Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **419,** 2703–2724 (2012) doi:10.1111/j.1365-2966.2011.19918.x

A new automatic method to identify galaxy mergers – I. Description and application to the Space Telescope A901/902 Galaxy Evolution Survey⋆

Carlos Hoyos,[1,2]† Alfonso Aragón-Salamanca,[1] Meghan E. Gray,[1] David T. Maltby,[1] Eric F. Bell,[3] Fabio D. Barazza,[4] Asmus Böhm,[5] Boris Häußler,[1] Knud Jahnke,[6] Shardha Jogee,[7] Kyle P. Lane,[1,8] Daniel H. McIntosh[9] and Christian Wolf[8]

[1]*School of Physics and Astronomy, The University of Nottingham, University Park, Nottingham NG7 2RD*
[2]*Departamento de Física Teórica, Facultad de Ciencias, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain*
[3]*Department of Astronomy, University of Michigan, 500 Church Street, Ann Arbor, MI 48109, USA*
[4]*Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland*
[5]*Institute of Astro and Particle Physics, University of Innsbruck, Technikerstr. 25/8, A-6020 Innsbruck, Austria*
[6]*Max-Plank-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[7]*Department of Astronomy, University of Texas at Austin, 1 University Station C1400 RLM 16.224. Austin, TX 78712-0259, USA*
[8]*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH*
[9]*Department of Physics, University of Missouri-Kansas City, 5110 Rockhill Road, Kansas City, MO 64110, USA*

Accepted 2011 September 28. Received 2011 August 25; in original form 2011 May 10

ABSTRACT

We present a new automatic method to identify galaxy mergers using the morphological information contained in the residual images of galaxies after the subtraction of a smooth Sérsic model. The removal of the bulk signal from the host galaxy light is done with the aim of detecting the much fainter and elusive minor mergers. The specific morphological parameters that are used in the merger diagnostic suggested here are the residual flux fraction (*RFF*) and the asymmetry of the residuals [*A*(Res)]. The new diagnostic has been calibrated and optimized so that the resulting merger sample is very complete. However, the contamination by non-mergers is also high. If the same optimization method is adopted for combinations of other structural parameters such as the Concentration, Asymmetry, clumpineSs (CAS) system, the merger indicator we introduce yields merger samples of equal or higher statistical quality than the samples obtained through the use of other structural parameters. We investigate the ability of the method presented here to select minor mergers by identifying a sample of visually classified mergers that would not have been picked up by the use of the CAS system, when using its usual limits. However, given the low prevalence of mergers among the general population of galaxies and the optimization used here, we find that the merger diagnostic introduced in this work is best used as a *negative* merger test, that is, it is very effective at selecting non-merging galaxies. In common with all the currently available automatic methods, the sample of merger candidates selected is heavily contaminated by non-mergers, and further steps are needed to produce a clean merger sample. This merger diagnostic has been developed using the *Hubble Space Telescope*/ACS *F*606*W* images of the A901/902 multiple cluster system ($z = 0.165$) obtained by the Space Telescope A901/902 Galaxy Evolution Survey team. In particular, we have focused on a mass- and magnitude-limited sample ($\log M/\mathrm{M}_{\odot} > 9.0$, $R_{\mathrm{Vega,Total}} \leq 23.5$ mag) which includes 905 cluster galaxies and 655 field galaxies of all morphological types.

Key words: methods: data analysis – galaxies: clusters: individual: A901/902 – galaxies: evolution – galaxies: interactions – galaxies: structure.

⋆Based on observations made with the NASA/ESA *Hubble Space Telescope*, obtained at the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy (AURA), Inc., under NASA contract NAS 5-26555. These observations are associated with programme GO10395.

†E-mail: carlos.hoyos@nottingham.ac.uk

© 2011 The Authors
Monthly Notices of the Royal Astronomical Society © 2011 RAS

1 INTRODUCTION

Mergers are the most extreme type of galaxy interaction, as the final product of a merger event can be totally different from the original objects involved. Considerable efforts have been devoted towards the understanding of the physical processes that regulate galaxy mergers starting from the very early work of Spitzer & Baade (1951) or the seminal simulational works presented in Toomre & Toomre (1972) and Toomre (1977). These works made it clear that, even though the stars rarely collide with each other during a merger process, such episodes can have dramatic consequences for the gaseous component of the galaxies involved. Later works such as those of Barnes & Hernquist (1991), Barnes & Hernquist (1992), Barnes & Hernquist (1996), Barnes (2002), Bournaud, Jog & Combes (2005), Wetzstein, Naab & Burkert (2007), Springel & Hernquist (2005), Bournaud, Duc & Emsellem (2008), Hopkins et al. (2009), Stewart et al. (2009) and Chilingarian et al. (2010) have helped address specific issues of merger processes such as the internal structure of the remnants, the relevance of the orbital parameters or the impact of the gas fraction on the possible regeneration of galactic discs after a merger episode.

However, the study of mergers is not only relevant because of the physics involved. The evolution of the massive early-type galaxies that populate the red sequence cannot be explained using passive evolution models only, and mergers have been found to play a key role in their evolution. In particular, the evolution of the luminosity function and colours of galaxies since $z \simeq 1.0$ observed in COMBO-17 (Classifying Objects by Medium-Band Observations in 17 Filters) (Wolf et al. 2003; Bell et al. 2004) and the phase 2 of the Deep Extragalactic Evolutionary Probe (DEEP) survey (Faber et al. 2007) suggest that the merger episodes have a huge impact on the evolution of early-type galaxies, increasing the stellar mass by a factor of 2 over the last 8 Gyr. More recently, the importance of mergers has been highlighted in Robaina et al. (2010), who conclude that the evolution of massive, red galaxies depends strongly on their merging history. Further studies have helped ascertain the impact of mergers in specific aspects of the evolution of red galaxies, including masses (van Dokkum et al. 2010), sizes (see Trujillo et al. 2006, 2007; Giavalisco, Ravindranath & Daddi 2007; Buitrago et al. 2008; van Dokkum et al. 2008, among others) and velocity dispersions (Cenarro & Trujillo 2009).

The identification of mergers in deep astronomical images is thus a very important issue for galaxy evolution studies, and the huge number of galaxies observed in modern surveys creates the need of reliable automated merger detection mechanisms. A reliable merger identification technique is the key element in the calculation of the merger fraction. The merger fraction is defined as the fraction of galaxies with a recognizable ongoing merger episode that is found in any given (often mass-limited) sample. It is the first step towards the *comoving* merger rate, which is the number of merger *events* per Mpc^3 Gyr.

Several automatic identification techniques have been developed to single out mergers from non-interacting galaxies. These methods use morphological criteria (Concentration, Asymmetry, clumpineSs, hereinafter CAS, and Gini coefficient and M_{20} parameter, hereinafter G–M_{20}, systems or other techniques, Conselice 2003; Lavery et al. 2004; Lotz, Primack & Madau 2004; Cassata et al. 2005; Conselice, Rajgor & Myers 2008; Jogee et al. 2008, 2009; Lotz et al. 2008; Wen et al. 2009), kinematical and spatial close pairs (Patton et al. 2000, 2002; Lin et al. 2004; De Propris et al. 2007; Lin et al. 2008; Robaina et al. 2010), or even the correlation function (Bell et al. 2006; Masjedi et al. 2006). The morphological techniques are based on the fact that the objects involved in a merger episode will be gravitationally disturbed. The CAS system measures these specific aspects of the surface brightness distribution of galaxies in order to identify mergers. Objects with high asymmetries are usually taken to be mergers by this method. On the other hand, the G–M_{20} system measures whether the galaxies appear to be shredded or not, since both the G and the M_{20} numbers measure whether and how is the light concentrated in any given object. In this system, the more shredded galaxies are selected as mergers. The pairing techniques look for pairs of galaxies whose relative positions and velocities should be conducive to strong interactions in a relatively short time-scale after observation. Each of these methodologies are sensitive to different time-scales, mass ratios, orbital parameters, and gaseous content of the galaxies involved. For instance, Conselice (2006) and Lotz et al. (2008) conclude that the CAS parameters are sensitive to roughly a time-scale of $(0.4–1.0) \times 10^9$ yr, while the time-sensitivity of the pairing techniques depends on the projected separation between the galaxies.

This work contributes to the morphological automated detection of mergers. Here, the morphological parameters *of the residual images* after the subtraction of a smooth Sérsic model (see Sérsic 1963 or Graham & Driver 2005, for a definition of this profile) are explored. This should, at least in principle, better reveal the impact of the gravitational interaction on the morphology of galaxies. This was done with the aspiration of detecting minor mergers. An isolated galaxy will, with time, adopt an approximately symmetric profile, whereas an interacting galaxy will appear to be more asymmetric. The removal of an intrinsically symmetric profile such as the Sérsic model, which could be regarded as a quiescent, underlying galaxy, will more clearly expose the asymmetric signature of the light from an interacting galaxy. Thus, the structure of the residuals is investigated with the aim of finding the combination of structural parameters that produces merger samples of better statistical quality. This optimization step is done in an unambiguous way, by using an objective criterion to grade the performance of the diagnostics tried. The specific criterion used here encourages completeness at the expense of a fairly high contamination by non-mergers, and the resulting merger sample needs to be cleaned afterwards.

The layout of this paper is as follows. Section 2 presents the observational data used together with the galaxy samples selected to derive and then test the proposed method. Section 3 describes the data-processing techniques employed, and the structural parameters used. Section 4 presents the objective method introduced here to determine what combination of structural parameters produces the merger sample of highest statistical quality. The precise definition of the 'statistical quality' of a sample that is selected from a parent population is given in Section 4.1. Section 5 presents a visual analysis of the objects selected as potential mergers by the method presented here, focusing on the contamination of the resulting sets of merger galaxies by non-mergers (Section 5.1) and on the ability to detect minor mergers (Section 5.2). Finally, Section 6 presents the conclusions of this work.

2 DATA

The data used to illustrate this method are provided by the *Hubble Space Telescope* (*HST*)/ACS *F*606*W* observations that were obtained as part of the Space Telescope A901/902 Galaxy Evolution Survey (STAGES)[1] (Gray et al. 2009). The STAGES is a

[1] http://www.nottingham.ac.uk/astronomy/stages

multiwavelength project that was designed to explore the impact of environment on galaxy evolution. Its main target is a multiple cluster system located at $z \simeq 0.165$ that harbours different environments with different densities. At the distance of this multiple cluster system, 1 arcsec corresponds to 2.83 kpc. This survey also includes X-ray *XMM–Newton*, ultraviolet *GALEX*, infrared *Spitzer*, spectroscopic 2dF, Giant Metrewave Radio Telescope, and optical COMBO-17 observations.

The *HST*/ACS STAGES observations form an 80-tile mosaic that covers almost 30×30 arcmin^2 in the $F606W$ filter, with an average exposure time of around 2 ks. The observations were reduced using an output pixel scale of 0.03 arcsec, and a *pixfrac* of 0.8, in order to keep the point spread function (PSF) ellipticity as stable as possible for weak-lensing studies. The PSF full width at half-maximum (FWHM) is 3.12 pixels. The point source completeness limit of these images is $F606W(\mathrm{AB}) = 28.5$ mag. These observations are deep enough so as to probe most of the luminosity function of galaxies in clusters, as it is possible to recover reliable structural information for galaxies up to an absolute magnitude of $M_{F606W,\mathrm{AB}} = -15.0$ mag. This limit is just 2 mag brighter than the most luminous globular clusters, and is typical of dwarf elliptical systems.

2.1 Sample selection and morphologies

This work makes use of a mass- and magnitude-selected sample ($9.0 \leq \log M/\mathrm{M}_{\odot}$, $R_{\mathrm{Vega}} \leq 23.5$ mag) which is very similar to the sample used in Maltby et al. (2010). The galaxies are all in the $0.05 \leq z_{\mathrm{Phot}} \leq 0.30$ redshift range, with a relative dearth of sources in the lower end of this interval. The mass limit ensures that the sample is complete in stellar mass for both the blue cloud and the red sequence, as was shown in Borch et al. (2006). The magnitude limit guarantees reliable visual morphologies, since all the sources show extended images in the *HST*/ACS data. The sample includes 1560 galaxies distributed among both the field and the cluster environments. The sample can also be divided into four different morphological classifications labelled as 'E', 'S0', 'Sp' and 'Oth'. The 'E' bin is made of a course of elliptical systems, the 'S0' bin gathers the lenticular galaxies, the 'Sp' bin comprises the spiral galaxies, and the 'Oth' bin includes a mixture of irregulars, compact, and highly disturbed sources that do not fit into any of the other galaxy classes. The 'Oth' bin includes 235 irregular galaxies, 41 disturbed galaxies and 13 compact sources. This bin is thus dominated by irregular systems. This sample was selected for this work due to the following reasons:

(i) *Redshift uniformity*. Most of the galaxies are found in the fairly narrow $0.16 \leq z \leq 0.30$ range, and only 7 per cent of the objects are in the $0.05 \leq z \leq 0.16$ interval. The look-back time is thus approximately the same for most objects and no age-related systematic effects are expected. The STAGES *HST*/ACS observations will also probe similar rest-frame wavelength intervals redwards of the 4000-Å break, observing the same stellar populations. The spatial resolution is approximately the same for most sources, too. There is no need to apply *morphological k*-corrections or distance corrections to the structural parameters. It is important to keep in mind, though, that the total *k*-correction spread could be relatively large, ranging from 0.0 mag for the nearest blue objects to 1.0 mag for the farthest red galaxies. The $(1 + z)^4$ cosmological dimming could also introduce a bias of around 0.5 mag between the high-redshift objects and those located at the cluster's distance. However, the sample used in this work is a mass-limited sample, and the faint magnitude cut applied is only made in order to ensure that all the relevant structural parameters can be measured reliably. This makes it possible to estimate the fraction of sources that could be affected by the *k*-correction and cosmological dimming biases previously mentioned. This is done by adding 1.5 mag to the observed magnitudes and calculating the fraction of sources that would have been excluded from the sample in this situation. This simple estimate indicates that 90 per cent of the galaxies in the current sample should be free from these systematics.

(ii) *Variety of environments and morphologies*. This sample includes a good number of galaxies of several morphological types residing in very different environments. This makes it ideally suited to explore whether the performance of our merger indicators is sensitive to these variables or not.

All the objects of this sample also have good morphological information, which will be presented in detail in a forthcoming paper (Gray et al., in preparation). The morphological catalogue was created using visual classification from seven STAGES team members who were first trained on a consistent subsample of previously classified galaxies. All 5090 galaxies with $R < 23.5$ mag and $z_{\mathrm{phot}} < 0.4$ were classified blindly without the knowledge of their cluster or field membership. Galaxies were randomly assigned to the classifiers who used a revised Hubble T-type scale and the weighting scheme described in Lane et al. (2007). Each galaxy in the sample was classified by three people, while a subset of 786 bright galaxies previously studied in Wolf, Gray & Meisenheimer (2005) received classifications from all seven.

In addition to the revised Hubble T type, classifiers were also able to note further information about the galaxy structure within certain well-defined parameters. The dynamical state involved an interpretation of the probable cause of any observed disruption, where the possibilities were a tidal interaction with a neighbour (I); tidal interaction suggesting a merger (M); tidal feature without obvious cause (T); and chaotic systems (C). The degree of disruption was quantified by a disturbance parameter which was allowed values of 0, indicating little or no disturbance; [1,2], indicating moderate/strong asymmetry (e.g. an H II region); and [3,4], showing a moderate or strong distortion (e.g. a tidal tail).

The precise sample definition and its breakdown can be summarized as given below:

(i) *The cluster sample*. This sample is defined by a redshift interval $z_{\mathrm{Phot}} = [0.17 - \Delta(R), 0.17 + \Delta(R)]$, where the photometric redshifts have been calculated from the COMBO-17 data, and the half-width
$\Delta(R) = \sqrt{0.015^2 + 0.0096525^2 \times [1 + 10^{0.6 \times (R_{\mathrm{Tot}} - 20.5)}]}$
is allowed to vary with apparent R-band magnitude. This redshift interval varies with apparent brightness due to the higher precision of the COMBO-17 photometric redshifts for the brighter sources (see Gray et al. 2009, for a more detailed explanation.). The photo-z half-width distribution was normalized so the completeness was $\geq$90 per cent at any magnitude. For the faint end, there is also some sample contamination from field galaxies. This was calculated using the counts of the smooth models seen in fig. 14 of Gray et al. (2009). When estimating the contamination, the field distribution was assumed to be consistent with the average galaxy counts $N(z, R)$ outside the cluster. The average contamination is 30 per cent. This subsample thus contains cluster galaxies with good photometry for which the *HST*/ACS observations show an extended source. This cluster sample contains 905 galaxies.

(ii) *The field sample*. This includes galaxies in the redshift intervals $z_{\mathrm{Phot}} = [0.05, 0.14]$ and $z = [0.22, 0.30]$. In this subsample, only

HST/ACS extended sources with good photometry are included. This field sample contains 655 galaxies.

