detected in all ALH bands introduces only a small selection bias affecting the sample of red galaxies at $z > 0.4$ and magnitudes <22 in the $F613W$ band, a range in which our method is not able to efficiently select ETs (see Section 4).

(iii) *Magnitude selection*. We selected only objects with magnitudes ≤23.0 in the $F613W$ filter, and with magnitude errors <0.5. Above this magnitude limit, the reliability of signal-to-noise (S/N) measurements, photo-z estimations, and geometrical extended/point-source classifications decrease significantly. Selection of the $F613W$ filter is explained in Section 3.2.2. In comparison with previous criteria, magnitude selection is the most restrictive one.

(iv) *Flag tests*. In the sample selected through previous conditions, 57 per cent of sources are 'good detections' (SExtractor FLAG parameter = 0), 27 per cent are possibly blended sources (FLAG = 2), 15 per cent, plus blending, have close neighbours or bad pixels (FLAG = 3), while <1 per cent have other FLAG values. We classified morphologically all galaxies independently of their FLAG parameter, including close/interacting systems. However, the final statistics (Sections 5 and 6), and the published catalogue (Appendix A) include only sources with FLAG values 0 and 2 (see Section 4).

The final selected sample to be morphologicaly classified has 43 665 galaxies.

3 MORPHOLOGICAL CLASSIFICATION OF GALAXIES

3.1 General methodology

The main tool used in this work to estimate morphologies is GALSVM, a non-parametric support vector machine (SVM) based code (Huertas-Company et al. 2008, 2009, 2011). Basically, GALSVM uses a training set of local galaxies with known visual morphologies to train the SVM that is then applied to the data set to be classified (see Section 3.2.1). Galaxies from the training sample are redshifted and scaled in luminosity to match the magnitude counts and redshift distribution of the ALH sample, resampled with the ALH pixel scale (∼0.222 arcsec pixel^{-1}), and finally dropped

[7] http://alhambrasurvey.com/

[8] Kron-like elliptical aperture magnitude [see Bertin (2009) and Holwerda (2009) SExtractor manuals at https://www.astromatic.net and http://mensa.ast.uct.ac.za/∼holwerda/Site/Source_Extractor.html, respectively].

[9] Peak surface brightness above background (see the manuals given in footnote 8).

[10] Defines the redshift confidence limits, where galaxies with higher ODDS have a more secure redshift estimations (see Benítez 2000, for more information).

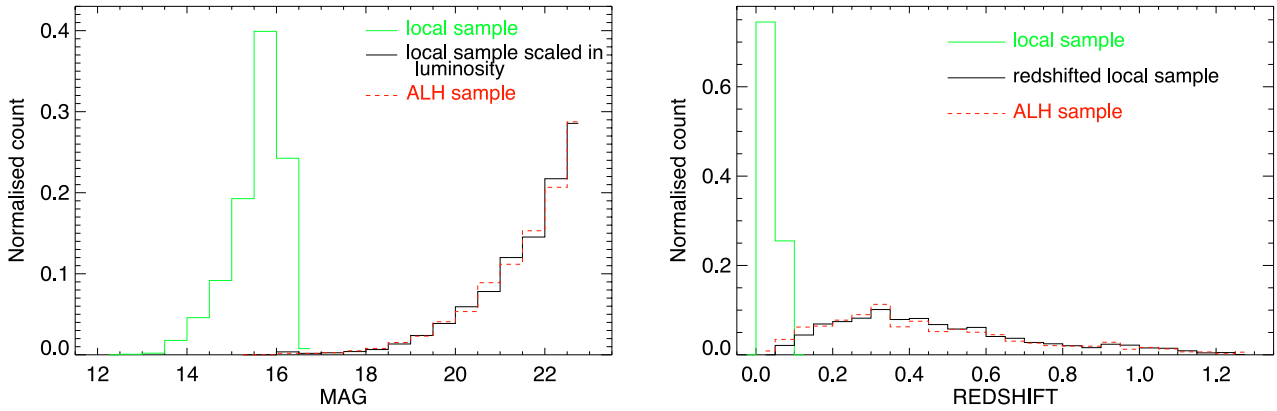

Figure 2. Magnitude and redshift normalized distributions of local sample used by GALSVM (green solid lines). This sample was scaled in luminosity and redshifted (black solid lines) to match the distributions of the ALH sample (red dashed lines). Shown example corresponds to distributions in the ALH-7 field.

in a real ALH background. Fig. 2 shows an example of redshift and magnitude distributions of local sample before and after being redshifted and scaled in luminosity (green and black solid lines), and corresponding distributions of the ALH-7 sample (red dashed lines).

We measured seven morphological parameters on this simulated data set (and on the ALH sample later on), and use them simultaneously to train the vector machine:

(i) Ellipticity
(ii) Abraham concentration index (CABR) – ratio between the fluxes at 30 and 90 per cent of the radius (Abraham et al. 1996)
(iii) Conselice–Bershady concentration index (CCON) – ratio between circular radii containing 20 and 80 per cent of the total flux (Bershady, Jangren & Conselice 2000)
(iv) Gini (GINI) – cumulative distribution function of galaxy's pixel values (Abraham et al. 2003)
(v) Asymmetry (ASYM) – measures the degree of asymmetry in the light distribution (Conselice et al. 2000)
(vi) Smoothness (SMOOTH) – measures the relevance of small-scale structures (Conselice et al. 2000), and
(vii) M_{20} moment of light (M20) – flux in each pixel multiplied by the squared distance to the centre of the galaxy, summed over the 20 per cent brightest pixels of the galaxy (Lotz et al. 2004).

Even though several of these parameters appear to be redundant, SVM were specially designed to be robust to redundancies in the feature space (see e.g. Huertas-Company et al. 2008). The set of parameters used to obtain the morphological classification was tested in Pović et al. (2012). Mentioned parameters were measured for both: local sample (redshifted and scaled in luminosity) and the ALH sample that we want to classify. Using vector machine, and comparison between the GALSVM classification of local galaxies and their original, visual one, we can then classify ALH galaxies. The output of the classification step is a probability value to be in a given class for each object. The probability takes values from 0 to 1 (99.9 for unclassified objects). The robustness of the classification and the sensitivity to the training set is estimated by repeating the classification several times through Monte Carlo runs (hereafter MC) with slightly different training sets (see Huertas-Company et al. 2011, for more details). We describe in the following the specific configuration used in this work.

3.2 GALSVM CONFIGURATION AND CLASSIFICATION

3.2.1 Training sample of local galaxies

We use a local sample of 3000 visually classified galaxies ($0.01 \leq z \leq 0.1$) taken from the Nair & Abraham (2010, hereafter N&A) catalogue which contains ∼14 000 galaxies from the SDSS. The number of galaxies used for training is selected as a trade-off between classification accuracy and computing time. The computing time to train the SVM with the current algorithm is indeed very sensitive to the size of the training data set. We therefore chose objects randomly and select the same number of ET (elliptical and lenticular) and LT (spiral and irregular) galaxies, to avoid problems related to unbalanced data sets, which is required for GALSVM to work properly. We checked that the randomly selected subsample is representative in terms of general properties (colours, magnitudes, etc.) of the complete data set. Fig. 3 shows these comparisons for the g-band magnitude, redshift, $g - r$ colour, N&A morphological classification and inclination (for LT galaxies only). Moreover, we also compared magnitudes, redshifts, and colours of ET and LT galaxies in both, selected and full samples. As can be seen from normalized distributions, the properties of randomly selected and full N&A samples are consistent in all plots, as stated by Kolmogorov–Smirnov (hereafter KS) tests which proves that no selection biases are introduced.

N&A classification was obtained in the SDSS g-band images, while our classification was carried out in the $F613W$ band (corresponds to SDSS r band; see Section 3.2.2). Taking into account that we are performing broad classification, separating all sources into ETs and LT galaxies, the difference between the g and r bands inspected visually is insignificant for our work. For each morphological type from N&A catalogue, we selected randomly 50 sources and we checked (between 4 people) their images in all five SDSS bands. In all cases, we do not see any differences in galaxy structures between the g, r and i bands. On the other hand, we do see significant changes between any of these three bands and u or z filters.

The selected local data set is then redshifted, scaled in luminosity and dropped in the ALH fields. To that purpose, GALSVM requires the apparent magnitude and photo-z distributions of all sources. To improve statistics, we used for each field the mag and photo-z distributions of the four CCDs after checking that these distributions are completely consistent with those of individual fields (individual CCDs). Before dropping the mock galaxies they are re-sampled

