The impact from survey depth and resolution on the morphological classification of galaxies

M. Pović1*, I. Márquez1, J. Masegosa1, J. Perea1, A. del Olmo1, C. Simpson2, J. A. L. Aguerri3, B. Ascaso4,1, Y. Jiménez-Teja5, C. López-Sanjuan6, A. Molino7,1, A. M. Pérez-García3,8, K. Viironen6, C. Husillos1, D. Cristóbal-Hornillos6, C. Caldwell2, N. Benítez1, E. Alfaro1, T. Aparicio-Villegas5,1, T. Broadhurst9,10, J. Cabrera-Caño11, F. J. Castander12, J. Cepa8,3, M. Cerviño1,3, A. Fernández-Soto13,14, R. M. González Delgado1, L. Infante15, V. J. Martínez16,14, M. Moles6,1, F. Prada1, and J. M. Quintana1

1Instituto de Astrofísica de Andalucía (IAA-CSIC), Granada, Spain
2Astrophysics Research Institute, Liverpool John Moores University, Liverpool, UK
3Instituto de Astrofísica de Canarias (IAC), La Laguna, Tenerife, Spain
4GEPI, Paris Observatory, Paris, France
5Observatório Nacional, Rio de Janeiro, Brazil
6Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Teruel, Spain
7Instituto de Astrofísica, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, São Paulo, Brazil
8Departamento de Astrofísica, Facultad de Física, Universidad de la Laguna, La Laguna, Spain
9Department of Theoretical Physics, University of Basque Country, Bilbao, Spain
10IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
11Facultad de Física. Departamento de Física Atómica, Molecular y Nuclear, Universidad de Sevilla, Sevilla, Spain
12Institut de Ciencies de l’Espai. IEEEC/CSIC, Barcelona, Spain
13Instituto de Física de Cantabria (CSIC-UC), Santander, Spain
14Unidad Asociada Observatori Astronòmic (IFCA - UV), Valencia, Spain
15Departamento de Astronomía, Pontificia Universidad Católica, Santiago de Chile, Chile
16Observatori Astronòmic de la Universitat de València, Valencia, Spain

Accepted ??, Received ??; in original form ??

ABSTRACT

We consistently analyse for the first time the impact of survey depth and spatial resolution on the most used morphological parameters for classifying galaxies through non-parametric methods: Abraham and Conselice-Bershady concentration indices, Gini, M20 moment of light, asymmetry, and smoothness. Three different non-local datasets are used, ALHAMBRA and SXDS (examples of deep ground-based surveys), and COSMOS (deep space-based survey). We used a sample of 3000 local, visually classified galaxies, measuring their morphological parameters at their real redshifts (z ∼ 0). Then we simulated them to match the redshift and magnitude distributions of galaxies in the non-local surveys. The comparisons of the two sets allow to put constraints on the use of each parameter for morphological classification and evaluate the effectiveness of the commonly used morphological diagnostic diagrams. All analysed parameters suffer from biases related to spatial resolution and depth, the impact of the former being much stronger. When including asymmetry and smoothness in classification diagrams, the noise effects must be taken into account carefully, especially for ground-based surveys. M20 is significantly affected, changing both the shape and range of its distribution at all brightness levels. We suggest that diagnostic diagrams based on 2 - 3 parameters should be avoided when classifying galaxies in ground-based surveys, independently of their brightness; for COSMOS they should be avoided for galaxies fainter than F814 = 23.0. These results can be applied directly to surveys similar to ALHAMBRA, SXDS and COSMOS, and also can serve as an upper/lower limit for shallower/deeper ones.

Key words: surveys; galaxies: morphology; galaxies: fundamental parameters;
1 INTRODUCTION

Morphology is one of the main characteristics of galaxies, and the morphological classification has been central to many advances in the picture of galaxy formation and evolution. Different correlations between morphology and other galaxy properties have been studied, including the relation with stellar mass (e.g., Deng 2013), colour (e.g., Strateva et al. 2001), Hogg et al. 2003; Bell et al. 2003; Baldry et al. 2004; Weiner et al. 2005; Cirasulu et al. 2005; Melbourne et al. 2007; Cassata et al. 2007; Pović et al. 2013), nuclear activity (e.g., Blanton et al. 2003; Kelm et al. 2005), environment (e.g., Cassata et al. 2007), black hole mass (e.g., Kormendy & Richstone 1995; McLure et al. 2000; Kelm et al. 2005), luminosity (e.g., Blanton et al. 2003; Cirasuolo et al. 2005; Melbourne et al. 2007; Cassata et al. 2007), galaxy optical and near-infrared properties (e.g., Adams 1977; Heckman 1978; Ho, Filippenko, & Sargent 1995; Kauffmann et al. 2003; Choi et al. 2009; Pierce et al. 2007; Pović et al. 2012; Gabor et al. 2009), X-ray properties (e.g., Hickox et al. 2009; Pović et al. 2009a,b).

The methods of morphological classification of galaxies can be separated into three groups: i) visual (e.g., Lintott et al. 2008; Nair & Abraham 2010; Baillard et al. 2011; Lintott et al. 2011; Willett et al. 2011b), for classifying the nearby and well resolved galaxies, ii) parametric, based on the galaxy physical properties (e.g. de Vaucouleurs 1948; Sérsic 1963; Peng et al. 2002; 2010; Simard et al. 2002; 2011; de Souza et al. 2004; Burden et al. 2012) and mathematical parameters, where assuming an analytic model for fitting the galaxy (e.g. Kelly & McKay 2004; 2005; Ngnai et al. 2009; Andrae et al. 2011a,b; Jiménez-Teja et al. 2012), they perform the decomposition of well resolved galaxies and provide the properties of different structures, and finally, iii) non-parametric, which does not assume any particular analytic model.

The non-parametric methods are based on measuring the different galaxy quantities that correlate with the morphological types, i.e. colours (e.g., Strateva et al. 2001), spectral properties (e.g. Humason 1931; Morgan & Mavall 1957; Baldwin et al. 1981; Folkes et al. 1996; Sánchez-Almeida et al. 2010), or light distribution (e.g. Doy et al. 1993; Abraham et al. 1994; 1996; 2003; Bershady et al. 2000; Conselice et al. 2000; Graham et al. 2001; Lotz et al. 2004; Yamauchi et al. 2005). The non-parametric methods are less time-consuming in comparison with other methods, and can provide an easy and fast separation between regular and irregular (or disturbed) sources and/or early- and late-type galaxies, down to intermediate redshifts (∼1.5), or higher if dealing with space based, or good seeing ground-based data.

Over the past years different morphological diagrams have been applied, relating galaxy light concentration and asymmetry or smoothness parameters either to classify galaxies (Abraham et al. 1994; 1996; 1999; Conselice et al. 2000; Bershady et al. 2000; Graham et al. 2001a; Conselice et al. 2006; Cassata et al. 2007; Pović et al. 2009a,b; Deng 2013), or to select merger candidates (Lotz et al. 2010a,b; Urrutia et al. 2008; Villforth et al. 2014). However, some of the previous works showed that these parameters can depend significantly on e.g. the aperture definition for measuring the galaxy flux/radius (Strateva et al. 2001; Graham et al. 2001b; Lisker 2008) and/or on the resolution and signal-to-noise ratio (hereafter S/N; Conselice et al. 2000; Graham et al. 2001b; Lisker 2008; Huertas-Company et al. 2009; Andrae et al. 2011a,b; Pović et al. 2012; Cibinel et al. 2013; Carollo et al. 2013; Petty et al. 2014). Different trends of these parameters can be observed in relation to the data quality, spatial resolution and depth. Since a reliable morphological classification is essential for galaxy formation and evolution studies, it is crucial to disentangle how strong is this impact on each morphological parameter when dealing with different data sets. In most of previous works this observational bias was analysed either for a particular parameter, either for a particular observational condition (survey), or for a particular sample of galaxies. In this work we present, for the first time, a systematic study, where in a consistent way the analysis was carried out for all parameters, on large sample of galaxies, and in relation with different observational conditions.

We analysed the observational bias that might affect each of the six commonly used morphological parameters: Abraham concentration index, Gini coefficient, Conselice-Bershady concentration index, M20 moment of light, asymmetry index, and smoothness. We studied how this bias depends on the spatial resolution and magnitude/redshift distributions of galaxies, and how it affects the diagnostic diagrams used to classify galaxies. We used a visually classified sample of local galaxies and simulated them to map the observational conditions of different ground- and space-based deep surveys: the Advanced Large Homogeneous Area Medium Band Redshift Astronomical survey (Moles et al. 2008, ALHAMBRA), the Subaru/XMM-Newton Deep Survey (Furusawa 2008, SXDS), and the Cosmos Evolution Survey (Scoville et al. 2007, COSMOS). With this analysis our main goal is to measure a set of morphological parameters in the real (local) conditions and in the simulated conditions of non-local surveys in order to study how the spatial resolution and data depth affect each parameter and the commonly used morphological diagrams. We emphasize that is our aim to classify the galaxies in the three selected non-local surveys, neither reclassify local galaxies once they were scaled to map the conditions of ALHAMBRA, SXDS, and COSMOS. Comparing the parameters of local galaxies measured before and after shifting them in redshift and magnitude, we were able to put constraints on the main diagrams, observing how the position and the shape of the regions typical of early- and late-type galaxies change in the local and non-local conditions. Moreover, analysing the morphological diagrams in ground-based and space-based surveys, and at different magnitude cuts, we quantified how strong is the impact from spatial resolution and survey depth, respectively. The results obtained in this work can be applied to surveys similar to ALHAMBRA, SXDS, and COSMOS.

The paper is organized as follows: in Sec. 2 we described the used local and high-redshift samples and the corresponding surveys. The applied methodology is described in Sec. 3. The results obtained for local and simulated galaxies are showed in Sec. 4. In Sec. 5 we discussed how the impact from data depth and spatial resolution affects the morphological parameters and commonly used diagnostic diagrams for galaxy classification. Finally, we summarized our results in Sec. 6.

