Advanced morphological galaxy classification: a comparison of observed and simulated galaxies

K. M. Hambleton,[1*] B. K. Gibson,[1,2,3] C. B. Brook,[1] G. S. Stinson,[1,4] C. J. Conselice,[5] J. Bailin,[4,6] H. Couchman[4] and J. Wadsley[4]

[1]*Jeremiah Horrocks Institute, University of Central Lancashire, Preston PR1 2HE*
[2]*Department of Astronomy & Physics, Saint Mary's University, Halifax, Nova Scotia B3H 3C3, Canada*
[3]*Monash Centre for Astrophysics, School of Mathematical Sciences, Monash University, Clayton, VIC 3800, Australia*
[4]*Department of Physics and Astronomy, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4M1, Canada*
[5]*University of Nottingham, School of Physics & Astronomy, Nottingham NG7 2RD*
[6]*Astronomy Department, University of Michigan, 830 Dennison Building, 500 Church Street, Ann Arbor, MI 48109-1042, USA*

ABSTRACT

Encoded within the morphological structure of galaxies are clues related to their formation and evolutionary history. Recent advances pertaining to the statistics of galaxy morphology include sophisticated measures of concentration (C), asymmetry (A) and clumpiness (S). In this study, these three parameters (CAS) have been applied to a suite of simulated galaxies and compared with observational results inferred from a sample of nearby galaxies. The simulations span a range of late-type systems, with masses between $\sim 10^{10}$ and $\sim 10^{12}$ $M_{\odot}$, and employ star formation density thresholds between 0.1 and 100 cm^{-3}. We have found that the simulated galaxies possess comparable concentrations to their observed counterparts. However, the results of the CAS analysis revealed that the simulated galaxies are generally more asymmetric, and that the range of clumpiness values extends beyond the range of those observed. Strong correlations were obtained between the three CAS parameters and colour $(B - V)$, consistent with observed galaxies. Furthermore, the simulated galaxies possess strong links between their CAS parameters and Hubble type, mostly in-line with their observed counterparts.

Key words: methods: numerical – galaxies: evolution – galaxies: structure.

1 INTRODUCTION

Classification in astronomy has led to many advances which have revealed important information about the Universe in which we live. For example, the classification of stars by their colours and brightnesses in the Hertzsprung–Russell (HR) diagram led to an understanding of stellar structure and evolution. The HR diagram continues to be employed as a means to determine the ages and metallicites of stellar populations, whether they are simple stellar populations, as in globular clusters, or in composite populations, such as dwarf galaxies (Dolphin et al. 2003; Monelli et al. 2010; Williams et al. 2010). Even when it is impossible to resolve stars, entire galaxies are classified by their colours and magnitudes. However, galaxies display resolved structure and morphology that stars do not, and using this morphology can reveal much about the evolution of galaxies.

Early methods used for the classification of galaxies were based on morphology. The Hubble (1926) Classification System is a method used to categorize galaxies by their morphology, and although it has been the dominant morphological tool since the mid-1920s, over time it has proven to be deficient in several areas. A prominent disadvantage to classifying galaxies based on their morphological features alone is that it is subjective with respect to distance (or resolution) and inclination. Furthermore it has led to the grouping of a wide range of asymmetric galaxies as simply 'irregular'.

As both ground- and space-based imaging has improved, and we probe higher redshifts, it has become apparent that classifying galaxies as 'irregular' means ignoring a vast amount of morphological information. de Vaucouleurs (1959) attempted to improve upon the Hubble (1926) Classification System by introducing 'later types' as a subclass of galaxies; however, this system still refers to local, axisymmetric galaxies as reference points, and, again, was based solely on morphology.

Morgan & Mayall (1957) noted a correlation between the spectra of galaxies and the dominance of their central bulge component.

*E-mail: kmhambleton@uclan.ac.uk

Morgan (1958) felt that a shortcoming of the Hubble classification scheme was that it was not a strong indicator of spectral class. In order to rectify this, Morgan (1958) devised a galaxy classification system based upon the central light concentration compared to its overall light distribution. This concentration classification scheme involved analysing the flux of galaxies to determine the degree of concentration of the central component. The purpose of this scheme was to study the colour–concentration correlation in an attempt to identify galactic evolutionary states. Due to the concentration parameter's direct relation with spectral class and type of stellar population, the concentration parameter was also found to be indicative of formation history and properties, such as velocity dispersion, galaxy size, luminosity and (more recently) central black hole mass (Graham, Trujillo & Caon 2001).

Extending Morgan's pioneering work, Conselice, Bershady & Jangren (2000) proposed a new statistical measure of asymmetry, in conjunction with a revised parametrization of Morgan's concentration index. This classification system was based on the desire to classify galaxies quantitatively at a range of redshifts. Asymmetry was defined by rotating a galaxy $180°$, about a central axis and finding the absolute sum of the normalized residuals. This method measures the high-frequency structure within each galaxy, whilst also considering the symmetry. The measure of asymmetry focuses primarily on morphological shape, and has a strong correlation with colour, congruent with Morgan's concentration parameter. As such, this property has a correlation with both star formation and merger history.

Conselice (2003) introduced one additional morphological measure sensitive to high spatial frequency clumpiness. This parameter is defined by comparing a galaxy to a smoothed image of itself. In doing so, the clumpiness parameter quantifies galaxy morphology in a manner which reflects star-forming regions and evolutionary history. This parameter distinguishes between varying types of galaxies and thus, partially, resembles Hubble's classification system. However, the clumpiness parameter quantifies galaxy morphology with respect to intrinsic characteristics and not morphology alone.

Conselice (2003) combined the high spatial frequency clumpiness parameter with both the asymmetry index and the concentration index to complete what is now called the CAS classification system. Underlying the CAS system's empirical nature lies the fundamental physics which governs its individual C, A and S, components. The CAS system highlights the intrinsic nature of galaxies by considering their light distributions and spectral class.

The observable properties of galaxies are a function of their merger histories, masses and environments, each of which are reflected in the CAS system. It is fitting to apply this new classification method to the theoretically predicted properties of galaxies and use it as a gauge to determine where our current understanding of galaxy formation and evolution is correct and where it is deficient. Galaxy simulations are an integral tool that can help interpolate between known stages of formation history, and thus further our understanding of galaxy evolution.

Within the observational community, the CAS system has become one of the most widely-used measures of galaxy morphology (Hernández-Toledo et al. 2008; Bertone & Conselice 2009). Since the completion of the combined CAS system, it has been applied to numerous galaxies including the 113 galaxies observed by Frei et al. (1996). However, it has not yet been applied systematically to high-resolution computational galaxy simulations as a method of comparison between observations and simulations. In this work, we develop a software package patterned on the Conselice et al. (2000) and Conselice (2003) CAS system, calibrated it on an optimized

training set, and applied it to high-resolution galaxy simulations. We show that the simulated galaxies exhibit both similarities to and differences from real galaxies, and highlight the intrinsic physical processes that are responsible for the differences.