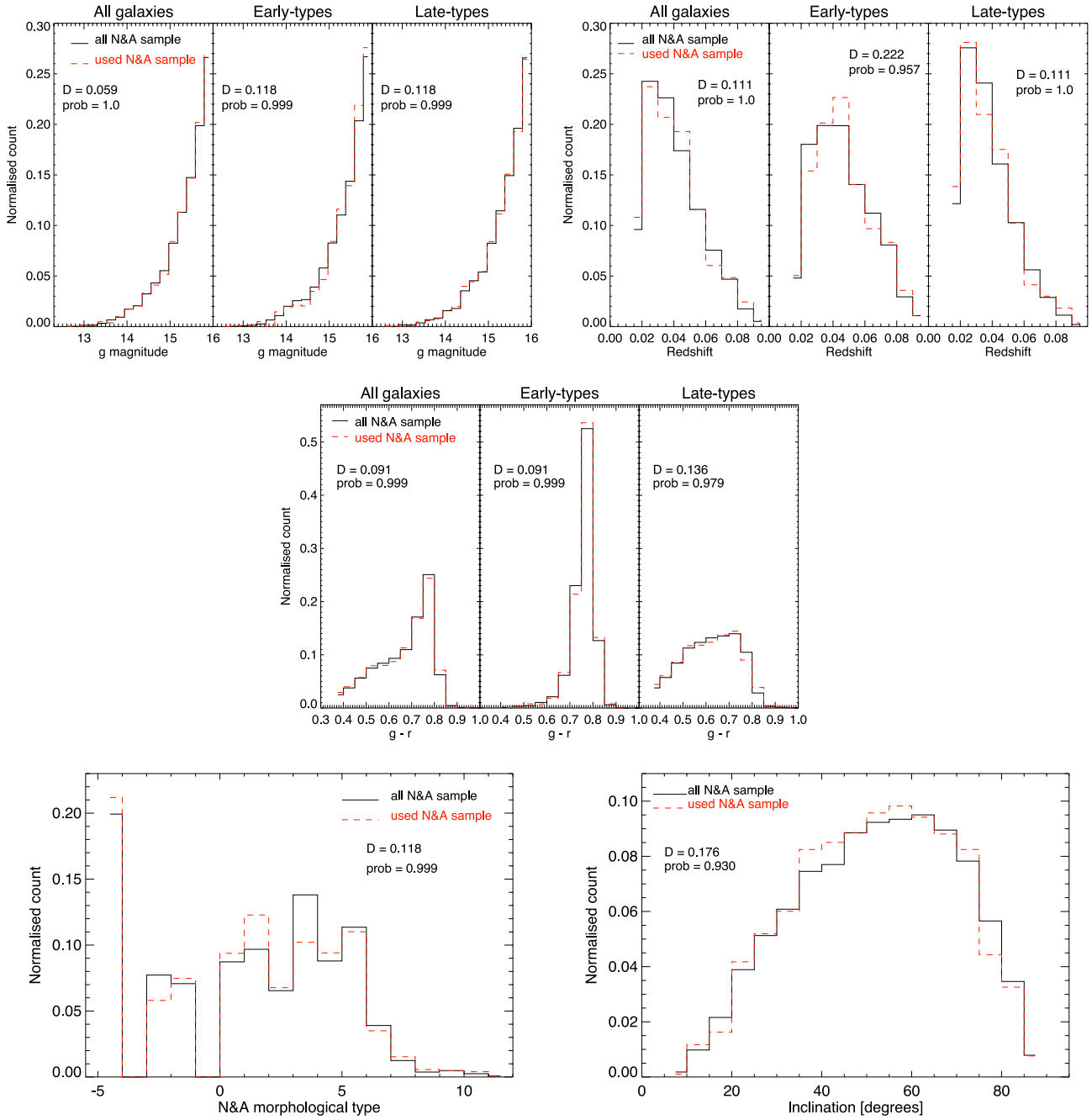

Figure 3. Comparison between the properties of the full N&A sample (black solid lines) and 3000 galaxies (red dashed lines) selected randomly from the full one to be used in our morphological classification. g-band magnitude (top left), redshift (top right) and $g - r$ colour (middle) distributions are represented for all galaxies (left subdiagram), ETs (middle subdiagram) and LTs (right subdiagram). We also compared morphological types (for all galaxies; bottom left) and inclinations (for LT galaxies only; bottom right) distributions. Maximum deviation (D) and probability that two distributions are the same (prob; takes values between 0 and 1, where small values show that the cumulative distribution function of selected sample is significantly different from the full sample) are results of the KS statistics, showing in all plots that properties of our selected N&A sample are completely consistent with the full one.

with the same pixel scale than the ALH galaxies and convolved with a PSF to match the same spatial resolution. In order to cope with the k-correction, we select the SDSS filter which is closer to the wavelength the ALH filter is probing given the redshift of the galaxy. Surface brightness dimming is taken into account when scaling the galaxy in flux since we empirically match the magnitude counts of the ALH survey (we do not introduce any size evolution though). Concerning the noise, we make the hypothesis that the noise from SDSS galaxies is negligible compared to the noise of galaxies at higher redshift (since we are using very bright galaxies). By dropping the galaxy in a real background, we expect to reproduce at best all the different noises from the real images.

We assumed that there is no change in galaxy properties between the local and high-redshift samples (e.g. luminosity–morphology dependence). This assumption might be strong for our redshift range (∼0–1), but can be justified since we are classifying all galaxies into two broad morphological types. Moreover, we only used morphological parameters in our classification, excluding luminosity.

This minimizes significantly the luminosity–morphology dependence; however, it does not eliminate it completely (e.g. still higher redshift sample will contain more luminous sources, besides, for the redshifted training sample we are trying to reproduce the magnitude distribution of ALH sources assuming that luminosity/morphology relations are similar). To minimize even more the luminosity effects on our morphological classification, we forced the GALSVM to select ∼50 per cent of ET and LT local galaxies in each training MC run.

Although, we used 3000 local galaxies in this work, a larger sample was constructed from the N&A and EFIGI[11] (Baillard et al. 2011) samples, with the total of 15 036 sources down to redshifts $z \leq 0.1$. The EFIGI visual classification was used at redshifts $z < 0.01$, including all galaxies with good morphological classification, and excluding dwarf objects, while N&A galaxies range between $0.01 \leq z \leq 0.1$. At the moment, this is the most complete local sample that can be used to test the morphology of high-redshift galaxies with the GALSVM code, including for each source the list of astrometric, photometric, redshift, morphological type, size parameters, poststamps, masks and PSF images in all five SDSS bands.

3.2.2 GALSVM *classification*

The final classification is performed in the ALH $F613W$ band, since the S/N ratio is higher in this filter, as shown in Aparicio-Villegas et al. (2010). For each GALSVM run, we set the quality parameters that correspond to used ALH image: filter name, central wavelength, full width half-maximum, pixel scale, sigma, zero-point and saturation level. Given the typical resolution of the ALH survey (mainly around 1.0 arcsec, see Table 1) and the depth, we have restricted in this work to two broad morphological classes: ET and LT. For each source in the final selected sample (see Section 2.1), we compute seven morphological parameters described in Section 3.1.

Since the sample is fully dominated by faint objects given the shape of the magnitude counts (Fig. 4), when fainter objects are included in the classification, the algorithm will be optimized to classify these galaxies and the fraction of misclassified bright objects might significantly increase. To avoid this effect, we performed the morphological classification with six increasing magnitude cuts: ≤20.0, ≤21.0, ≤21.5, ≤22.0, ≤22.5 and ≤23.0. That way, bright galaxies are classified using a training set made only of bright galaxies. These cuts also correspond roughly to increasing redshifts: ≤0.7, ≤0.9, ≤1.0, ≤1.2, ≤1.2 and ≤1.3.

The final probability for each galaxy is computed as the average of the output probability of 15 MC independent runs. Again, this number is an empirical trade-off between computing time and accuracy, and is a result of previous tests carried out with 10, 15 and 20 runs. In each MC run, GALSVM selects randomly a balanced set of 2000 out of 3000 input local galaxies. The result of each run is a probability for a ALH galaxy to be ET, where for small probability values increases the possibility to be an LT. At the end, for each ALH source we compute the final probability to be ET (hereafter $p_{\rm E}$) as average value of 15 probabilities, and the probability error as the scatter of the distribution. Obviously, since we deal with a two-class problem only, the probability for the source to be LT (hereafter $p_{\rm L}$) is simply $p_{\rm L} = 1 - p_{\rm E}$.

[11] http://www.astromatic.net/projects/efigis

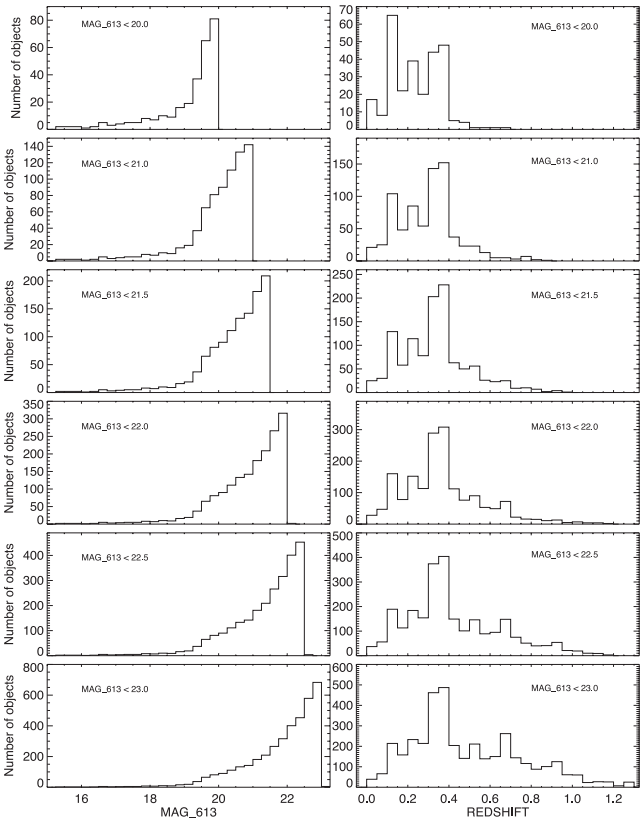

Figure 4. From top to bottom: distributions of $F613W$ magnitudes (left) and corresponding redshifts (right) in the ALH-4 field for six magnitude cuts used in our morphological classification: ≤20.0, ≤21.0, ≤21.5, ≤22.0, ≤22.5 and ≤23.0.

4 CALIBRATION OF THE MORPHOLOGICAL CLASSIFICATION USING COSMOS/*HST* DATA

The classification is calibrated in the ALH-4 field using the classification of the same objects observed with *HST*/Advanced Camera for Surveys (ACS) in the COSMOS survey. The morphological classification in COSMOS used here was carried out with GALSVM by Huertas-Company et al. (2009), and is publicly available through the Tasca et al. (2009) catalogue.[12] The classification separates galaxies into three classes (early, late and irregulars) but does not have probabilistic information though since it was done with an older version of GALSVM.

By using COSMOS as the reference sample, we explicitly neglect the classification errors of the galaxies in COSMOS. This choice is justified since we focus here on bright galaxies with only two morphological classes. We also neglect the eventual morphological drift between the morphologies in COSMOS ($F814W$) and ALH ($F613W$). Again, since we are splitting our galaxies in two classes we do not expect a significant effect. This assumption is indeed confirmed through extensive tests on the SDSS data set. We compared different morphological parameters of local N&A galaxies between the r and i bands, without finding any significant differences in two bands, independently of the redshift and the Hubble type.

Fig. 5 shows the distributions of $p_{\rm E}$ derived in the ALH-4 field for three morphological types classified in the COSMOS/*HST* survey

[12] http://irsa.ipac.caltech.edu/data/COSMOS/tables/morphology/cosmos_morph_tasca_1.1.tbl

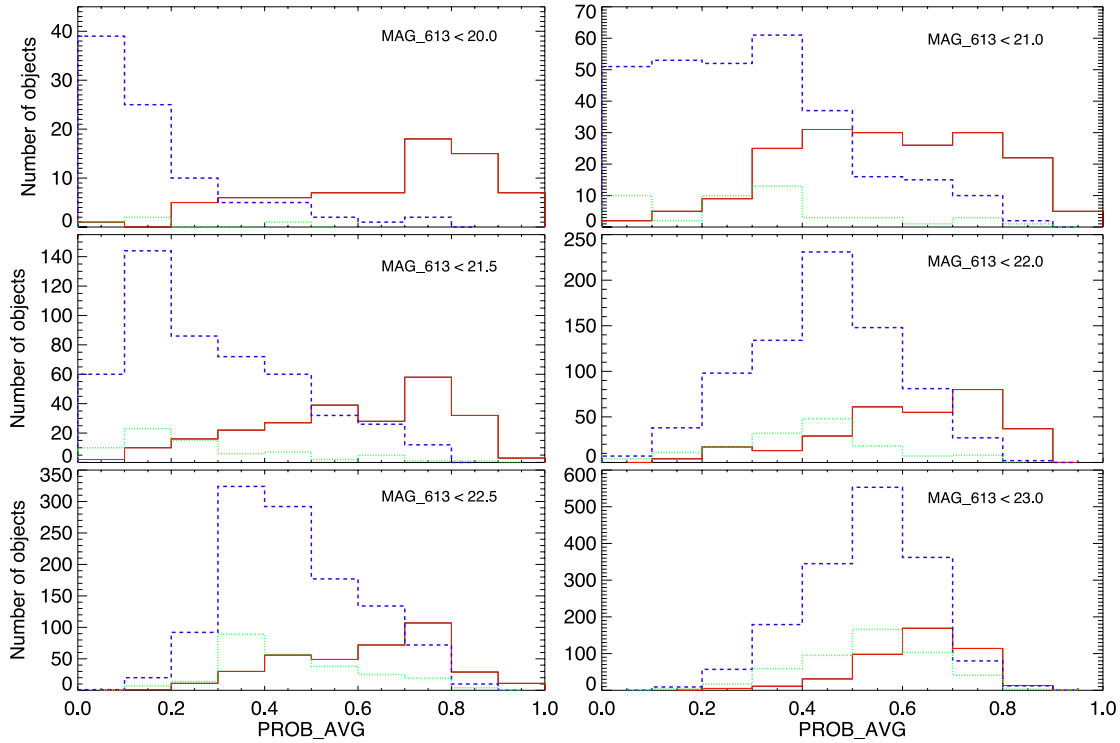

Figure 5. Comparison between the GALSVM classification in the ALH-4 and COSMOS surveys for different $F613W$ magnitude cuts: ≤20.0 (top left), ≤21.0 (top right), ≤21.5 (middle left), ≤22.0 (middle right), ≤22.5 (bottom left) and ≤23.0 (bottom right). In each plot, we compare the distributions of averaged probabilities (PROB_AVG) measured in the ALH-4 field for three morphological types classified in the COSMOS field using the *HST*/ACS imaging data (Huertas-Company et al. 2009; Tasca et al. 2009): elliptical/S0 (red solid lines), spiral (blue dashed lines) and irregular (green dotted lines) galaxies.