We assume the following cosmological parameters throughout the paper: ΩΛ = 0.7, Ωμ = 0.3, and H0 = 70 km s⁻¹ Mpc⁻¹. Unless otherwise specified, all magnitudes are given in the AB system (Oke & Gunn 1983).

¹ Hereafter, when using the expression 'observational bias' we will refer to the impact of spatial resolution and data depth on analysed morphological parameters.
Morphological classification of galaxies: impact from survey depth and resolution

2 THE DATA

2.1 Local sample
To test how the observational bias affects the analysed morphological parameters, we used a sample of 3000 local galaxies at $0.01 \lesssim z \lesssim 0.1$ (with a mean redshift of 0.04), observed in the Sloan Digital Sky Survey (SDSS) Data Release 4 (DR4) down to an apparent extinction-corrected magnitude of $g < 16$, and visually classified by Nair & Abraham (2010; hereafter N&A), using the $g$ and $r$ bands. The number of local galaxies is selected as a compromise between the computing time and the classification accuracy, since the computing time to train the Support Vector Machine (SVM) through the galSVM code used in this work (see Sec. 3) is totally dependent and very sensitive to the size of the training data set (Huertas-Company et al. 2008, 2009). The galaxies were selected randomly out of ~14,000 sources contained in the N&A catalogue, making sure that the selected sub-sample is representative in terms of the general properties of the whole data set (see Pović et al. 2013, and their Fig. 3): $g$ band magnitude, redshift, $g-r$ colour, morphological classification, and inclination in the case of the late-type galaxies. On the other hand, Figure 1 shows the redshift, $g$-band magnitude, and N&A morphological classification (T-Type in their work) distributions of the selected local sample. As can be seen from the T-Type histogram, the selected sample occupies all range of morphologies, from elliptical to irregular galaxies. In this work, all analyses were performed dividing galaxies into three morphological groups:

1) early-type galaxies (hereafter ET) - with T-Type $\leq 0$, including elliptical (E0, E+), lenticular (S0-, S0, S0+) and S0/a galaxies from the N&A classification.

2) early spirals (hereafter LT_et) - with $0 < $T-Type $\leq 4$, including Sa, Sab, Sb, and Sbc, and

3) late spirals and irregular galaxies (hereafter LT_lt) - with T-Type $> 4$, including Sc, Scd, Sd, Sdm, Sm, and Im galaxies.

We stress that throughout this work, we do not re-classify local galaxies and every time we specify the morphological type, it refers to that from N&A classification. Here we want to measure a set of the most used morphological parameters in real conditions of the local sample and in the simulated conditions of non-local surveys in order to evaluate how the spatial resolution and data depth affect each parameter. ET, LT_et, and LT_lt galaxies represent 45%, 38%, and 17% of the selected local sample, respectively, completely consistent with the whole N&A sample.

Irregular galaxies constitute only $< 1\%$ of all sources in the N&A catalogue. With such a small population, we were not able to provide reliable statistical analyses in comparison to the other two broader groups, EII/S0 and spirals. This is the reason why we do not study irregulars separately, and divide instead all late-type galaxies into two subgroups: earlier and later. In our LT_lt defined group, galaxies classified as Irr by N&A form only 2% of the whole population, being therefore insignificant for affecting the results represented in this paper.

2.2 Non-local samples
In this Section we describe the properties of non-local surveys used for studying the impact from spatial resolution and data depth on the morphological classification of galaxies. Since our work is statistical, to select the surveys we considered the following criteria:

- a large number of detected galaxies, either a deep survey covering a smaller area, and/or a shallower survey covering a large area,

- surveys with accurate photometry as well as accurate photometric redshift measurements available, and

- surveys with different properties in terms of spatial resolution and depth.

Finally, as a result of these criteria, we selected three non-local surveys for our analysis: deep ground-based surveys ALHAMBRA and SXDS (ALHAMBRA being shallower than SXDS, but with a large covered area and accurate photometric and photometric redshift measurements, and SXDS as an example of the deepest available ground-based data), and COSMOS as an example of deep space-based surveys. We had to take into account the number of selected surveys, since for each of them the applied computational procedure is time consuming. In the following sections we describe each of these surveys. Their basic properties are summarised in Table 1. We can not compare directly the results obtained between the
three analysed non-local surveys: first, because they have different depth, so the analysed magnitude cuts are different (see below), and secondly they have different photometric information available, so the bands that we used are different as well. We tested each survey separately, and tried to represent the obtained results depending on the properties of each one.

2.2.1 ALHAMBRA

The Advanced Large Homogeneous Area Medium Band Redshift Astronomical survey (ALHAMBRA; Moles et al. 2008) is a photometric survey that imaged ∼ 3.5 deg² of the sky, in eight different fields, through 20 optical and 3 near-infrared (NIR) filters. The average spatial resolution of the ALHAMBRA images is ∼ 1 arcsec (not higher than 1.6, and ranging mainly between 0.8 and 1.2 arcsec in the F613W band; see Table 1 in Pović et al. 2013, for more details). More than half million sources were detected down to the magnitude limit $r = 25.0$

The Bayesian Photometric Redshift code (BPZ2.0) was used to measure the photometric redshifts (Benitez et al. 2000, Molino et al. 2014), obtaining the accuracy for galaxies brighter than 22.5 and with magnitudes 22.5–24.5 in the constructed F814W band of $\sim \delta z/(1 + z) = 0.011$ and $\sim \delta z/(1 + z) = 0.014$, respectively. All images and catalogues are available through the ALHAMBRA webpage.

ALHAMBRA sample selection

In this work we used the same sample as described in Pović et al. (2013), whereby we will provide here only a brief summary. We selected only sources with the BPZ quality parameter ODDS > 0.2, expecting no more than 3% of photometric redshift outliers (Molino et al. 2014). For details about the galaxy-star separation see Pović et al. (2013). And finally, we selected only objects with magnitudes $< 23.0$ in the F613W band, taking into account that above this magnitude limit S/N ratio decreases significantly, leading to less accurate photometric redshift estimations and geometrical galaxy-star classifications. The F613W photometric band was selected due to its higher signal-to-noise (S/N) ratio in comparison with other filters (Aparicio-Villegas et al. 2010). The final selected sample has more than 43,000 galaxies.

2.2.2 SXDS

The Subaru/XMM-Newton Deep Survey (SXDS; Sekiguchi et al. 2004) is a large, multiwavelength, and one of the deepest optical ground based surveys, covering five continuous, rectangular subfields, and a total area of 1.22 deg². All subfields were observed in five broadband filters BVR, $i'z'$, detecting in each more than 800,000 sources down to the limiting magnitudes $B = 28.4$, $V = 27.8$, $R = 27.7$, $i' = 27.7$, and $z' = 26.6$ at 3σ and a typical seeing of 0.8. All images and photometric catalogues are described in Furusawa (2008), and can be downloaded from the SXDS webpage.

The photometric redshift distribution of each sample of galaxies was determined using multi-band photometry of objects in the K-band-selected catalogue of Hartley et al. (2013), having a depth of $K = 24.3$. In addition to the $uBVri'z'$ and Spitzer/IRAC channels 1 and 2 using the method of Simpson et al. (2013). Photometric redshifts and object classifications were derived using the code EAZY (Brammer et al. 2008), using a development of the method described in Simpson et al. (2013) and Caldwell et al. (in prep). Measured photometric redshifts have accuracy of $\delta z/(1 + z) = 0.031$.

SXDS sample selection

The sample we selected in the SXDS field is based on the available photometric redshifts. Therefore, we used only those galaxies that overlap with the K-band survey, instead of the entire SXDS field. The survey covers 0.77 deg², where the overlap region with the SXDS field is ∼ 0.63 deg², covering the entire central subfield and fractions of the side ones. We discarded all galaxies with poor photometric redshift fits, having $\chi^2 < 100$. To separate between point-like and extended sources, we applied the classification described above (Simpson et al. 2013 and Caldwell et al. in prep). Moreover, we selected all galaxies with $i' < 24.5$, which is the band that we will use in the following analysis, and the limit at which our sample is complete, taking into account the available redshifts and photometry. The final selected sample consists on more than 68,000 galaxies. Since the photometric redshift selection is based on the K-band, we might be missing the bluest sources detected in the SXDS. We did not perform a detailed analysis with respect to this, but we estimated by cross-matching i- and K-band catalogues an upper limit of < 9% of these sources, which should not affect significantly our results.

2.2.3 COSMOS

The Cosmic Evolution Survey (COSMOS; Scoville et al. 2007) is a deep, 2 deg² multiwavelength survey. In this work we use the observations from the HST ACS survey, where more than $1.2 \times 10^6$ sources were detected down to a limiting magnitude 26.5 in the filter F814W. The ACS data reduction and images are described in Koekemoer et al. (2007), while the photometric catalogue is presented in Leauthaud et al. (2007).

The photometric redshifts used in this work are presented in Ilbert et al. (2009), and were measured using the Le Phare code and photometric information from 30 broad, intermediate, and narrow-band filters from UV, optical, NIR and MIR bands. The obtained measurements show accuracy of $\sim \delta z/(1 + z) = 0.007$ for galaxies brighter than $i' = 22.5$. At magnitudes fainter than $i' < 24.0$ the accuracy is $\sim \delta z/(1 + z) = 0.012$, while for the very faint sample ($i' < 25.0$) the accuracy drops significantly with $\sim 20\%$ of outliers. All images and catalogues used in this work are available from the COSMOS webpage.

COSMOS sample selection

The COSMOS 2 deg² field is covered with 81 HST/ACS images (Koekemoer et al. 2007). Taking into account the number and size (20,480 pixels square) of the images, and the computing time that we would need to process all the data with the galSVM code (Sec. 3), we are finally using an area of 0.5 deg². Considering the depth of the COSMOS data, for the study presented in this paper, the selected area provides us with the statistical information completely comparable with the other two surveys. Moreover, we
checked that the magnitude and redshift distributions of galaxies in the selected field correspond to the total ones. Our main criteria for the COSMOS sample selection is again based on the accuracy of the photometric-redshifts. We selected only sources with \( z < 24.0 \), since the photo-z accuracy decreases significantly at fainter magnitudes, as described above. For the galaxy-star separation we used the classifications presented in both [Leauthaud et al. (2007)] and [Ilbert et al. (2009)]. Finally, the selected sample has more than 107,000 galaxies.