This leaves a final sample of 655 field galaxies (100 Es, 60 S0s, 318 Sps, 177 Oths) and 905 cluster galaxies (192 Es, 216 S0s, 383 Sps, 114 Oths). As before, the 'Oth' galaxies represent a mixture of irregulars, compact, and highly disturbed systems. This classification is dominated by irregular galaxies. The STAGES morphological catalogue has good morphological information for all objects in this sample. The source detection runs presented in Section 3.1 and the structural parameter analysis presented in Section 3.2 yielded structural information for a total of 1537 sources. Thus, there is structural information available for 98.5 per cent of the full sample.

3 DATA PROCESSING

This section presents the SExtractor[2] (Bertin & Arnouts 1996) runs and subsequent morphological analysis carried out here. The main ingredients this work needs are an error or weight image and a mask image. The error image is needed by both SExtractor and GALFIT (Peng et al. 2002) in order to modulate the importance of each pixel in the photometry and morphological properties. The mask image, which is created by SExtractor, is later needed by GALFIT so that it fits galaxies in a way designed to identify mergers. The basic idea is to produce a simple Sérsic model that is as focused on the main body of the target galaxies as possible. The putative mergers should not therefore have any effect on these fits and will be either simultaneously fitted or masked out.

This Sérsic model should be an appropriate description of the surface brightness distribution of elliptical galaxies. However, it has to be noted that in the case of lenticular and spiral galaxies, bulge+disc decompositions (B+D) would produce better descriptions. However, there are a number of reasons that make simple Sérsic fits preferable.

(i) A simple Sérsic profile will remove a fraction of the smooth, symmetric signal from any given galaxy, regardless of its morphology. This statement remains true for disc galaxies.

(ii) The goal of this work is to develop an automatic and fork-free method. The use of B+D decompositions together with simple Sérsic fits would require an additional step in order to decide what functional form describes best each galaxy. There are objective ways to do this, such as the Bayesian method presented in den Brok et al. (2011) or the Akaike information criterion, but this is well beyond the scope of this work.

(iii) B+D decompositions need to be re-examined for their internal consistency, as their final results might depend on the initial point in parameter space chosen for GALFIT minimization.

Thus, it is not possible to use the SExtractor and structural catalogue presented in Gray et al. (2009) because the source detection and fits shown there were made with the goal of creating a robust catalogue of structural parameters derived from single Sérsic fits. Instead, the SExtractor configuration used in this work (see Section 3.1) is fairly sensitive to the detection of faint, small features, as it is set to detect the faintest sources that can be said to be detectable in the images used. More importantly, the deblending parameters are selected so that SExtractor will tell apart sources with up to a 3 mag brightness difference, incorporating much of the `1:3` to `1:10` mass ratio range defining minor mergers, assuming that their mass-to-light ratios (M/Ls) are similar. These SExtractor parameters are thus not only intended to separate the objects involved in a major merger, with luminosity ratios between `1:1` and `1:3`, but are also geared towards the deblending of the objects involved in a minor-merger episode, with luminosity ratios between `1:3` and `1:10`. Also, the very sensitive detection threshold employed makes it more likely that the less luminous segment of any given deblended object will be engulfed by the Kron aperture of the larger segment. This is key in the current analysis, since it is hypothesized here that such less luminous segments could be the less luminous galaxies in merger episodes. However, these smaller segments could just be H II regions or simply objects along the same line of sight, which will naturally lead to a contamination of the merger sample. This will be discussed in Section 5.1. The larger segments in deblended sources would then be the more luminous galaxies. The Kron aperture (Kron 1980) is defined to contain a specific fraction of the light of a galaxy. Such fraction depends on the intrinsic profile of galaxies, but it ranges from 90 per cent for the steepest profiles to 95 per cent for exponential discs. Both percentages depend on the observed surface brightness, though in the sense that SExtractor misses more flux from the dimmest objects. See Hammer et al. (2010) for a description of this systematic error.

The larger and more luminous object is then fitted by a smooth Sérsic model which is created using GALFIT according to the rules explained in Section 3.2. This model is then subtracted from the original image in order to estimate how the image would look without the more luminous galaxy of a merger event. This *residual* image still contains most of the signal from the less luminous sources that were found within the observed Kron aperture of the larger object. This ensures that these smaller sources will have a great impact on the structural properties of the *residual* images. This is indeed one interesting property of these *residual* images that will be exploited.

The structural parameters of both the original image and of the residual image are then calculated within the aforementioned Kron aperture. The specific morphological parameters used are given in Section 3.3. These are mostly based on previous works. The structural parameters of the residual image within the Kron aperture are then expected to be very sensitive to the smaller and less luminous member of the merger.

3.1 SExtractor runs

SExtractor was run twice on each of the 80 tiles that compose the STAGES *HST*/ACS *F*606*W* mosaic. The first pass was performed for the sole purpose of obtaining an empirical map of the background variance and of the filtered Poissonian signal. Such images can be produced by SExtractor as *output* images. These images were then combined to produce an appropriate weight image to be used in the second pass. The first SExtractor run is presented in Section 3.1.1, and it can be safely skipped by users providing their own error or weight images. The second run is presented in Section 3.1.2, which presents the SExtractor configurations used to produce the final source lists and mask images.

3.1.1 Preliminary SExtractor *run*

The most relevant SExtractor parameters used in the first run are summarized in Table 1. Table 1 also presents the main results obtained from this first SExtractor run, averaged over the 80 tiles. The average background σ is an estimate of the noise that is observed in the image areas in which there are no galaxies. This is set

[2] Version 2.5.0

Table 1. SExtractor parameters used for the first pass of this code on the *HST*/ACS *F606W* images. This first run was used to obtain an empirical estimate of the final error image. The main results obtained from this first pass are also presented in this table.

DETECT_MINAREA	7	WEIGHT_TYPE	BACKGROUND
DETECT_THRSHLD	0.75	BACK_TYPE	AUTO
ANALYSIS_THRSHLD	0.75	BACK_SIZE	256
FILTER_FWHM	3.0	BACK_FILTERSIZE	5
DEBLEND_NTHRESH	32	BACKPHOTO_TYPE	LOCAL
DEBLEND_MINCONT	0.005	BACKPHOTO_THICK	64
Average background σ ($\sigma_{\rm Bkg}$) =	3.73×10^{-3}	Average effective gain ($\langle G_{\rm eff}\rangle$) =	1.6×10^{3}

by the readout noise of the analog-to-digital converters of the ACS, the Poissonian noise of the sky background, and the use of the MULTIDRIZZLE technique. It was estimated as the average of the median values of the background *RMS* images created by SExtractor in this first pass. The average effective gain $\langle G_{\rm eff}\rangle$ roughly measures the growth in the photometric errors caused by the intrinsic Poissonian nature of photon-counting measurements. It was measured using the background-subtracted filtered frame and a very rough empirical estimate of the per pixel *RMS* image obtained by means of a loose adaptation of the method presented in Grazian et al. (2006), which shows precise formulae to calculate the *RMS* of an image in the case of noise correlation. The effective gain image is then

$$\langle G_{\rm eff}\rangle = \frac{I}{RMS^2}, \tag{1}$$

where I is the background-subtracted filtered image, and *RMS* is the empirically derived uncertainty.

The $\langle G_{\rm eff}\rangle$ number is yielded by the value of this image in the brighter areas of the images dominated by the Poissonian noise.

The measured value of $\langle G_{\rm eff}\rangle$, reported in Table 1, agrees very well with its expected value of around 1445, which can be theoretically estimated as

$$G_{\rm eff} = g \times T \times \left(\frac{0.03}{0.05}\right)^2 \simeq 1445.0 \tag{2}$$

where $G_{\rm eff}$ is the effective gain, g is the original detector's gain (2, for the STAGES observations), T is the total exposure time, and the fraction is the ratio between the effective areas of the pixels before and after the operation of the MULTIDRIZZLE technique. The latter value of 1445.0 was adopted for use in the second run of SExtractor. The measured value of $G_{\rm eff}$ was then merely used as a sanity check to ensure that the errors in the input image indeed behave as expected.

The first SExtractor run thus produces an estimate of the background $\sigma_{\rm Bkg}$ and a background-subtracted filtered frame containing the Poissonian signal S. These were combined according to the usual CCD error equation in order to obtain a weight image for use in the second SExtractor run:

$$\text{Weight} = \frac{1}{\sigma^2_{\rm Bkg} + S/G_{\rm eff}}, \tag{3}$$

where Weight is the final weight image, $\sigma_{\rm Bkg}$ is the background *RMS* image, S is the background-subtracted filtered image created by SExtractor, and $G_{\rm eff} = 1445.0$. The final weight image was later processed using the WEIGHTWATCHERS[3] code to ensure that problematic pixels with either zero exposure time (typical of the image edges) or with saturated signal were assigned zero weight. Less than 2 per cent of the pixels had to be discarded in this final procedure. WEIGHTWATCHERS also produces a flag image that was used in the second pass of SExtractor. This final weight frame is then used in the second run of SExtractor to suppress the detection of objects in low-weight pixels and give appropriate pixel weights for the photometry.

[3] See `http://www.astromatic.net/software/weightwatcher`.

3.1.2 *Second* SExtractor *run*

This second pass was used to obtain the final segmentation image and a SExtractor catalogue, which are key to the GALFIT analysis of the images. The segmentation images separate object pixels from background pixels, and provide basic photometric information to be used as initial conditions for GALFIT. It is highlighted here that the *input photometric catalogue* of target sources is not produced by this SExtractor run. The list of objects to fit and study is defined in Section 2.1, and SExtractor is run here with the sole purpose of obtaining basic photometric information about this pre-defined sample. The new configuration used is presented in Table 2.

As it can be seen from Table 2, the new configuration is fairly aggressive. It is more sensitive than the 'hot' configuration used in Gray et al. (2009). The minimum nominal integrated signal-to-noise ratio (S/N) of the detections is 8.3. This is 1.7 times higher than the usual S/N limit of 5.0 which is usually accepted for a point source detection. The new SExtractor parameters were chosen this way because, to zeroth order, the effect of noise correlation is an artificial

Table 2. SExtractor parameters used for the second pass of this code on the *HST*/ACS *F606W* images. This second pass makes use of the empirical error image created using the *output* images of the first SExtractor run. The flag image created by WEIGHTWATCHERS was also used to exclude pixels with saturated signal or zero effective exposure time.

DETECT_TYPE	CCD	FLAG_TYPE	OR
DETECT_MINAREA	12	KRON_FACT	2.5
THRESH_TYPE	RELATIVE	MIN_RADIUS	3.5
DETECT_THRSHLD	0.80	DEBLEND_NTHRESH	32
ANALYSIS_THRSHLD	0.80	DEBLEND_MINCONT	0.01
FILTER_FWHM	3.0 (G)	CLEAN	Y
BACK_TYPE	AUTO	CLEAN_PARAM	1.0
BACK_SIZE	256	STARNNW_NAME	default.nnw
BACK_FILTERSIZE	3	MASK_TYPE	CORRECT
BACKPHOTO_TYPE	LOCAL	INTERP_TYPE	ALL
BACKPHOTO_THICK	64	INTERP_MAXXLAG	16
WEIGHT_TYPE	MAP_WEIGHT	INTERP_MAYYLAG	16
WEIGHT_GAIN	N	SATUR_LEVEL	40 000.0
BACK_TYPE	AUTO	MAG_ZEROPOINT	26.49113
BACK_FILTERSIZE	3	PIXEL_SCALE	0.03
BACK_SIZE	256	GAIN	1445.0
BACKPHOTO_TYPE	LOCAL	SEEING_FWHM	0.106
BACKPHOTO_THICK	64		

© 2011 The Authors, MNRAS **419,** 2703–2724
Monthly Notices of the Royal Astronomical Society © 2011 RAS

increase in the σ image with respect to the inverse square root of the weight image. Thus, a stronger signal is needed in order to spawn a genuine detection. Given the number of images that were multidrizzled together and the target resolution of the *HST*/ACS images used, experience shows that the nominal S/N has to be multiplied by around 0.7, as is shown in Hammer et al. (2010) and Casertano et al. (2000). This is the reason behind the additional factor of 1.7. The second SExtractor configuration makes use of a Gaussian filter with a FWHM of 3 pixels, and a minimum detection area of 12 pixels. This effectively removes spurious noise peaks, which are generally much smaller than the instrumental FWHM. As for the deblending parameters, the ones shown in Table 2 would, in principle, make SExtractor separate objects with a factor of 100 difference in flux. However, the value of the `DEBLEND_NTHRESH` parameters also affects the deblending process. With the choice of 32 deblending thresholds, the child detections tend to bear a larger fraction of the total flux than just 1 per cent of the parent source. This was done this way with the idea of separating minor mergers in a very late stage, with a mass ratio close to `1:10`. The SExtractor parameters used in the second pass also ensure that the source catalogue will include detections all the way up the real, almost point source detection limit of the images. At the same time, the minimum nominal S/N required to foster a detection and the `DETECT_MINAREA` value adopted ensure that the number of spurious sources caused by *the* multidrizzle *algorithm and noise peaks* be kept at a minimum. This configuration, however, is open to the inclusion of spurious extended sources caused by *statistical fluctuations.* Also, this configuration makes SExtractor include pixels with poor signal within the isophotal area of the detections which guarantees that the sky portion of the segmentation image is free from most of the flux originating from the uncovered sources. This also has the important effect of making the Kron apertures as large as they could possibly be, given the data used. This maximizes the probability of a small galaxy in a minor merger event falling within the Kron aperture of the more luminous object. This is particularly important since the structural parameter measurements shown in Section 3.3 are performed over this Kron aperture. On average, SExtractor found 34 600 objects in each of the frames. This number is to be compared with the *total* number of sources found by Gray et al. (2009), which is 75 805 *in all 80 tiles.* This is explained by the fact that the SExtractor configuration used in Gray et al. (2009) was optimized to find and fit $R_{\rm ap} \leq 24.0$ mag counterparts from a previous catalogue obtained from the *R*-band COMBO-17 data, while the goals of the SExtractor catalogue used in this work encourage the detection of faint sources near the brighter ones. This does not mean that all the sources found by this second SExtractor run are legitimate, bona fide detections. In fact, specific simulations indicate that the majority of objects with `MAG_ISO` dimmer than 27.0 mag are spurious objects caused by statistical fluctuations, while the majority of the objects with `MAG_ISO` brighter than 27.0 mag are real sources. This is not a problem for the target sources studied in this work, whose `MAG_ISO` are all brighter than 24.0 mag. The fitting scheme used here, which is presented in Section 3.2, just discards these faintest detections and therefore they will not have any impact on the actual fits produced for the much brighter targets of interest. At the same time, this SExtractor configuration will detect *and more importantly, isolate* sources in the $24.0 < F606W(\mathrm{AB}) < 27.0$ mag magnitude interval, which are much more likely to be real objects. The comparison between the number of detections obtained here and in the Gray et al. (2009) work merely reflects that the configuration used in this work is much more sensitive to the smallest features, which increases the chance of including spurious detections in the SExtractor catalogue.

3.2 Galaxy fitting: the GALP-HYT wrap scripts

The main contribution of this work is that it begins to explore whether the morphological information contained in the *residual* images of galaxies can be used to assess if a galaxy is involved in a merger episode. Such residual images are created by subtracting smooth models of the target galaxies from the original images while leaving most of the signal from other, possibly interacting objects in these *residual* images. There are a number of codes capable of producing models of galaxies in astronomical images by performing two-dimensional model fits to their surface brightness distributions. The most commonly used ones are GIM2D (Simard 1998) and GALFIT. Even though these codes are different in their specific details, their basic principles are very similar. They both try to minimize a possibly weighted χ^2 value that depends on the structural parameters of the object being fitted. This residual sum of squares is formed from the difference between the observational data and a trial function that is created according to a user-supplied set of rules. Chief among these rules are an error image, a mask image and a PSF image. The error image regulates the relative weight that the different pixels should be given. Thanks to this image, it is possible to prevent saturated pixels from having any weight in the figure-of-merit that the codes are set out to minimize. This image can also be used to reduce the impact of areas with lower or no exposure time in mosaic images created using the MULTIDRIZZLE technique. The mask image can be used to modify how the different fitting codes treat the different areas of the input image. For instance, it can help the fitting codes to tell what pixels belong to the target object being analysed, and what pixels belong to other objects and should therefore be discarded. The PSF image is also a key ingredient in the structural analysis of the surface brightness distribution of galaxies. It is also possible to constrain the parameter space region that the codes are allowed to explore.

