

Galaxy And Mass Assembly (GAMA): The unimodal nature of the dwarf galaxy population

Smriti Mahajan,^{1,2*} Michael J. Drinkwater², S. Driver^{3,4}, Lee S. Kelvin⁵, A. M. Hopkins⁶, I. Baldry⁷, S. Phillipps⁸, J. Bland-Hawthorn⁹, S. Brough⁶, J. Loveday¹⁰, Samantha J. Penny¹¹, A.S.G. Robotham³

¹Indian Institute for Science Education and Research Mohali- IISERM, Knowledge City, Manauli, 140306, Punjab, India

²School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia

³International Centre for Radio Astronomy Research (ICRAR), University of Western Australia, Crawley, WA 6009, Australia

⁴Scottish Universities' Physics Alliance (SUPA), School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews, KY16 9SS, UK

⁵Institut für Astro- und Teilchenphysik, Universität Innsbruck, Technikerstrabe 25, 6020 Innsbruck, Austria

⁶Australian Astronomical Observatory, PO Box 915, North Ryde, NSW 1670, Australia

⁷Astrophysics Research Institute, Liverpool John Moores University, IC2, Liverpool Science Park, 146 Brownlow Hill, Liverpool, L3 5RF

⁸Astrophysics Group, School of Physics, University of Bristol, Bristol BS8 1TL, UK

⁹Sydney Institute for Astronomy, School of Physics A28, University of Sydney, NSW 2006, Australia

¹⁰Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK

¹¹School of Physics, Monash University, Clayton, Victoria 3800, Australia

ABSTRACT

In this paper we aim to (i) test the number of statistically distinct classes required to classify the local galaxy population, and, (ii) identify the differences in the physical and star formation properties of visually-distinct galaxies. To accomplish this, we analyse the structural parameters (effective radius (R_{eff}), effective surface brightness within R_{eff} ($\langle\mu\rangle_e$), central surface brightness (μ_0), and Sérsic index (n)), obtained by fitting the light profile of 432 galaxies ($0.002 < z \leq 0.02$; *Viking Z*-band), and their spectral energy distribution using multi-band photometry in 18 broadbands to obtain the stellar mass (M^*), the star formation rate (SFR), the specific SFR (sSFR) and the dust mass (M_{dust}), respectively.

We show that visually distinct, star-forming dwarf galaxies (irregulars, blue spheroids and low surface brightness galaxies) form a unimodal population in a parameter space mapped by $\langle\mu\rangle_e$, μ_0 , n , R_{eff} , SFR, sSFR, M^* , M_{dust} and $(g - i)$. The SFR and sSFR distribution of passively evolving (dwarf) ellipticals on the other hand, statistically distinguish them from other galaxies with similar luminosity, while the giant galaxies clearly segregate into star-forming spirals and passive lenticulars. We therefore suggest that the morphology classification scheme(s) used in literature for dwarf galaxies only reflect the observational differences based on luminosity and surface brightness among the apparent distinct classes, rather than any physical differences between them.

Key words: galaxies: dwarf; galaxies: evolution; galaxies: fundamental parameters; galaxies: general; galaxies: statistics; galaxies: structure

1 INTRODUCTION

The visual appearance of a galaxy is the manifestation of its star formation history, stellar content and, dynamic and kinematic properties. As a result, visual morphology of a galaxy can not only provide important clues to its formation mechanisms, but the longevity of specific features such as the spiral structure and bars could potentially be used to understand secular evolution.

It is however a known fact that galaxy properties are modu-

lated by their environment, such that red, bulge-dominated, passively evolving galaxies are mostly found in the dense cores of galaxy clusters and groups, while blue, star-forming, disk galaxies appear more frequently in less dense environments (e.g. Dressler 1980). But can a galaxy evolve from being disk to spheroidal away from the harsh environment of clusters and groups? To learn about the mechanisms responsible for converting one galaxy population into another, and, any evolutionary links between visually distinct galaxies, it is important to get a complete description of galaxies which are evolving independent of the impact of dense environment.

* E-mail: smritimahajan@iisermohali.ac.in

Historically, galaxies were distinguished by visual classification (Hubble 1926), supported by quantitative description of their stellar light profile (Reynolds 1920). But since Hubble's tuning fork diagram, we have come a long way in classifying galaxies based on spectroscopic or photometric data obtained in multiple wavebands. After morphology, the most popular classification scheme for galaxies is based on luminosity. Most studies based on wide-angle survey data such as those obtained by the Sloan Digital Sky Survey (SDSS), define galaxies fainter than $M_r \sim -18$ mag as dwarfs. One of the key characteristics of low-luminosity galaxies however, is their low surface brightness which is correlated with their luminosity albeit with considerable scatter (Caldwell 1983; Binggeli, Sandage, & Tarenghi 1984). However, in some cases like the giant low surface brightness (LSB) galaxy Malin I (Bothun et al. 1987), and galaxies having extended ultraviolet disks (e.g. Thilker et al. 2007), the need to distinguish between low-surface brightness and low-luminosity galaxies might be more critical than immediately intuitive.

The literature on low-luminosity, low surface brightness galaxies is further complicated by incorporation of extensive jargon over the decades. Although dwarf ellipticals (dEs), (dwarf) irregulars ((d)Irr), dwarf spheroidals (dSph), dwarf spirals or lenticulars (dS, dS0), blue compact dwarfs (BCDs), ultra-compact dwarfs (UCDs), and little blue spheroids (LBS) all have their glorious presence in the extensive literature, it is often not obvious whether the difference between these subpopulations of galaxies extend beyond their morphology, or if one class is related to another or their giant counterparts.