3 METHODOLOGY

Non-parametric methods of galaxy classification are usually applied either when dealing with high redshift and/or low resolution data, where the galaxy decomposition and profile fitting becomes impossible, or in large surveys, where due to the high number of detected sources it is necessary to use more automated and statistically based approaches. We test how much the morphological parameters described in Sec. 3.1 (basically related with the distribution of light within the galaxy and its shape) are sensitive to depth and spatial resolution, and how this introduced observational bias might affect the final galaxy classification. To do this, we used a sample of local galaxies with available detailed visual classification (see Sec. 3.1), and measured their morphological parameters in two cases:

1) at their real redshifts (magnitudes), i.e. at \( z \sim 0 \) (see Sec. 3.2), and
2) moving them to higher redshifts (therefore to fainter magnitudes) and lower resolution, to simulate the conditions of galaxies on deep, ground- and space-based non-local surveys (see Sec. 3.3).

In both cases all parameters are measured in a completely consistent way, using their definitions as described in the following section. Finally, we compared the results obtained in cases 1) and 2) to quantify the impact from data depth and spatial resolution on each morphological parameter, for each of the three analysed surveys (see Sec. 3.2 and 3.3).

3.1 Description of the tested morphological parameters

The six morphological parameters described here are by large, the most commonly used to distinguish between ET and LT galaxies, and to classify perturbed galaxies and interacting/merging systems. In the following definitions, when necessary, the galaxy centre is determined by minimizing the asymmetry index \( \text{Abraham et al.}[1996] \), see below), while the total flux is defined as the one contained within 1.5 times the Petrosian radius \( (r_p) \), measured by SExtractor; see Huertas-Company et al. 2008).

- **Abraham concentration index** (hereafter \( \text{CABR} \); [Abraham et al. 1996]) - measured as the ratio between the flux at 30% of the Petrosian radius (\( F_{30} \)) and the total flux (\( F_{tot} \)):
  \[
  \text{CABR} = \frac{F_{30}}{F_{tot}}. \tag{1}
  \]
  It gives values from 0 to 1, where higher values correspond to higher central light fractions.

- **Gini coefficient** (hereafter \( \text{GINI} \); [Abraham et al. 2003; Lotz et al. 2004]) - presented as the cumulative distribution function of galaxy pixel \( i \) values:
  \[
  \text{GINI} = \frac{1}{|X|n(n-1)} \sum_{i=1}^{n} (2i-n-1)|X_i|, \tag{2}
  \]
  where \( n \) is the total number of pixels in a galaxy, \( X \) the pixel flux value, and \( |X| \) the mean over all the pixel flux values, using as an aperture 1.5 \( r_p \). Usually, it correlates with CABR, having also values from 0 to 1, where higher ones correspond to galaxies with higher central concentrations. However, unlike CABR, it can distinguish between galaxies with shallow light profiles and those with the light concentrated in a few pixels, but outside the galaxy centre.

- **Conselice-Bershady concentration index** (hereafter \( \text{CONC} \); [Bershady et al. 2000]) - measured as the logarithm of the ratio of the circular radii containing 80% and 20% of the total flux:
  \[
  \text{CONC} = 5\log(r_{80}/r_{20}). \tag{3}
  \]
  In general, lower CONC values correspond to lower fractions of light in the central region.

- **M20 moment of light** (hereafter \( \text{M20} \); [Lotz et al. 2004]) - measured as the flux (\( f_i \)) in each pixel \( (x_i, y_i) \) multiplied by the squared distance to the centre \( (x_c, y_c) \) of the galaxy, summed over the 20% brightest pixels, and normalized by the total second-order moment (\( M_{tot} \)):
  \[
  M_{20} = \log\left( \frac{\sum_i M_i}{M_{tot}} \right), \quad \text{while} \quad \sum_i f_i < 0.2 M_{tot}, \tag{4}
  \]
  where
  \[
  M_{tot} = \sum_i f_i ((x_i - x_c)^2 + (y_i - y_c)^2). \tag{5}
  \]
  It shows the light distribution of central sources, off-centre star clusters, bars, spiral arms, etc.

- **Asymmetry** (hereafter \( \text{ASYM} \); [Abraham et al. 1996]) - measured by subtracting the galaxy image rotated by 180° from the original image, taking also into account the background:
  \[
  \text{ASYM} = \sum_{i,j} \left( |I(i,j) - I_{180}(i,j)|/2 \right) - \sum_{i,j} \left( |B(i,j) - B_{180}(i,j)|/2 \right). \tag{6}
  \]

| Survey     | tot. A/used A [deg^2] | Sp. Res. [arcsec] | Lim_mag/Sel. sam. (band) | Num of gal | mag_sim | z_sim |
|------------|------------------------|------------------|--------------------------|------------|---------|-------|
| ALHAMBRA   | 4.0/∼ 3.0              | 1.1              | 25.0/23.0 (F613W)        | > 43,000   | < 20.0  | ∼ 1.0 |
| SXDS       | 1.22/∼ 0.65            | 0.8              | 27.7/24.5 (′)            | > 68,000   | < 21.0  | ∼ 2.2 |
| COSMOS     | 2.0/∼ 0.5              | 0.09             | 26.5/24.0 (F814W)        | > 107,000  | < 21.0  | ∼ 1.6 |

* Column description: Survey - selected high-redshift surveys; tot. A/used A - the total covered area of each survey, and the area that covers the data used in this work; Sp. Res. - the average survey spatial resolution; Lim. mag / Sel. sam. (band) - limiting survey magnitude, and limiting magnitude of the selected sample in the correspondent band; Num. of gal - the final number of selected galaxies used in this work to obtain the magnitude and redshift distributions in each survey, and move later the local galaxies to these conditions (see Sec. 3 and 4); mag_sim and z_sim - magnitude cuts and the corresponding redshifts (for 95% of the sample) used to simulate the local galaxies;
where I(i, j) is the flux in the (i, j) pixel position on the original image, I_{mag}(i, j) is the flux in (i, j) on the image rotated by 180°, and B(i, j) and B_{mag}(i, j) are fluxes of the background in the original and rotated image, respectively. Defined in this way, this parameter quantifies the degree to which the light of a galaxy is rotationally symmetric. It gives values from 0 to 1, where the higher ones correspond to more asymmetric galactic shapes.

- Smoothness or clumpiness (hereafter SMOOTH: Conselice et al. [2000]) - quantifies the degree of small-scale structure. To measure SMOOTH, the original galaxy image is smoothed out with a boxcar of a given width and then subtracted from the original image:

\[
\text{SMOOTH} = \sum \frac{|I(i, j) - I_{\text{smooth}}(i, j)|}{\sum I(i, j)} - \sum \frac{|B(i, j) - B_{\text{smooth}}(i, j)|}{\sum I(i, j)}
\]

where I(i, j) (B(i, j)) present again the flux (background) in the pixel position (i, j) on the original image, while I_{smooth}(i, j) and B_{smooth}(i, j) are the flux and the background, respectively, on the image smoothed by a boxcar of width 0.25r_p. The residual image provides with information about possible regions of clumpiness, quantifying the level of small-scale structures.

3.2 Methodology applied on the local sample

We measured the reference morphological parameters of the local sample, at their real redshifts z ∼ 0, in three different SDSS bands: g, r, and i. This is needed in order to compare in a consistent way the morphologies of local and simulated galaxies in Sec. 3.2 and 3.3, since the corresponding k-correction will imply a different filter for such a comparison. We used the function implemented in the galSVM public code of galaxy classification (Huertas-Company et al. [2008, 2009]) for measuring the parameters. The input information that the code needs, related with the source position, size, ellipticity parameters, and background, was obtained by running SExtractor in all three bands.

3.3 Methodology applied at higher redshifts

To simulate the conditions of non-local surveys (case 2 in Sec. 3), we used again the galSVM code (Huertas-Company et al. [2008, 2009]), performing the following three steps for each of the three analysed surveys:

- First, we constructed the correspondent magnitude and redshift distributions of non-local galaxies (see Sec. 3.2 and Fig. 7). We randomly redshifted and scaled in luminosity the selected sample of local galaxies to match the distributions of non-local ones (see Tab. 1 for the used photometric band in each survey). In this way, when scaling the galaxy in flux, the surface brightness dimming was directly taken into account. Moreover, we re-sampled the local galaxies with the corresponding pixel-scale for each selected non-local survey, and convolved with its PSF to match the spatial resolution.

- Second, we dropped the simulated galaxies, obtained in the first step, into the real background of the high-redshift survey images. We made sure that the galaxies are dropped into empty regions of sky, minimising the chance of superposition with high-redshift foreground/background sources. To detect the empty sky regions, the code makes use of corresponding SExtractor segmentation images. Moreover, although the probability of overlapping is related with the survey properties, we’ve been strict, dropping only 10 galaxies per image. This step is then repeated until dropping all simulated galaxies, generating every time new images. With all this, we make the probability of overlapping and merger confusion practically negligible (for more information see Huertas-Company et al. [2008]). Finally, by dropping the local galaxy in a real background of the high-redshift sources, we expect to reproduce the noise from the real images.