In Section 2 we describe our simulated galaxy samples. In Section 3 we discuss the three structural parameters of the CAS system, concentration (C), asymmetry (A) and clumpiness (S). We then compare the simulations with the empirical Frei et al. (1996) sample of galaxies in Section 4. The results are discussed in Section 5 and conclusions drawn in Section 6.

2 SIMULATED GALAXY SAMPLE

To compare with the morphological parameters inferred from observed galaxies, we used three samples of simulated galaxies. Each sample was generated using the gravitational N-Body + smoothed particle hydrodynamics (SPH) code GASOLINE (Wadsley, Stadel & Quinn 2004). We briefly outline the simulations here, but refer the reader to the detailed descriptions found in Brooks et al. (2009), for what will henceforth be called the 'UW sample', Governato et al. (2010), for the 'Dwarf sample', and Stinson et al. (2010) for the McMaster Unbiased Galaxy Simulations ('MUGS) sample'.

Regardless of the aforementioned 'sample' from which a specific simulation was drawn, each individual 'zoom-style' simulation was run within a *WMAP*-3 ΛCDM cosmological framework (Spergel et al. 2007) and consisted of a central high resolution region (centred on the target halo) embedded within a lower resolution cosmological volume. Star formation and supernova feedback was computed using the recipes outlined by Stinson et al. (2006), with the primary difference between the three samples being the adopted star formation density threshold (n_{th}). For the UW sample, the canonical n_{th} of $0.1\,\mathrm{cm}^{-3}$ was employed; for the MUGS galaxies, n_{th} was increased to $1\,\mathrm{cm}^{-3}$; for the Dwarf galaxy sample, $n_{th} = 100\,\mathrm{cm}^{-3}$ was adopted. The latter simulations represent the first successful realizations of essentially bulgeless disc galaxies (Governato et al. 2010). The masses of the MUGS and UW samples are comparable to that of the Milky Way ($\sim 10^{12}\,\mathrm{M}_\odot$) and, therefore, also the representative massive galaxies in the Frei et al. (1996) sample, while those of the Dwarf sample are of the order $\sim 10^{10}\,\mathrm{M}_\odot$ (more akin to the Small Magellanic Cloud). We include one additional dwarf (labelled DG1LT) which was simulated using the same parent dark matter halo as one of the canonical dwarf simulations (DG1), but employed a lower star formation threshold of $0.1\,\mathrm{cm}^{-3}$ (similar to the UW sample).

We used SUNRISE (Jonsson 2006), a Monte Carlo radiative transfer code, to generate a posteriori, mock images of the simulations, incorporating the effects of dust attenuation and spanning a range of viewing angles from edge- to face-on perspectives, as listed in Table 1.

All galaxies are presented at inclinations of both $0°$ and $90°$, and in most cases also at a uniform distribution of intermediate inclinations. Within this distribution there are some correlations due to the intrinsic features of each galaxy that are maintained at all inclinations; e.g. in Fig. 11, DG2 is clumpy at all inclinations. We view each galaxy from a range of inclinations, which allows us to compare them to observed galaxies, which are viewed at a range of inclinations. Furthermore, it increases the size of the sample given the limited number of galaxy simulations available at sufficient resolution. Subsequently we have multiple copies of the same physical system, thus the images are not truly independent and care should be taken not to overinterpret any statistical conclusions. In Figs 1–4, SDSS *r*-band images of representative edge-on and face-on

Table 1. CAS parameters for simulated galaxies, including values obtained at different inclinations. The names of the galaxies were defined by the teams that simulated them (see text for details); the specific names are arbitrary for the purposes of this study.

| Galaxy	Inclination (°)	$C(r)$	$A(r)$	$S(r)$
UW sample				
Gal1	0	3.49	0.351	0.485
	90	3.26	0.312	0.680
h277	0	4.96	0.215	0.126
	90	3.82	0.161	0.508
MW1024	0	3.97	0.180	0.181
	90	3.65	0.180	0.357
Dwarf sample				
DG1	0	3.38	0.385	0.378
	45	3.38	0.232	0.453
	90	3.47	0.369	0.667
DG1LT	0	2.80	0.267	0.259
	30	3.01	0.281	0.253
	45	2.87	0.275	0.303
	60	3.18	0.260	0.380
	90	4.06	0.231	0.761
DG2	0	3.66	0.574	0.217
	30	3.17	0.570	0.193
	45	3.19	0.568	0.228
	60	3.18	0.552	0.197
	90	3.17	0.376	0.699
DG3	0	3.28	0.315	0.139
	30	3.31	0.317	0.126
	45	3.29	0.314	0.134
	60	3.22	0.311	0.177
	90	3.48	0.200	0.664
MUGS sample				
g1536	0	4.51	0.158	0.078
	45	4.46	0.174	0.084
	90	4.11	0.123	0.201
g21647	0	3.49	0.162	0.104
	45	3.31	0.150	0.061
	90	3.49	0.173	0.314
g22437	0	3.71	0.240	0.258
	30	3.49	0.193	0.186
	45	3.34	0.209	0.159
	60	3.31	0.222	0.146
	90	3.45	0.209	0.192
g24334	0	3.57	0.235	0.362
	30	3.44	0.235	0.424
	45	3.44	0.230	0.406
	60	3.36	0.215	0.393
	90	3.54	0.246	0.457
g25271	0	3.63	0.106	0.007
	45	3.63	0.086	0.007
	90	4.06	0.148	0.100
g3021	0	3.71	0.206	0.277
	30	3.76	0.216	0.324
	45	3.84	0.235	0.393
	60	3.91	0.234	0.407
	90	4.03	0.209	0.419
g5664	0	3.16	0.146	0.359
	45	2.93	0.170	0.330
	90	3.39	0.142	0.726
g7124	0	4.51	0.077	0.000
	45	4.00	0.099	0.029
	90	4.22	0.142	0.001
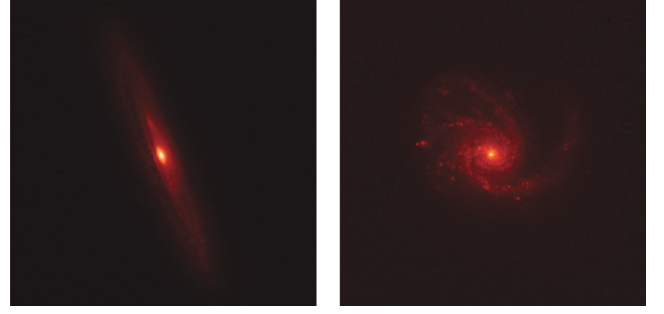
Figure 1. Left-hand panel: NGC 4216, an edge-on disc galaxy from the Frei et al. (1996) sample in the SDSS *r* band. Right-hand panel: NGC 4254, a face-on disc from the Frei et al. (1996) sample in the SDSS *r* band.
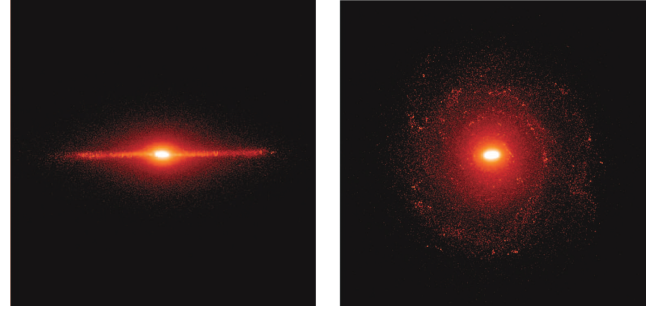
Figure 2. Left-hand panel: an image of g1536, a disc galaxy from the MUGS sample, as viewed edge-on in the SDSS *r* band. Right-hand panel: g1536, a disc galaxy from the MUGS sample, as viewed face-on in the SDSS *r* band.
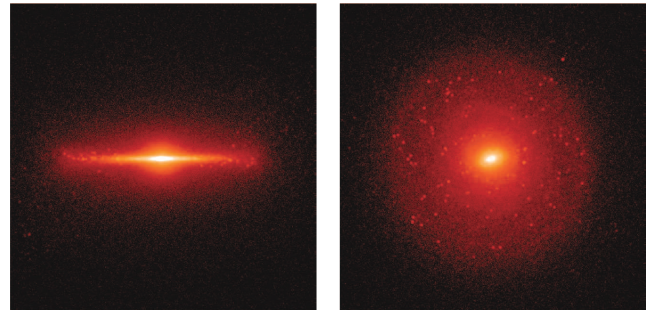
Figure 3. Left-hand panel: an image of MW, a disc galaxy from the UW sample, as viewed edge-on in the SDSS *r* band. Right-hand panel: an image of MW, a disc galaxy from the UW sample, as viewed face-on in the SDSS *r* band.
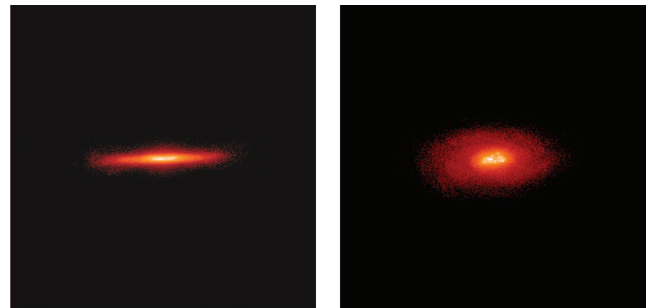
Figure 4. Left-hand panel: an image of DG3, a disc galaxy from the Dwarf sample, as viewed edge-on in the SDSS *r* band. Right-hand panel: an image of DG3, a disc galaxy from the Dwarf sample, as viewed face-on in the SDSS *r* band.