with GALSVM code: E/S0 (red solid lines), spirals (blue dashed lines) and irregulars (green dotted lines). Following these distributions, we look for the probability threshold ($p^{\rm E}_{\rm th}$, $p^{\rm L}_{\rm th}$) to apply to the ALH classification so that the resulting classification contains less than 10 per cent contaminations from the neighbouring morphological type. To that purpose, for a given probability threshold $p^{\rm E}_{\rm th}$,$p^{\rm L}_{\rm th}$ we define the following parameters:

(i) *True positives (tp):* galaxies with $p_{\rm E} > p^{\rm E}_{\rm th}$ ($p_{\rm L} > p^{\rm L}_{\rm th}$) in ALH which are classified ET (LT) in COSMOS;
(ii) *True negatives (tn):* galaxies with $p_{\rm E} < p^{\rm E}_{\rm th}$ ($p_{\rm L} < p^{\rm L}_{\rm th}$) in ALH which are classified LT (ET) in COSMOS;
(iii) *False positives (fp):* galaxies with $p_{\rm E} > p^{\rm E}_{\rm th}$ ($p_{\rm L} > p^{\rm L}_{\rm th}$) in ALH which are classified LT (ET) in COSMOS;
(iv) *False negatives (fn):* galaxies with $p_{\rm E} < p^{\rm E}_{\rm th}$ ($p_{\rm L} < p^{\rm L}_{\rm th}$) in ALH which are classified ET (LT) in COSMOS.

The purity (P, fraction of well-classified objects among all objects classified in a given class) and the completeness (C, fraction of well-classified objects among all objects really belonging to a given class) are therefore defined as follows:

$$P = 1 - \frac{fp}{fp + tp} \quad (1)$$

$$C = \frac{tp}{fn + tp}, \quad (2)$$

$p^{\rm E}_{10}$ and $p^{\rm L}_{10}$ are therefore the thresholds to apply so that the contamination is lower than 10 per cent ($P > 0.9$). Table 2 shows the values obtained for the different magnitude cuts and for the two morphological classes. Note that for ET galaxies fainter than 22, the contamination of LT galaxies is always above 10 per cent.

Table 2. p_{10} threshold for selecting ET and LT galaxies (p10 sample) in different magnitude bins.

Mag. limit	20.0	21.0	21.5	22.0	22.5	23.0
$p^{\rm E}_{10}$	> 0.6	> 0.7	> 0.75	> 0.8	–	–
$p^{\rm L}_{10}$	> 0.7	> 0.7	> 0.6	> 0.5	> 0.5	> 0.5

Hereafter, a sample classified with p_{10} thresholds we will call p10 sample. This is a sample that we release in paper and which we analyse in the following sections, as we explain in Section 5. However, in the same way we obtained the p_{10} thresholds, we measured also the thresholds at higher contamination levels. We compare statistics in Section 5, while the full catalogue of whole >44 000 selected sources can be obtained through direct contact.

Each probability threshold, has a corresponding completeness, which is reported later in Table 4. For the p10 sample, the completeness varies from ∼70 per cent for the brightest objects to ∼30 per cent at the faint end. These values become slightly worse if galaxies from CCD3 are included, since it is degraded compared to the others.[13] All statistics are carried out using only 'good detections', i.e. FLAG = 0 (see Section 2.1). However, a similar accuracy is obtained when 'blended' (FLAG = 2) sources are included. Therefore, in the following sections and in the published catalogue (Appendix A) we use both FLAG = 0 and FLAG = 2 sources.

As a sanity check, we performed a visual inspection of p10 classified sample in the ALH-4 field (overlaps with COSMOS). We used

[13] http://www.caha.es/CAHA/Instruments/LAICA

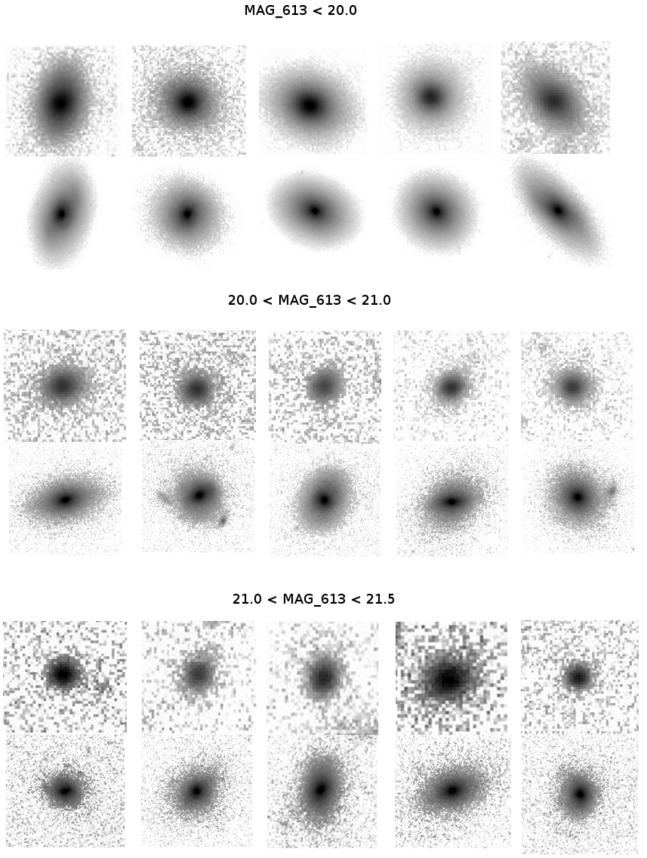

Figure 6. Sample of ET galaxies in different magnitude bins, observed in the ALH-4 (top images, *F*613*W* band) and COSMOS/*HST* (bottom images, *F*814*W* band) surveys.

the *HST*/ACS images of all ETs (86 in total) and of 300 randomly selected LTs (20 per cent of the total p10 LT sample classified in the ALH-4 field). Visual classification was performed by five persons looking each object individually. By each classifier we measured the population of possible misclassified sources (within each type), and then we obtained the averaged one: 14.0 ± 6.0 and 7.8 ± 4.8 per cent for ET and LT classified sources, respectively, which is in agreement with the statistics presented above. Figs 6 and 7 show a sample of selected ET and LT galaxy images, respectively, in different magnitude bins. For each ALH source (top images), we show also the corresponding *HST*/ACS image.

5 MORPHOLOGICAL CLASSIFICATION IN ALH

After calibrating our morphology, we then run in the consistent way the GALSVM code in whole ALH survey. In seven ALH fields, we set the GALSVM configuration files, and for all detections we measured morphological parameters and averaged probabilities. Given the number of fields/observations and magnitude cuts (see Section 3.2.2), we obtained 288 catalogues.

In total, with GALSVM we obtained the classification for 85 per cent of selected sources (with different levels of contamination), while 15 per cent stayed unclassified. Sources for which we were able to measure morphological parameters have different levels of contamination by other morphological type. We analysed the unclassified sources in the ALH-4 field (overlaps with COSMOS) by using the *HST*/ACS images, detecting that ∼80 per cent of sample are interaction/merger candidates. In addition, ∼40 per cent of unclassified sources were detected with the CCD3, which in general has lower S/N in comparison with the other three CCDs (as already mentioned above). The population of ET/LT sources we might be missing for being unclassified with GALSVM is about 3 per cent.

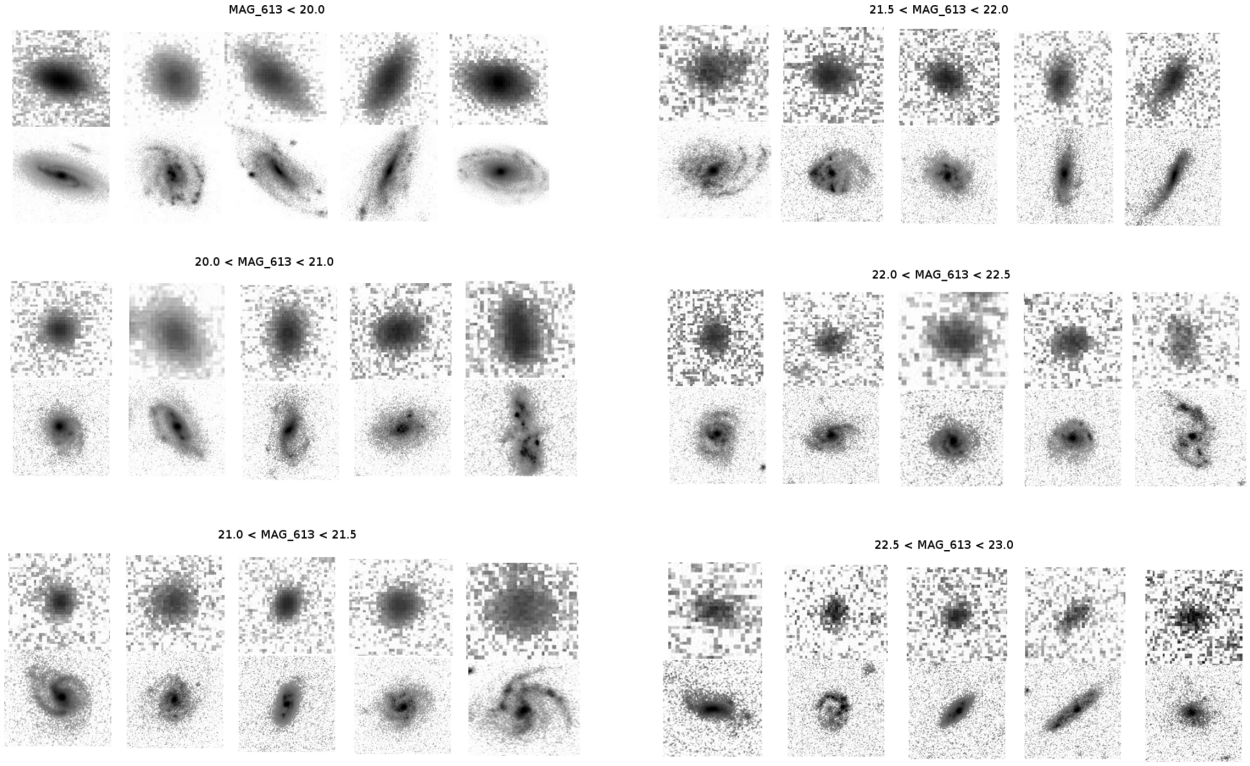

Figure 7. Same as in Fig. 6, but for galaxies classified as LT.