- Third, we measured the morphological parameters described in Sec. 3.1. We took care of the k-correction effect introduced by cosmological redshift, and depending on the band selected in each survey, and also depending on the redshift to which the local galaxy was shifted, we measured the morphological parameters using the corresponding SDSS rest-frame band image. In the ALHAMBRA survey, since we are using the F613W AUTO magnitudes, the SDSS r band images were used for galaxies simulated to \(z_{\text{sim}} < 0.13\), g band for the redshift range \(0.13 - 0.5\), and u band for higher redshifts. In the SXDS survey, using the \(i'\) magnitudes, the \(r\) band was used for redshifts \(z_{\text{sim}} < 0.4\), the \(g\) band for \(0.4 < z_{\text{sim}} < 0.86\), and the \(u\) band for \(z_{\text{sim}} > 0.86\). Finally, in the COSMOS survey, using the F814W AUTO magnitudes, the \(i\) band was used for the lowest redshifts \(z_{\text{sim}} < 0.18\), the \(r\) band for \(0.18 < z_{\text{sim}} < 0.5\), the \(g\) band for \(0.5 < z_{\text{sim}} < 1.0\), and the \(u\) band for \(z_{\text{sim}} > 1.0\). Since the filter efficiency for the SDSS u band is much lower than that for g, r, and i bands (Gunn et al. [1998]), and taking into account that we are analysing three broad morphological groups (ET, LT, et, and LT, lt) as described in Sec. 2.1 (instead of studying a finer morphology), we used the g band images instead of u when necessary. Figure 2 shows an example of \(u, g, r,\) and i images of four galaxies from our local sample with different morphologies; note the poor information available in the \(u\) image in comparison to the other bands. Finally, since the morphological classification directly depends on the source brightness, in each survey we analysed the observational bias at the three different magnitude cuts: \(\text{mag}_{\text{cut}} < 20.0\), \(\text{mag}_{\text{cut}} < 21.5\), and \(\text{mag}_{\text{cut}} < 23.0\) in the F613W band in ALHAMBRA; \(\leq 21.0\), \(\leq 23.0\), and \(\leq 24.5\) in the \(i'\) band in SXDS; and \(\leq 21.0\), \(\leq 23.0\), and \(\leq 24.0\) in the F814W band in COSMOS. In all surveys we chose the first magnitude cut such that we have a sufficient number of sources to perform our analysis. The cut at which our selected photometric/photometric redshift sample is complete, is the last one. Finally, we chose a third cut intermediate to these two. Figures 2, 4, and 5 show examples of local galaxies, with different morphologies, after being scaled to the conditions of the ALHAMBRA, SXDS, and COSMOS surveys, respectively. In each survey, we show the galaxies being redshifted to the corresponding magnitude cuts, as explained above, providing for each cut the simulated values of magnitude and redshift. These images represent the output of the first step explained in this section, that we use in the second step to drop them in the real background of the non-local survey images, and to measure their morphological parameters in the third step. The colour images were gathered from the EFIGI project, by combining the g, r, i images (Baillard et al. [2011]), while the original images were downloaded from the SDSS DR4 database and correspond to the images used by N&A in their visual classifi-
ground-based data. Therefore, in our following analysis we discuss the logical classification of galaxies, and especially when dealing with the elements where the noise is reliable. The results based on ASYM and SMOOTH are more stable to noise and where our galSVM measure.

Noise effect

With respect to the noise, since we are using very bright local galaxies, galSVM assumes that the noise from the galaxies themselves is negligible. Therefore, the noise coming from the galaxy itself once it has been dropped into the non-local image is not treated by the code. We tested how strong this effect could be on the measured morphological parameters. We selected 30 galaxies, 10 from each morphological group defined in Sec. 2.1, and we measured the morphological parameters in the case of ALHAMBRA and COSMOS in two cases: 1) without the above noise added, like in our analysis, and 2) adding the random Poisson noise, typical of each survey, to the scaled local images (before dropping them into the background of non-local surveys). We then compared the results obtained in these two cases for the whole samples, and in relation with the morphological types. We found insignificant differences of the above noise on the mean surface brightness (<1%), and on the parameters CABR, CCON, GINI, and M20 (<6% and <5% in ALHAMBRA and COSMOS, respectively, at the faintest magnitude cuts), but more significant differences in the case of ASYM and SMOOTH. In the case of ALHAMBRA, the differences for these two parameters go up to ~40% and ~50%, respectively. In the case of COSMOS, differences for ASYM are 13%, 19%, and 57% at F814W ≤ 21.0, ≤ 23.0, and < 24.0, respectively. For the corresponding values are 15%, 35%, and 50%. From this result, we suggest that for the parameters ASYM and SMOOTH the noise effect should be taken into account when using them in morphological classification of galaxies, and especially when dealing with ground-based data. Therefore, in our following analysis we discuss all results only for the four parameters CABR, GINI, CCON, and M20, that are more stable to noise and where our galSVM measurements are reliable. The results based on ASYM and SMOOTH are provided only in the case of COSMOS, for the first two magnitude cuts, where the noise effect is lower.

4 RESULTS

4.1 Morphological parameters of the local sample at z ~ 0

For each galaxy, we measured the morphological parameters defined in Sec. 3.1 in the g, r, and i bands, as was described in Sec. 3.2. To test how sensitive are morphological parameters to the used wavelengths and if somehow it could affect our further analysis, for each parameter, we compared its measurements obtained in the three photometric bands, finding, in general, a good linear correlation in all cases. Figure 6 shows the comparison of all six parameters between g and r (panel a), and g and i bands (panel b). In both cases, the concentration parameters, CABR, CCON, and GINI, show the best linear correlations, having similar Pearson Correlation Coefficient (PCC) ~ 0.9. For M20, ASYM, and SMOOTH, the linear correlation is also conserved in both cases, however they are more sensitive to the selected band in comparison with the concentration indexes, and show lower PCC values (between ~ 0.7 and ~ 0.85). They also show a slightly higher dispersion in panel b, in comparison with panel a, with a difference in the PCC of ~ 0.05. M20 is the most sensitive to the selected band, with PCC = 0.733 between the g and r bands, and PCC = 0.688 between the g and i measurements. The number of catastrophic outliers, placed in the ‘wings’ above and below the linear correlations is ~ 10% in both panels. We checked the properties of these sources (magnitude, size, and redshift), without finding any significant difference in comparison with the whole sample. We only found a small difference in relation with the morphological type, with 6% more LTs in the wings. To have an idea about how these outliers could affect the results in our following analysis and if there are some important differences that we should study in more detail, we performed a test by excluding them from the whole sample and checking again the corresponding distributions of morphological parameters (see Sec. 3.2 and Figs. 6b and 10), without finding any significant differences. However, this was done only to test the properties of outliers, but the all analysis and results presented in this paper were obtained using the whole sample, outliers included. For the comparison of morphological parameters between near-infrared and optical bands see Huertas-Company et al. (2014).

4.2 Morphological parameters of the simulated sample

To measure the morphological parameters of the simulated sample, we followed the methodology described in Sec. 3.3. We first moved randomly the local sample to map the magnitude and redshift distributions of galaxies in each high-redshift survey, at each selected magnitude cut. Figure 7 shows the corresponding magnitude and redshift distributions of the simulated galaxies in case of ALHAMBRA (panel a), SXDS (panel b), and COSMOS (panel c). Since the magnitude cut introduces also a redshift selection, we are dealing at the same time with different redshift ranges in the three surveys (as summarized in Table 1). After simulating the local galaxies and dropping them into the real background of high-redshift surveys, we run again SExtractor to measure the same parameters as we did at z ~ 0. Finally, we measured the morphological parameters of simulated galaxies, again in the same way as we did at z ~ 0. As mentioned above, in each survey, depending on the used photometric band and the redshift to which the selection is lower.

Figure 2. An example of E11, Sa, Sc, and Sdm galaxies (from top to bottom) seen in four SDSS bands, u, g, r, and i (from left to right).
the galaxy was moved, we used the corresponding rest-frame SDSS images to carry out all the measurements.

In each non-local survey, we had to repeat all the procedure a number of times, for each separate image, in order to cover the analysed survey areas (see Table 1). In case of ALHAMBRA, we were dealing with 48 sub-fields (4 CCD images per pointing, with 12 pointing in total (1 or 2 per each field); see Pović et al. 2013). In SXDS we repeated the procedure in each of five SXDS fields (see Sec. 2.2.2), while to cover the selected area of COSMOS we run the code on 17 images. Moreover, in all surveys we repeated the measurements three times, since having selected three magnitude cuts (see Table 1). Since each time we run galSVM on one image, 3000 simulated galaxies were distributed randomly to the corresponding redshift and magnitude distributions of the non-local samples, for each galaxy we obtained a number of measurements that is equal to the number of used images. Therefore, the final catalogues contain 144000, 15000, and 51000 simulated galaxies in the case of ALHAMBRA, SXDS, and COSMOS, respectively. Each catalogue contains the columns with morphological parameters measured at higher-redshifts in the three magnitude cuts, besides the corresponding reference measurements (obtained at z ∼ 0). Once again, when assigning the reference value to the simulated one, we took care of the k-correction, assigning the value measured in the corresponding rest-frame band (that will depend on the photometric band we used for analysis in each non-local survey, and on the simulated redshift). The reader should keep this in mind for understanding easier the distributions of comparison sample of local galaxies, represented in figures from 11 to 17 in Sec. 4.3.

The obtained morphological parameters of simulated galaxies are represented with the filled-line histograms in Fig. 8, 9, and 10 that reproduce the conditions of the ALHAMBRA, SXDS, and COSMOS surveys, respectively. In each figure are shown the distributions of CABR, GINI, and CCON concentration indexes, and the galaxy was moved, we used the corresponding rest-frame SDSS images to carry out all the measurements.
Morphological classification of galaxies: impact from survey depth and resolution

Figure 4. Same as Figure 3, but in the SXDS survey. (A colour version of this figure is available in the on-line journal)

M20, asymmetry, and smoothness (from left to right panels) for three morphological groups: ET (top plots), LT$_{et}$ (middle plots), and LT$_{lt}$ (bottom plots), defined in Sec. 2.1. For each morphological group, we represent the distribution of parameters at the three analysed magnitude cuts. In these figures we also represent the distributions of the reference morphological parameters of local galaxies, measured in Sec. 4.1, with the filled-colour histograms.