galaxies drawn from the observational (Fig. 1) and simulated (Figs 2–4) samples are shown.

3 THE CAS STRUCTURAL PARAMETERS

The three CAS parameters were recreated after Conselice et al. (2000) and Conselice (2003), and calibrated on the Frei et al. (1996) sample of nearby galaxies in the 'r' and 'R' bands. Prior to the computation of each parameter, each galaxy was background subtracted by initially taking the median count value of a section at the edge of each galaxy's CCD frame, and subtracting this from each pixel within the array. The surface brightness profile was then inspected to ensure that the correct level of background was removed.

Prior to computing the CAS indices, a characteristic radius for each galaxy was determined. Conselice et al. (2000) recommends the use of the characteristic η radius which is 1.5 $r_{\mathrm{Petrosian}}$. The Petrosian radius is defined as the radius r at which the ratio of the local surface brightness (in an annulus from 0.8 to 1.25r) to the mean surface brightness within r is 0.2. This is found by calculating the aforementioned ratio for radii within an annulus 0.8–1.25r.

3.1 Concentration (C)

The concentration parameter is a measure of how concentrated (or diffuse) the central bulge component is with respect to the total flux of the galaxy. Concentration has a strong correlation with colour, and is related to other intrinsic features such as velocity dispersion and mass (Graham et al. 2001). The concentration of a galaxy is determined by locating the radii of the circular apertures that contain 20 per cent ($r_{0.2}$) and 80 per cent ($r_{0.8}$) of the total flux of the galaxy, and then taking their ratio:

$$C = 5 \log \frac{r_{0.8}}{r_{0.2}}. \tag{1}$$

3.2 Asymmetry (A)

The asymmetry parameter (A) is a measure of how symmetric (or asymmetric) a galaxy is with respect to the total flux of the galaxy. Galactic asymmetries are predominantly caused by star formation, mainly in underlying structure such as spiral arms. Asymmetry is also a strong indicator of merger history, as demonstrated by Conselice (2003). The asymmetry parameter initially involves locating the centre of a galaxy image, which is defined as the global minimum of asymmetry. The centre was obtained using the centring algorithm specified by Conselice et al. (2000), which involves an iterative trial-and-error process. An initial guess for the centre of the galaxy is made and the asymmetry for the specified pixel, and the surrounding eight pixels is determined. The pixel corresponding to the lowest asymmetry is then redefined to be the centre. In order to locate the global asymmetry minimum, which corresponds to the centre, it is required that the process be repeated until the central pixel is also the pixel with the lowest value of asymmetry.

The asymmetry itself is calculated by rotating the image 180° about the defined central point and subtracting the rotated image from the initial image, in order to determine the absolute value of the residuals. Following this, the residuals are divided by the quantity corresponding to the summation of all the pixels in the galaxy. Subsequently, a similar method is then applied to a small section, taken from a corner of the image which contains background only. This quantity is then multiplied by a scalefactor so that it is comparable to the size of the galaxy image. This quantity is divided by the summation of the total flux of the original galaxy image. This is done in order to determine the asymmetry of the background so that it may be discounted from the galactic asymmetry. The calculation for the asymmetry is as follows:

$$A = \frac{\sum I - I_{180}}{\sum I} - \frac{\sum B - B_{180}}{\sum I}, \tag{2}$$

where I is the image flux, I_{180} is the rotated image flux, B is the background and B_{180} is the rotated background.

3.3 Clumpiness (S)

The measure of high spatial frequency clumpiness (S) is directly related to a galaxy's star formation and high-frequency spatial power. The clumpiness is defined as the patchiness of the high-frequency light distribution within a galaxy (Conselice 2003). It is computed by subtracting a degraded-resolution version of the galaxy image from the original galaxy image. As previously with the asymmetry parameter, it is required that the method is repeated on a section of sky background in order to reduce the effect of noise. Again, a corner of the CCD containing background only is selected, replicated and smoothed, so that an identical procedure may be carried out on both the background and the galaxy image.