Table 3. Population of sources classified as ET or LT with different levels of contamination by other type.

Contamination level	Population
<10 per cent	**61 per cent**
< 20 per cent	73 per cent
< 30 per cent	74 per cent
< 40 per cent	77 per cent
< 50 per cent	82 per cent
> 50 per cent	18 per cent

For each magnitude bin we then applied the p_{10} thresholds, obtained in the *HST*/ACS calibration phase to ensure a 10 per cent contamination level (p10 sample; see Table 2). With this contamination, in all fields we keep classified 22 051 galaxies, 61 per cent of all classified sources, as shown in Table 3 (table also provides the summary on other sources classified as ET/LT with higher contamination levels). Of those, 1640 and 10 322 sources are classified as ET (photo-z < 0.5) and LT (photo-z < 1.0), respectively, down to magnitudes 22.0 in the $F613W$ band. For this magnitude limit, the fraction of our ET galaxies respect to LTs is ∼16 per cent [consistent with the results obtained by Huertas-Company et al. (2009) in the COSMOS field using the same method of galaxy classification]. In addition, for magnitude range $22.0 < F613W \leq 23.0$ we classified other 10 089 LT galaxies with the same level of contamination of 10 per cent and down to photo-zs of 1.3. For the rest of the sample, the contamination is higher (including also classified sources with very close companions (FLAG ≥ 3) which we excluded from the p10 sample), as shown in Table 3, being in general affected by the resolution of our data and by the worst quality of CCD3 data. Table 4 shows the number of objects and completeness of each morphological type in each magnitude bin. Besides the information about p10 sample, we also provide the statistics at higher contamination levels. As expected, the level of contamination is directly related with the magnitude, where fainter objects were classified with poorer probabilities. At higher contamination levels we have problems classifying ET sources, since many LTs start to mix with ETs giving worst probability distributions. For LT galaxies, the contamination does not go above 20 per cent; so in general, if we are able to detect LT galaxy structures and classify the galaxies, the contamination will always be low. When this is not the case, we have a contamination of ET sample with LT sources. Finally, for the p10 galaxies we expect to have a contamination by other type lower than 10 per cent, and contamination of stars below 1 per cent down to magnitudes 22.0 and 5–7 per cent between magnitudes 22.0 and 23.0. (see Section 2.1). This low-contamination catalogue is the one we are releasing with this paper. Description of all columns and a sample of catalogue for the first five objects are available in Appendix A. The full p10 catalogue of 22 051 galaxies is available in the electronic version of this paper and through the ALH website http://alhambrasurvey.com/.

We do not treat mergers specifically in our classification. However, we do minimize their population by excluding from our p10 classified sample all sources with SExtractor FLAG parameter >2 (include objects with close neighbours). They do enter in the total selected sample, but we excluded them from all analysis presented in this paper (see Section 4). Moreover, we saw that many interacting systems stayed unclassified with GALSVM. As mentioned above, ∼80 per cent of all unclassified sources show clear signs of distortions, close companions, are edge-on systems, or a mixture. Finally, we checked the *HST*/ACS images of all p10 galaxies that overlap with the COSMOS field, in order to quantify the population of possible interactions. From independent inspections carried out by five persons, we obtained the averaged population of 15.0 ± 2.5 per cent merger/interaction candidates between the galaxies classified as LT.

5.1 Selection effects

Since the final catalogue is obtained after applying arbitrary probability cuts, it might be affected by non-trivial selection effects which we investigate in the following from the point of view of redshift and size distributions.

(i) *Redshift*. Figs 8 and 9 show for each analysed magnitude bin, the redshift distributions of ET and LT galaxies, respectively, of the full sample as compared to the final p_{10} sample defined in the previous section. For a total sample (black solid lines), the classification is directly obtained from the Tasca et al. (2009) catalogue

Table 4. Number and completeness of ET and LT galaxies, classified in each magnitude bin in relation with the contamination level. Beside each number we also represent the percentage of galaxies respect to the total number of sources in the observed contamination category (last column).

	F613W magnitude	≤20.0	20.0–21.0	21.0–21.5	21.5–22.0	22.0–22.5	22.5–23.0	
	Contamination ≤ 10 per cent	**715 (44 per cent)**	**549 (33 per cent)**	**219 (13 per cent)**	**157 (9 per cent)**	**–**	**–**	**1640**
	$\mathbf{C_{<10}}$	**0.7**	**0.4**	**0.4**	**0.3**	**–**	**–**	
	Contamination ≤ 20 per cent	1103 (28 per cent)	1841 (47 per cent)	622 (16 per cent)	252 (6 per cent)	112 (3 per cent)	–	3930
ET	$\mathbf{C_{<20}}$	0.7	0.4	0.4	0.3	0.1	–	
	20 < Contamination ≤ 50 per cent	–	–	706 (21 per cent)	1609 (47 per cent)	1081 (32 per cent)	–	3396
	$\mathbf{C_{20-50}}$	–	–	0.6	0.5	0.4–0.5	–	
	Contamination > 50 per cent	–	–	–	492 (7 per cent)	1877 (28 per cent)	4297 (65 per cent)	6666
	$\mathbf{C_{>50}}$	–	–	–	0.7	0.6–0.7	0.8	
	Contamination ≤ 10 per cent	**1082 (5 per cent)**	**2818 (14 per cent)**	**2689 (13 per cent)**	**3733 (18 per cent)**	**4828 (24 per cent)**	**5261 (26 per cent)**	**20 411**
	$\mathbf{C_{<10}}$	**0.7**	**0.7**	**0.7**	**0.7**	**0.6**	**0.3**	
	Contamination ≤ 20 per cent	1720 (7 per cent)	3432 (15 per cent)	3032 (13 per cent)	3659 (16 per cent)	4966 (22 per cent)	6177 (27 per cent)	22 986
LT	$\mathbf{C_{<20}}$	0.7	0.7	0.7	0.7	0.6	0.3	
	20 < Contamination ≤ 50 per cent	–	–	–	–	–	–	
	$\mathbf{C_{20-50}}$	–	–	–	–	–	–	
	Contamination > 50 per cent	–	–	–	–	–	–	
	$\mathbf{C_{>50}}$	–	–	–	–	–	–	

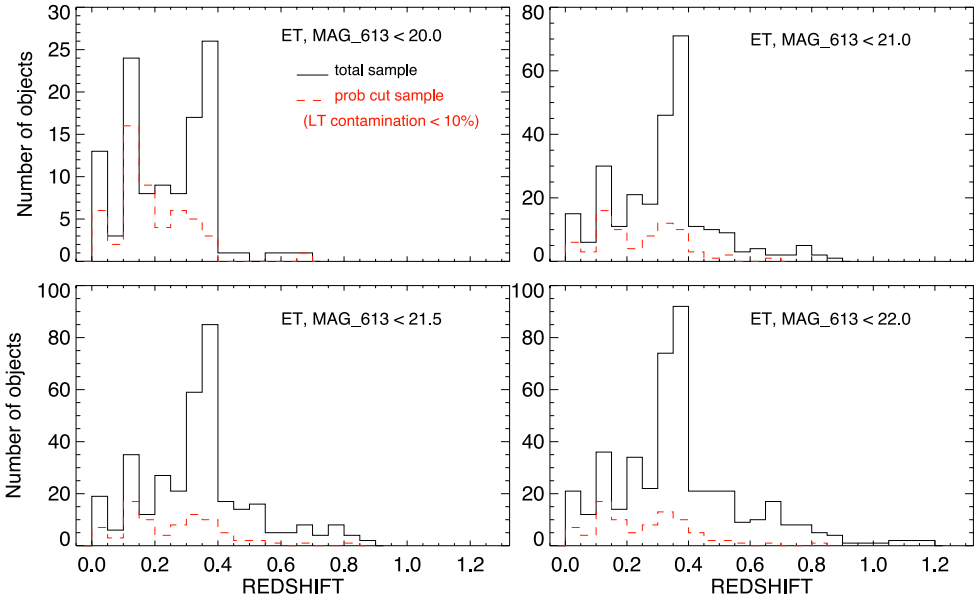

Figure 8. Redshift distributions of total ET samples (black solid lines), selected through the magnitude cuts defined in Section 3.2.2, and of p_{10} morphologically classified ET sample (red dashed lines), selected after applying the probability cuts obtained in the COSMOS/*HST* comparisons. Distributions are compared for four magnitude cuts, since above 21.5–22.0 the contamination of LT objects is too significant. On the other hand, for LT galaxies (Fig. 9) we compare the redshift distributions for all six magnitude cuts.

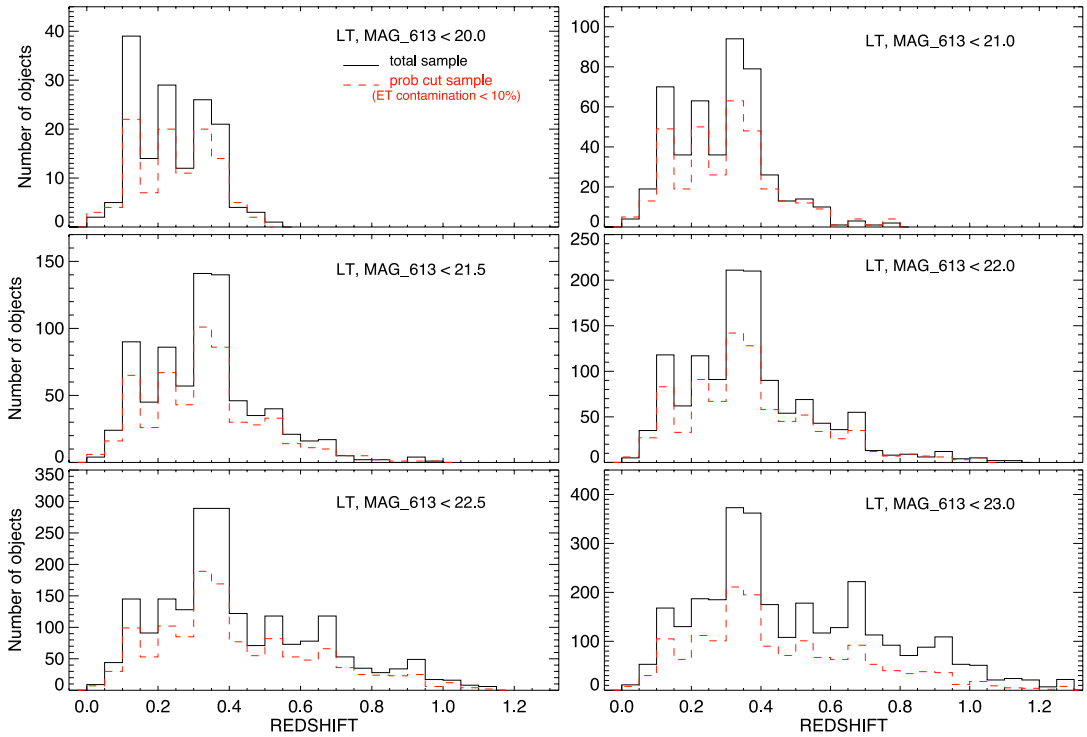

Figure 9. Same as Fig. 8, but for LT galaxies observed in six magnitude bins.