For all morphological parameters we can observe important differences between the reference (local) values and those obtained after moving the local galaxies to the conditions of non-local surveys, indicating the influence that the observational bias from spatial resolution and depth can have on them. To quantify this bias, we measured the difference between the median values of the local and simulated distributions, normalised with respect to the median value of the local sample:

$$\frac{\langle \text{local} \rangle - \langle \text{simulated} \rangle}{\langle \text{local} \rangle}.$$
The obtained values are expressed in % in each plot, together with the corresponding range of interquartile range (IQR; between parentheses), measured as the difference between the third and first quartiles. In the following, we describe the obtained differences, considering as significant those within $\geq 2$ IQR.

**CABR**

In ALHAMBRA conditions, CABR showed significant differences in the case of the faintest ETs and brightest LTs ($\sim 30\%$ within 2 IRQ), with galaxies appearing less and more concentrated, respectively. In the conditions of SXDS we did not find any differences (all within 1 IQR). In COSMOS, we obtained differences of $\sim 20\% - 36\%$ (within 2 IRQ) at the brightest magnitude cuts for all three morphological types, with simulated galaxies appearing more concentrated. In ALHAMBRA, 2 - 3\% of sources showed invalid measurements (with values outside the typical range of this parameter, between 0 and 1), independently of the morphological type; in SXDS, this went between 3 - 5\% for ET and LT$_{et}$, and 7 - 10\% for LT$_{lt}$ galaxies; while in COSMOS, 2 - 3\% for ET and LT$_{et}$, and 8 - 10\% for LT$_{lt}$ galaxies.
Morphological classification of galaxies: impact from survey depth and resolution

Figure 6. Comparison between the six morphological parameters for the galaxies in the local sample measured in the g and r bands (panel a), and g and i bands (panel b). In each panel we compared, from top to bottom, and from left to right: Abraham concentration index, asymmetry parameter, Gini coefficient, smoothness, Conselice-Bershady concentration index, and M20 moment of light. In all plots the red solid line presents the perfect linear correlation between the two measurements.

(A colour version of this figure is available in the on-line journal)

GINI

In ALHAMBRA, this parameter showed differences of 10 - 20\% (2 IQR) at the first magnitude cut for all types, and at the second magnitude cut for LT_It galaxies. In SXDS, the difference for LT_It galaxies was observed at all three magnitude cuts (15 - 18\% within 2 IQR). The same was obtained for LT_It sources but with slightly

Figure 7. Magnitude (left plots) and redshift (right plots) distributions of simulated galaxies for the conditions in ALHAMBRA (panel a), SXDS (panel b), and COSMOS (panel c) surveys. In each survey, they are represented for three magnitude cuts, from top to bottom: for F613W < 20.0, F613W < 21.5, and F613W < 23.0 in the ALHAMBRA; i < 21.0, i < 23.0, and i < 24.5 in the SXDS; and F814W < 21.0 (top), F814W < 23.0 (middle), and F814W < 24.0 in the COSMOS.
higher differences. In the case of COSMOS, the only differences were observed at the first magnitude cut: $\sim 20\%$ (2 - 3 IQR) for all three types. In all cases where the differences were detected, we obtained negative values, showing that the simulated galaxies would appear more concentrated than they really are. With respect to the measured values of simulated galaxies that are outside of the valid ranges, the populations are in agreement with those for the CABR parameter.

**CCON**

A different behavior of CCON was detected for the conditions of the analysed ground- and space-based surveys. In the case of COSMOS, we did not detect any differences, independently of the magnitude cut and morphological type. This index showed to be especially sensitive to the spatial resolution in comparison to CABR and GINI. In the case of ALHAMBRA and SXDS, the highest differences are detected for ET galaxies at the highest magnitude cuts, $\sim 30\%$ at 3 IQR and 4 IQR. In all cases where differences were observed in the ground-based surveys, lower concentrations appeared for simulated galaxies in comparison with the reference (local) values, independently on the magnitude cut and morphological type. Again, the population of the obtained invalid measurements is consistent with those presented for CABR.

**M20**

As described in Sec. 3.2, this parameter is sensitive to galaxy structures as the central light from the bulge, bars, spiral arms, strong star formation regions, etc. Therefore, for local ET galaxies, it mainly maps the central light, having a much narrower distribution than in the case of LT, or even more in the case of LT galaxies, where the distribution is broader (filled-colour histograms in Figs. 8, 9, and 10). However, when moved to higher-redshift surveys the distribution of this parameter changed significantly (both, shape and range; filled-line histograms), and showed in all cases the information coming mainly from the centre of the galaxy, and independently on the morphological type. In ALHAMBRA, the highest differences are of order $> 30\%$ within 2 - 3 IQR, for all three types. In SXDS conditions, for ET galaxies M20 suffered similar changes as in ALHAMBRA (at higher magnitude cuts), except for LT galaxies sources, where no difference were found respect to local sample. In COSMOS conditions, M20 shows higher concentrations for simulated galaxies in comparison with the local ones, affecting only LT galaxies, with differences $\leq -31\%$ (within 2 IQR). The population of sources with invalid measurements of M20 is consistent with that for CABR.

**ASYM**

This parameter showed the highest differences in the case of ALHAMBRA (even up to $\sim 80\%$ within 3 IQR at the faintest magnitude cut $F613W \leq 23.0$), affecting LT galaxies in the sense of obtaining more symmetric galaxies in comparison with their reference ($z \sim 0$) values. In COSMOS, all three morphological types suffered similar shifts ($\sim 30\%$ within 2 IQR), being the strongest one at the lowest magnitude cut $F814W \leq 21.0$, but showing positive values in comparison to ALHAMBRA, and so producing more asymmetric galaxies in simulated conditions. The population with invalid measurements was higher ($< 20\%$) than in the case of CABR, GINI, and CCON. All these measurements were obtained without taking into account the noise introduced by the galaxy itself once being dropped into the background of non-local surveys. However, as showed in Sec. 3.3 in the case of ASYM and SMOOTH this noise could have a significant effect on these parameters, and the differences measured there should be added to those measured in this section. Therefore, we found these parameters unreliable for the use at higher redshifts in the case of ground-based surveys. In the space-based surveys similar to COSMOS, it becomes unreliable for the use in the case of galaxies fainter than $F814W = 24.0$.

**SMOOTH**

Without taking into account the noise studied in Sec. 3.3, the differences in SMOOTH in the case of ALHAMBRA were detected only for LT galaxies, being lower than 36\% (within 2 IQR); in the case of SXDS, no differences were observed between the simulated and local samples; nevertheless, in COSMOS conditions it showed 40 - 50\% differences (within 2 - 3 IQR) at the brightest magnitude cuts $F814W \leq 21.0$, resulting in galaxies to appear more clumpy. Again, to these differences those related to the inclusion of noise (see Sec. 3.3) should be added. Moreover, this parameter showed to be significantly more affected with both spatial resolution and depth, when considering the number of invalid measurements of the simulated sample (up to 50\% and 20\% at the faintest magnitude cuts in the case of ALHAMBRA and COSMOS, respectively). Taking all this into account, we found this parameter especially sensitive to all, noise, spatial resolution, and data depth, and therefore unreliable for the use in the morphological classification of galaxies in the case of ground-based surveys, and for faintest sources in space-based surveys.

### 4.3 Behaviour of morphological diagnostic diagrams

In this section we analyse how the observational bias, discussed in the previous section, affects some of the commonly used diagnostic diagrams, that allow to separate between the ET and LT galaxies, to select disturbed galaxies, and/or interacting (merging) systems. To do this, we analysed the simulated sample in the conditions of the three non-local surveys at different magnitude cuts, and compared to the corresponding local sample. We can directly use the obtained diagrams to quantify how strong is the bias for each of these diagrams when taking into account the survey observational properties: basically the spatial resolution (comparing the results between the ground-based and space-based surveys) and survey depth (observing the changes at different magnitude cuts).

As already mentioned in Sec. 3.3, we analysed diagrams based on CABR, GINI, CCON, and M20 in all surveys, since these parameters are more stable to the noise effects. Figures 11, 12, 13, 14, and 15 show the relation between CCON and CABR, CCON and GINI, CCON and M20, GINI and CABR, M20 and CABR, and finally M20 and GINI, respectively. In each figure are shown the results obtained for the conditions in the three surveys: ALHAMBRA (top), SXDS (middle), and COSMOS (bottom). For each survey we plot the locus of the three analysed morphological groups, ET (red contours), LT (blue contours), and LT (green contours), obtained once local galaxies (bottom plots) were moved to higher-redshift conditions (top plots). Although initially we simulated the same sample of 3000 local galaxies to the conditions of the three non-local surveys, in some plots of the reference (local) values small differences can be observed, which is the reason why we represent them in each survey and for each magnitude cut. To understand these differences it has to be considered that, first, the local measurements correspond to the rest-frame band of the simulated sample, and hence slightly changes depending on the magnitude cut and survey (see Sec. 4.2). Secondly, the simulated galaxies with invalid measurements of morphological parameters
Figure 8. Comparison between the morphological parameters of the sample of local galaxies measured at their real redshift (filled-colour histograms, representing the reference galaxy properties) and in conditions that correspond to the galaxies from the ALHAMBRA survey (filled-line histograms). Six parameters are compared, from left to right: Abraham concentration index, Gini coefficient, Conselice-Bershady concentration index, M20 moment of light, asymmetry, and smoothness. Top) three magnitude cuts of the ET galaxies, middle) LT.et galaxies, and bottom) LT.lt galaxies. In all diagrams the numbers represent the difference between the median value of the reference (local) and simulated distributions, normalised with the median of the local sample (expressed in %), and the corresponding interquartile range. (A colour version of this figure is available in the online journal)
### Figure 9.
Same as Fig. 8, but moving the local sample to map the conditions of the SXDS survey.

(A colour version of this figure is available in the online journal)
Figure 10. Same as Fig. 8 but moving the local sample to map the conditions of the COSMOS survey.

(A colour version of this figure is available in the online journal)
(discussed in the previous section) are excluded in all plots, and therefore the corresponding local galaxies are also excluded from the bottom diagrams. ASYM and SMOOTH were only used in the case of COSMOS and its first two magnitude cuts. Figure 17 shows some of the commonly used relations, representing the relation between CABR and ASYM (top plots), CCON and SMOOTH (middle plots), and finally ASYM and GINI (bottom plots). In all three plots, we represent the same comparisons as in Fig. 11 to 15.