This work uses the GALFIT code. The setup with which GALFIT runs is created by a PYTHON code called GALP-HYT[4] written by CH. The GALFIT setup used is designed to create and subtract a smooth model of the primary or target sources, while leaving the signal of nearby objects largely intact in the residual images. In this context, the target sources are the galaxies that are being studied, and close sources are the objects found within the Kron aperture of the primary galaxy. Given the SExtractor parameters used, these close sources could be the less luminous galaxies in a minor merger event, or one of the protagonists in a major-merger episode. Thus, GALFIT is configured to produce a residual image in which only the model for the primary source has been removed. This is done in order to guarantee that the Kron aperture in the residual images will contain most of the information from the merged sources, while being free from the effect of the target galaxy. The GALP-HYT code is now briefly described.

For each science image or tile it is given, it takes the following information as input:

(i) *The science image itself*. These are the same images used in the structural parameter catalogue presented in Gray et al. (2009).

[4] Pronounced Galp-Hit.

(ii) *The weight image that* SExtractor *used during its second pass*. This is converted into a σ image by taking its inverse square root. In this process, pixels with zero weight are given a very large value of σ.

(iii) *The* SExtractor *list of sources and segmentation images.* These are created during the second pass of the detection software.

(iv) *A list of primary targets*. This list of targets is given in a separate text file, one object per line. In this file, targets are identified by their SExtractor number IDs.

(v) *A PSF imagelet*. This was again taken from the work done in Gray et al. (2009). This is a PSF imagelet that was built from many different non-saturated stars found across a number of tiles and hence its S/N is very high.

For each object in the list of targets, the SExtractor catalogue is examined and objects are classified with respect to the primary object according to the following set of rules:

(i) *The target object itself*. It is fitted using a single Sérsic model with a free floating disciness–boxiness 'C0' parameter. This extra freedom allows the model for the target object to take into account a larger fraction of the symmetric, undisturbed signal of the primary source. At the same time, this extra parameter does not complicate the interpretation of the fits since it is unlikely to introduce important degeneracies. Also, the evolution of 'C0' as a function of the number of minor-merger events is a prediction of the Bournaud et al. (2005); ,2008) models.

(ii) *Objects up to 2 mag fainter than the target source whose centres lie within the Kron aperture of the target source and are tagged as 'A'*. In this context, the relevant magnitudes are the `MAG_ISO` magnitudes. The 2 mag difference was used since this is approximately the expected magnitude difference between the galaxies involved in a minor merger with a mass ratio of `10:1`, assuming that the M/Ls are similar. These sources are fitted with a single Sérsic profile with elliptical isophotes. The centres of the 'A' components are fixed to the values found by SExtractor. This constraint prevents degenerate fits, as the 'A' components will not blend and shift towards the primary object.

(iii) *Objects more than 2 mag fainter than the target source whose centres lie within the Kron aperture of the target source and are tagged as 'B'*. Again, the magnitudes used in this criterion are the `MAG_ISO` magnitudes. These sources are fitted as simple exponential models with elliptical isophotes. This ansatz is less flexible than the one used for the 'A' sources, but this is justified by the fact that there is less information available for these dimmer sources. The use of exponential profiles also helps to increase the execution speed. Again, the centres of the 'B' components are directly taken from the values found by SExtractor in order to prevent degenerate fits.

(iv) *Objects up to 3 mag fainter than the target source outside the Kron aperture of the target source and are tagged as 'E'* (*for external*). As before, `MAG_ISO` measurements were used. These objects are fitted using simple exponential discs. Objects more than 3 mag fainter than the source of interest are ignored. The centres, ellipticities and position angles of the 'E' components are again taken from the SExtractor catalogue.

The segmentation image produced by SExtractor is also modified so that the pixels belonging to the target source and to the 'A' objects are given a value of 0. This ensures that these pixels are taken into account by GALFIT when calculating its figure-of-merit. Pixels belonging to 'B' and 'E' objects are not nulled, and thus the pixels contained within their `ISOAREAS` have no weight in the fit. Only the extended tails of the 'B' and 'E' objects that go beyond their `ISOAREAS` are fitted by the exponential profiles. It is important to note that there will be many objects $\geq$3 mag fainter than the target source that will not fit any of the three categories above. These faint objects are not fitted in any way and have no effect on the fits. This is important because the integrated magnitudes of the sources of interest are all brighter than $F606W(\mathrm{AB}) = 24.0$ mag, and therefore the large number of spurious sources with magnitudes fainter than $F606W(\mathrm{AB}) = 27.0$ mag will have no effect on the fits. The sky is left as a free floating parameter. This is justified by a number of reasons. First of all, the SExtractor configuration used guarantees that the isophotal apertures of the detected objects reach to fairly faint surface brightnesses, and thus the sky pixels in the segmentation image contain very little residual signal from the detected objects. Also, the masking process implies that the majority of the flux from the 'B' and 'E' objects is not taken into account by the fits and only their tails are fitted using exponential profiles. This naturally means that whatever their contribution might be to the average sky level affecting the target object, it will be approximately corrected for by the exponential fits. Finally, the 'A' objects are fully fitted. For these reasons, it is appropriate to leave the sky as a free floating parameter.

The size of the fitting box is defined in the following two steps. In the first step, the maximum between 150 pixels and the circularized Kron diameter is used. In the second and final step, this area is expanded so that it encompasses the centres of all 'E' objects whose Kron apertures intersected the first box.

The total number of additional sources that have to be fitted in conjunction with a primary target is around 20. It is highlighted here that each SExtractor detection is fitted by a simple, solid profile. In each case, the target object is fitted by a single Sérsic model, and the majority of the remaining components correspond to 'E' objects and do not merge with the main body of the target galaxy. This number of extra components is not unusual (see e.g. Häussler et al. 2007). The main target galaxy is fitted using a single Sérsic model. In most cases, the majority of the remaining components are external (E) objects whose main purpose is to help GALFIT compute a good sky value. The A and B components are the putative minor mergers, and their role is to allow the Sérsic profile to provide a good fit to the main body of the target galaxy.

Although this is very expensive in terms of CPU time, the GALPHYT code was written to be run in high-performance computers (HPCs) with thousands of CPUs, as it features a semi-intelligent built-in system of organizing its internal data flow. It was run in the HPC of the University of Nottingham.[5] Initial values and constraints are taken from the source catalogue created by SExtractor.

3.3 Additional structural parameters: the CAS indices and the G–M_{20} system

One of the common ways to tackle the automated detection of mergers is based on the signatures that merger events leave on the morphology of galaxies. As shown in Conselice (2003) and Papovich et al. (2005), the structures of galaxies contain important information about their past star formation modes, and they can also shed light on their interaction history. These ideas are used to identify mergers according to their morphological properties. In this work, the structural properties of the residual images after the

[5] See `http://www.nottingham.ac.uk/hpc/`

subtraction of the Sérsic model described in Section 3.2 are explored and used as merger diagnostics.

There are two systems which are currently in use for the morphological identification of mergers. The first one is the use of the CAS system, which was first introduced in Bershady, Jangren & Conselice (2000) and Conselice (2003). This system has been used in many works aimed at the study of merger fraction in many different contexts (see, for instance, Conselice et al. 2009; Jogee et al. 2009; López-Sanjuan et al. 2009).

This system makes use of three different indices:

(i) *C*: The concentration index '*C*' measures to what extent the light in the galaxy is concentrated towards its centre. It is defined as

$$C = 5.0 \times \log\left(\frac{r_{80}}{r_{20}}\right), \tag{4}$$

where r_{80} is the circular radius containing 80 per cent of the total light from the galaxy, and r_{20} is the radius of the circular aperture that encloses 20 per cent of the total light of the target galaxy.

The concentration index '*C*' takes values between 2.2 and 5.0. If it is calculated for simple Sérsic models, then it depends mainly on the Sérsic index and then on the ellipticity.

(ii) *A*: The asymmetry '*A*' measures to what extent any given image changes under a 180° rotation around the point that minimizes the asymmetry of that image. It is defined as

$$A = \left(\frac{\sum_{i,j}|I_{i,j} - I^{180}_{i,j}|}{\sum_{i,j}|I_{i,j}|}\right) - \left(\frac{\sum_{i,j}|B_{i,j} - B^{180}_{i,j}|}{\sum_{i,j}|I_{i,j}|}\right), \tag{5}$$

where $I_{i,j}$ represents the original image, and $I^{180}_{i,j}$ is a 180°-rotated version of the original image. In the same manner, $B_{i,j}$ is a patch of background, and $B^{180}_{i,j}$ is a 180°-rotated version of this patch of background. This contribution from the background is minimized independently in the same manner. This asymmetry measurement is defined even for images whose average value is 0.0, as it is normalized to the sum of the *absolute* values of the fluxes from each pixel. The rotation centre is optimized so that the value of the first term in the subtraction is a minimum. Here, the rotation centre is allowed to lie at most 9 pixels away from the SExtractor-defined centre. However, when calculating the asymmetry of the residual images, the rotation centre is only allowed to move 4 pixels. The second term in the subtraction also undergoes this optimization process and it removes the contribution to the asymmetry from the background. The asymmetry index of real images of galaxies can take values between 0.0 and approximately 0.8. Most objects have asymmetries lower than 0.2, though. The asymmetry index of the residual images of galaxies after the subtraction of a single Sérsic model ranges between 0.4 and 1.6.

(iii) *S*: The clumpiness 'S' quantifies the fraction of light in a galaxy that is contained in clumpy distributions. Large values of *S* imply that the light in the galaxy is accumulated in few, distinct structures. Low values of *S* indicate that the light distribution is smooth. It is defined as

$$S = 10 \times \left\{\left[\frac{\sum_{i,j}(I_{i,j} - I^{\sigma}_{i,j})}{\sum_{i,j} I_{i,j}}\right] - \left[\frac{\sum_{i,j}(B_{i,j} - B^{\sigma}_{i,j})}{\sum_{i,j} I_{i,j}}\right]\right\}, \tag{6}$$

where $I_{i,j}$ again represents the original image, and $I^{\sigma}_{i,j}$ represents a blurred version of it, which is produced by convolving the original image with a two-dimensional circular Gaussian kernel with a typical dispersion of σ. It is usually correlated with the size of the target galaxy. The residual image after this subtraction only includes signal that is included in high-frequency features of the galaxy. Also, the convolution procedure is applied to a blank patch of sky in the image. This ensures that the contribution from the background noise is discounted from the final value of *S*. The '*S*' parameter can take values between −0.5 and 1.5, although it depends on the size of the convolution kernel used.

The second system is based on the use of the *G* and M_{20} parameters, which were originally introduced by Abraham, van den Bergh & Nair (2003) and Lotz et al. (2004). The Gini coefficient *G* measures the light concentration, like the *C* parameter, but it is insensitive to any particular centre. It is calculated according to the following formula:

$$G = \left[\frac{1}{|\bar{f}| \times n \times (n-1)} \sum_{i=1}^{n} (2 \times i - n - 1) \times |f_i|\right], \tag{7}$$

where *n* is the number of pixels, f_i is the flux observed in the *i*th resolution element, and the sum is made in ascending order of fluxes, so that $f_{i-1} \leq f_i \leq f_{i+1}$. This *G* index tells whether the light is evenly distributed among the different resolution elements of an image. The *G* index has a value of 0.0 for flat light distributions, and it has a value of 1.0 for light distributions in which all the light is contained in a single pixel. In practical terms, the Gini coefficient of real galaxy images lies in the [0.35, 0.85] interval.

The M_{20} parameter is based on the second-order moment of the light distribution $M_{\rm Tot}$. It is defined as

$$M_{20} = \log\left(\frac{\sum_{i=1\in A}^{K:\sum_{l=1\in A}^{K} f_l = 0.2 \times L_{\rm Tot}} f_i \times \left[(x - x_{\rm c})^2 + (y - y_{\rm c})^2\right]}{M_{\rm Tot}}\right), \tag{8}$$

where *A* is the aperture within which this number is obtained, f_i is the flux of the *i*th resolution element, $L_{\rm Tot}$ is the total apparent luminosity contained in the aperture used, $x_{\rm c}$ and $y_{\rm c}$ are the coordinates of the barycentre of the light distribution for which the index is being calculated, and $M_{\rm Tot} = \sum_{i\in A} f_i \times [(x - x_{\rm c})^2 + (y - y_{\rm c})^2]$. In this definition, and in contrast with the definition of *G*, $f_{i-1} \geq f_i \geq f_{i+1}$.

M_{20} measures how far from the galaxy centre it is possible to find the brightest features of the surface brightness distribution of the light. The M_{20} number can go from −3.0 for very concentrated objects to −0.4 for objects with shredded light distributions.

In addition to the aforementioned indices, the residual flux fraction (*RFF*, see Hoyos et al. 2011) is used. It is defined here as

$$RFF = \frac{\sum_{i,j\in A}|I_{i,j} - I^{\rm GALFIT}_{i,j}| - 0.8 \times \sum_{i,j\in A}\sigma_{{\rm Bkg}i,j}}{\sum_{i,j\in A} I^{\rm GALFIT}_{i,j}}, \tag{9}$$

where *A* is the particular aperture used to calculate this index, and $I^{\rm GALFIT}$ is the model created by GALFIT.

The *RFF* as defined here measures the fraction of the signal contained in the residual image that cannot be explained by fluctuations of the background. The 0.8 factor included in the definition of the *RFF* ensures that the expectation value of the *RFF* of a purely Gaussian noise error image of constant variance (as opposed to a spatially varying variance) is 0.0. This fact arises from the following integral:

$$0.8 = \sqrt{\frac{1}{2\pi}} \int_{-\infty}^{\infty} |x| \times {\rm e}^{-x^2/2}\, {\rm d}x, \tag{10}$$

which calculates the expectation value of the absolute value of a Gaussian random variable.

In this work, the structural indices presented above were calculated for three different images. These are the imagelets containing the target object cropped by GALFIT, the simple Sérsic model created

by GALFIT, and the *residual* image obtained by subtracting the second image from the first one. It is stressed that the latter image is only stripped from the signal of the main body of the target galaxy. The indices were also calculated for an artificial image simulating the background noise that affects the primary galaxy, which will be described later on. This image is needed to calculate the background terms in the *A* and *S* indices, but it is also used to estimate the errors in the derived morphological parameters.

In all cases, the aperture used was the Kron aperture calculated by SExtractor. This aperture was chosen for the following reasons:

(i) This aperture is designed to trace an elliptical region in the image in which the contribution in flux from the source of interest to which it is associated is either dominant or notable. Outside the Kron aperture, there is still flux from the primary source, but it is heavily affected by noise. It is therefore pointless to go much further since morphological perturbations of the target galaxy at these levels will be impossible to measure reliably.

(ii) The Kron aperture misses a definite fraction of the light from most common profiles. This fraction depends on the intrinsic shape or profile of the target source and on its effective surface brightness as well. Thus, the Kron aperture is a S/N-matched aperture that grabs a more or less constant fraction of the total flux for objects with similar S/N, with a slight dependence on its profile.

(iii) This aperture has a radius that is typically 60 per cent larger than the Petrosian radius defined in Petrosian (1976). This is little bit smaller than the typical aperture of choice in which the structural parameters defined above are calculated in other studies, which is twice the Petrosian radius. The choice of the Kron aperture will therefore provide higher S/N measurements of the morphological parameters, losing only a minimal amount of information about the outer structure of the studied objects.

The whole Kron aperture is used to calculate the CAS, *G* and M_{20} numbers for the real image, the model and the background noise frame. However, when calculating the structural parameters for the residual image, a small area 3 pixels in diameter is removed from the centre of the Kron ellipse so that the indices are not biased by uncertainties in the fit, PSF mismatches and resampling problems in these complicated regions. The latter aperture in which the centre is excluded is also used to calculate the *RFF* and *S* structural parameters. Also, the rotation axis for the calculation of the asymmetry parameter of the *residual* images is not allowed to drift from the optimal centre found for the original image by more than 4 pixels. In all cases, the background was subtracted from the images using the sky value yielded by GALFIT. This was done with the purpose of minimizing the impact of the background terms on the *A* and *S* numbers.

For the calculation of the *S* parameter, a Gaussian kernel with $\sigma = 0.2 \times R_{\rm K}$ was used. In this expression, $R_{\rm K}$ is the radius of the Kron aperture calculated by SExtractor.

The background noise image that is needed in order to calculate the background terms of the *A* and *S* numbers was created from the *RMS* image that GALFIT used. It was created on an object-by-object basis, so that each object has an individualized background noise image. The first step is to change the data number values of the pixels in the *RMS* image that, according to the segmentation image created by SExtractor, have been flagged as object[6] pixels. Their new value is then set to the median value of the *remaining* sky pixels in the *RMS* image. This modified image is then multiplied by a white noise image with $\sigma = 1.0$. This final step creates an image that is a good representation of the underlying noise that affects the measurements. It would be ideal if this image included correlated noise, but since the target galaxies are all larger than the error correlation length, this image was deemed sufficient.

[6] This includes all objects, not only the galaxy of interest.