(classification described in Huertas-Company et al. 2009), putting all their 'ellipticals' as ET, and spirals and irregulars as LT. Sample marked with red dashed lines represent ALH p_{10} sample, obtained through COSMOS/*HST* comparisons (criteria summarized in Table 2). The redshift distributions of the LT galaxies are consistent in both samples which indicate that there is no redshift selection bias. However, for ET galaxies, the redshift distribution of the p_{10} sample peaks at lower values (0.2–0.4) than the full sample. This reflects the fact that ET galaxies at high redshift cannot be identified properly because we start being seriously affected by the spatial resolution. The reduction of contamination therefore results in a sample of ET galaxies with $z < 0.5$. We will discuss this selection bias in terms of stellar mass and absolute magnitudes in Section 6.2.

(ii) *Size*. Fig. 10 (left-hand panel) shows the comparison between the normalized distributions of galaxy size at 90 per cent of flux for

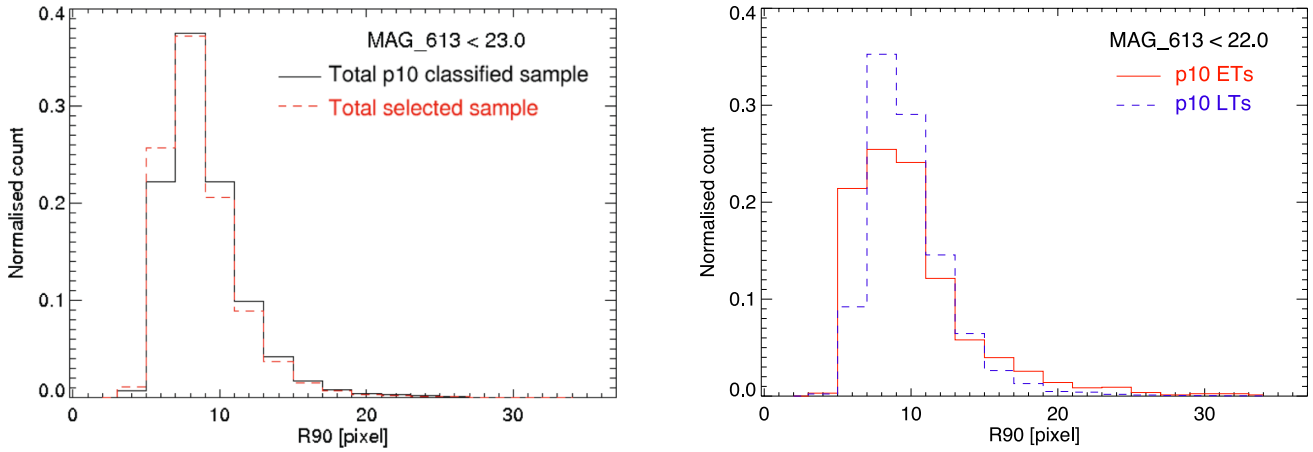

Figure 10. Normalized distributions of radius at 90 per cent of flux for: total selected (red dashed lines) and p_{10} classified (black solid lines) samples (left-hand panel), and ET (red solid lines) and LT (blue dashed lines) galaxies (right-hand panel).

the total selected sample of >40 000 sources (see Section 2.1) and total p_{10} classified sample of 22 051 galaxies (see above), down to magnitudes 23.0 in the $F613W$ band. Size distributions of ET and LT galaxies classified with <10 per cent contamination are shown in the right-hand panel of the same figure, down to magnitudes 22.0 in the $F613W$ band (where we still have classification into both types). We can confirm that down to the magnitude limit of 23.0 studied in this paper, we do not find any significant influence of galaxy size on to the morphological classification. KS statistic shows that the size distributions of total selected and p_{10} ET and LT classified samples are consistent, having the KS probability parameter of 0.9994 (out of 1.0).

5.2 Properties of ET sample

We compared different properties between the p10 ET sample and galaxies classified as ET with higher levels of contamination by LT type. Fig. 11 represents these comparisons, and shows the normalized distributions of apparent $F613W$ magnitude, redshift, stellar mass and radius at 90 per cent of flux. As can be seen, galaxies classified as ET with higher contamination levels (>20 per cent) are in general fainter and more distant, and have higher population of sources with lower stellar masses and sizes in comparison with the p10 ET sample. We performed the KS statistical test comparing (for all properties) p10 with other ET classified samples, obtaining in all cases that distributions are significantly different. Using the *t*-means statistical test, for all sources, except when comparing stellar mass and size between p10 and 10 per cent < Contamination ≤20 per cent samples, we found that p10 and other samples have significantly different means.

6 MAIN PROPERTIES OF THE MORPHOLOGICAL CATALOGUE

In this section, we test the general properties of p10 galaxies classified as ETs or LTs and check that they populate the expected regions in classical relations (e.g. morphological diagrams, colour–magnitude and colour–stellar mass relations).

6.1 Morphological parameters

The p_{10} thresholds and hence the morphological classification were calibrated in the ALH-4 field, and extrapolated to the other fields. We checked in Fig. 12 that the classified galaxies on the full survey populate the expected regions in the morphological planes. The central panels, from left to right and from top to bottom, show relations between the smoothness and Gini; M20 moment of light and Gini; asymmetry and Abraham concentration index; M20 moment of light and Abraham concentration index; Gini and Abraham concentration index; and finally, smoothness and M20. In all plots, we present $F613W \leq 22.0$ p10 sample, where ET galaxies are marked with red diamonds, and LT with blue triangles. Top and right-hand panels of each diagram, represent normalized distributions of corresponding parameters for ET (red solid lines) and LT (blue dashed lines) galaxies. In addition, violet contours in the central diagrams and dotted violet histograms represent the distribution of $22.0 < F613W \leq 23.0$ p10 galaxies classified as LT (see Section 5). In all central plots and histograms, there is a clear separation between the two morphological types. It is well known that, on average, ET galaxies present higher central light concentrations (CABR, GINI and CCON), but lower asymmetries and small-scale structures (SMOOTH, ASYM, M20) than LTs. These expected trends are observed in the full sample which confirms that our p_{10} morphological classification is reliable in all ALH fields. Areas marked with black dot–dashed lines on the central plots define the regions populated by ∼80 per cent of ET and ∼20 per cent of LT galaxies in our sample, down to magnitudes 22.0 in the $F613W$ band. For each morphological parameter, we report in Table 5 the criteria obtained from normalized distributions to isolate most of the ET and LT galaxies, and also indicate for each region the expected population of both classes.

6.2 Absolute magnitudes, stellar masses and colours

We measured the *K*-corrections by means of the IDL routine KCORRECT (Blanton et al. 2007). We implemented all 23 ALH filters and their response files into the code, fitting our 23 point SED with about 500 available spectral templates (see Blanton et al. 2007, for more information about the templates and SED fitting). 91, 98 and 93 per cent of sources have the difference between the original magnitudes and those recovered from the final KCORRECT fits below 0.3, 0.2 and 0.2 in three observed bands, respectively (or 98, 100 and 100 per cent if the difference is <0.5), showing a good quality of our *K*-corrections.

We then obtained the rest-frame magnitudes, absolute magnitudes, rest-frame colours and luminosities of all sources presented

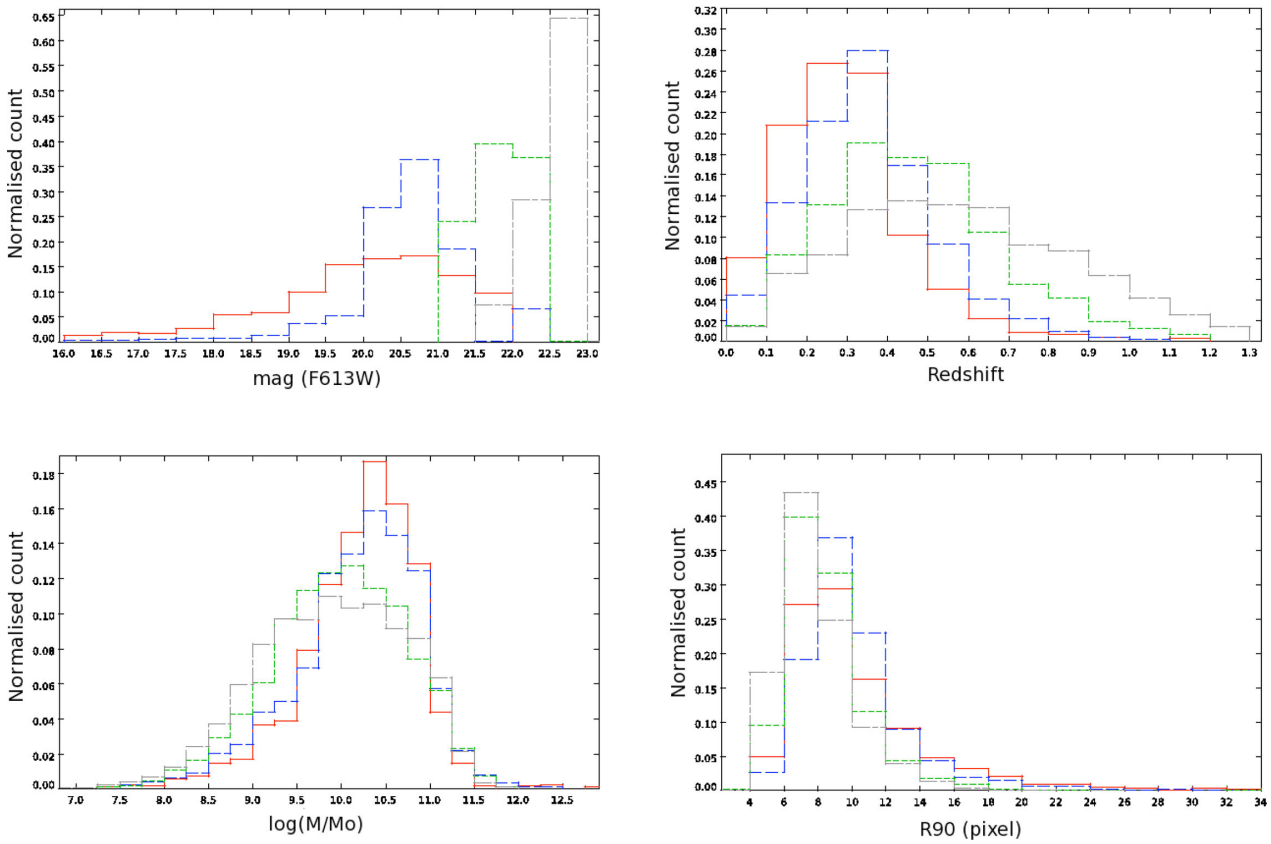

Figure 11. Normalized distributions of $F613W$ apparent magnitude (top left), redshift (top right), stellar mass (bottom left) and radius at 90 per cent of flux (bottom right) of ET galaxies classified with different levels of contamination: ≤10 per cent (red solid line), 10–20 per cent (blue dashed line), 20–50 per cent (green dotted line) and >50 per cent (grey dot–dashed line).