In all figures, when observing the reference parameters for local galaxies, we can clearly separate the regions typically occupied by ET, LT<et, and LT<lt galaxies, as expected. On the other side, we can observe how the position and the shape of the same regions change once we go to fainter magnitudes (higher redshifts). We measured the level of contamination for the highest density population (50%) of each analysed morphological group (ET, LT<et, and LT<lt) with the other two types, for the morphological diagrams presented in Fig. 11 to 16. We applied the Tukey’s Five Number Summary statistic (Tukey 1977), and measured for each parameter, and each morphological type: the sample minimum, first quartile, median value, third quartile, and sample maximum. The first and third quartiles, present the median values of the lower half and the upper half of the sample, respectively, with respect to the median value of the total sample. Therefore, measuring the first and the third quartile of each parameter on the particular diagnostic diagram, we are able to determine the regions that correspond to the highest population (50% of sources) for ET, LT<et, and LT<lt around their median values. From this, for each morphological type, we can estimate the level of contamination by the other two types, within their densest regions: e.g., contamination of the densest ET region with the densest region of the LT<et galaxies (as a number of LT<et galaxies in the overlapping area, normalized with the number of ET galaxies), and/or with the densest region of the LT<lt galaxies, and vice-versa. Table 2 shows these measurements for simulated galaxies in ALHAMBRA (top table), SXDS (middle table), and COSMOS (bottom table), for their corresponding magnitude cuts. These measurements represent a quantification of the bias represented in the diagrams 11 to 16 in all the three surveys, and in Fig 17 in the case of COSMOS. We can again observe that the contamination levels in the morphological types increase with the data depth in all three surveys, being however significantly lower in the case of COSMOS.

5 DISCUSSION

5.1 Individual parameters

As seen in Sec. 4.4 the morphological parameters show complex behaviors when going from one survey to another, from one magnitude cut to another, and when dealing with different morphological types. In principle, they suffer changes due to the impact of spatial resolution and survey depth, and depending on how strong these two effects are, they might cause:

- only a re-distribution of the light within the galaxy, without changing much the total flux, resulting in a weak impact, and/or
- the loss of the galaxy low surface brightness structures and the change of the total flux, hence producing a more significant impact.

However, how the distributions of morphological parameters behave in each survey when going from brighter to fainter magnitude cuts (from lower to higher redshifts), depends on their definitions. In general, for weak impacts of spatial resolution and data

Figure 11. Relation between the Conselice-Bershady and Abraham concentration indexes, in the three analysed non-local surveys: ALHAMBRA (top plots), SXDS (middle plots), and COSMOS (bottom plots). In each survey, top rows represent the morphological parameters of simulated sample obtained after moving the local galaxies to higher redshifts, considering three magnitude cuts (the first and the last column showing the lowest and the highest analysed magnitude cut, respectively), while bottom rows show the corresponding reference (local) values. The red, blue, and green contours represent ET, LT<et, and LT<lt galaxies, respectively.

(A colour version of this figure is available in the online journal)
depth, the galaxies might appear as more concentrated, asymmetric and clumpy, while for stronger impacts, when the galaxy total flux can be affected significantly, the galaxies might appear as less concentrated, more symmetric, and less clumpy. To trace the real impact from spatial resolution and depth on the parameters in Fig. 8, 9, and 10, it is not only important to take into account the difference between the simulated and real values (and its standard deviation), but also the sign of that difference. In general, for the concentration parameters (CABR, GINI, CCON), the distributions will be the most affected when the real differences have positive values (showing that the simulated galaxies are less concentrated in comparison with the real ones); negative in case of M20, ASYM, and SMOOTH due to the use of the logarithm scale in our analysis.

5.2 Diagnostic Diagrams: impact from spatial resolution and data depth

To evaluate the separate effect of resolution and data depth, we compare two cases: first, the results obtained for the two ground-based surveys, where both of them are affected with the spatial res-
5.2.1 The effect of survey depth in deep ground-based surveys: ALHAMBRA and SXDS

From the morphological diagrams and from the Table 2 for the deep ground-based surveys, like ALHAMBRA and SXDS, diagrams combined with M20 show to be less effective in separating galaxies. On the other side, as seen in all figures, survey depth is crucial when classifying galaxies. In the case of ALHAMBRA, at its lowest magnitude cut (F613W ≤ 20.0), although in all diagrams the position/shape of locus of the three morphological types change significantly in comparison with the same sample at z ~ 0, in most of diagrams we are still able to distinguish between the ET and LT galaxies, or at least the regions of the highest source densities do not overlap. Observing for each type their contours enclosing 50% of galaxies, we found a ∼ 5 - 25% contamination of ET with LT_et sources, and a ∼ 3 - 35%, of LT_et with ET sources, where the contamination level depends on the used diagram, with...
limits of Lotz et al. (2008) to distinguish between the normal galaxies and
lines were used to classify sources at $0.2$ and red) provided this separation for a sample of local galaxies, while solid
tions, mergers occupy the regions above the green lines, ETs (LTs) below
mergers, and between ET and LT galaxies, respectively. In their classifica-
ing between the three regions increases significantly, but still for
fainter galaxies (second magnitude cut F613W
the highest one. Similar conclusions are obtained for classifying
est contamination (see Sec. 3.3). For more explications see the caption of Fig. 11. (A colour version of this figure is available in the online journal)

Figure 16. Same as Fig. 11 but representing the relation between the M20 moment of light and Gini coefficient. The green and red lines represent the limits of Lotz et al. (2008) to distinguish between the normal galaxies and mergers, and between ET and LT galaxies, respectively. In their classifications, mergers occupy the regions above the green lines, ETs (LTs) below the green lines and above (below) the red ones. Dotted lines (both green and red) provided this separation for a sample of local galaxies, while solid lines were used to classify sources at $0.2 < z < 0.4$.

(A colour version of this figure is available in the online journal)

GINI vs. CABR and CCON vs. CABR diagrams showing the lowest contamination ($< 5\%$), and CCON vs. M20, and M20 vs. GINI the highest one. Similar conclusions are obtained for classifying LTs, but with higher contamination by ET sources. When including fainter galaxies (second magnitude cut F613W $< 21.5$), the mixing between the three regions increases significantly, but still for

Figure 17. Relation between the Abraham concentration index and asymmetry (top), Conselice-Bershady concentration index and smoothness (middle, top), asymmetry and smoothness (middle, bottom), and asymmetry and Gini coefficient (bottom) in COSMOS, in the two analysed magnitude cuts, where the noise effect on ASYM and SMOOTH parameters is still insignificant (see Sec. 3.3). For more explications see the caption of Fig. 11. (A colour version of this figure is available in the online journal)
parameters measured within COSMOS conditions is still present, account when dealing with the photometric data and morphological spatial resolution is the main limitation that should be taken into from one magnitude cut to another, LT overlapping of the ET and LT regions like in the case of ground- and SXDS selected bands, we still don’t observe the complete diagrams the contamination rate changes between 2 - 25% at the central wavelength of the i’ band, of ~760 nm, this would corre- sponds to even higher magnitudes in ALHAMBRA), are higher than those obtained at the lowest analysed magnitude cut in ALHAMBRA (F613W ≤ 20.0), but lower than the second magnitude cut (F613W ≤ 21.5). Moreover, at the highest magnitude cut in the SXDS (i’ ≤ 24.5), we obtained ~ 10 - 80% lower contaminations, depending on the type, than in the case of the last magnitude cut in ALHAMBRA (F613W ≤ 23.0). The combination of other parameters with M20 (Fig. 13) and 16, which seems to be the most affected with the resolution, again shows to be the most ineffective. Finally, we should notice that either the effectiveness in separating between ET and LT changes from one diagram to another and from one magnitude cut to another, LT galaxies in all cases, in both surveys, suffer very high levels of contamination (> 80% in most of cases) from other two morphological types.

5.2.2 The effect of resolution. COSMOS as a reference

Since the spatial resolution affects little the measurements of the morphological parameters in the case of COSMOS data, in comparison with the results obtained in ALHAMBRA and SXDS surveys, we can evaluate how much impact does it have when classifying galaxies. In all diagnostic diagrams (Fig. 11 to 16) the loci of galaxies simulated to the conditions of COSMOS, irrespec- tive of their morphological type, change significantly less at all magnitude cuts than in case of ALH and SXDS. Moreover, the diagrams that include ASYM and SMOOTH parameters provide the additional information, and can be used up to F814W = 24.0 magnitudes. However, the data are still affected by the survey depth, as can be seen in all figures when comparing the diagrams at different magnitude cuts. When including only the brightest simulated galaxies in the sample (F814W ≤ 21.0) the locus of the three morphological groups basically does not change in comparison with that of the corresponding local values, and the contamination levels between ET and LT galaxies are ≲ 10%. This picture starts to change when including fainter galaxies, down to F814W ≤ 23.0, and even more at F814W ≤ 24.0. Depending on the diagram the contamination rate changes between 2 - 25% at F814 ≤ 23.0, and 6 - 35% for ET and 20 - 70% for LT galaxies at F814 ≤ 24.0. However, even at the highest analysed magnitude cut, which corresponds to even fainter magnitudes in ALHAMBRA and SXDS selected bands, we still don’t observe the complete overlapping of the ET and LT regions like in the case of ground-based surveys, and the contamination levels are still below those obtained for the brightest ALHAMBRA/SXDS magnitudes.

Although the impact from survey depth on the morphological parameters measured within COSMOS conditions is still present, it is significantly lower than in the case of deep ground-based surveys like ALHAMBRA and SXDS. We can assume that the spatial resolution is the main limitation that should be taken into account when dealing with the photometric data and morphological classification of galaxies.