Table 3. Typical absolute errors in the structural parameters used in this work. The left-hand columns present the typical errors for the structural parameters of the original galaxy image, while the right-hand columns show the uncertainties for the structural parameters of the residuals. The errors above the horizontal line are the random errors, while the entries below the horizontal line are the systematic uncertainties for the *A*(Obj) index.

σC(Obj)	1.0×10^{-4}	σC(MDL)	1.0×10^{-4}
σA(Obj)	4.0×10^{-2}	σA(Res)	4.0×10^{-2}
σS(Obj)	6.0×10^{-2}	σS(Res)	6.0×10^{-2}
σG(Obj)	1.0×10^{-3}	σG(Res)	1.0×10^{-3}
σM_{20}(Obj)	1.0×10^{-3}	σM_{20}(Res)	1.0×10^{-3}
σRFF	2.0×10^{-2}	...	...
ΔA(Obj)	6.0×10^{-2}	...	...

The noise image was also used to calculate the errors in the structural parameters. This was done by using this frame to recreate 10 realizations of the original image. The structural parameters were then recalulated for this set of realizations and an error estimate is obtained by the very simple prescription of removing the smallest one and the largest one. According to Chebyshev's inequality, this produces an interval whose upper limit is 2.3–3.0σ. The actual value depends on the assumed underlying distribution. Here, the error distribution is assumed to be Gaussian for all structural parameters, and robust estimates of the 1σ uncertainty are obtained by doubling the Chebyshev error estimate.

The first term of the *A* index was also calculated for the GALFIT model. This was used to estimate the systematic uncertainties in the asymmetries of the real galaxies, since they should be zero in the model image. Deviations from zero thus reflect inaccuracies in the minimizing algorithm. These deviations amount to less than 0.05 in almost all cases.

Table 3 presents the typical errors in all the derived parameters. Although each object should have its own error for all the derived quantities, Table 3 gathers the median error for all the sources. It is seen that the typical errors are very small. This is caused by the fact that the Kron apertures deployed by SExtractor include of the order of 10^4 pixels, and hardly ever less than 10^3 pixels. The expected errors in the structural parameters are therefore very small. The table entries above the horizontal line in Table 3 are the statistical uncertainties. The table entries below the horizontal line gather the systematic uncertainties for the asymmetry indices, derived by calculating the asymmetry indices for the GALFIT models.

4 MERGER IDENTIFICATION USING STRUCTURAL PARAMETERS OF THE RESIDUALS

As it was mentioned in Section 1, the merger fraction is defined as the fraction of galaxies with an ongoing merger episode that is found in any given sample, which is often selected as a mass-limited one. This is a fundamental issue to galaxy evolution studies and therefore the automated identification of mergers is of great importance. This section presents the main contribution from this paper. It shows that the use of the structural parameters *of the residual images* allows us to identify mergers like the use of the structural

parameters of the galaxies themselves. It also shows that the merger samples obtained using the properties of the residual images are of better or comparable statistical quality than the samples that are culled using the morphological information of the original images. The precise definition of the statistical quality of a sample, which is given in Section 4.1, is used together with the set of galaxies defined in Section 2.1, which were morphologically classified by the STAGES team. This key ingredient allows us to obtain a sample of visually detected mergers, which will be used as a training set for the method presented in this section. Many studies have used the morphological or structural properties of galaxies to estimate the merger fraction. Usually, these works make use of the CAS system or of the G–M_{20} system. Each of these systems has its own pros and cons. For instance, Conselice et al. (2008) concludes that the G–M_{20} system discovers more mergers than the CAS methodology, although it also picks up more interlopers. It is, however, the case that all the structural approaches that have been presented so far have only taken advantage of the structural parameters of the real, direct images of galaxies. In addition, the CAS system typically recovers a fairly high fraction (50–70 per cent in Jogee et al. 2009; Heiderman et al. 2009) of visually classified mergers, but it is also significantly contaminated by dusty, highly inclined non-interacting galaxies. The latter galaxies have low-level asymmetries caused by star formation episodes, as was noted in Jogee et al. (2009).

4.1 Statistical quality of samples

In science, one often confronts the problem of finding an algorithm or method to select a sample of items from a larger parent population with the condition that the selected items have to satisfy some requirements of scientific interest. However, one rarely has a mechanism to retrieve *all* the items in the parent population that satisfy the needed requirements, and it is also very unlikely that the method is able to retrieve the required items *only*. One is then forced to speak about the sensitivity and specificity of the selection process.

The sensitivity, also known as the recall ratio, is defined as

$$r = \frac{\#\text{True Positives}}{\#\text{True Positives} + \#\text{False Negatives}}. \quad (11)$$

This is more commonly known as the completeness in the astronomical literature.

The specificity is defined as

$$p = \frac{\#\text{True Negatives}}{\#\text{True Negatives} + \#\text{False Positives}}. \quad (12)$$

In the above definitions, a 'True Positive' is a recovered item that did indeed present the required properties. A 'False Negative' is an item that was not retrieved by the culling algorithm but did present the needed properties. The latter errors usually reflect an excessive skepticism. A 'True Negative' is an item that was rightfully rejected by the selection process since it did not have the required properties. A 'False Positive' is an item that was incorrectly picked up by the sampling algorithm, but that does not have the properties of interest.

The sensitivity and the specificity can be combined into a single number, known as the F score, F_β (van Rijsbergen 1979). This is a measure of sample purity, and it is just a weighted harmonic average of r and p:

$$F_\beta = \frac{(1+\beta^2) \times p \times r}{(\beta^2 \times p + r)}, \quad (13)$$

where β is a control parameter that regulates the relative importance of r with respect to p. This is a user-supplied value that depends on the particular goals of the test.[7] In this work, a value of $\beta = 1.25$ is used, which can be thought of as weighing completeness more than the *lack of* contamination. The use of this value will be justified in Section 4.4. This choice leads to a galaxy sample that contains most mergers from that training set, although the corresponding contamination is rather high. This will be further discussed in Section 5.1.

The F score is used in this study in order to grade the performance of a number of merger diagnostics at separating a merger sample from its parent population. Galaxies undergoing a merger episode play the role of the 'items presenting the required properties' discussed above, and the parent population used here is of course the galaxy sample defined in Section 2.1.

In this context, a 'merger diagnostic' is defined as a two-dimensional diagram in which the parent population of galaxies described in Section 2.1 is presented. In these plots, the horizontal axis is one structural parameter and the vertical axis is another morphological parameter, both selected from the set of indices described in Section 3.3. In these diagrams, merger galaxies should preferentially occupy specific regions. For instance, mergers should have large asymmetries and higher than average values of G. This is exploited by searching for the *best* border that separates mergers from other galaxies in each of these diagrams. The 'border' of a diagnostic is defined as a second-order polynomial in the horizontal coordinate that maps to the vertical coordinate.[8] Although it would be possible to use the same technique with more than two parameters, this would make it more difficult to interpret the resulting three-dimensional parameter spaces. For this reason, the merger diagnostics considered in this work are simply two dimensional.

Galaxies are then classified into four different types, depending on the side of the border in which they fall and on whether they are involved in a merger or not:

(i) Mergers that fall in the merger side of the border are the 'True Positives'.

(ii) Mergers that do not fall in the merger side of the border are called 'False Negatives'.

(iii) Non-mergers that, however, fall in the merger side of the border are regarded as 'False Positives'.

(iv) Non-mergers that do not fall in the merger side of the border are of course 'True Negatives'.

In all the above definitions, the merger side of the border is to be understood as the zone in the diagram in which the majority of mergers exist.

The *best* border is then the border that maximizes the F score. In this step, mergers serve as buoys or perhaps better as a training set in order to find the *best* border. This maximization is done by means of the Amoeba algorithm explained in Press et al. (1988), using the polynomial coefficients as the problem parameters.

The method is schematically presented in Fig. 1. Both axes represent a dummy structural parameter, which could be any one of the structural parameters defined in Section 3.3. The large dots represent merging systems and the small dots are non-mergers. The thick line is a *best* border found by the F-score maximization algorithm.

[7] The reader can probably work out what is the value of β used by the managers of airport security screenings, where False Positives represent a small additional test but False Negatives have disastrous consequences.

[8] The G–M_{20} method, in which galaxies with too high values of G for their M_{20} values are classified as mergers, is a basic example of this approach.

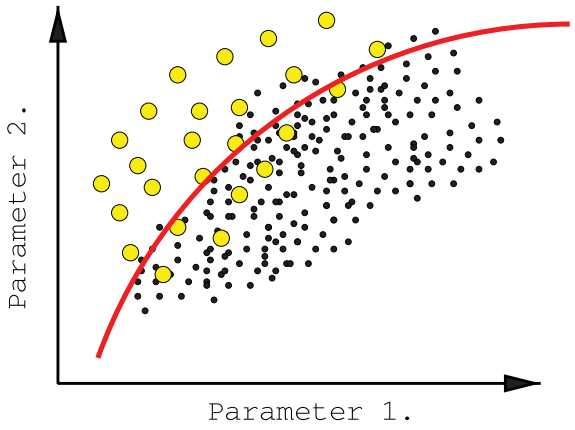

Figure 1. Schematic graphical explanation of the F-score method used. The horizontal axis is one of the structural parameters calculated either for the original images or for the *residual* images. The vertical axis is another morphological parameter, regardless of its nature. The larger dots represent the galaxies that were marked as mergers in the STAGES morphological catalogue, and the smaller dots are galaxies that were not classified as mergers. The thick line is then the border, which is fine-tuned by the Amoeba algorithm so that most of the large dots lie above it, while at the same time, most of the small points are located underneath it. In this representation, the merger side of the border is clearly the zone of the diagram above the border.

This is a second-order polynomial in the `Parameter #1` that maps into the `Parameter #2` dimension.

Section 4.3 presents this optimization for a number of merger diagnostics. Some of the merger diagnostics presented are taken from the literature, while others are introduced there for the first time. The optimization process used tries to maximize the F-score number, which is derived from the r and p statistics. A meaningful assessment of the resulting contamination requires an accurate estimate of the merger fraction in the galaxy sample used here. The contamination ratios for the best performing diagnostics are presented in Section 5.1.

4.2 The training set for the F-score technique

The next question is then what objects from the parent population of galaxies are to be taken as *true mergers* and therefore used as the training set for the F-score maximization technique presented in Section 4.3. To this end, the morphological information presented in the STAGES morphological catalogue is used.

For the purposes of this study, galaxies classified as mergers by at least two of the visual observers of the STAGES team are considered as mergers. This subset includes 39 objects and constitutes the training set. On the other hand, sources that were classified as mergers by one or none of the STAGES observers were then considered as non-mergers. This subset is made of 1498 sources. This subset includes 83 sources that were regarded as mergers by only one of the STAGES visual classifiers. These objects were not included in the training set in order to obtain a more robust merger training sample.

Fig. 2 is an image atlas showing images of all the galaxies from this training set. This figure contains both the angular and physical scales of each inset, the STAGES `ID` of each source, the environment (cluster 'C', and field 'F'), and the morphological type (elliptical 'E', lenticular 'S0', spiral 'Sp' and 'Oth'). Each inset also includes the number of observers that agreed on whether that particular galaxy is a merger or not.

Fig. 2 shows that most of the 39 galaxies used as the training set are indeed bona fide mergers. It is also true that a small fraction might be more marginal. The latter galaxies need to be included in the merger set because one of the main goals of this work is to explore whether the analysis of the structural parameters of the residuals can be used to detect minor mergers reliably.

4.3 Comparison between merger diagnostics

Given the wealth of structural parameters that have been calculated for the galaxies in the parent population, it is better to start presenting how the F-score method works for the G(Obj)–M_{20}(Obj) plane. This is a very well known merger diagnostic that has already been used in the literature. In this subsection, the image for which the structural parameters are calculated is denoted inside parentheses after the name of the structural parameter itself. 'Obj' means that the morphological parameter was obtained in the original image, 'MDL' refers to the Sérsic model, and 'Res' implies that the parameter was obtained for the residual image.

Fig. 3 shows the G(Obj)–M_{20}(Obj) plane used as a merger diagnostic. The large symbols are the galaxies that were marked as mergers by the STAGES observers,[9] and the smaller symbols show the galaxies that were not regarded as mergers. The cyan-filled circles represent irregular objects, blue squares give the location of spiral galaxies, black triangles represent lenticular systems and red diamonds denote elliptical galaxies. This panel shows the 'best' border as a thick, green solid line. The black dashed line is just the initial border that the Amoeba algorithm is given to start its iterations. This is just a rough guess, given the location of the larger symbols in the diagram. The final *best* border the Amoeba algorithm obtains does not depend on the initial guess as long as this initial guess is reasonable. The polynomial that defines the *best* border is also given within the figure, together with the resulting completeness, contamination and F-score values. Here, the completeness is defined as the number of clear visual mergers above the border, divided by the total number of visual mergers (i.e. the sensitivity).

It is seen that the F-score maximization algorithm has been able to find most of the merging systems that constitute the training set. The statistical quality of this sample is $F = 0.77$. The completeness is $r = 0.79$, and the contamination by objects not classified as mergers is $1 - p = 0.28$. It is interesting to note that this method rejects most of the lenticular objects from the training sample.

The next merger diagnostic presented is the A(Obj)–RFF plane. This merger indicator is motivated in the more common A(Obj)–S(Obj) diagnostic, using RFF instead of the clumpiness since these two quantities are very similar. Fig. 4 presents this test. This is the first test that makes use of the structural parameters of the residual images. In this test, the initial border is A(Obj) $= 0.30$, simply because the asymmetry is expected to bear the highest predictive power in this test.

Fig. 4 shows that the statistical quality of the sample obtained using this criterion has improved significantly with respect to the results achieved by the G(Obj)–M_{20}(Obj) diagnostic. The sample purity is $F = 0.85$, the sensitivity is $r = 0.92$, and the specificity is $p = 0.76$. Also, Fig. 4 shows a clear correlation between A(Obj) and RFF. It might be objected that the usual limit in the asymmetry

[9] This is thus the training set the Amoeba algorithm will use to find the *best* border to separate mergers from non-mergers.

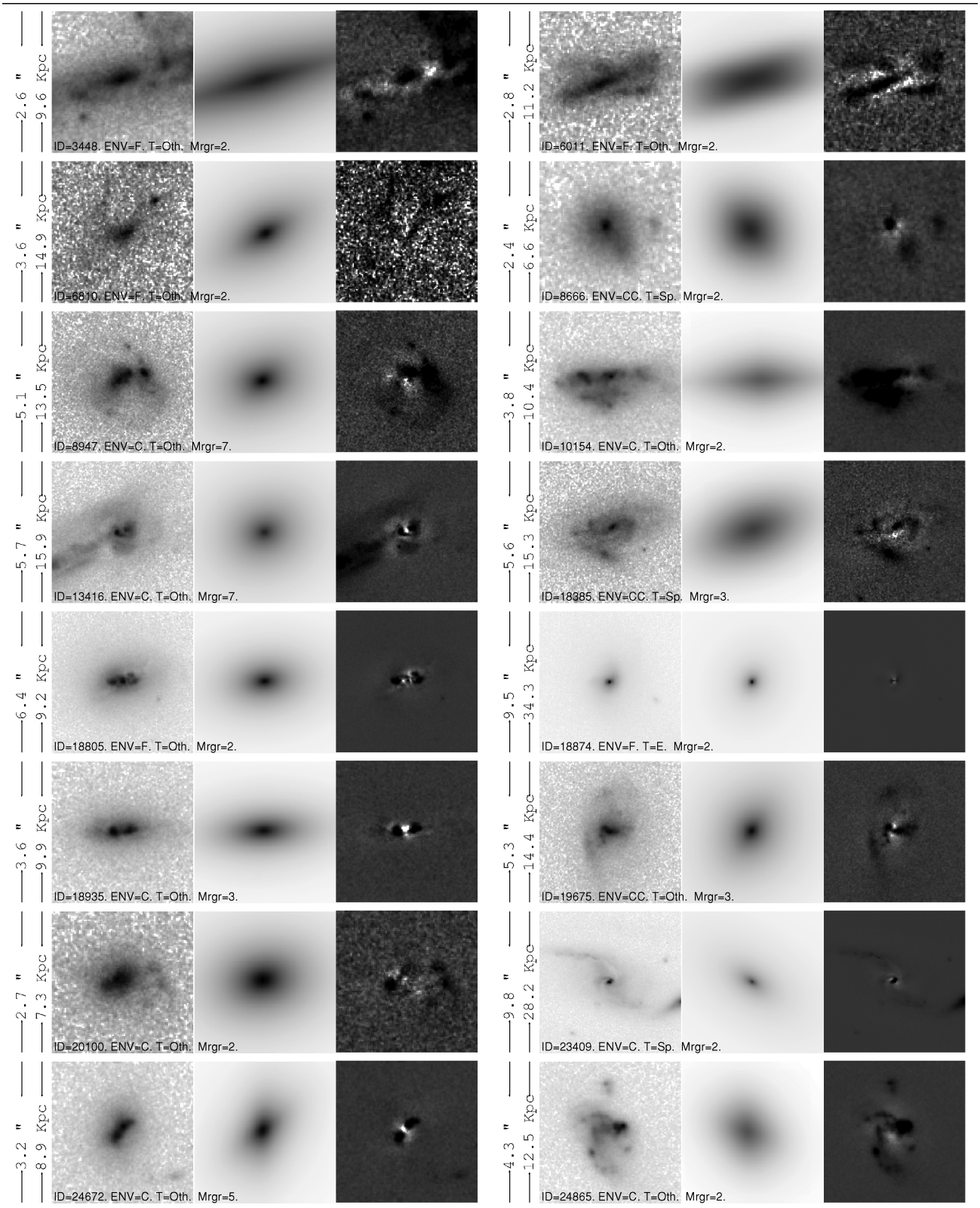

Figure 2. Image atlas of the merger training set. The COMBO-17 ID is included in each panel. This figure shows the objects which were classified as mergers by at least two of the STAGES team visual observers. Each panel shows three different insets. The first image is the direct image, the second image is the model created by GALFIT, and the third image is the residual image. The latter image is shown with an inverted look-up table whose dynamic range is 15 per cent that of the other two images. This is done in order to enhance the visibility of the fainter features in the residual images.