As the central bulge component is a prominent feature in the simulated galaxies, it contributes to a large portion of the high-frequency structure and hence clumpiness of the simulated galaxies. As it is already a known issue that simulated galaxies tend to have large central bulge components, with respect to their observed counterparts, it was decided that the standard central region specified by Conselice (2003) would be extended to include the total central bulge component of each galaxy. Furthermore, the inferred total flux of the galaxies would also exclude the central region, in order to avoid any bias due to the varying sizes of the bulge components of the galaxies. The size of the central bulge component was determined from the light profile of the galaxy by considering the radius at which the profile changed from $\ln (I) = r^{1/4}$, to exponential decay. The image was then examined to ensure that the correct area had been removed, prior to the computation of the clumpiness parameter.

The calculation for the clumpiness is as follows:

$$S = 10 \times \left(\frac{\sum I - I_{\mathrm{s}}}{\sum I} - \frac{\sum B - B_{\mathrm{s}}}{\sum I} \right), \tag{3}$$

where I is the image flux, I_{s} is the smoothed image flux, B is the background and B_{s} is the smoothed background. Further details on the derivation of the clumpiness parameter are provided in Appendix A.

4 COMPARISON OF OBSERVATIONS WITH SIMULATIONS

The CAS parameters form a well-established system that has been used on occasion to analyse galaxy simulations, although not in any systematic manner. For example, Conselice (2006) applied the system to dark matter merger-remnant simulations, and Lotz et al. (2008) who extended this to the analysis of equal-mass gas-rich disc merger remnants. Our unique contribution to this area is, for the first time, an application to a suite of simulations realized within a fully cosmological framework. Such a comparison should enable underlying physical differences to be revealed and allow for scaling relations to be tested using simulations in a new and unique manner. The main purpose of this paper is to consider the morphological

details of the galaxies produced in simulations, and uncover new information regarding the differences between observed and simulated galaxies. To avoid simply re-identifying known issues and biases, we have modified slightly the underlying CAS parameters, as explained now.

4.1 Concentration (C)

The overproduction of the centrally concentrated bulge component of most simulated disc galaxies (e.g. Stinson et al. 2010, and references therein; cf. Governato et al. 2010), is a long-standing problem associated with cosmological simulations of galaxy formation. By definition, the concentration parameter is a proxy for the dominance of the central region, simulation or not, although consideration of the surface brightness profiles suggested that the basic definition not be altered here. Instead, we concentrate here on the relationship between the size of the central component and disc of each galaxy, and compare this with observations.

Another discrepancy between observed and simulated galaxies is, not surprisingly, their noise content. Unlike the other CAS parameters, the measure of concentration does not include a background subtraction stage. This is due to the negligible effects of noise on the measure of concentration. For this reason, no adjustments were made to the concentration parameter regarding background noise.

4.2 Asymmetry (A)

The measure of asymmetry considers the residuals produced when a galaxy image is subtracted from the same image that has been rotated by 180°. The central bulge component of a galaxy is generally symmetric, and for this reason no adaptation was required to account for the excessive central bulge component of the simulated galaxies. However, it was discovered that the effects of noise were critical, as the background component was comparable in magnitude to the component due to the galaxy itself. It was imperative to ensure that this aspect of the algorithm was correct, as simulated galaxies contain essentially zero background noise compared to their observed counterparts, the latter of which contain varying amounts of noise depending upon exposure time, sky background and distance. For this reason, any discrepancies concerning the computation of the background would lead to an inaccurate comparison.

In order to test this hypothesis, an artificial array of noise was created, and the galaxies analysed both before and after the application of this noise array. We replicate the signal-to-noise ratio values of the Frei et al. (1996) sample by first pixelating the simulated images to the same resolution, then scaling the flux to be comparable to that in the images of the real galaxies, and finally adding a normally distributed noise array of sufficient amplitude to reproduce the signal-to-noise ratio of the observed galaxies in Frei et al. (1996). Furthermore, Conselice et al. (2000) demonstrated that accurate asymmetry calculations can only be performed, without corrections, on galaxies with signal-to-noise ratio greater than 100. For this reason, each galaxy was tested individually to ensure that the application of these characteristic noise values led to signal-to-noise ratios in excess of 100.

A comparison of the asymmetry results derived from the simulated galaxies, both before and after the application of noise, can be seen in Fig. 5. It was found that the inferred asymmetry values differed only marginally after the application of noise (in large part because of the stringent signal-to-noise ratio limit imposed), although the impact varied for each galaxy and appeared to be more substantial for more asymmetric galaxies. Whilst there is some
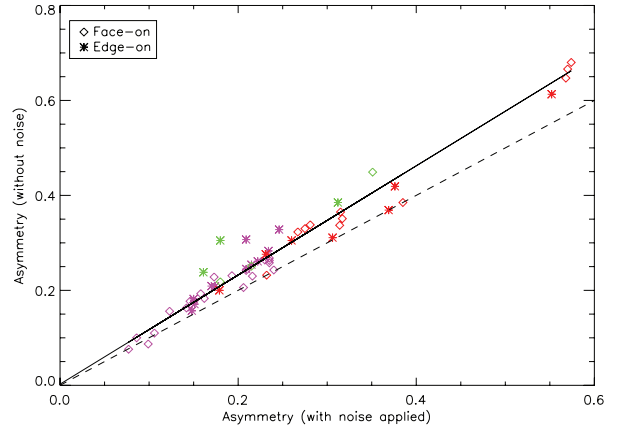
Figure 5. Relationship between the asymmetry of simulated galaxies both before and after the application of noise. The solid line represents the least squares regression line for which the equation is $y = 1.15x - 0.002$, and the dashed line corresponds to the one-to-one relation. Here, the points are coloured to represent the different galaxy samples: MUGS = purple, UW = green and Dwarf = red. The choice of symbol corresponds to galaxy inclinations of 0°–45° ('face-on': diamonds) and 45°–90° ('edge-on': asterisks).

variation visible in Fig. 5 due to the application of noise, it is not significant. However, to ensure that a fair comparison was made between observed and simulated galaxies, it was decided that noise would be applied to the simulated galaxies prior to the computation of the asymmetry parameter.

4.3 Clumpiness (S)

When initially computed, the central bulge component dominated the clumpiness calculation for the simulated galaxies, suggesting that the simulations were excessively clumpy with respect to their observed counterparts. Prior to calculating the clumpiness, Conselice (2003) forces the central region, which corresponds to 0.05 times the Petrosian radius, to zero, to ensure that any unresolved region of the central bulge component is excluded. For the purpose of this study this region has been extended to encompass the total central bulge component of each galaxy, thus allowing for the comparison of observed and simulated galaxies outside the bulge region, which we know is overproduced in these simulations. By making this alteration, the clumpiness of the disc region only is calculated, excluding any high-frequency structure associated with the central bulge component.