in the catalogue we described above. Moreover, from the KCORRECT SED fits, we obtained the stellar masses and star formation histories, using the Chabrier (2003) stellar initial mass function. We do not have any mid-infrared nor far-infrared data as an input information for SED fitting, the reddest band we have is K. This means that, following the well-known mathematical relation between the observed and rest-frame wavelengths ($1 + z = \lambda_{\rm observed}/\lambda_{\rm emited}$), we can count with our K-corrections and measured absolute magnitudes and stellar masses down to redshifts ∼1.3 in all 20 optical bands, and down to ∼0.8 and ∼0.4 in the J and H NIR bands, respectively. We compared obtained stellar masses with those measured with BPZ code. The typical scatter between masses is 0.22 dex for all p_{10} classified galaxies, or 0.17 dex for ETs and 0.21 dex for LTs. Stellar masses in all ALH surveys measured with BPZ will be released in a companion paper (Molino et al. 2013).

Fig. 13 shows the $F613W$ band absolute magnitude–redshift (left-hand panel) and stellar mass–redshift (right-hand panel) relations of ET and LT galaxies (with $F613W \leq 22.0$). Notice the lack of ETs at higher redshifts which is a consequence of the selection effects discussed in Section 5.1. Compared to COSMOS/*HST* data, we trace the same population of LT galaxies (see figs 1 and 2 in Tasca et al. 2009, and fig. 2 in Pannella et al. 2009). However, for ET galaxies, we are incomplete including lower redshifts. We cannot distinguish ET sources from LTs fainter than −18.0 or −20.0 in the $F613W$ band at redshifts 0.2 or 0.5, respectively.

In Fig. 14, we tested the distributions of our p10 $F613W \leq 22.0$ ET (red diamonds and solid lines) and LT (blue triangles and dashed lines) selected galaxies on standard colour–magnitude and colour–stellar mass diagrams. We measured the rest-frame colour between the $F458W$ and $F892W$ ALH bands, which corresponds to approximately central wavelengths of standard Johnson B and SDSS z bands. This colour is compared with the absolute magnitude in the $F458W$ band (left-hand panel) and stellar mass (right-hand panel). We also present, for both morphological types, the histograms with normalized distributions of corresponding parameters, including also the p10 LT galaxies selected in the magnitude range $22.0 < F613W \leq 23.0$. In all diagrams, we can see a clear bimodal distribution between the 'red sequence' and 'blue cloud' galaxies, widely discussed in the literature (e.g. Strateva et al. 2001; Bell et al. 2003; Hogg et al. 2003; Cassata et al. 2007; Melbourne et al. 2007, are only some of them), where the majority of ET and LT galaxies are located, respectively. This bimodal distribution of galaxies on both colour–magnitude and colour–stellar mass diagrams is sometimes used to distinguish between the ET and LT populations of galaxies (e.g. Bell et al. 2003; Faber et al. 2007; Franzetti et al. 2007). However, it was shown that a fraction of spirals and irregulars with high extinctions or quenched stellar formations can reside in the red sequence, or earlier types with bursts of star formation in the blue cloud (e.g. Bell 2008; Williams et al. 2009; de la Torre et al. 2011; Oteo et al. 2013a,b). Black horizontal line in Fig. 14 shows the regions where 70 per cent of our $F613W \leq 22.0$ ET and LT p10 galaxies lie in the red sequence (B458–z892 > 1.12) and in the blue cloud (B458–z892 < 1.12; 85 per cent of LT galaxies selected with magnitudes $22.0 < F613W \leq 23.0$), respectively. Through the colour–stellar mass relation in Fig. 14 (left-hand panel), we defined

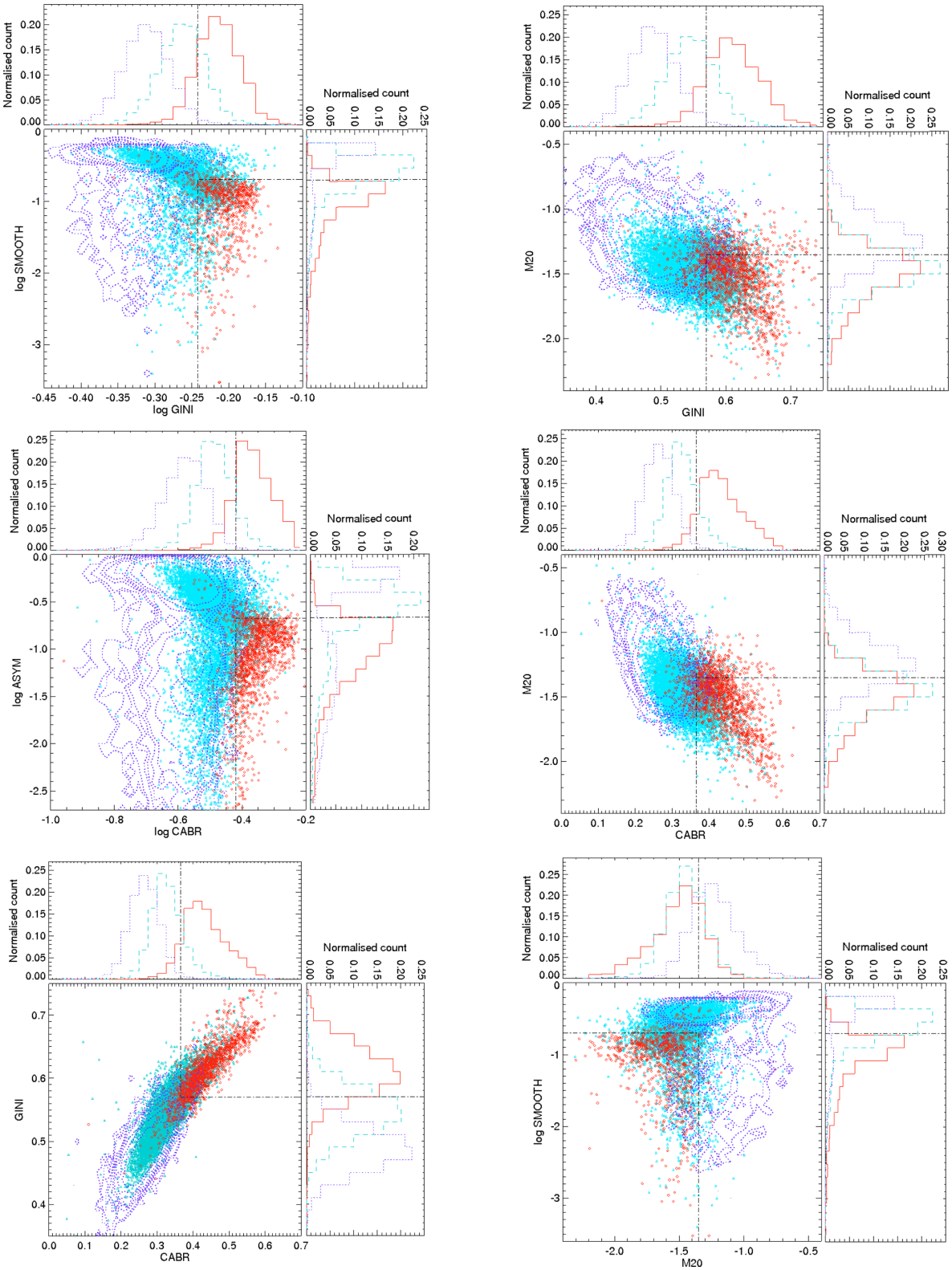

Figure 12. Central plots: morphological diagrams representing the relation between the logarithms of smoothness and Gini coefficient (top left), M20 moment of light and Gini (top right), logarithms of asymmetry and Abraham concentration indices (middle left), M20 moment of light and Abraham concentration index (middle right), Gini and Abraham concentration index (bottom left), and finally, logarithm of smoothness versus M20 moment of light (bottom right). In all plots, of p10 sources classified down to $F613W \leq 22.0$, ET galaxies are marked with red diamonds, and LT with blue triangles (darker and brighter symbols, respectively, when printed in black and white). Boxes marked with dot–dashed black lines define the locus of ∼80 and ∼20 per cent of galaxies from ET and LT samples, respectively. Top and right histograms: normalized distributions of corresponding parameters represented on the central plots of ET (red solid lines) and LT (blue dashed lines) galaxies. Dot–dashed black lines shown on the histograms are values that separate the majority of sources in two classes. Green dotted contours and histograms represent the distributions of p10 LT galaxies with magnitudes $22.0 < F613W \leq 23.0$.

Table 5. Distribution of morphological parameters of ET and LT galaxies.

Parameter	ET criteria	ET pop.	LT criteria	LT1* pop.	LT2** pop.
GINI	>0.57	85 per cent	<0.57	74 per cent	98 per cent
log SMOOTH	<−0.70	88 per cent	>−0.70	46 per cent	22 per cent
M20	<−1.35	78 per cent	>−1.35	25 per cent	75 per cent
CABR	>0.36	92 per cent	<0.36	79 per cent	99 per cent
log ASYM	<−0.67	86 per cent	>−0.67	62 per cent	44 per cent
CCON	>2.40	75 per cent	<2.40	84 per cent	95 per cent

* $F613W \leq 22.0$
** $22.0 < F613W \leq 23.0$.