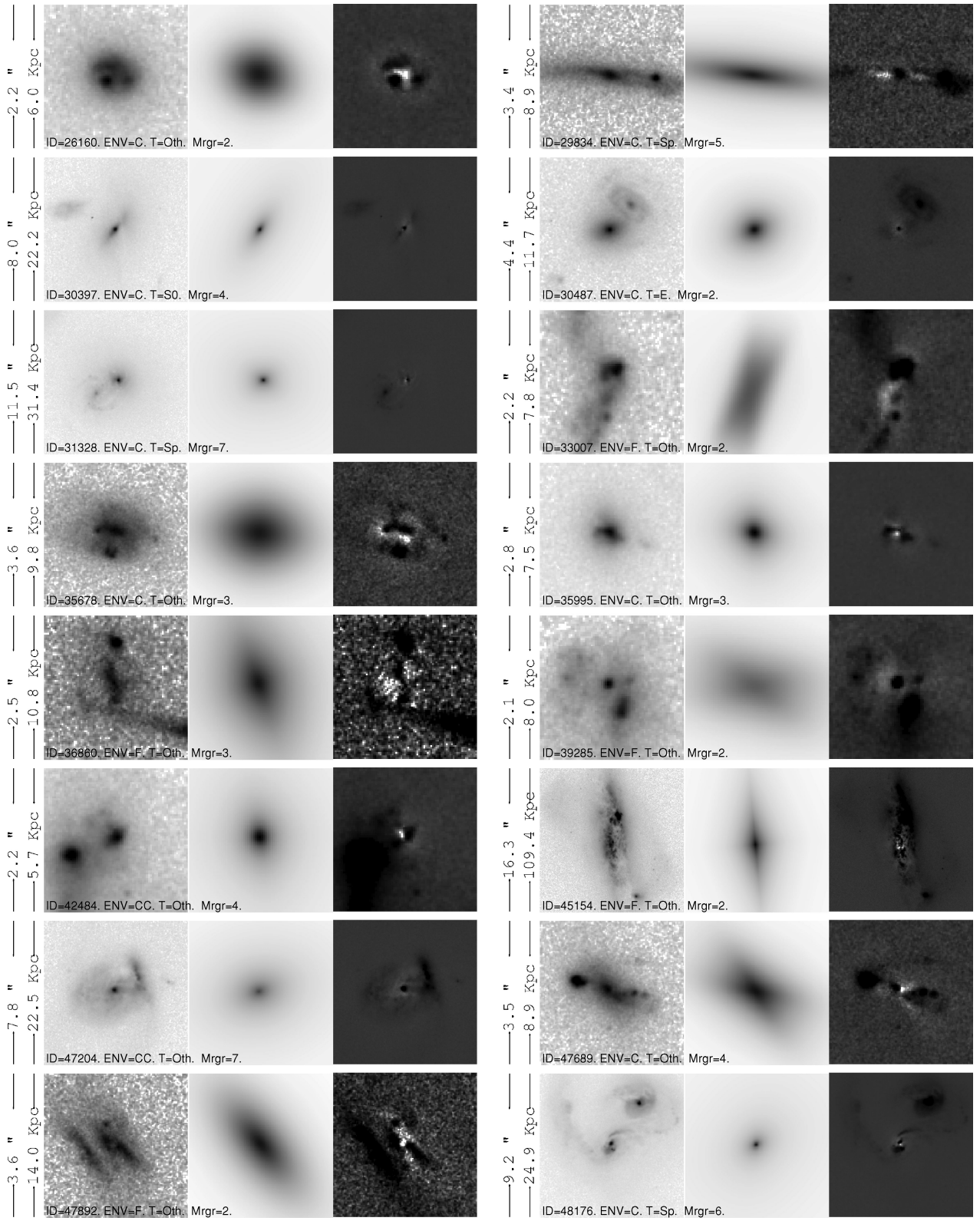

Figure 2 – *continued*

introduced in Conselice (2003) is $A(\mathrm{Obj}) > 0.35$, while the limit suggested by this test is $A(\mathrm{Obj}) > 0.20$. This is not only caused by the training sample used, but by the current choice of the β parameter, which is designed to weight completeness more than specificity. Had these conditions been different, the resulting *best* borders would have changed. This method therefore imperatively requires an objective calibration in order to produce meaningful, physically motivated borders and hence reliable merger fractions.

© 2011 The Authors, MNRAS **419,** 2703–2724
Monthly Notices of the Royal Astronomical Society © 2011 RAS

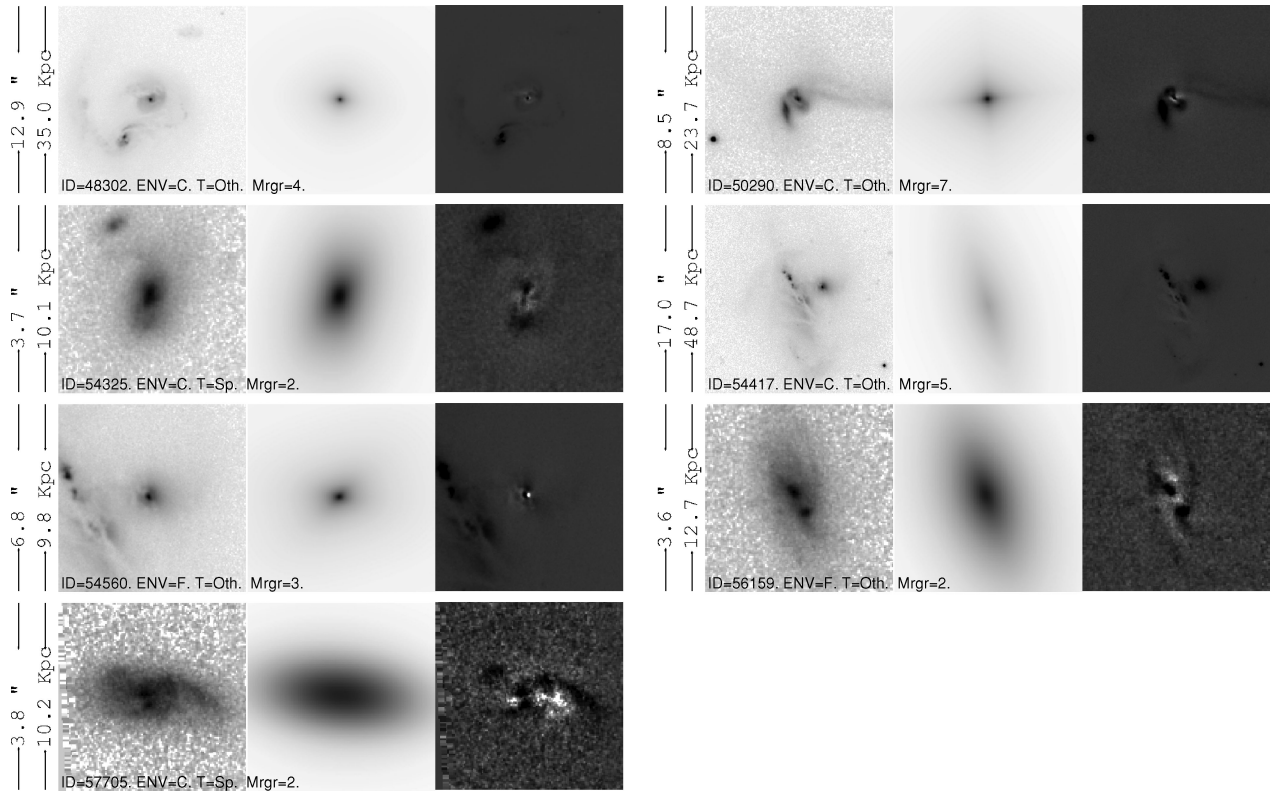

Figure 2 – *continued*

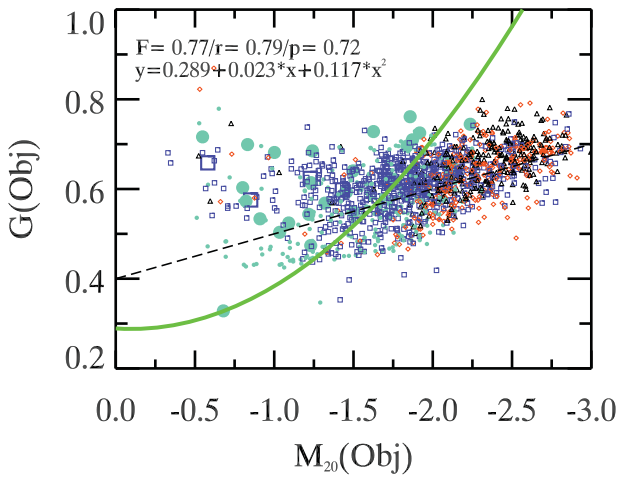

Figure 3. The well-known G(Obj)–M_{20}(Obj) merger criteria applied to the STAGES sample. The large symbols are sources that were marked as mergers by the STAGES team observers. The blue squares show spiral galaxies, black triangles denote lenticular systems and red diamonds present elliptical galaxies. Irregular and disturbed systems are represented as the beige-filled circles. The green line is the *best* border found by the Amoeba algorithm, and the black dashed line is the initial guess this algorithm is given. The optimal value of the F-score, together with the sample completeness and specificity, is given within the figure.

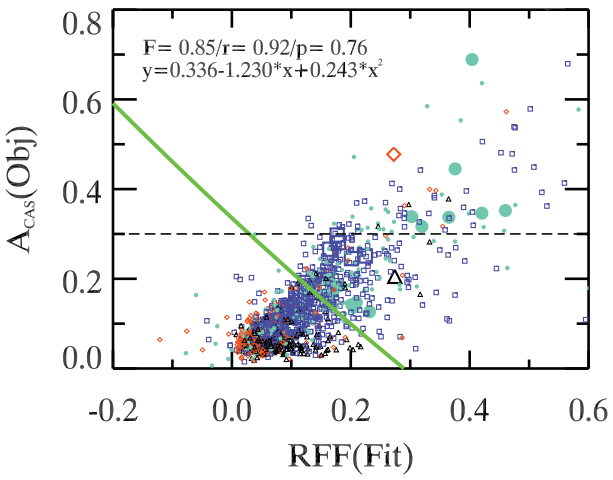

Figure 4. The A(Obj)–*RFF* merger diagnostic. Symbols and information as in Fig. 3.

This objective calibration will be produced in a forthcoming paper, using full-fledged N-body simulations of galaxy mergers. Also, it has to be borne in mind that the $A(\mathrm{Obj}) > 0.35$ criterion is tuned to detect major mergers, and one of the aims of this work is to improve the morphological detection of minor-merger episodes. Fig. 4 also shows that the F-score maximization algorithm has indeed found that the 'best' border is very different from the initial, flat guess. In particular, the Amoeba algorithm has discovered the correlation between A(Obj) and *RFF*, and takes advantage of it by converging towards a line that cuts the correlation in a perpendicular way. It is also tempting to think that the use of B+D decompositions could produce better overall fits for the S0 and Sp galaxies, leading to lower *RFF* numbers. This in turn could reduce the number of False Positives for these sources, increasing the potential of this structural parameter as a merger diagnostic. This will be explored in the future.

The results shown in Figs 3 and 4 make it possible to think that a good merger diagnostic could be put to the test by combining G(Obj) with A(Obj). Fig. 5 presents this investigation, which confirms the

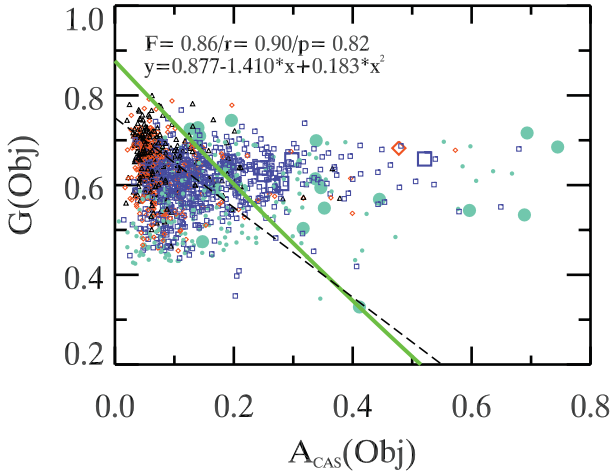

Figure 5. This figure presents the G(Obj)–A(Obj) plane used as a merger diagnostic. Symbols and information as in Fig. 3.

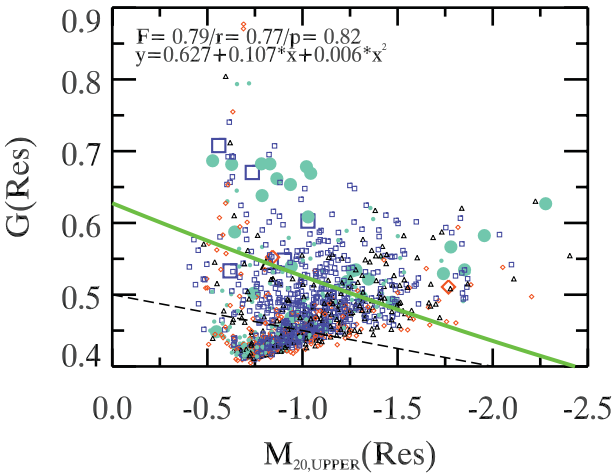

Figure 6. The G(Res)–M_{20}(Res) plane as a merger diagnostic. Symbols as in Fig. 3.

previous ideas with a sample purity of $F = 0.86$, a recall $r = 0.90$ and a fairly high specificity $p = 0.82$. The algorithm has correctly identified that the *best* border in this plane is a diagonal, as expected. The final border is found to be very close to the initial border tried by the minimization algorithm.

The following merger criterion considered continues with the exploration of the morphological parameters of the residual images. Fig. 6 shows the G(Res)–M_{20}(Res) merger test plane. This diagnostic was motivated by two main ideas:

(i) If the galaxy that was fitted and removed by the smooth Sérsic model was indeed involved in a merger episode in a very late stage, then the residual image should expose the effect of the fainter component as a bump. This bump will then be easily detected in G(Res) because it will form a reduced number of high-intensity pixels, surrounded by a large number of pixels with very low intensity values with an average value of 0.0. This will boost the value of G(Res). If, on the other hand, the galaxy is well described by the smooth model, then the pixel intensity of the residuals will all cluster around 0.0, and G(Res) will be very close to 0.5.

(ii) In the case of the horizontal axis, which is M_{20}(Res), the situation is very similar. If there is an off-centre bump in the residual image, it will probably be among the 20 per cent brightest pixels of the residual image. This will enhance the value of M_{20} with respect to the situation in which no substructure is seen in the residuals. In the latter situation, the brightest 20 per cent pixels of the residuals will be randomly distributed in the image, albeit with a preference for the central values.

Fig. 6 confirms all these expectations. It is seen that the sample purity is $F = 0.79$, with $r = 0.77$. The specificity is $p = 0.82$. It is also clearly seen that the predicting power of this diagnostic is mostly associated with G(Res). The results of this test suggest the use of the G(Res)–A(Obj) plane as a merger diagnostic. This idea is presented in Fig. 7. The correlation between the two quantities presented is evident from this figure. This alone indicates that G(Res) is bound to be a good merger tracer, even when used by itself. For this test, the sample purity is $F = 0.85$, with a good completeness $r = 0.90$ and a better specificity $p = 0.78$. This merger indicator is then comparable to the G(Obj) with A(Obj) indicator.