5.3 Diagnostic diagrams: implications for morphological classification of galaxies

For intermediate- and high-redshift galaxy samples, and large data sets, typical for current (and future) deep photometric surveys, the application of automated non-parametric methods, that use the parameters analysed within this work, is one of the ways to obtain the information about the morphology. In some of the latest works, the classification based on non-parametric methods was done using a set of parameters simultaneously, therefore decreasing the impact of spatial resolution and depth (e.g. Huertas-Company et al. 2009, 2011; Pović et al. 2012, 2013; Cibinel et al. 2013; Carollo et al. 2013; Cotini et al. 2013; Holwerda et al. 2014). However, for classifications based on only one, two, or three parameters, special attention should be taken in relation to the magnitude and redshift distribution of the analysed sample, and the data quality.

The ASYM index was used in some of previous works to select mergers and disturbed galaxies (Lotz et al. 2010; Villforth et al. 2014). ASYM is especially sensitive to the survey noise, as showed in Sec. 3.3 and to both, the spatial resolution and depth, as showed in Sec. 4.2. In surveys similar to COSMOS, for sources fainter than F814 ≤ 23.0, we again do not recommend the use of a single criterion based on ASYM when classifying them morphologically, since the contamination levels of 50% highest density regions of ET and LT rise to above 40%. In the case of ground-based surveys, we do not recommend the use of this single criterion in any case (except for maybe the brightest galaxies), since the contamination levels are much higher, taking the noise effect especially into account; we suggest instead the simultaneous use of multiple diagnostic diagrams.

Over the past ten years, the M20 - GINI diagram (Fig. 16) found wide applications. It was mainly used in identifying galaxy mergers and peculiar (disturbed) galaxies (e.g. Lotz et al. 2004; Pierce et al. 2005; Lotz et al. 2008; Conselice et al. 2008; Mendez et al. 2011; Chen et al. 2009; Stott et al. 2013; Petty et al. 2014; Hung et al. 2014); often it was used in the analysis where different types of galaxies, e.g. ULIRGs, dust obscured galaxies, star-forming galaxies, or quiescent galaxies, were compared with control samples (Bussmann et al. 2009; Pentericci et al. 2010; Lanyon-Foster et al. 2012; Wang et al. 2012; Chung et al. 2013; Böhm et al. 2014); or for selecting ET and LT galaxies (Lotz et al. 2004, 2008).

From our results, in ALHAMBRA and SXDS surveys even for the brightest analysed sources (F613 ≤ 20.0 and i’ ≤ 21.0, respectively) the contamination levels between the ET and LT galaxies in M20 - GINI diagram are about 30%. In COSMOS however, at the lowest magnitude cut F814 ≤ 21.0 (z ≤ 1) the contamination is insignificant (≤ 10%, except for the youngest spiral/irregular galaxies), but already at the second magnitude cut F814 ≤ 23.0 (z ≤ 1.5) the contamination is > 20%. We represented in Fig. 16 the limits established by Lotz et al. (2004, 2008) to separate between normal ET and LT galaxies and major mergers, in the local universe (dotted lines), and at 0.2 ≤ z ≤ 0.4 (solid lines), using the All-wavelength Extended Groth strip International Survey (AEGIS) HST/ACS data. The solid-lines criterion (non-local sample) was later applied by Lotz et al. (2008, 2010) to z ≤ 1.2 sources to select major mergers, reporting a population of
Table 2. Expected level of contamination for densest regions (50% of sources) of ET, LT, and LT galaxies in some of the most used morphological diagrams in the ALHAMBRA (top), SXDS (middle), and COSMOS (bottom).
~ 15% of misclassified sources, and was used in other works, as mentioned above. From the established limits we can see again how sensitive the diagram is to spatial resolution and data depth, changing significantly between the ground- and space-based surveys. While in COSMOS the contamination by normal galaxies in the region assigned for major mergers, does not exceed the values measured by the authors (even at higher redshifts), in the analysed ground-based surveys the contamination can be > 70%.

When applying the M20 - GINI diagram to classify galaxies and select mergers, we highly recommend to use it always with additional criteria when dealing with ground-based data, among which: a combination with other morphological diagrams, a careful selection of the control sample and evaluation of the diagram effectiveness, and/or application of probabilistic approaches. In space-based surveys like COSMOS to distinguish between the ET and LT sources additional criteria should be taken into account at least for galaxies fainter than F814 = 23.0.

Finally, different combinations of 2 - 3 parameters analysed in the previous sections, were used in many studies to distinguish between the ET and LT galaxies, and also to study the structure of galaxies with other properties, like SFR, mass, environment, nuclear activity, etc (e.g. Pierce et al. 2007; Cassata et al. 2007; Bussmann et al. 2009; Pović et al. 2009a, b; Pentericci et al. 2010; Chung et al. 2013; Böhm et al. 2014). The combination between the concentration parameters (CABR and/or CCON) and ASYM was also used for selecting ET and LT sources (Abraham et al. 1994, 1996; Conselice et al. 2003; Huertas-Company et al. 2007; Pović et al. 2009a) while in some cases the combination between the GINI and CCON was used to study the host properties of AGN and quasars (e.g. Urrutia et al. 2008). The CCON - ASYM - SMOOTH diagram was used in classifying both galaxies and mergers (Conselice et al. 2003; Cassata et al. 2007; Conselice et al. 2008). We found SMOOTH to be especially sensitive to both spatial resolution and depth, and the most unstable parameter of all analysed. Moreover, together with ASYM, we also found it to be more sensitive to noise, as showed in Sec. 3.3. As already mentioned, we do not recommend the use of these two parameters in the ground-based surveys if the noise effects are not taken into account carefully. In surveys similar to COSMOS, they can be used in combination with other parameters but for sources brighter than F814 = 24.0. The combination of only ASYM and SMOOTH gives high contamination levels even in the case of COSMOS (as showed in Fig. 17 and Table 2).

We suggest the consultation of Table 2 and Fig. 11 to 17 when using diagnostic diagrams in morphological studies, especially if they are based on 2 - 3 parameters. We discussed in Sec. 4.3 the effectiveness of a particular diagnostic diagram in the ground-based surveys and noted that in cases similar to both ALHAMBRA and SXDS, special care has to be taken into account when classifying galaxies at all magnitude cuts (redshifts). Our suggestion when using ground-based data is to avoid the diagnostic diagrams based on 2 - 3 parameters, and to apply instead all morphological parameters simultaneously, and to use the statistical approaches providing a probability for a galaxy to belong to a given morphological type (e.g., approaches similar to that used in the galSVM code and tested in Huertas-Company et al. 2008). Diagrams that include M20 moment of light showed to be more affected by spatial resolution, and less effective in separating galaxies. Moreover, the noise should be taken into account if using ASYM and SMOOTH parameters.

In space-based surveys similar to COSMOS, even at the highest magnitude cuts (F814 ≤ 24.0), in all diagnostic diagrams the contamination levels are significantly lower than at the brightest magnitudes in the ground-based surveys. The observational bias is insignificant for all parameters at magnitudes brighter than F814 = 21.0 (except for the youngest spiral and irregular galaxies), as shown in Sec 4.3; however when going to fainter magnitudes the contamination increases. Therefore, even when dealing with data sets similar to COSMOS, we remind that eventual contamination levels are present at fainter magnitudes; the diagnostic diagrams that show to be more effective in this case are CCON - CABR, CCON - GINI, CCON - SMOOTH, and/or CABR - ASYM. To apply again, when possible, different diagnostics will minimise the effect of the observational bias.

5.4 Results from Linear Discriminant Analysis

In order to merge all the analysed morphological parameters, and to provide an additional and statistical test of our previous results, a Linear Discriminant Analysis (LDA; as implemented by Pedregosa et al. 2011) was performed. The LDA method allows us to find the linear combination of measured parameters that best separates two or more classes of objects or events. We tested the combination of the four morphological parameters (CABR, GINI, CCON, and M20) to distinguish between ET, LT, and LTt galaxies in the three analysed non-local surveys (ALHAMBRA, SXDS, and COSMOS). Knowing that ASYM and SMOOTH are found to be affected by noise, we performed the LDA without these two parameters. However, we also performed a LDA including all the six parameters for comparison and to check the effect of noise since both ASYM and SMOOTH parameters resulted to be the ones more significantly affected by noise, mainly in the two ground-based surveys, as we described in Sec. 4.3.

The training phase of the LDA was obtained by fitting the coefficients over the measured parameters of the galaxies of the local sample. We found that all the possible discrimination is based on two eigenvectors (λ1 and λ2) both being linear combinations of the four parameters entering in the classification. The first one, λ1, is
dominated by GINI, CON, and CABR, whereas the second one, $\lambda_2$, is based on CABR, GINI, and M20. Almost all the power of the discrimination arise from the $\lambda_1$ eigenvector whereas $\lambda_2$ only slightly helps to the classification of LT$\rightarrow$ET galaxies, as it can be seen in Fig. 18 where the distributions of both $\lambda_1$ and $\lambda_2$ for the three morphological types are shown. The three concentration indexes show to be more stable, discriminating better between different morphological types in comparison with M20. An attempt to improve the LDA was also carried out by doing a second LDA iteration including only those galaxies that show the same input and predicted output morphology. We did not find any significant change, and therefore did not proceed further with this additional step. As a result of the LDA of the local sample we found that about 30% of the objects show a different input morphology than the one predicted by the resulting eigenvectors of the LDA. This fact could be caused by errors in the measured parameters or even in the original classification. To illustrate the results obtained by the LDA we present in Figure 18 the histogram of the predicted morphologies after applying the eigenvector discrimination over the local sample, where for each output morphological type (ET, LT$\rightarrow$ET, LT$\rightarrow$LJt) we show the contribution of galaxies with different input morphology. As it clearly can be seen, whereas ET galaxies are well classified, LT$\rightarrow$ET spread between adjacent input types, and LT$\rightarrow$LJt are split between both input late-types.