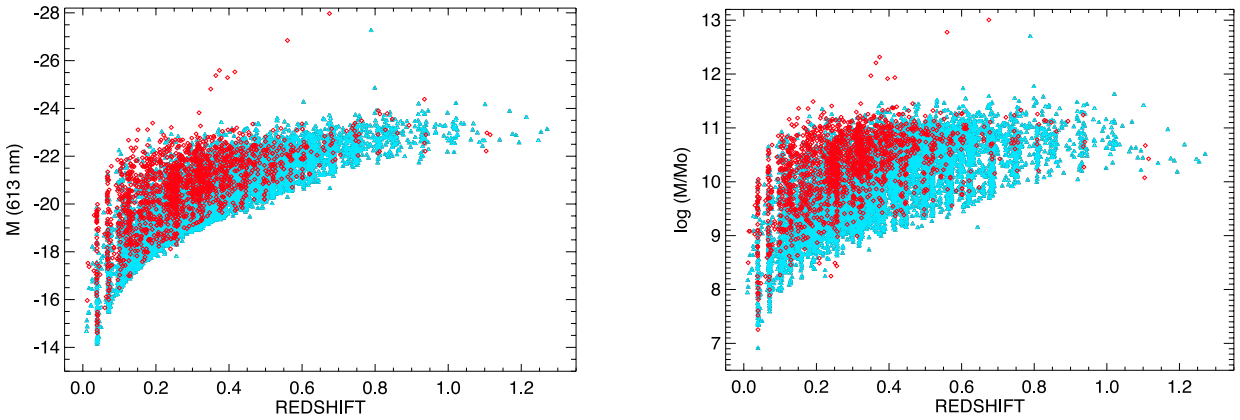

Figure 13. Relation between the redshift and $F613W$ absolute magnitude (left), and logarithm of stellar mass in solar mass units (right) for ET (red diamonds) and LT (blue triangles) galaxies down to magnitudes $F613W \leq 22.0$.

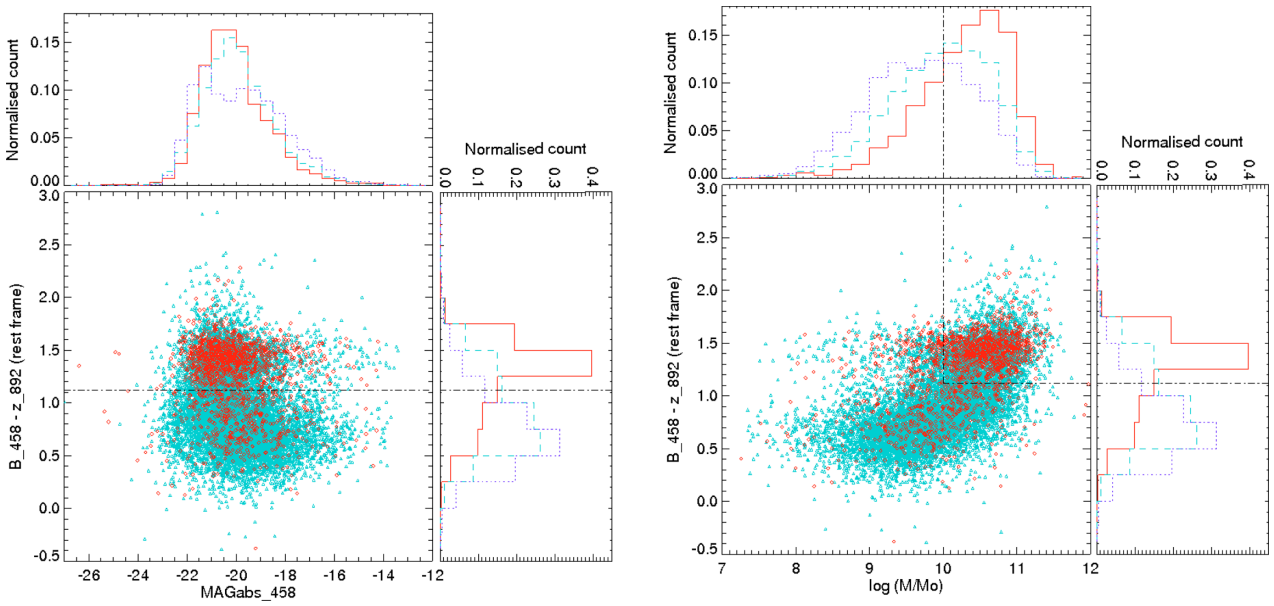

Figure 14. Left: relation between the rest-frame $B - z$ colour and absolute magnitude in $F458W$ band (central diagram) of ET (red diamonds) and LT (blue triangles) galaxies down to magnitudes $F613W \leq 22.0$. To estimate the colour we used the information from the $F458W$ and $F892W$ ALH bands. Histograms present the normalized distributions of compared parameters: absolute magnitude (above the central plot) and colour (to the right of the central diagram) for ET (red solid lines) and LT (blue dashed lines) sources. Dotted violet histograms show the distributions of LT galaxies with magnitudes $22.0 < F613W \leq 23.0$. Right: relation between the rest-frame $B - z$ colour and stellar mass (central diagram). Histograms show the normalized distributions of analysed parameters, as in previous case. All symbols and lines have the same significance as in left diagram.

the locus where ∼60 per cent of our ET galaxies are located, having B458–z892 colour >1.12 and stellar masses log $M/M_o > 10.0$. On the other hand, ∼25 and ∼10 per cent of $F613W \leq 22.0$ and $22.0 < F613W \leq 23.0$ LT galaxies reside in this region.

7 SUMMARY AND CONCLUSIONS

We presented the morphological classification of >40 000 galaxies in the ALH survey (in seven fields), classifying all galaxies in two broad morphological classes: ETs and LTs. With this paper we

release the low-contamination catalogue of 22 051 galaxies classified with the contamination lower than 10 per cent. We classified 1640 and 10 322 ET (down to redshifts ∼0.5) and LT (down to redshifts ∼1.0) galaxies, respectively, with magnitudes $F613W \leq 22.0$. In addition, for magnitude range $22.0 < F613W \leq 23.0$, we classified other 10 089 LT galaxies with redshifts ≤1.3. ALH is a photometric survey, having for all detections observations in 23 optical and NIR bands, with a large number of detected sources (>670 000) down to photometric completeness of SDSS $r \sim 25.0$, large covered area and precisely measured photometric redshifts (obtained through SED fittings of 23 point spectra), ALH is an ideal survey for tracing the cosmic variance and cosmic evolution. However, morphological properties and classification of galaxies are crucial for any kind of galaxy formation and evolution studies. Low-contamination, highly populated morphological catalogue presented in this paper, together with the precise measured photometric redshifts and photometric properties of the ALH survey, presents an important addition to other data sets for studying the morphological properties of extragalactic sources and their evolution (taking into account the variety of covered ALH fields). Some of these studies include: the evolution of ETs and LTs down to redshifts ∼0.5 and ∼1.3, respectively, their star formation histories, morphological properties and evolution of active galaxies and their comparison with non-active galaxies, as well as statistical comparisons between the morphological and SED fitting classifications. Using the ground-based data, the Galaxy Zoo survey classified morphologicaly ∼900 000 galaxies down to redshifts ∼0.25 (Lintott et al. 2011). On the other hand, the COSMOS survey provides morphological classification of >200 000 galaxies (Tasca et al. 2009) at high redshifts. With this work we provide the astronomical society with the additional morphological information of >22 000 high-redshift galaxies, but observed in seven different fields which may have an important constraints on cosmic variance and galaxy evolution studies.

To select the sample for morphological classification, we first separated galaxies from point-like sources, then we selected sources with good photo-z measurements (observed in all filters and with BPZ_ODDS parameter > 0.2), and finally, we selected sources photometrically, considering only detections with magnitudes ≤23.0 in the $F613W$ band and photometric errors <0.5 (to make sure to deal with reliable extended versus point-like source selections, photometric and photo-z measurements). The final selected sample has 43 665 sources.

We used the GALSVM code in our classification, testing the morphology through a sample of 3000 visually classified local galaxies. We redshifted these galaxies and scaled them in luminosity, to reproduce the redshift and magnitude distributions of our ALH sources. Morphological classification was carried out in the $F613W$ band, the most efficient one of all 20 optical bands, and for six magnitude cuts, down to 23.0. For each ALH galaxy, we measured seven morphological parameters and the averaged $p_{10}^{\rm E}$ probability that the galaxy is ET (probability that the galaxy is LT is then $p_{10}^{\rm L} = 1 - p_{10}^{E}$), obtained through 15 MC simulations.

Our classification is calibrated against the COSMOS field, using the *HST*/ACS images. We used this calibration to determine the probability cuts (in each of six magnitude cuts) to select ET and LT galaxies with the contamination lower than 10 per cent. With the obtained probability cuts, we can recover ∼70 per cent of ET galaxies down to magnitudes 20.0, 30–40 per cent down to 21.5, and 20–30 per cent down to 22.0 in the $F613W$ band. On the other hand, for LT galaxies, we recover ∼70 per cent down to magnitudes 22.0, ∼60–70 per cent, down to 22.5, and ∼30 per cent down to 23.0. We tested our classification in whole ALH survey through different morphological diagrams and general ET and LT relations (e.g. colour–magnitude and colour–stellar mass diagrams), obtaining the expected distributions in all of analysed relations.

The complete, low-contamination (<10 per cent) catalogue of 22 051 galaxies provides all measured morphological parameters, averaged probabilities, morphological types, magnitudes, physical sizes and redshifts. The catalogue is available in the electronic version of this paper and through the ALH webpage http://alhambrasurvey.com/, while the description of columns and the small example of five sources are presented in the appendix.

ACKNOWLEDGEMENTS

We thank the referee Chris Lintott for constructive comments which improved the paper significantly. This research was supported by the Junta de Andalucía through projects PO8-TIC-03531 and TIC114, the Spanish Ministry of Economy and Competitiveness (MINECO) through projects AYA2006-14046, AYA2010-15169, AYA2010-22111-C03-02, AYA2011-29517-C03-01, and the Generalitat Valenciana through project GV/Prometeo 2009/064. MP acknowledges financial support from JAE-Doc program of the Spanish National Research Council (CSIC), co-funded by the European Social Fund. Based on observations collected at the Centro Astronómico Hispano Alemán (CAHA) at Calar Alto, operated jointly by the Max-Planck Institut fur Astronomie and the Instituto de Astrofísica de Andalucía (CSIC). The CEFCA is funded by the Fondo de Inversiones de Teruel, supported by both the Government of Spain (50 per cent) and the regional Government of Aragón (50 per cent). In this work, we made use of Virtual Observatory Tool for OPerations on Catalogues And Tables (TOPCAT). We thank the developers of Ubuntu, PYTHON, NUMPY, SCIPY and MATPLOTLIB for making their work public and available to all scientific community.