The next natural merger criterion that the conclusions drawn from Figs 6 and 4 lead to explore is the G(Res)–*RFF* plane, shown in Fig. 8. These two parameters have been selected because *RFF* correlates with A(Obj). It is therefore expected that this indicator will also be able to separate mergers from other galaxies. G(Res) was used because Fig. 6 shows it could single out mergers almost as

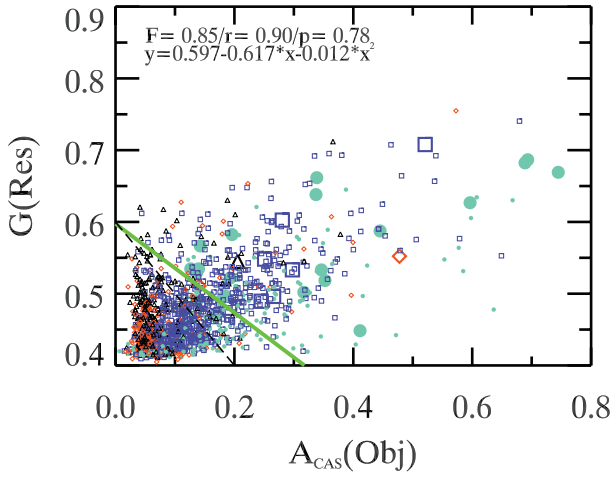

Figure 7. The G(Res)–A(Obj) plane as a merger diagnostic. Symbols as in Fig. 3.

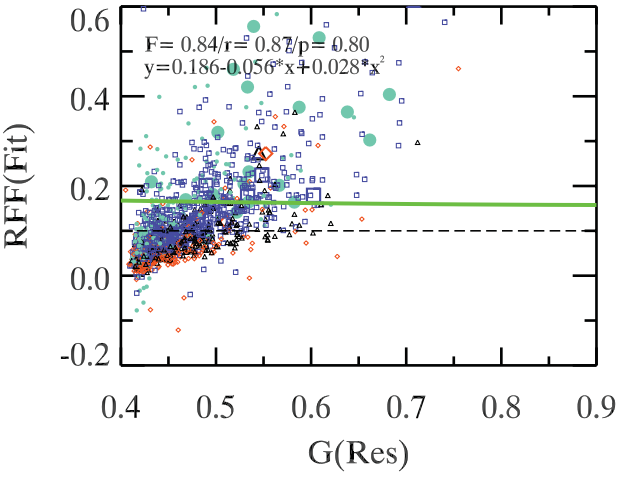

Figure 8. The *RFF*–G(Res) plane used as a merger criterion. Symbols as in Fig. 3.

a stand-alone indicator. Fig. 8 shows that *RFF* is indeed correlated with *G*(Res), and that this correlation could be exploited to identify mergers. The sample quality is $F = 0.84$, with a specificity $p = 0.80$. The completeness is $r = 0.87$. It is, however, the case that the *best* border in Fig. 8 is approximately horizontal, indicating that the Amoeba algorithm does not take advantage of the correlation between *RFF* and the Gini index of the residuals.

This work would not be complete without studying the CAS parameter space. Fig. 9 explores the *A*(Obj)–*S*(MDL), *A*(Obj)–*C*(MDL) and *A*(Obj)–*C*(Obj) planes. In all the panels in Fig. 9, the vertical axis is *A*(Obj). However, the horizontal axis of the upper panel is the clumpiness *S*(Obj), the horizontal axis of the middle panel is *C*(MDL) and the horizontal axis of the lower panel is *C*(Obj). In all these cases, the initial guess line is obviously motivated by the classical $A(\mathrm{Obj}) > 0.35$, $A(\mathrm{Obj}) > S(\mathrm{Obj})$ criterion. It is seen that the classical CAS criterion works very well, achieving a sample purity $F = 0.86$, with a good recall $r = 0.95$ and a fairly high specificity $p = 0.76$. It is also seen that the Amoeba algorithm has found that the clumpier mergers also need to be more asymmetric in order to be classified as such. Comparison between the *A*(Obj)–*C*(MDL) and *A*(Obj)–*C*(Obj) planes highlights that the use of a smooth model to calculate *C*(MDL) is slightly beneficial for merger detection. This comes from the higher sample purity obtained with the *A*(Obj)–*C*(MDL) plane. For the latter test, $F = 0.82$, with a specificity $p = 0.68$. The completeness is $r = 0.95$. This is clearly a success of the *A*(Obj) parameter. The contamination rate obtained by the use of the *A*(Obj)–*S*(Obj) plane, as presented here, will be shown in Section 5.1.

It is interesting to complete the analysis of the CAS parameter space by deriving the values of the *r*, *p* and *F* parameters for the $A(\mathrm{Obj}) \geq 0.35$, $A(\mathrm{Obj}) > S(\mathrm{Obj})$ criterion, which is a frequently used criterion for CAS-based major-merger studies. Using these limits, $r = 0.21$, $p = 0.98$ and $F_\beta = 0.30$. It is clear that with this criterion the merger sample obtained discards the majority of non-mergers, but is very incomplete. The contamination of the latter sample will be presented in Section 5.1, but it has to be borne in mind that the optimization used in this work aims towards completeness, while the goal of the $A(\mathrm{Obj}) \geq 0.35$, $A(\mathrm{Obj}) > S(\mathrm{Obj})$ limits is a clean sample of major mergers. The two diagnostics cannot be compared directly, then.

The results presented so far suggest that asymmetry is a very good merger indicator. It is also the case that *G*(Res) and *RFF* can also be used as merger indicators by themselves. In the spirit of this work, the planes studied will include the asymmetry *of the residuals* *A*(Res). This morphological parameter is explored because the asymmetry of a system is expected to be boosted after the subtraction of an intrinsically symmetric profile such as the Sérsic profile used here. Fig. 10 shows the success of this approach, in particular, in its lower panel.

Both diagnostics shown in Fig. 10 are found to work very well. In particular, the *RFF*–*A*(Res) diagnostic produces the best sample purity of all the diagnostics tested here that exploit the structural parameters of the residuals, with a very high completeness. As expected, the asymmetry of the residuals is boosted with respect to the asymmetry of the original images. The *best* border for the *RFF*–*A*(Res) diagnostic is no longer horizontal, implying that *A*(Res) has enough predicting power so as to curve the *best* line. The *best* border line of the *RFF*–*A*(Res) test appears to be more curved than that of the *G*(Res)–*A*(Res) merger test. This is consistent with the previous conclusion that *G*(Res) could also be used as a stand-alone merger diagnostic. The purity of the sample obtained using the *G*(Res)–*A*(Res) indicator has $F = 0.84$, with a recall $r = 0.90$ and a specificity

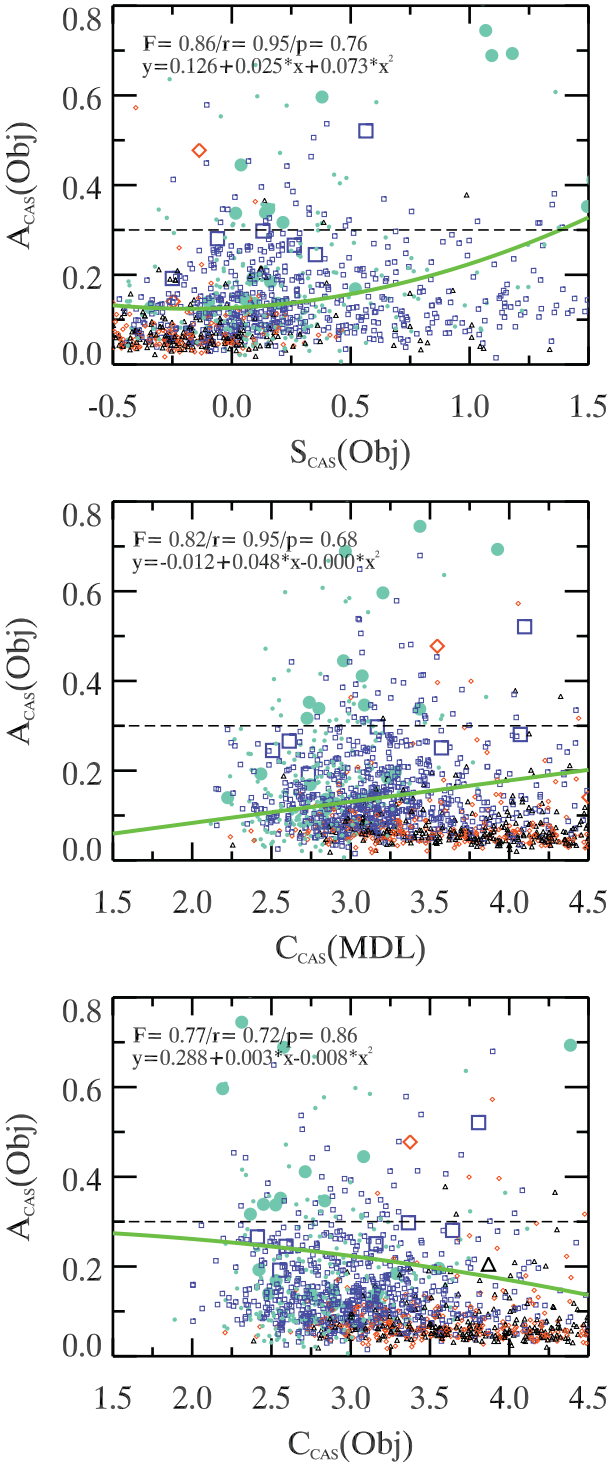

Figure 9. Upper panel: the *A*(Obj)–*S*(Obj) plane as a merger diagnostic. Middle panel: the *A*(Obj)–*C*(Mdl) indicator. Lower panel: the *A*(Obj)–*C*(Obj) merger test. Symbols as in Fig. 3.

$p = 0.78$. The purity of the sample obtained using the *RFF*–*A*(Res) indicator has $F = 0.86$, with $r = 0.92$ and $p = 0.78$. This is then the best test, and this confirms that the structural parameters of the residuals can indeed be competitive if used as merger diagnostics. The corresponding contamination stemming from the *RFF*–*A*(Res) diagnostic will be shown in Section 5.1, where a sample of visually classified mergers will be used to establish the merger prevalence.

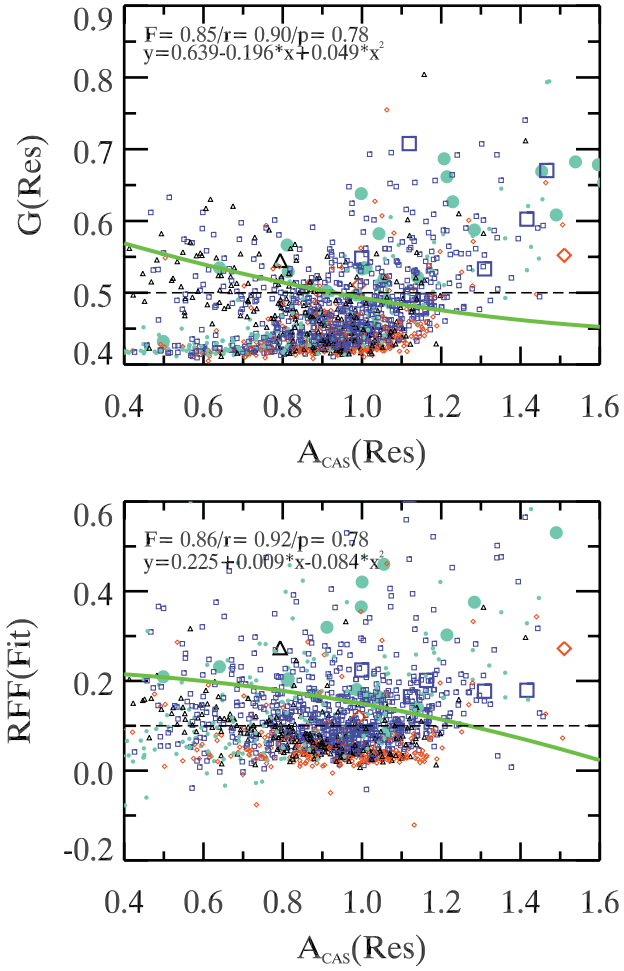

Figure 10. Upper panel: the *A*(Res)–*G*(Res) merger test. Lower panel: the *A*(Res)–*RFF* plane as a merger diagnostic. Symbols as in Fig. 3.

Whether the use of these parameters can be used to probe deeper into the luminosity function in order to detect mergers with mass ratios *larger* than 10:1 or if these parameters allow us to trace the merger event up to later stages in which the less luminous galaxy is almost engulfed by the host will be presented in a forthcoming paper.

The quality of the merger samples obtained using the *RFF*–*A*(Res) test is thus the best one of all the residual-based merger diagnostics that have been explored in this work. The discussion presented in Section 5 will be based on the merger sample obtained using this merger test. Also, Section 5.2 shows some objects that are thought to be involved in minor-merger events that have been detected in the *RFF*–*A*(Res) plane as mergers that would have been missed by the usual CAS criterion $A(\mathrm{Obj}) \geq 0.35, A(\mathrm{Obj}) > S(\mathrm{Obj})$.

Table 4 gathers a summary of the different merger diagnostics used and their respective statistical performances, for convenience.

4.4 The effect of the β parameter

The *F*-score maximization technique used in this work is simply a way to select a number of objects from a parent population whose structural properties are similar to those of the training set. The method is thus only as good as the training set. The β parameter in the definition of F_β is the ingredient used by this statistic to decide whether or not the structural properties of an object are close enough to the structural properties of the objects in the training set to be considered as a merger candidate. The β parameter is therefore determined by the scientific needs of the sampling process. Fig. 11 presents a recalculation of the 'best' border of the *A*(Res)–*RFF* diagnostic using two different values of β, $\beta = 2.0$ (upper panel) and $\beta = 0.5$ (lower panel).

Comparison of Fig. 11 with Fig. 10 clearly indicates that the β parameter has a decisive impact on the sample of potential mergers obtained by the Amoeba algorithm. If completeness is considered

Table 4. Summary of the different merger diagnostics tried and their statistical performance.

Diagnostic	Figure	*F* score	*r*	*p*
G(Obj)–M_{20}(Obj)	3	0.77	0.79	0.82
A(Obj)–*RFF*	4	0.85	0.92	0.76
G(Obj)–*A*(Obj)	5	0.86	0.90	0.82
G(Res)–M_{20}(Res)	6	0.79	0.77	0.82
G(Res)–*A*(Obj)	7	0.85	0.90	0.78
RFF–*G*(Res)	8	0.84	0.87	0.80
A(Obj)–*S*(Obj)	9	0.86	0.95	0.76
A(Obj)–*C*(Mdl)	9	0.82	0.95	0.68
A(Obj)–*C*(Obj)	9	0.77	0.72	0.86
A(Res)–*G*(Res)	10	0.85	0.90	0.78
A(Res)–*RFF*	10	0.86	0.92	0.78

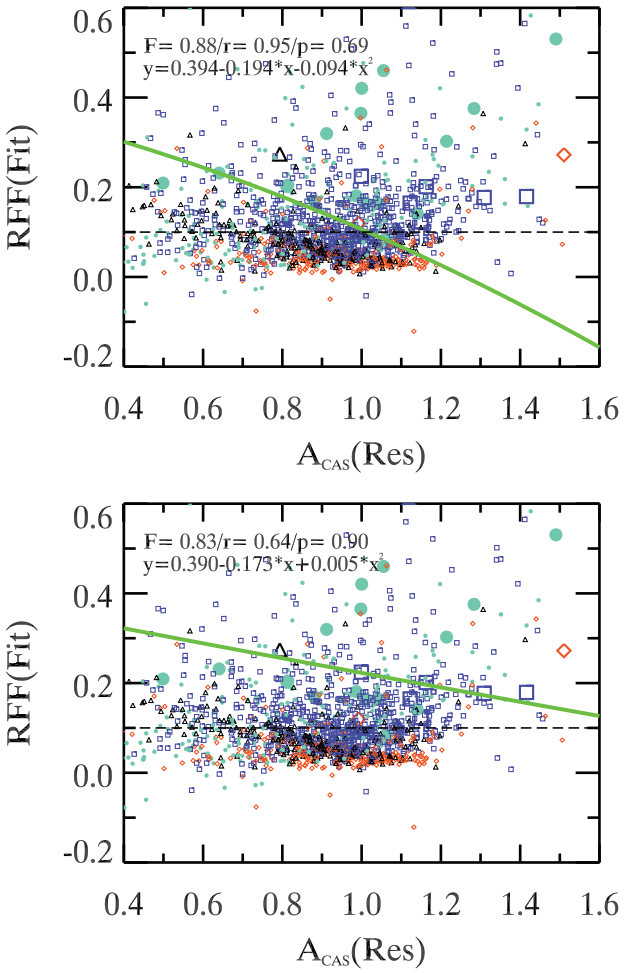

Figure 11. Re-calculation of the *A*(Res)–*RFF* merger diagnostic using $\beta = 2.0$ (upper panel) and $\beta = 0.5$ (lower panel). Symbols as in Fig. 3.

to be much more important than specificity, as is done in the upper panel of Fig. 11, the Amoeba algorithm increases the recall ratio by 3 per cent with respect to the value achieved in Fig. 10. This amounts to the inclusion of just an additional object from the training sample, at the price of littering the sample with 150 non-mergers. This decreases the precision by 10 per cent. On the other hand, the lower panel of Fig. 11 shows what happens if specificity is weighted more than completeness. It is first seen that the new merger diagnostic excludes a higher number of non-mergers, raising the precision by 12 per cent. This comes at the expense of missing 11 objects from the training set. The choice of $\beta = 1.25$ adopted in Section 4.1 is thus a good compromise between these two options.