As the area of the central bulge differs for each galaxy, in order to avoid any bias due to bulge size, it was then decided that the central bulge component would also be excluded from the summation of the total flux which is present in the denominator of the equation for the computation of the clumpiness (see Section 3.3).

The application of noise had a negligible effect upon the values generated by the clumpiness algorithm. This suggests that the subtraction algorithm for the clumpiness parameter is more than adequate for minimizing the impact of noise on its calculation. For this reason it was deemed unnecessary to apply noise to the simulated galaxies in the computation of the clumpiness.

5 DISCUSSION

For each simulation, surface brightness profiles, for a range of inclinations, were generated. The simulations were 'pixelated' to a scale corresponding to 100 pc pixel⁻¹, which, as shown by

Conselice et al. (2000), appears adequate for CAS analysis. The Frei et al. (1996) sample of observed, nearby galaxies, used in this study, all possess resolutions comparable to (or better than) the same 100 pc pixel^{-1}, and hence are again ideal for use with the CAS system. As all the galaxies used in this study are above the threshold value of $\sim$1 kpc pixel^{-1}, where the accuracy of the CAS parameters begins to decline (Conselice et al. 2000), a comparison between simulated and observed galaxies can be made with a high degree of confidence.

The Frei et al. (1996) sample of galaxies contains only nearby high surface brightness galaxies. It is not meant to be a complete and/or unbiased compilation, as it mostly excludes irregular and low surface brightness dwarf galaxies, which contribute significantly to the population of local galaxies. Similarly, the sample of simulated galaxies is incomplete in the sense that it does not contain any irregular or elliptical galaxies, resulting in a restricted range of colours; this impacts particularly on the red end of the distribution, where we lack counterparts to the Frei et al. (1996) sample. This does place constraints on the ability to compare observed and simulated galaxies, and is considered during this analysis by incorporating Hubble type in many of the figures.

5.1 Concentration

In Fig. 6, both observed and simulated galaxies demonstrate a strong correlation between colour and concentration. This trend naturally arises for observed galaxies where early-type galaxies are redder in colour and also tend to be more centrally concentrated, whereas later types are bluer due to ongoing star formation and tend to be less centrally concentrated. These values are well quantified by Hubble type and hence, for the observed galaxies, it can be seen that the different symbols that represent the various Hubble types are contained within well-defined regions. The visual appearance of a trend among the simulated galaxies comes more from the systematic offset between the Dwarf sample in red and the MUGS and UW samples in purple and green. While there is a trend within the MUGS sample of galaxies, it is not nearly as visually striking.

It is evident from Fig. 6 that the simulated galaxies tend to be bluer than the observed galaxies, both in general and with respect to their Hubble type. The main cause of this blueward offset is the presence of excessive late-time star formation in the simulated discs.
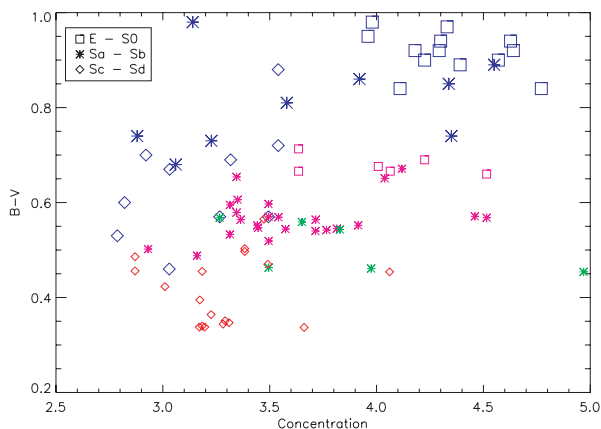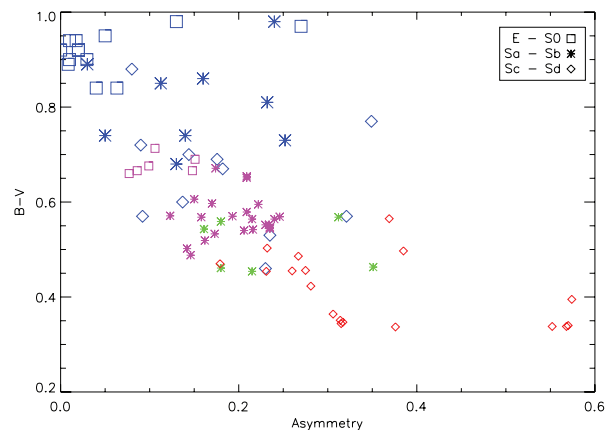
Furthermore, there are a lack of elliptical galaxies in the simulated sample, with the earliest galaxies being lenticular. However, in a relative sense, there is a definite relationship between colour and Hubble type, as expected.

The simulated galaxies in Fig. 6 show a similar range of concentrations to those observed for real galaxies, and also appear to be divided into groups by Hubble type. Furthermore, the simulated galaxies demonstrate a positive linear trend that, although offset due to colour differences, is comparable to the trend demonstrated by the observed galaxies. Concentration is also related to surface brightness; it is expected that the low concentration values generated by the majority of the simulated dwarf galaxies is primarily due to their low surface brightness. Furthermore, the dwarf galaxies in this sample are all younger disc-like galaxies which are observed to be generally more diffuse. Other factors which influence the concentration parameter include the mass and velocity dispersion, both of which are low for the dwarf galaxy sample.

The MUGS sample of simulated galaxies in Fig. 6 appears to extend to higher values of concentration than both the UW sample of Milky Way-like galaxies and the Dwarf sample. This is most likely due to the presence of earlier type discs within the unbiased MUGS sample (cf. with the UW sample, which were chosen a priori to more closely resemble later type discs). Moreover, the MUGS sample of galaxies have a higher density threshold for star formation (1 cm^{-3}), which could lead to preferential star formation in the denser (central) regions of the galaxy, although this behaviour is not seen in dwarf galaxy simulations (Governato et al. 2010). The MUGS sample also has a greater average circular velocity (150–250 km s^{-1}), compared to the other simulated samples of galaxies, which is a parameter known to correlate with concentration (Graham et al. 2001).

5.2 Asymmetry

In Fig. 7, the observed galaxy sample demonstrates a strong anti-correlation between asymmetry and colour. This correlation stems from the bluer, younger galaxies generally containing spiral arms with dust lanes and star-forming regions, and the older, elliptical galaxies generally being devoid of structure. The simulated galaxies follow an approximately identical trend to that of the observed galaxies, however, due to the bluer nature of the selected simulations, they form a blue extension of the observed galaxy trend. Conselice (2003) showed that there was a distinct region in the

Figure 6. Relationship between concentration and colour ($B - V$) for both observed and simulated galaxies. The points are coloured to represent the different galaxy samples: observed = blue (and slightly larger in size relative to the other symbols), MUGS = purple, UW = green and Dwarf = red. The symbols correspond to the classical Hubble types, as noted in the inset.