REFERENCES

Abraham R. G., Valdes F., Yee H. K. C., van den Bergh S., 1994, ApJ, 432, 75
Abraham R. G., van den Bergh S., Glazebrook K., Ellis R. S., Santiago B. X., Surma P., Griffiths R. E., 1996, ApJS, 107, 1
Abraham R. G., van den Bergh S., Nair P., 2003, ApJ, 588, 218
Aguerri J. A. L., Iglesias-Paramo J., Vilchez J. M., Muñoz-Tuñón C., 2004, AJ, 127, 1344
Aguerri J. A. L., Iglesias-Páramo J., Vílchez J. M., Muñoz-Tuñón C., Sánchez-Janssen R., 2005, AJ, 130, 475
Andrae R., Jahnke K., Melchior P., 2011a, MNRAS, 411, 385
Andrae R., Melchior P., Jahnke K., 2011b, MNRAS, 417, 2465
Aparicio-Villegas T. et al., 2010, AJ, 139, 1242
Baillard A. et al., 2011, A&A, 532, 74
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJS, 149, 289
Bell E. F., 2008, ApJ, 682, 355
Benítez N., 2000, ApJ, 536, 571
Benítez N. et al., 2004, ApJS, 150, 1
Benítez N. et al., 2009, ApJ, 692, L5
Bershady M. A., Jangren A., Conselice C. J., 2000, AJ, 119, 2645
Bertin E., 2009, SExtractor v.13 User's Manual
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Blanton M. R., Roweis S., 2007, AJ, 133, 734

Casata P. et al., 2007, ApJS, 172, 270
Castander F. J., 1998, Ap&SS, 263, 91
Chabrier G., 2003, PASP, 115, 763
Conselice C. J., Bershady M. A., Dickinson M., Papovich C., 2000, ApJ, 529, 886
Cristóbal-Hornillos D. et al., 2009, ApJ, 696, 1554
de la Torre S. et al., 2011, MNRAS, 412, 825
de Souza R. E., Gadotti D. A., dos Anjos S., 2004, ApJS, 153, 411
de Vaucouleurs G., 1948, Ann. Astrophys., 11, 247
Faber S. M. et al., 2007, ApJ, 665, 265
Fasano G. et al., 2012, MNRAS, 420, 926
Fioc M., Rocca-Volmerange B., 1997, A&A, 326, 950
Folkes S. R., Lahav O., Maddox S. J., 1996, MNRAS, 283, 651
Franzetti P. et al., 2007, A&A, 465, 711
Hogg D. W. et al., 2003, ApJ, 585, 5
Holwerda B. W., 2009, Source Extractor for Dummies
Hubble E. P., 1926, ApJ, 64, 321
Hubble E. P., 1936, Realm of the Nebulae. Yale University Press, New Haven
Huertas-Company M., Rouan D., Tasca L., Soucail G., Le Fèvre O., 2008, A&A, 478, 971
Huertas-Company M. et al., 2009, A&A, 497, 743
Huertas-Company M., Aguerri J. A. L., Bernardi M., Mei S., Sánchez Almeida J., 2011, A&A, 525, 157
Humason M. L., 1931, ApJ, 74, 35
Jiménez-Teja Y., Benítez N., 2012, ApJ, 745, 150
Kelly B. C., McKay T. A., 2004, ApJ, 127, 625
Kelly B. C., McKay T. A., 2005, AJ, 129, 1287
Lintott C. J. et al., 2008, MNRAS, 389, 1179
Lintott C. J. et al., 2011, MNRAS, 410, 166
Lotz J. M., Primack J., Madau P., 2004, AJ, 128, 163
Matute I. et al., 2012, A&A, 542, 20
Melbourne J., Phillips A. C., Harker J., Novak G., Koo D. C., Faber S. M., 2007, ApJ, 660, 81
Méndez-Abreu J., Aguerri J. A. L., Corsini E. M., Simonneau E., 2008, A&A, 478, 353
Moles M. et al., 2008, AJ, 136, 1325
Molino A. et al., 2013, MNRAS, preprint (arXiv:1306.4968)
Morgan W. W., Mayall N. U., 1957, PASP, 69, 291
Nair P. B., Abraham R., 2010, ApJS, 186, 427 (N&A)
Ngan W., van Waerbeke L., Mahdavi A., Heymans C., Hoekstra H., 2009, MNRAS, 396, 1211
Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
Oteo I. et al., 2013a, MNRAS, 433, 2706
Oteo I. et al., 2013b, MNRAS, preprint (arXiv:1306.1121)
Pannella M. et al., 2009, ApJ, 701, 787
Peng C. J., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266
Peng C. J., Ho L. C., Impey C. D., Rix H.-W., 2010, AJ, 139, 2097
Pović M. et al., 2012, A&A, 541, 118
Prieto M., Gottesman S. T., Aguerri J.-A. L., Varela A.-M., 1997, AJ, 114, 1413
Prieto M., Aguerri J. A. L., Varela A. M., Muñoz-Tuñón C., 2001, A&A, 367, 405
Reynolds J. H., 1920, MNRAS, 80, 746
Sánchez Almeida J., Aguerri J. A. L., Muñoz-Tuñón C., de Vicente A., 2010, ApJ, 714, 487
Scarlata C. et al., 2007, ApJS, 172, 406
Scoville N. et al., 2007, ApJS, 172, 1
Sérsic J. L., 1963, BAAA, 6, 41
Simard L. et al., 2002, ApJS, 142, 1
Simard L., Mendel J. T., Patton D. R., Ellison S. L., McConnachie A. W., 2011, ApJS, 196, 11
Strateva I. et al., 2001, AJ, 122, 1861
Tasca L. A. M. et al., 2009, A&A, 503, 379
Trujillo I., Aguerri J. A. L., Gutiérrez C. M., Cepa J., 2001, AJ, 122, 38
Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, ApJ, 691, 1879
Wuyts S., Labbé I., Schreiber N. M. F., Franx M., Rudnick G., Brammer G. B., van Dokkum P. G., 2008, ApJ, 682, 985

APPENDIX A: ALH MORPHOLOGICAL CATALOGUE

In this section, we describe the high-quality, two-class morphological catalogue in the ALH survey, with the total contamination lower than 10 per cent. Catalogue contains morphological, photometric, size and photometric redshift information of 22 051 galaxies. Of those, 1640 and 20 411 were classified as ETs and LTs, down to magnitudes 22.0 and 23.0 and photometric redshifts of $\sim$0.5 and $\sim$1.3, respectively.[14] Table A1 shows an example of the format and content of the catalogue for five sources. The catalogue is available in the electronic addition of this paper or through the ALH website http://alhambrasurvey.com/. The column entries are as follows:

(i) Column 1 (ID): identification number.

(ii) Column 2 (FIELD_P_CCD): ALH field, pointing and CCD.

(iii) Column 3 (ID_phot): identification number in the photometric catalogue (Husillos et al., in preparation); equal to NUMBER parameter in the original catalogue.

(iv) Column 4 (ID_zphot): identification number in the photometric redshift catalogue (Molino et al. 2013); equal to ID parameter in the original catalogue.

(v) Columns 5 and 6 (RA, Dec.): equinox J2000.0 right ascension and declination in degrees of the centroid.

(vi) Column 7 (FLAGS): SExtractor FLAG parameter (Bertin & Arnouts 1996) contained in the photometric catalogue.

(vii) Columns 8 and 9 (m_{458}, errm_{458}): apparent magnitude in the $F458W$ band and its error.

(viii) Columns 10 and 11 (m_{613}, errm_{613}): apparent magnitude in the $F613W$ band and its error.

(ix) Columns 12 and 13 (m_{892}, errm_{892}): apparent magnitude in the $F892W$ band and its error.

(x) Column 14 (logR50): logarithm of radius at 50 per cent of flux in kpc.

(xi) Column 15 (logR90): logarithm of radius at 90 per cent of flux in kpc.

(xii) Column 16 (MUMEAN): mean surface brightness measured by GALSVM.

(xiii) Column 17 (ELLIPTICITY): ellipticity parameter measured by SExtractor.

(xiv) Column 18 (ASYM): asymmetry index measured by GALSVM, defined as in Conselice et al. (2000).

(xv) Column 19 (CABR): Abraham concentration index, measured by GALSVM and defined as in Abraham et al. (1996).

(xvi) Column 20 (Gini): Gini coefficient measured by GALSVM and defined as in Abraham et al. (2003).

(xvii) Column 21 (SMOOTH): Smoothness of the source, measured by GALSVM and defined as in Conselice et al. (2000).

(xviii) Column 22 (M20): Moment of light at 20 per cent, measured by GALSVM and defined as in Lotz et al. (2004).

(xix) Column 23 (CCON): Conselice–Bershady concentration index, measured by GALSVM and defined as in Bershady et al. (2000).

(xx) Columns 24 and 25 (p^{E}_{10}_AVG and errp^{E}_{10}_AVG): averaged probability that the galaxy is ET, and its error. The probability is measured from other 15 probabilities (see Section 3.2.2). It takes

[14] In case you need less-strict catalogue, with the contamination above 10 per cent, or the total one with 43 665 sources down to magnitudes 23.0, as well as any additional information presented in the paper but not available in the published catalogue, please contact us at mpovic@iaa.es.

Table A1. Morphological catalogue of 22 051 galaxies classified in the ALH survey with the contamination <10 per cent.

ID	FIELD_P_CCD	ID_phot	ID_zphot	RA (degrees)	Dec. (degrees)	FLAGS
m_{458}	errm$_{458}$	m_{613}	errm$_{613}$	m_{892}	errm$_{892}$	logR50 (kpc)
logR90 (kpc)	MUMEAN	ELLIPTICITY	ASYM	CABR	GINI	SMOOTH
M20	CCON	p^E_{10}_AVG	errp^E_{10}_AVG	CLASS	REDSHIFT	
1	ALH2_p1_ccd1	8895	81421100234	37.436 48	1.264 531	0
23.0	0.098	22.04	0.022	21.067	0.084	0.723
0.985	24.5309	0.2143	0.2686	0.3075	0.532	0.0
−1.2281	2.1533	0.3034	0.0417	LT	0.666	
2	ALH2_p1_ccd1	9202	81421100256	37.523 216	1.263 539	2
23.565	0.231	22.411	0.04	21.052	0.125	0.971
1.356	25.5367	0.4043	0.8212	0.2458	0.4558	0.633
−0.9595	2.3292	0.4649	0.1535	LT	0.598	
3	ALH2_p1_ccd1	9353	81421100262	37.532192	1.263552	2
23.033	0.122	22.323	0.032	21.524	0.16	0.879
1.125	25.3686	0.1158	0.0211	0.2054	0.4731	0.0
−0.9625	1.8086	0.3041	0.0437	LT	0.681	
4	ALH2_p1_ccd1	9453	81421100274	37.582 558	1.262 976	0
22.042	0.069	20.741	0.012	19.914	0.056	0.749
1.094	24.1509	0.1518	0.1118	0.3918	0.5719	0.1679
−1.4883	2.6371	0.8955	0.0948	ET	0.421	
5	ALH2_p1_ccd1	8568	81421100275	37.398 445	1.263 138	2
22.036	0.051	21.104	0.012	20.657	0.079	0.564
0.876	24.0459	0.4405	0.053	0.3176	0.5584	0.1136
−1.5274	2.2498	0.138	0.0389	LT	0.258	

values from 0 to 1, where small values indicate that the galaxy is LT. We measure $p^{\rm L}_{10}$ probability as $p^{\rm L}_{10} = 1 - p^{\rm E}_{10}$.

(xxi) Column 26 (CLASS): final morphological class, after applying the probability cuts described in Section 4.

(xxii) Column 27 (REDSHIFT): photometric redshift (Molino et al. 2013).

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table A1. Morphological catalogue of 22 051 galaxies classified in the ALH survey with the contamination <10 per cent (http://mnras.oxfordjournals.org/lookup/suppl/doi:10.1093/mnras/stt1538/-/DC1).

Please note: Oxford University Press are 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.