It is highlighted here that no single value of β can be considered to be 'correct'. The value of this parameter is set by the goals of the test. If, for instance, the objects selected by the method are to be the targets of a spectroscopic follow-up programme where targeting only true mergers is deemed essential, then it would be advisable to give a higher weight to specificity.

5 VISUAL ASSESSMENT OF THE MERGER SAMPLES OBTAINED

As it is clear from the previous discussion, the method presented in this work is optimized to be very complete, hopefully detecting minor mergers, thanks to the use of the structural parameters of the residual images. It is therefore needed to establish both the contamination by non-mergers and the success in recovering minor mergers. The study of the contamination by non-mergers is gathered in Section 5.1. On the other hand, Section 5.2 presents a number of examples of minor-merger candidates that have been detected through the use of the structural parameters of the residuals that would otherwise have been missed by the traditional CAS diagnostics.

5.1 Contamination by non-mergers

The different galaxy samples obtained by the *sheer* use of the various indicators that have been put to the test in Section 4.3 need to be further evaluated. It is clear that not all the objects that test positive (i.e. sources that fell in the merger side of the merger tests used) to these criteria can be mergers. In particular, contamination is the key statistic that needs to be included in this appraisal. This section presents the non-merger contamination of the galaxy sets obtained from the use of the merger diagnostics that attained the highest F-score values in Section 4.3.

The contamination is defined as

$$C = \frac{\#\text{Non-mergers that, however, test positive}}{\#\text{All positives.}}, \qquad (14)$$

where the objects that test positive are those in the merger side of the 'best' borders calculated above. The denominator of this fraction includes both mergers and non-mergers, which implies that the contamination depends on the merger fraction in a non-linear way.

The next step is thus oriented towards obtaining a complete and accurate estimate of the fraction of mergers found in the parent population studied. The total number of good mergers is here calculated by adding together the objects included in the training set of galaxies used above and a number of merger galaxies that were recovered during an additional visual classification which will be described below. This further observational classification is justified because the STAGES morphological catalogue used to define the training set of objects for the F-score maximization technique is a general morphological catalogue that is not designed to split the parent population into mergers and non-mergers. In particular, the STAGES observers were not specifically looking for the *minor* mergers whose detection is the goal of this study. This additional study thus serves to check whether minor mergers do share the structural properties of major mergers, which is the main assumption behind the use of the F-score number as a diagnostic discriminator.

The new visual assessment examines the set of galaxies obtained by the blind application of the A(Res)–*RFF* diagnostic as shown in Fig. 10. This merger test was shown in Section 4.3 to yield the highest F score and specificity numbers of all the diagnostics that make use of the structural parameters of the *residual* images and is therefore more likely to produce a clean list of mergers. This set of galaxies is made of $36 = 0.92 \times 39$ objects from the original training set and $332 = (1 - 0.78) \times 1498$ objects that were not included in the training set. The latter systems are the False Positives involved in the calculation of the p statistic. The majority of these objects (282) have zero merger marks in the STAGES morphological catalogue and 50 of them have only one merger mark. These simple statistics motivated us to examine more closely the 332 objects not included in the original training set that, however, fell in the merger side of the 'best' border in the A(Res)–*RFF* plane. The purpose of this further investigation is to establish whether or not those objects could be mergers that escaped the original assessment of the STAGES team observers. This will allow an accurate and non-parametric determination of the merger prevalence in the parent population of galaxies studied. To this end, four of the authors of this paper (AA-S, CH, EFB and MEG) re-inspected the 332 False Positives together with 332 randomly selected True Negatives as a control sample. These four independent assessments were then combined into a single trinary outcome, splitting these 664 sources into three different sets: (i) clear mergers; (ii) clear non-mergers; and (iii) the dubious cases. These 664 objects were then placed in the A(Res)–*RFF* plane. This is presented in Fig. 12.

Fig. 12 immediately shows a correlation between the location of the mergers that have been recovered by the new visual merger re-assessment and the *RFF*–A(Res) 'best' border. It is seen that the majority of the recovered mergers and dubious mergers lie above

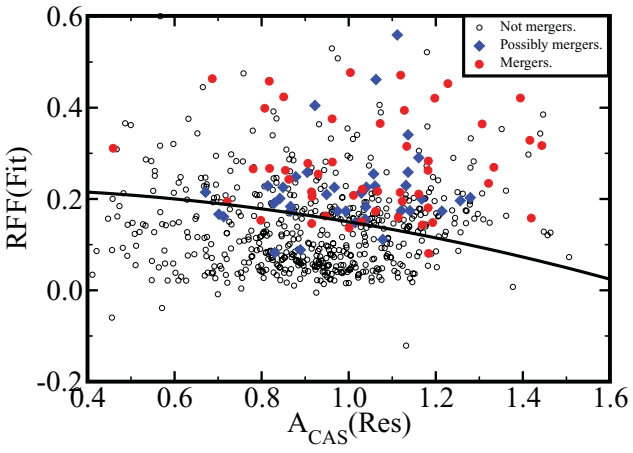

Figure 12. Visual assessment of the galaxy sample obtained using the *RFF*–A(Res) merger test. The red solid dots represent the locations of the new best merger cases (55 points), blue diamonds give the locations of the dubious cases (41 points), and the open black points show the locations of the non-mergers (568 points). The black line shown is the same as shown in Fig. 10.

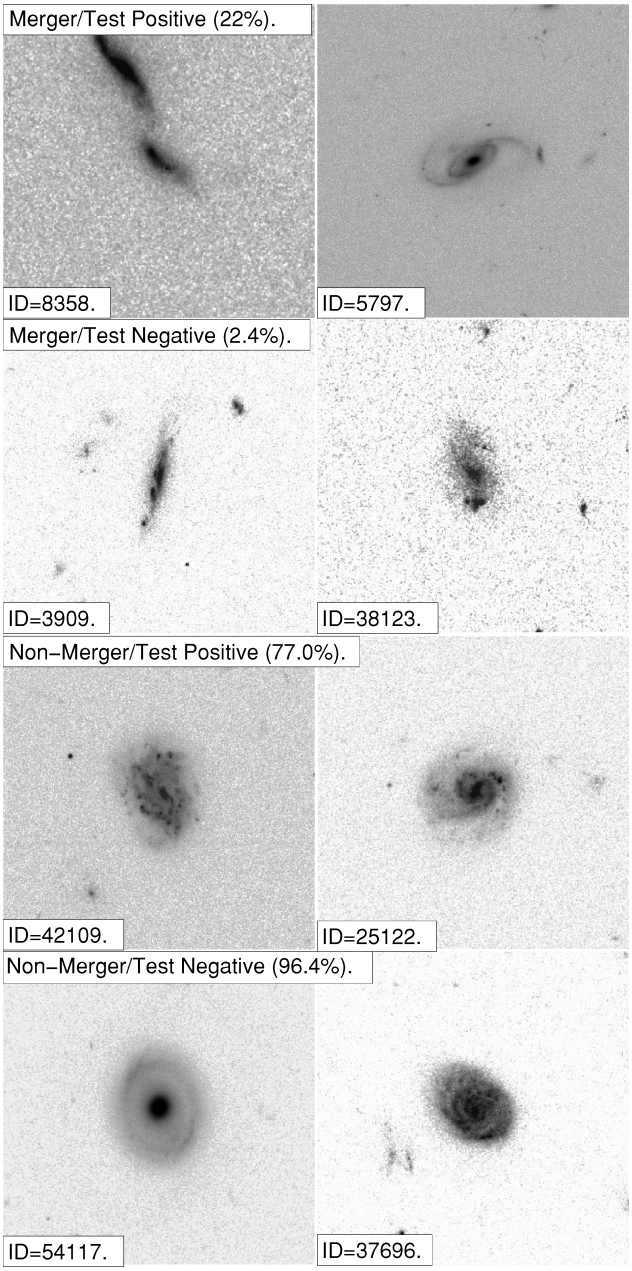

Figure 13. Examples of galaxies examined in the new visual assessment. COMBO-17 IDs are shown for each galaxy. The percentages given reflect the frequencies of the four different possibilities shown.

the diagnostic line. It is also highlighted that *none* of the objects shown in Fig. 12 belongs to the original training set.

Fig. 13 presents eight of the galaxies that were re-inspected during this further classification. This figure presents four rows of two objects each, marked with their respective COMBO-17 IDs. The upper row shows the objects that tested as positives to the *RFF*–*A*(Res) diagnostic and were subsequently classified as mergers in the new visual assessment. The second row presents visual mergers that, however, tested negative to the diagnostic. The latter objects are very rare. The third row is made of non-mergers that nevertheless tested positive in the automated test, and the bottom row presents the non-mergers that fell below the 'best' border in Fig. 10. The percentages given in each row are the frequencies with which each of the different possibilities appears. These percentages do not add up to 100 per cent because the galaxies that were classified as dubious mergers are not shown.

The new visual inspection has recovered a total of 55 clear merger systems together with 41 dubious cases. The remaining 568 objects were classified as non-mergers. It is interesting to note that 36 per cent of the 55 clear merger cases had just one merger mark in the original STAGES morphological catalogue, 18 per cent of the new possible mergers were considered as mergers by only one of the STAGES classifiers, and only 6 per cent of the non-mergers had received a single merger mark. This adds up to a total of 60 objects with only one merger mark in the whole pool of 664 objects inspected. A total of 50 sources out of these 60 objects are found above the 'best' border defined in Fig. 10 for the *RFF*–*A*(Res) merger test. In addition, the *RFF*–*A*(Res) merger test as presented in Fig. 10 detects 93 per cent of the recovered clear merger cases, 88 per cent of the new dubious mergers, and 43 per cent of the non-mergers.

Thus, the total number of objects selected by the *RFF*–*A*(Res) test can be broken down in the following way:

(i) $36 = 0.92 \times 39$ objects included from the training set.

(ii) $51 = 0.93 \times 55$ new mergers recovered by the new visual inspection.

(iii) An unknown number of merger systems between 0 and $36 = 0.88 \times 41$ that were classified as dubious mergers by the visual re-assessment. Under the assumption of a flat probability distribution, this number could be represented as 18 ± 10.

(iv) A total of approximately $260 = 332 - 57 - 19$ non-mergers.

Therefore, the total number of bona fide mergers above the *RFF*–*A*(Res) 'best' border is 105, and the total number of objects of all classes above this line is $368 = 332 + 36$. The final contamination ratio is 71 per cent $= (368 - 105)/368$, which is fairly high since it means that 70 per cent of all the objects set aside by the blind use of the diagnostic are non-merger contaminants. This number is close, but conceptually different from the fraction of non-mergers that are found to test positive during the additional visual classification, which is 74 per cent.

In the same way, the non-merger contamination associated with the *A*(Obj)–*S*(Obj) test as shown in Fig. 9 is 71 ± 3 per cent $= [397 - \{37 + (360/286) \times [49 + (0.5 \pm 0.3) \times 28]\}]/397$. This calculation takes into account that there are $397 = 39 \times 0.95 + 1498 \times (1 - 0.76)$ objects above the 'best' border defined in Fig. 9 and that not all the dubious mergers will indeed be mergers. The latter consideration is again made assuming a flat probability distribution. This contamination that affects the sample selected by this method is seen to be fully compatible with the contamination that is calculated for the *RFF*–*A*(Res) test. If a similar analysis is carried out for the *A*(Obj)–*S*(Obj) criterion using the traditional CAS limits $A(\mathrm{Obj}) \geq 0.35$, $A(\mathrm{Obj}) \geq S(\mathrm{Obj})$, the contamination is 50 per cent $= 19/(19 + 17)$. Thus, the contamination by non-mergers found in this set of galaxies is lower than the one found for the other two merger diagnostics explored, but its completeness is obviously much lower.

The next step in this analysis is then the study of the negative detections, which focuses on the 1169 sources below the 'best' border line presented in Fig. 10. These 1169 galaxies include three galaxies from the original training set (these are the 'False Negatives' in the *F*-score analysis). These 1169 galaxies will also include an indeterminate number of mergers that can be estimated from the visual re-assessment results by multiplying the fraction of mergers found in the 332 objects from the control sample by the total number of objects below the 'best' line (1169). The new visual classification discovered four clear merger and five dubious cases in the 332

objects in the control set. If the latter sources are given a weight of 0.5, the fraction of mergers below the 'best' line of the *RFF*–*A*(Res) test is 2.2 ± 0.4 per cent $= \{3 + (1169/332) \times [4 + (0.5 \pm 0.3) \times 5]\}/1169$, where the error interval again assumes a flat probability distribution for the dubious mergers. If the galaxies from the original training set and the newly identified dubious mergers are not included in this calculation, the fraction drops to 1 per cent $= [(1169/332) \times 4]/1169$, which is the number given in Fig. 13 for the negative contamination ratio of the visual re-assessment. The negative contamination ratio is then very low, indicating that this technique is *very* powerful as a negative merger test. Furthermore, if this negative contamination ratio is derived for the *A*(Obj)–*S*(Obj) test as presented in Fig. 9, the percentage is 4.0 ± 1 per cent $= \{2 + (1140/378) \times [6 + (0.5 \pm 0.3) \times 13]\}/1140$, which is compatible but slightly worse than the result for the *RFF*–*A*(Res) test. The latter calculation uses that there are 378 objects in the total pool of objects inspected during the second visual classification that fell below the 'best' border line shown in the upper panel of Fig. 9. Finally, the corresponding negative contamination for the traditional CAS criterion is 5 per cent $= 77/(1424 + 77)$. This is compatible with the negative contamination for this diagnostic as presented in Fig. 9.

In summary, despite the fact that the contamination ratio is fairly high for the positive detections, it is very low for the negative detections. The above considerations therefore lead us to conclude that the *RFF*–*A*(Res) *minor*-merger diagnostic presented in Fig. 10 works best as a *negative* test. In particular, it could be possible to use this automated technique with large-area surveys such as the APM (Maddox et al. 1990), 2dFGRS (Folkes et al. 1999), SDSS (Abazajian et al. 2009), UKIDSS (Dye et al. 2006), KIDS,[10] VIKING,[11] GEMS (Rix et al. 2004), COSMOS (Koekemoer & Scoville 2005), DES[12] or LSST (Ivezic et al. 2008) by calibrating the new 'best' border lines to the new observational conditions and wavebands using a reduced and manageable number of objects within a pre-defined redshift window up to a certain magnitude. These new 'best' border lines should then be used to split the target galaxy population. The technique illustrated here would then produce two different sets of galaxies. One set would be almost completely free of mergers and the other set would include the overwhelming majority of mergers. The latter sample would need further purification in order to produce a clean sample of mergers. This re-calibration step is particularly needed in the case of surveys including *U*-band observations. In this case, even very minor wet mergers will leave a larger impact on the general appearance of their host galaxies because the youngest stars will be much clearly seen. This step could also help understand the effect of the photometric band chosen on the optimization process used here. It is clear that the additional pruning cannot be done using the structural properties of the galaxies only, since these have been fully exploited here. In addition, one clear conclusion from this visual assessment which will be strengthened in Section 5.2 is that even a visual inspection cannot unambiguously tell whether a particular object is involved in a merger episode or is just the product of a by-chance alignment or an H II region; additional information such as colours or kinematical information obtained using Integral Field Units spectrographs needs to be included. Note that this problem affects all automated methods based on structural parameters that we have explored in this paper and are commonly used in the literature. This issue is merely a consequence of the empirical existence of visually classified mergers with undisturbed morphologies, which was also observed in Heiderman et al. (2009). The above considerations make it advisable to seek an inclusive criterion, with a high recall ratio, rather than a specific one, which is reflected in the choice of the β parameter.

[10] http://www.astro-wise.org/projects/KIDS/
[11] http://www.astro-wise.org/Public/viking10.pdf
[12] http://www.darkenergysurvey.org

5.2 The search for minor mergers

As it has been mentioned, the main driver behind the very high completeness that the method presented here is tuned to achieve is the detection of *minor* mergers. It is therefore interesting to test the ability of the methodology introduced here to detect minor mergers. This is done by selecting a set of visually identified mergers that would not have been detected by the usual cuts that are applied to the *A*(Obj)–*S*(Obj) plane, which were, however, recovered by the *RFF*–*A*(Res) diagnostic. This set of galaxies is made of 69 galaxies. As a comparison, there are 17 visually classified mergers that would have been detected by the usual CAS diagnostic.