Figure 7. Relationship between asymmetry and colour ($B - V$) for both the observed and simulated galaxies. Symbols and colours are the same as in Fig. 6.
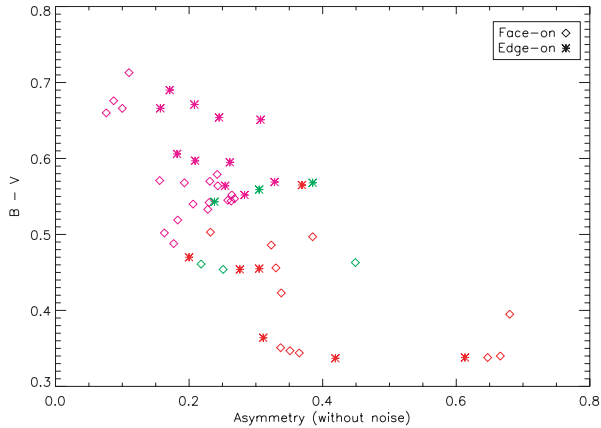
Figure 8. Relationship between the asymmetry (A) of simulated galaxies, without the application of noise, and colour $(B - V)$. Symbols and colours as in Fig. 5.

Figure 9. Relationship between clumpiness (S) and colour $(B - V)$ for both the observed and simulated galaxies. Symbols and colours as in Fig. 6.

bottom left corner of the colour–asymmetry plot where no galaxies were bluer and more symmetric. This region is well defined by the simulations which demonstrate a very strong correlation between colour and asymmetry.

The earlier types within the MUGS sample are redder in colour with respect to the other simulations, and tend to be more symmetric. Again there is a strong distinction between the various Hubble types, which is similar to the correlation demonstrated by the observed galaxy sample. However, due to the lack of redder simulated galaxies with low, approximately zero, asymmetry values, this trend does not extend to the upper left corner of Fig. 7, where the redder, symmetric, galaxies reside in nature. Again, this can be traced to the ongoing late-time star formation in the earlier type simulations which is not seen in nature. In Fig. 7 some of the dwarf simulated galaxies appear to be excessively asymmetric relative to the other simulations. The large variation of apparent flux levels within these systems is most likely the cause of this increased in asymmetry.

Conselice et al. (2000) showed that it is almost always the inclined galaxies that contribute to the scatter in the colour–asymmetry diagram and that the less inclined galaxies are systematically higher. In Fig. 8, the edge-on and face-on galaxies have different symbols, and from this it can be concluded that while the inclined, observed galaxies do tend to be somewhat more significant outliers, in detail, the scatter of the simulated galaxies is not a function of inclination angle. Also, the MUGS and UW galaxies that are less inclined do appear to be systematically higher in the plot, as found for the observed galaxies.

Noise was applied to the simulated galaxies to ensure that a fair comparison was made when considering the asymmetry. The addition of a noise array to the simulated galaxy images had the benefit of allowing for a more accurate comparison with the observed galaxies. In Fig. 8 the range of asymmetry values, without the addition of noise, can be seen. The absence of noise only appears to affect the outcome of the asymmetry parameter at greater asymmetry values by restricting the range of asymmetries.

5.3 Clumpiness

In Fig. 9, it can be seen that there is an inverse correlation between colour and clumpiness for both the observed galaxy sample and the MUGS simulated galaxy sample, although the anticorrelation of the MUGS galaxy sample is much weaker (and shifted in colour). The UW sample shows a positive correlation (although over a lim-

ited colour range) and the Dwarf sample shows no correlation. As the earlier type disc galaxies are generally devoid of star-forming regions and internal structure, and also tend to be red in colour, it follows that they would generally be located at the top left of Fig. 9, where the smoother, redder galaxies reside. Conversely, later type galaxies tend to be bluer and contain high-frequency residuals. For this reason, the anticorrelation is the expected correlation for Fig. 9.
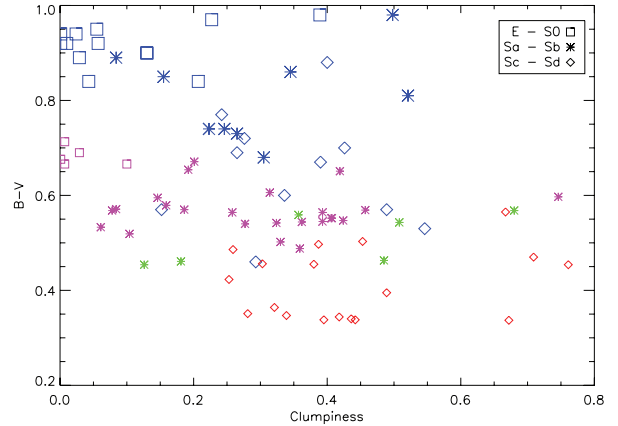
The simulated and observed galaxies demonstrate a strong correlation between clumpiness and Hubble type. Due to the lack of redder galaxies, as with asymmetry, the simulated galaxies do not reach the top left corner of Fig. 9. However, unlike asymmetry, there are several galaxies from the MUGS sample that possess approximately zero clumpiness, consistent with their observed counterparts.

In Fig. 9, it can be seen that there is a group of galaxies (mainly from the Dwarf sample, with the exception of one MUGS and one UW galaxy), that are excessively clumpy. All the simulated galaxies that are excessively clumpy also have inclination angles of 90° (see Table 1). Thus a study into the effects of dust attenuation on the clumpiness parameter was carried out. It was found that the clumpiness of the simulated galaxies was not affected by the contribution of the dust from SUNRISE (Jonsson 2006). It was therefore concluded that the Dwarf sample, and Gal1 from the UW sample, are simply intrinsically clumpy. In the case of the Dwarf sample, this is consistent with an analysis of the power spectrum of the cold gas associated with their interstellar media (Pilkington et al. 2011). This may be attributed to the relatively low surface brightnesses associated with the dwarfs, as this causes the high frequency residuals to dominate in the calculation of the clumpiness.

5.4 CAS comparison

The CAS parameters are based solely on the morphology of galaxies, but are directly related to many intrinsic qualities. By comparing CAS parameters amongst themselves, as opposed to CAS versus colour, sampling biases, such as the colours of stellar populations, may be avoided. This allows for the structures of the galaxies to be compared directly, which may further emphasize the differences between observed and simulated galaxies.