After obtaining the eigenvectors that better discriminate the morphology in the local sample, LDA transformations to the ($\lambda_1$ and $\lambda_2$) space were performed for each non-local sample (ALHAMBRA, S Dix, COSMOS). The results so obtained at the three magnitude cuts and for each non-local sample are presented in Table 3. In this table, for each survey, the first column represents the three corresponding magnitude cuts, the second column corresponds to the total number of galaxies (Ngal) at each magnitude cut and sample (i.e. galaxies with valid measured parameters after applying the filtering criteria explained in Sec. 4.3) while the score column represents the ratio between the galaxies having the same LDA output morphology as the input one normalized to the total number of galaxies Ngal: a score of 1 would indicate a perfect LDA classification. We also tested the results when including ASYM and SMOOTH. In Table 4, we present the results obtained from the LDA taking into account all morphological parameters. In general, the scores obtained after removing these two parameters (see Table 3) are found to be similar to those that would be obtained including also ASYM and SMOOTH, however the total number of valid galaxies increase significantly and this is more notable in the case of the ground base surveys (ALHAMBRA and S Dix). If ASYM and SMOOTH are excluded, the total number of galaxies with good morphology is found to be larger, thus the final classification get worst when including these two parameters as they are indeed affected by noise. Only in surveys similar to COSMOS, the inclusion of ASYM and SMOOTH would have a slightly positive net effect.

Table 3. LDA without SMOOTH and ASYM.

| Survey    | Ngal | Score ET | Score LT$\rightarrow$ET | Score LT$\rightarrow$LJt |
|-----------|------|----------|--------------------------|--------------------------|
| ALHAMBRA  | 20.0 | 1.9470   | 0.583                    | 0.417                    |
|           | 21.5 | 1.9406   | 0.510                    | 0.493                    |
|           | 23.0 | 1.9363   | 0.556                    | 0.444                    |
| COSMOS    | 21.0 | 49.989   | 0.549                    | 0.451                    |
|           | 23.0 | 50.150   | 0.601                    | 0.399                    |
|           | 24.0 | 48.519   | 0.505                    | 0.495                    |

Table 4. LDA using all parameters.

| Survey    | Ngal | Score ET | Score LT$\rightarrow$ET | Score LT$\rightarrow$LJt |
|-----------|------|----------|--------------------------|--------------------------|
| ALHAMBRA  | 20.0 | 114.271  | 0.593                    | 0.407                    |
|           | 21.5 | 88.186   | 0.525                    | 0.475                    |
|           | 23.0 | 66.978   | 0.479                    | 0.521                    |
| COSMOS    | 21.0 | 49.932   | 0.603                    | 0.397                    |
|           | 23.0 | 47.988   | 0.655                    | 0.345                    |
|           | 24.0 | 45.061   | 0.600                    | 0.400                    |

Figure 19. Histogram of the predicted morphology obtained from the LDA. For each predicted morphology, the number of galaxies of each different input morphological type are displayed. Colours and representations are the same as in Fig. 18. ET galaxies are found to be well classified, however LT$\rightarrow$ET spread between adjacent input types, and LT$\rightarrow$LJt galaxies are split between both late-types.

Table 5. LDA results by morphological type.

| Survey    | ET | LT$\rightarrow$ET | LT$\rightarrow$LJt |
|-----------|----|-------------------|-------------------|
| ALHAMBRA  | 20.0 | 0.86 | 0.51 | 0.02 |
|           | 21.5 | 0.47 | 0.74 | 0.010 |
|           | 23.0 | 0.17 | 0.55 | 0.41 |
| COSMOS    | 21.0 | 0.99 | 0.29 | 0.00 |
|           | 23.0 | 0.83 | 0.54 | 0.15 |
|           | 24.0 | 0.55 | 0.55 | 0.15 |

6 SUMMARY AND CONCLUSIONS

In this paper we present for the first time a systematic study of the impact from spatial resolution and depth on the six morphological parameters commonly used for galaxy morphological classification: Abraham concentration index, Gini coefficient, Conseilce-Bershady concentration index, M20 moment of light, asymmetry, and smoothness. We studied how strong the impact is for three data sets that have different spatial resolutions and depth: ALHAMBRA and S Dix, as an example of ground-based data, and COSMOS, as an example of deep space-based data. Our results correspond to maximum analysed redshift values of: 1.0, 2.2, and 1.6 in the three surveys, respectively, where the provided values describe the distribution of 95% of each sample. We used morphological type and magnitude cut in each survey, as resulting from the LDA with neither SMOOTH nor ASYM included. As can be seen, in all surveys, ET galaxies are better recognised at the brighter magnitude cuts, whereas LTs tend to be better recognised at fainter magnitude cuts.
a sample of 3000 early- and late-type local galaxies, with the available visual morphological classification, and we measured their parameters in 2 cases:

- first, the reference values, that correspond to the real galaxy redshifts (magnitudes), and
- second, the simulated values, obtained after moving the local galaxies to the observational conditions and magnitude/redshift distribution of three selected surveys.

Comparing the results obtained in these two cases we showed how each parameter changes in each survey, at three particular magnitude cuts. Finally, we analysed and quantified how the impact from spatial resolution and depth affects some of the most used morphological parameters, gave some suggestions for the galaxy/merger classification studies, and used the LDA statistical approach to analyse the most effective combination of parameters to distinguish between different types.

In the following, we summarise some of our main findings:

- All six analysed morphological parameters suffer from significant biases related to the spatial resolution and data depth.
- The impact of the spatial resolution on the morphology is much stronger in comparison with the data depth, being therefore the most responsible for changing the parameters in the ground-based surveys, making in general the galaxies to appear less concentrated and more symmetric.
- We stress that ASYM and SMOOTH are more sensitive to noise effects. Survey noise should be taken into account carefully when using these two parameters in morphological classification of galaxies, especially when dealing with ground-based data.
- M20 results to be also significantly affected in all surveys, changing both the shape and range of its distribution at all brightness levels. However, M20 - GINI is the most used diagram for selecting interacting (merging) systems. We highly recommend the use of this diagram simultaneously with other morphological parameters, when dealing with ground-based data sets, and in space-based surveys like COSMOS at least when studying sources fainter than \( F814 = 23.0 \).
- CCON shows to be more sensitive to the spatial resolution in comparison to GINI and CABR. It works therefore better for space-based data, whereas the other two concentration indexes behave better for ground-based surveys.
- In surveys similar to ALHAMBRA, when analysing the highest density regions of early- and late-type galaxies in the main morphological diagnostic diagrams, the impact from spatial resolution and data depth introduces contamination levels of 5 - 25% for ET and 3 - 35% for LT galaxies (depending on the diagram), for the brightest galaxies with \( F613 \leq 20.0 \). The diagrams that seem to work the best in this case are those that do not include M20 moment of light. At the faintest analysed magnitudes (\( F613 \leq 23.0 \)) the contamination levels increase significantly, being as high as 60 - 100%, making each of the diagnostics useless if used separately. Similar results are obtained in the case of SXDS, but at higher magnitude cuts (of order 1 - 2). Moreover, in both surveys, at all magnitude cuts, classification of LT galaxies suffers very high levels of contamination. Taking all this into account, when dealing with ground-based data sets, we suggest to avoid the use of 2 - 3 parameter diagnostic diagrams in morphological classification, and to apply instead the use of all morphological parameters simultaneously, and to use statistical approaches based on probability distributions that the galaxy is ET or LT.
- In space-based surveys similar to COSMOS, even at the highest magnitude cuts (\( F814 \leq 24.0 \)), the contamination levels are significantly lower than those at the brightest magnitudes in the ground-based surveys. The observational bias is insignificant for all parameters at magnitudes brighter than \( F814 = 21.0 \) (except for the LT galaxies). However, when going to fainter magnitudes the contamination increases, and depending on the diagram it changes from 2 - 25% at \( F814 \leq 23.0 \), to 6 - 35% for ET, and 20 - 70% for LT galaxies at \( F814 \leq 24.0 \). In surveys similar to COSMOS, we again suggest to use several diagnostics to classify galaxies when going to magnitudes fainter than \( F814 = 23.0 \).
- Through LDA analysis we obtained that the combination of CABR, GINI, and CCON parameters, is the most effective to distinguish between different morphological types.

The results presented in this paper can be directly applied to any survey similar to ALHAMBRA, SXDS and COSMOS, and also can serve as an upper/lower limit to take into account when classifying galaxies using shallower/deeper data sets.

ACKNOWLEDGMENTS

We thank Chris Lintott for accepting to review this paper, giving us constructive comments that in our believe improved the paper significantly. We also appreciate that he signed his report. We thank Kaz Sekiguchi and the SXDS team for informing us about the stage of the photometric redshift measurement. MP acknowledge financial support from JAE-Doc program of the Spanish National Research Council (CSIC), co-funded by the European Social Fund. This research was supported by the Junta de Andalucía through project TIC114, and the Spanish Ministry of Economy and Competitiveness (MINECO) through projects AYA2010-15169, AYA2013-42227-P, and AYA2013-43188-P. The work uses the observations collected at the Centro Astronómico Hispano Alemán (CAHA) at Calar Alto, operated jointly by the Max-Planck Institut für 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%) and the regional Government of Aragón (50%). In this work, we made use of Virtual Observatory Tool for OPerations on Catalogues And Tables (TOPCAT). IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

REFERENCES

Abraham, R. G., et al., 1994, ApJ, 432, 75
Abraham, R. G., et al., 1996, ApJS, 107, 1
Abraham, R. G., van den Bergh, S., Nair, P., 2003, ApJ, 588, 218
Adams, T. F., 1977, ApJs, 33, 19
Andrae, R., Jahnke, K., & Melchior, P., 2011, MNRAS, 411, 385
Andrae, R., Melchior, P., & Jahnke, K., 2011, MNRAS, 417, 2465
Aparicio-Villegas, T., et al. 2010, AJ, 139, 1242
Baillard, A., et al., 2011, A&A, 532, 74
Baldry, I. K., et al., 2004, ApJ, 600, 681
Baldwin, J. A., Phillips, M. M., & Terlevich, R., 1981, PASP, 93, 5
Barden, M., et al., 2012, MNRAS, 422, 449
Bell, E. F., et al., 2003, ApJS, 149, 289
Benítez, N., et al. 2000, ApJ, 536, 571
Benítez, N., et al. 2009, ApJ, 692, L5

24  Pović et al.