Fig. 14 presents a number of examples taken from the 69 visually classified galaxies that would have not been retrieved by the *A*(Obj)–*S*(Obj) diagnostic using its usually adopted limits. For each of the selected galaxies, this figure presents the COMBO-17 ID, its environment, its morphological type, and its *B*-band absolute magnitude. Each inset also presents an estimate of the contribution from the less luminous component to the total `FLUX_AUTO` SExtractor measurement, expressed as a fraction. This is calculated simply by dividing the flux enclosed in the aperture shown by the automatic flux measurement that is performed by SExtractor.

Fig. 14 shows that the *RFF*–*A*(Res) method presented here has detected galaxies that do present inhomogeneities in their light distribution. These inhomogeneities typically amount to 10 per cent of the total flux received from the parent galaxy, as estimated by the crude but robust flux estimate given by SExtractor. This confirms the potential validity of the approach presented in this paper to separate minor mergers from the whole parent population of galaxies. However, Fig. 14 also makes manifest the reason for the fairly high contamination by non-mergers that has been found in Section 5.1. It is very difficult, even with the eye, to tell apart galaxies undergoing a minor-merger episode from galaxies that are experiencing other local phenomena such as star formation in H II regions. This is particularly true in the case of very late minor mergers, in which the less massive galaxies have been almost entirely dissolved amidst the larger galaxies. However, for less evolved merger systems for which the less luminous object has not entirely lost its individuality, it is easier to separate mergers and non-mergers. Specifically, objects 20213 and 7479 represent two cases of objects selected by the *RFF*–*A*(Res) diagnostic that could indeed be star formation enhancements. Finally, Fig. 14 also shows that, although the structural merger diagnostic used here can indeed select galaxies with inhomogeneous light distributions, this method alone cannot identify which light clump is to be identified as the potential satellite. This is best seen in the case of ID = 40654.

6 CONCLUSIONS

We present a new structural merger diagnostic geared towards the structural detection of minor mergers which is entirely based on the morphological properties of the *residual* images of galaxies after the subtraction of a smooth Sérsic model. The new indicator makes use of the asymmetry of the residuals and of the *RFF* of the fit, both calculated over the Kron aperture of the galaxies. This diagnostic has been objectively proven to be able to produce merger samples of

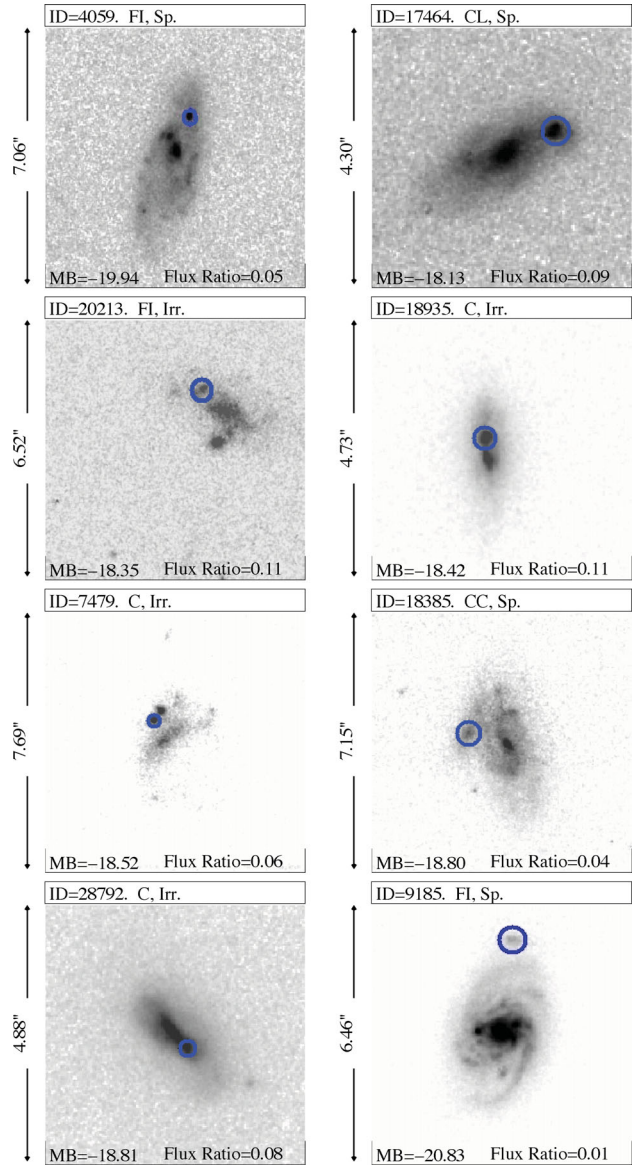

Figure 14. Visually classified minor-merger candidates found by the *RFF*–*A*(Res) diagnostic but missed by the CAS method adopting its usual limits. COMBO-17 IDs, environments, morphological types and *B*-band absolute magnitudes from the STAGES public catalogue are shown for each galaxy. The panels also give an indication of the angular extent of the image insets and the approximate fractional contribution of the light included in the circle to the total SExtractor automatic flux measurement. Objects 20213 and 7479 might be localized star formation episodes in irregular galaxies. Also, it is not straightforward to identify the putative satellite in object 40654, which presents an alternative satellite marked in red.

equal or better statistical quality than samples obtained using other well-established methods based on the morphological properties of the original images. In particular, objects with symmetric residuals for which *RFF* is larger than 0.2 or objects with more asymmetric residuals for which *RFF* is larger than 0.1 are very good candidates to be mergers. We have also found that the Gini index of the residuals could also produce merger samples of high statistical purity. In this case, objects for which the Gini index of the residual image calculated within the Kron aperture is higher than 0.5 are also good merger candidates.

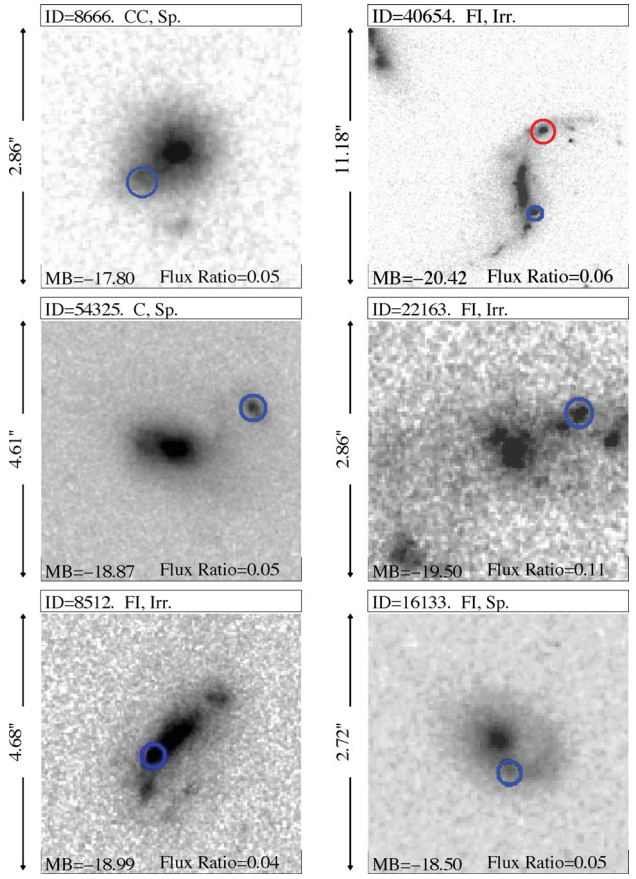

Figure 14 – *continued*

Using the structural parameters of the residuals and the limits provided by the *F*-score optimization process shown here, we have split the whole population of galaxies into two different sets of galaxies. The first set, sharing the structural trends and properties of the mergers included in the training set, is shown to contain the majority of *major* and *minor* mergers. The second set has been shown to be almost completely free of mergers, as exposed by the very low *negative* contamination rates. However, given the relative dearth of mergers among the general galaxy populations and the self-imposed goal of detecting the more elusive *minor* mergers, it turns out that the *RFF*–*A*(Res) diagnostic introduced in this paper works best as a negative merger test. In other words, it is very effective at selecting non-merging galaxies. In common with all the currently-available automatic methods, the sample of both *major*- and *minor*-merger candidates selected by our test is heavily contaminated by non-mergers, and further steps are needed to produce a clean merger sample from the first set of galaxies. Nevertheless, the methodology introduced in this paper can be very useful when applied to the huge data sets provided by modern large-area surveys such as the SDSS or UKIDSS. By visually classifying a relatively small and manageable number of galaxies, one can derive the resulting cuts and best borders which can then be applied to the whole sample. This would produce two statistically well defined sets of galaxies, a smaller one containing the vast majority of *both major* and *minor* mergers, and a much larger one almost completely devoid of them. To identify bona fide mergers, only the first set would need to be processed further to remove the non-mergers. This could be done either by visual inspection or by using additional information such as colours or three-dimensional spectra.

This work also suggests that the use of the structural parameters of the residual images could indeed be used as a tool to study the properties of minor mergers. We argue that this is due to the fact that by removing the bulk of the host galaxy light we might be able to detect much fainter merging galaxies and over much longer time-scales. This will be further studied by using *N*-body merger simulations in a forthcoming paper, which will eventually provide the way to link the structural parameters measured over the *HST*/ACS images with the underlying properties of the observed mergers.

ACKNOWLEDGMENTS

CH acknowledges financial support from the project 'Estallidos de Formación Estelar. Fase III', under the Spanish Ministerio de Ciencia e Innovación grant AYA2007-67965-C03-03. CH also thanks a Spanish MICINN postdoctoral grant. MEG acknowledges an STFC advanced Fellowship. SJ acknowledges support from the NASA LTSA grant NAG5-13063, NSF grant AST-0607748, and HST grants GO-11082 from the STScI, which is operated by the AURA, Inc., for NASA, under NASA contract NAS 5-26555.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Abraham R. G., van den Bergh S., Nair P., 2003, ApJ, 588, 218
Barnes J. E., 2002, MNRAS, 333, 481
Barnes J. E., Hernquist L. E., 1991, ApJ, 370, L65
Barnes J. E., Hernquist L., 1992, ARA&A, 30, 705
Barnes J. E., Hernquist L., 1996, ApJ, 471, 115
Bell E. F. et al., 2004, ApJ, 608, 752
Bell E. F., Phleps S., Somerville R. S., Wolf C., Borch A., Meisenheimer K., 2006, ApJ, 652, 270
Bershady M. A., Jangren A., Conselice C. J., 2000, AJ, 119, 2645
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Borch A. et al., 2006, A&A, 453, 869
Bournaud F., Jog C. J., Combes F., 2005, A&A, 437, 69
Bournaud F., Duc P.-A., Emsellem E., 2008, MNRAS, 389, L8
Buitrago F., Trujillo I., Conselice C. J., Bouwens R. J., Dickinson M., Yan H., 2008, ApJ, 687, L61
Casertano S. et al., 2000, AJ, 120, 2747
Cassata P. et al., 2005, MNRAS, 357, 903
Cenarro A. J., Trujillo I., 2009, ApJ, 696, L43
Chilingarian I. V., Di Matteo P., Combes F., Melchior A.-L., Semelin B., 2010, A&A, 518, A61
Conselice C. J., 2003, ApJS, 147, 1
Conselice C. J., 2006, ApJ, 638, 686
Conselice C. J., Rajgor S., Myers R., 2008, MNRAS, 386, 909
Conselice C. J., Yang C., Bluck A. F. L., 2009, MNRAS, 394, 1956
De Propris R., Conselice C. J., Liske J., Driver S. P., Patton D. R., Graham A. W., Allen P. D., 2007, ApJ, 666, 212
den Brok M. et al., 2011, MNRAS, 414, 3052
Dye S. et al., 2006, MNRAS, 372, 1227
Eliche-Moral M. C. et al., 2010, A&A, 519, A55
Faber S. M. et al., 2007, ApJ, 665, 265
Folkes S. et al., 1999, MNRAS, 308, 459
Giavalisco M., Ravindranath S., Daddi E., 2007, Nuovo Cimento B, 122, 1209
Graham A. W., Driver S. P., 2005, Publ. Astron. Soc. Australia, 22, 118
Gray M. E. et al., 2009, MNRAS, 393, 1275
Grazian A. et al., 2006, A&A, 449, 951
Hammer D. et al., 2010, ApJS, 191, 143
Häussler B. et al., 2007, ApJS, 172, 615
Heiderman A. et al., 2009, ApJ, 705, 1433
Hopkins P. F. et al., 2009, MNRAS, 397, 802
Hoyos C. et al., 2011, MNRAS, 411, 2439
Ivezic Z. et al. (for the LSST Collaboration), 2008, preprint (arXiv: 0805.2366)
Jogee S. et al., 2008, in José G., Funés S. J., Enrico M. C., eds, ASP Conf. Ser. Vol. 396, Formation and Evolution of Galaxy. Astron. Soc. Pac., San Francisco, p. 337
Jogee S. et al., 2009, ApJ, 697, 1971
Koekemoer A. M., Scoville N. Z., 2005, New Astron. Rev., 49, 461
Kron R. G., 1980, ApJS, 43, 305
Lane K. P., Gray M. E., Aragón-Salamanca A., Wolf C., Meisenheimer K., 2007, MNRAS, 378, 716
Lavery R. J., Remijan A., Charmandaris V., Hayes R. D., Ring A. A., 2004, ApJ, 612, 679
Lin L. et al., 2004, ApJ, 617, L9
Lin L. et al., 2008, ApJ, 681, 232
López-Sanjuan C., Balcells M., Pérez-González P. G., Barro G., García-Dabó C. E., Gallego J., Zamorano J., 2009, A&A, 501, 505
Lotz J. M., Primack J., Madau P., 2004, AJ, 128, 163
Lotz J. M., Jonsson P., Cox T. J., Primack J. R., 2008, MNRAS, 391, 1137
Maddox S. J., Efstathiou G., Sutherland W. J., Loveday J., 1990, MNRAS, 242, 43p
Maltby D. T. et al., 2010, MNRAS, 402, 282
Masjedi M. et al., 2006, ApJ, 644, 54
Papovich C., Dickinson M., Giavalisco M., Conselice C. J., Ferguson H. C., 2005, ApJ, 631, 101
Patton D. R., Carlberg R. G., Marzke R. O., Pritchet C. J., da Costa L. N., Pellegrini P. S., 2000, ApJ, 536, 153
Patton D. R. et al., 2002, ApJ, 565, 208
Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266
Petrosian V., 1976, ApJ, 209, L1
Press W. H., Flannery B. P., Teukolsy S. A., Vetterling W. T., 1988, Numerical Recipes in C: The Art of Scientific Computing. Cambridge Univ. Press, Cambridge
Rix H.-W. et al., 2004, ApJS, 152, 163
Robaina A. R., Bell E. F., van der Wel A., Somerville R. S., Skelton R. E., McIntosh D. H., Meisenheimer K., Wolf C., 2010, ApJ, 719, 844
Sérsic J. L., 1963, Bol. Asociacion Argentina Astron., 6, 41
Simard L., 1998, in Albrecht R., Hook R. N., Bushouse H. A., eds, ASP Conf. Ser. Vol. 145, Astronomical Data Analysis Software and Systems VII. Astron. Soc. Pac., San Francisco, p. 108
Spitzer L., Jr, Baade W., 1951, ApJ, 113, 413
Springel V., Hernquist L., 2005, ApJ, 622, L9
Stewart K. R., Bullock J. S., Wechsler R. H., Maller A. H., 2009, ApJ, 702, 307
Toomre A., 1977, in Tinsley B. M., Larson R. B., eds, Evolution of Galaxies and Stellar Populations. Yale University Observatory, New Haven, p. 401
Toomre A., Toomre J., 1972, ApJ, 178, 623
Trujillo I. et al., 2006, MNRAS, 373, L36
Trujillo I., Conselice C. J., Bundy K., Cooper M. C., Eisenhardt P., Ellis R. S., 2007, MNRAS, 382, 109
van Dokkum P. G. et al., 2008, ApJ, 677, L5
van Dokkum P. G. et al., 2010, ApJ, 709, 1018
van Rijsbergen C. J., 1979, Information Retrieval, 2nd edn. Butterworth-Heinemann Newton, MA, USA (ISBN:0408709294)
Wen Z. L., Liu F. S., Han J. L., 2009, ApJ, 692, 511
Wetzstein M., Naab T., Burkert A., 2007, MNRAS, 375, 805
Wolf C., Meisenheimer K., Rix H.-W., Borch A., Dye S., Kleinheinrich M., 2003, A&A, 401, 73
Wolf C., Gray M. E., Meisenheimer K., 2005, A&A, 443, 435

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