In Fig. 10, the observed galaxies from the Frei et al. (1996) sample demonstrate a general negative correlation between concentration and clumpiness. Galaxies that are smooth tend to be concentrated and those that are clumpy tend to be more diffuse. This correlation is generated because younger, disc galaxies tend to be both clumpy due
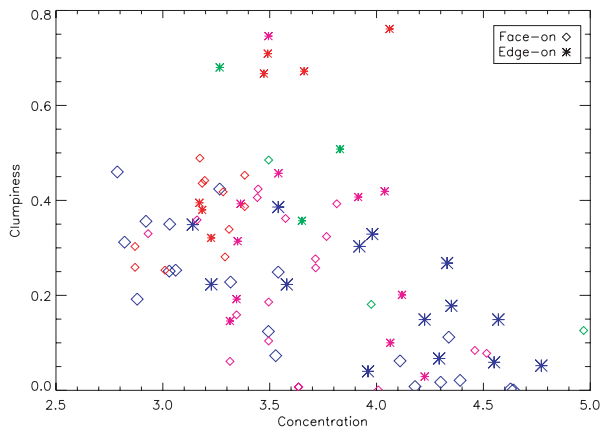
Figure 10. Relationship between the concentration (C) and clumpiness (S) parameters for both the observed and simulated galaxies. The points are coloured to represent the different galaxy samples. The Frei et al. (1996) sample of observed galaxies are coloured blue with larger symbols relative to the simulated samples. The MUGS sample are shown in purple, the UW sample in green and the Dwarfs represented with red symbols. The symbols denote the inclination angle, as in Fig. 5.
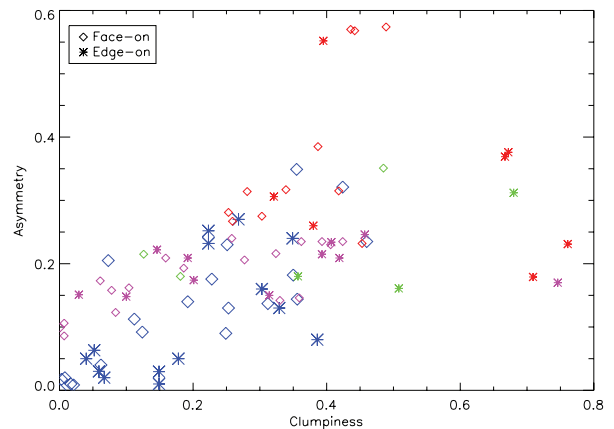

Figure 11. Relationship between clumpiness (S) and asymmetry (A) parameters for both observed and simulated galaxies. Symbols and colours as in Fig. 10.

to active star formation and diffuse, while older elliptical galaxies tend to be smoother and more dense.

The majority of simulated galaxies lie within the normal range of observed galaxies, in Fig. 10. It is apparent though that some of the simulations are excessively clumpy in comparison to their observed counterparts. As can be garnered from Table 1, all these simulated galaxies which are clumpy outliers only occur when viewed at an inclination of 90°. This group of galaxies display approximately average concentration values, whilst being excessively clumpy, in comparison to both their observed and simulated galaxy counterparts. In regards to the excessively clumpy dwarfs, we note that the high-frequency structure present in these simulated galaxies is consistent with the power spectrum analysis of the associated interstellar medium (ISM) by Pilkington et al. (2011).

In Fig. 10, the MUGS sample demonstrates an approximately equivalent trend to the observed galaxy sample, despite their excessive bulge sizes. This is unexpected, as the central bulge component is disregarded in the computation of the clumpiness, but is included for the concentration parameter. However, with respect to the Dwarf and UW samples, the MUGS sample contains galaxies that are generally more concentrated. It is likely that the greater mass, velocity dispersion, and higher star formation threshold of the MUGS sample all contribute to both the larger bulge sizes and concentration of these galaxies, with respect to the other simulated galaxy samples. Furthermore, it is likely that these characteristics produced the earlier type discs in the MUGS sample, which are expected to be less clumpy in nature.

In Fig. 11, it can be seen that there is a strong positive correlation between the clumpiness and asymmetry of observed galaxies. This correlation arises from the later type galaxies tending to be clumpy and asymmetric, due to their star forming regions and intrinsic structure, and earlier type galaxies tending to be devoid of structure and star-forming regions, which pertains to them being more smooth and symmetric. The simulated galaxies appear to follow the same correlation. However, the sample of highly symmetric observed galaxies are not reproduced within the simulations.

In Fig. 11, the Dwarf sample of galaxies form three groups: those which lie within the normal range, those which are excessively clumpy, and those that are excessively asymmetric. As previously

discussed, those galaxies that are excessively clumpy all have 90° inclination angles. Fig. 11 shows that the high-frequency spatial clumpiness of these galaxies is not directly related to asymmetry, as those galaxies that demonstrate excessive clumpiness possess a wide range of asymmetries with respect to other galaxies from the same sample.

The group of simulations which are seemingly highly asymmetric outliers at the top of Fig. 11 is in fact just one simulated galaxy (DG2) viewed at various inclinations. This particular galaxy has a large star forming complex at $z = 0$ which is offset from the centre of the galaxy (see Governato et al. 2010), explaining both the clumpiness and asymmetry of this galaxy.

In Fig. 12, there is an apparent negative correlation between asymmetry and concentration for both observed and simulated galaxies. This can be attributed to the fact that later type galaxies are both diffuse and asymmetric, and earlier type galaxies are generally dense with less internal structure. Within the Dwarf galaxy sample there is again a single galaxy viewed at various inclinations that demonstrates low concentration values coupled with extremely high asymmetry values. The large star-forming complex in this simulation has not resulted in a concentration being outside the region defined by observed galaxies.
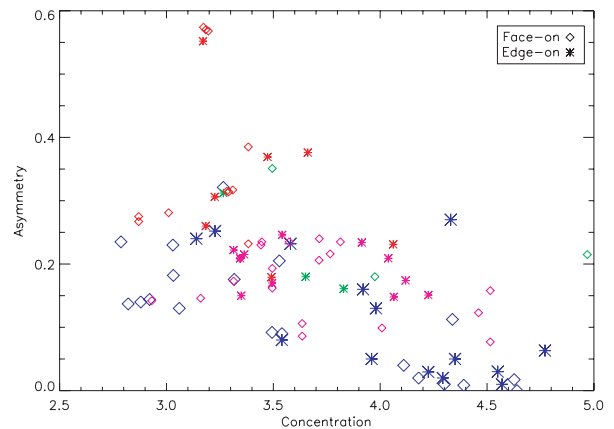
Figure 12. Relationship between concentration (C) and asymmetry (A) parameters for both observed and simulated galaxies. The points are coloured to represent the different galaxy samples. Symbols and colours as in Fig. 10.

6 CONCLUSION

We have applied measures of concentration (C), asymmetry (A) and clumpiness (S) to a sample of simulated galaxies which have a range of masses and morphologies, as well as to a sample of observed galaxies. We have explored the correlations between these parameters as well as between each of them with $B - V$ colour. In general, reasonable agreement was found between the simulated and observed populations in these relations, although some differences become apparent, as summarized below.

Overall, the trend generated by the observed galaxy sample, with respect to colour ($B - V$), was approximately replicated by the simulated galaxies, although there is a distinct lack of redder simulated galaxies, which is related to a well-established inability of these simulations to halt late-time star formation in early-type galaxies (Kawata & Gibson 2005).

It was established that Hubble type is strongly linked with the three CAS parameters for the both the observed and simulated galaxy samples. This is expected as later type galaxies, such as Hubble type Sc–Sd, have star-forming regions, resulting in significant internal spatial structure, which manifests itself in significant asymmetries and clumpiness, and possess low concentration values. Conversely, earlier type galaxies are smoother, more symmetric and more concentrated, due to their relative lack of star-forming regions, structure and their evolutionary histories.

The concentration parameter was found to be robust and easily replicated. The values obtained for the simulated galaxies demonstrated an approximately equal range of values compared with their observed counterparts. The fewer number of simulated galaxies with high concentration merely reflected the relative paucity of early-type galaxies compared with the observed sample.

It was found that the range of asymmetries of the simulated galaxies did not extend to as low values as those of the most symmetric observed galaxies. This may be related to the ongoing star formation in the simulated galaxies, the same process which results in a dearth of red galaxies in the simulations. It is generally the early-type observed galaxies which are highly symmetric. Nevertheless, it remains a challenge for simulators to create highly symmetric galaxies within a hierarchical structure formation paradigm.

A single dwarf galaxy simulation also shows excessive asymmetry compared to the observed sample. This is due to the existence of a large star-forming complex offset from the centre of that simulation. Detailed exploration of the resolution of star-forming complexes and star clusters within simulations is beyond the scope of this study, but it appears that the physical size of the star-forming complex in at least one of the simulations is larger than observed, at least within the Frei a et al. (1996) sample.

As it is a well-known issue that simulated galaxies have disproportionately large bulge components (Stinson et al. 2010), it was decided that a central region, pertaining to the size of the bulge component, would be removed for all galaxies in the calculation of the clumpiness. Following this, we found that the simulated and observed galaxies showed similar spreads in the values of clumpiness. The early-type disc galaxies from the MUGS sample, which have lower rates star formation occurring in the disc at $z = 0$, have low S, while the later type simulated galaxies continue to have high levels of star formation, and high values of S, consistent with their observed counterparts. Analysis of the clumpiness values generated by the simulated galaxies also demonstrated that several simulated galaxies were excessively clumpy when viewed at an inclination angle of 90°. At least in relation to the dwarfs, this is likely related

to high-frequency power of these galaxies, when analysed at an edge-on inclination, a characteristic that is not consistent with observations (cf. Pilkington et al. 2011, for the high-frequency analysis of the ISM of these dwarfs).

REFERENCES

Bershady M. A., Jangren J. A., Conselice C. J., 2000, AJ, 199, 2645
Bertone S., Conselice C. J., 2009, MNRAS, 396, 2345
Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, ApJ, 694, 396
Conselice C. J., 2003, ApJS, 147, 1
Conselice C. J., 2006, ApJ, 638, 686
Conselice C. J., Bershady M. A., Jangren A., 2000, ApJ, 529, 886
de Vaucouleurs G., 1959, Handbuch Phys., 53, 275
Dolphin A. E. et al., 2003, AJ, 126, 187
Frei Z., Guhathakurta P., Gunn J. E., Tyson J. A., 1996, AJ, 111, 174
Governato F. et al., 2010, Nat, 463, 203
Graham A. W., Trujillo I., Caon N., 2001, AJ, 122, 1707
Hernández-Toledo H. M., Vázquez-Mata J. A., Martínez-Vázquez L. A., Avila Reese V., Méndez-Hernández H., Ortega-Esbrí S., Núñez J. P. M., 2008, AJ, 136, 2115
Hubble E. P., 1926, ApJ, 64, 321
Jonsson P., 2006, MNRAS, 372, 2
Kawata D., Gibson B. K., 2005, MNRAS, 358, L16
Lotz J. M., Jonsson P., Cox T. J., Primack J. R., 2008, MNRAS, 391, 1137
Monelli M. et al., 2010, ApJ, 720, 1225
Morgan W. W., 1958, PASP, 70, 364
Morgan W. W., Mayall N. U., 1957, PASP, 69, 291
Pilkington K. et al., 2011, MNRAS, in press
Spergel D. N. et al., 2007, ApJS, 170, 377
Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074
Stinson G. S., Bailin J., Couchman H., Wadsley J., Shen S., Nickerson S., Brook C., Quinn T., 2010, MNRAS, 408, 812
Wadsley J. W., Stadel J., Quinn T., 2004, New Astron., 9, 137
Williams B. F. et al., 2010, ApJ, 709, 135

APPENDIX A: COMPUTATION OF THE HIGH SPATIAL FREQUENCY CLUMPINESS PARAMETER (S)

For the purpose of comparing simulated galaxies with those observed, the central bulge component of each galaxy was removed in the computation of the high-frequency spatial clumpiness. However, the method proposed by Conselice (2003) favoured the removal of a central region, amounting to 0.05 times the Petrosian

radius ($r_{\text{Petrosian}}$), for the purpose of excluding any unresolved regions of the image only. As the measure of high spatial frequency clumpiness specified by Conselice (2003) forms an integral part of the CAS system, the method of computation is presented here.

Initially the image is sky-subtracted, flat-fielded, and the characteristic radius determined. The radius is defined as $1.5 \times r_{\text{Petrosian}}$ (further details can be found in Bershady, Jangren & Conselice 2000). If the data are undersampled, the original image undergoes a boxcar smooth of 5×5 pixel prior to any further computation; this image is then redefined as the original image. The next step involves smoothing the image by a factor $r_{\text{Petrosian}}/6$, again using a boxcar smooth. Following this, a map of the residuals is created by subtracting the smoothed image from the original image. Within this residual image, all the negative pixels, and the pixels within the central region of $0.05 \times r_{\text{Petrosian}}$, are then forced to zero, leaving behind the high-frequency residuals only. The total flux from the residual map, within the specified radius, is then summed and divided by the total flux of the initial galaxy, within the specified radius. The value generated equates to the clumpiness of the galaxy including the clumpiness in the background.

To ensure that the results of the clumpiness parameter determination are unaffected by noise, further background correction is then required. This involves selecting an area from the image that contains background only and repeating the aforementioned procedure. Note that the background should also be smoothed by a boxcar smooth of 5×5, if the data are undersampled. Once selected, the area should be smoothed by a boxcar smooth of $r_{\text{Petrosian}}/6$, using the value of the Petrosian radius obtained from the galaxy image. A background residual map is then created by subtracting the smoothed background image from the original background image and forcing the negative pixels to zero. The background flux is then be summed and normalized to the the number of pixels that constitute the total galaxy, within the specified radius. This value is then divided by the summation of the total flux of the galaxy thus producing the clumpiness in the background. The clumpiness in the background is then subtracted from the aforementioned clumpiness in the galaxy, which is computationally defined in Section 3.3.

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