While dwarf ellipticals, which are predominantly found in rich galaxy clusters, have received considerable attention in the context of the dichotomy between dEs and giant ellipticals (e.g. Kormendy 1985; Graham & Guzmán 2003; Kormendy & Bender 2012; Graham 2013, and references therein), and, more recently for structure 'hidden' under the smooth radial profiles (Janz et al. 2012, 2013), other classes of low-luminosity galaxies have not been studied much beyond the local volume ($\lesssim 10$ Mpc). This is partly due to the lack of data and partly due to the low volume density of these galaxies, which are hard to observe and easily missed from observations focussing on interiors of galaxy clusters. Within the local volume however, dwarf galaxies, especially the dwarf irregulars have been well explored at the optical (e.g. Sharina et al. 2008; Herrmann, Hunter, & Elmegreen 2013) as well as the ultraviolet wavelengths (Hunter, Elmegreen, & Ludka 2010). It is however only with the advent of wide-angle sky surveys in the last decade or so that it is finally possible to measure a sufficient volume of the local universe to sample large numbers of low-luminosity stellar systems outside the extreme environments of clusters.

One such programme is the Galaxy and Mass Assembly survey (GAMA; Driver et al. 2011), which now provides a unique opportunity to analyse the properties of low-luminosity, low surface brightness galaxies because of its very high volume (almost $20\,000$ Mpc³) at low redshifts. The GAMA survey regions were not selected on the basis of environment. With photometric and spectroscopic information for nearby galaxies in hand, in this project we aim to:

- characterise the galaxies, especially dwarfs, residing away from clusters,
- compare the intrinsic and physical properties of visually distinct galaxies with an aim to find evolutionary links between them, and
- determine the minimum number of classes which are required

to separate galaxies, especially dwarfs, into statistically distinguishable populations.

At this point we emphasise that our unique sample (8-87 Mpc) helps in bridging the gap between studies of galaxies in the local volume ($\lesssim 10$ Mpc; Sharina et al. 2008; McConnachie 2012) and statistical analyses of galaxies in rich clusters such as Virgo (16.5 Mpc) and Coma (~ 100 Mpc).

We characterise the dataset used in this paper and various selection biases that may affect our analysis in the following section. In §3 we explain the methods we used for classifying our sample: (i) the visual classification scheme, (ii) the light profile fitting algorithm which is used to quantify the morphology of galaxies through structural parameters, and, (iii) the spectral energy distribution fitting procedure used to estimate the star formation and dust properties of galaxies. We discuss the relation between various structural parameters obtained from fitting the light profile in §4, and discuss the star formation and dust properties of our sample in §5. Various implications of our analysis on the stellar mass function of galaxies and the unimodal nature of galaxies are discussed in §6, along with the limitations of our analysis. We conclude with a summary of our findings in §7.

All distances, magnitudes and masses in this paper are calculated under the assumption of a Λ CDM concordance cosmological model with $H_0 = 70$ km s⁻¹ Mpc⁻¹, $\Omega_\Lambda = 0.7$ and $\Omega_m = 0.3$.

2 DATA AND SAMPLE SELECTION

2.1 GAMA spectroscopic data

The Galaxy and Mass Assembly (GAMA) survey is a combined spectroscopic and multi-wavelength programme that exploits various ground-based and space-borne observing facilities to study cosmology and galaxy evolution. The GAMA spectroscopic campaign is based on the Sloan Digital Sky Survey (SDSS; Data Release 7) imaging complete to $r = 19.8$. The survey, after completion of Phase II will have photometry in 20 wavebands, from far ultraviolet to radio, and, spectroscopic redshifts for $\sim 300,000$ galaxies ($z \lesssim 0.25$) in three equatorial regions, each measuring 12×5 square degrees and two southern fields of similar size (Driver et al. 2011).

The surface brightness completeness limit for the GAMA survey is mainly driven by the SDSS imaging, and is important for understanding some of the selection biases in our sample. Fainter than this completeness limit the imaging data becomes unreliable. Since the targets for the spectroscopic campaign for GAMA are selected in the r -band, in Figure 1 we show the number distribution of all the GAMA II sources with $r \leq 19.8$, SURVEY_CLASS > 1 and VIS_CLASS ≤ 1 or = 255 (Baldry et al. 2010). The VIS_CLASS parameter signifies the likelihood of an object for being a target for spectroscopic follow-up. Many of the low-redshift targets are classified as VIS_CLASS = 3, implying they are deblednded part of a galaxy, and are excluded from our analysis.

The bottom panel of Figure 1 shows the fraction of galaxies with reliable redshifts (quality of redshift, $nQ \geq 3$; Baldry et al. 2010) in each bin of surface brightness, in this sample. As expected, the completeness fraction drops towards fainter surface brightness and goes below 95% at $\mu_e > 24.2$ magnitude per square arcsecond. This is in agreement with the median 1σ limit of $\mu_e = 24.60$ magnitude per square arcsecond derived for GAMA I ($r_{complete} = 19.4$; Kelvin et al. 2012). Since we are mainly dealing with low-luminosity and low-surface brightness galaxies

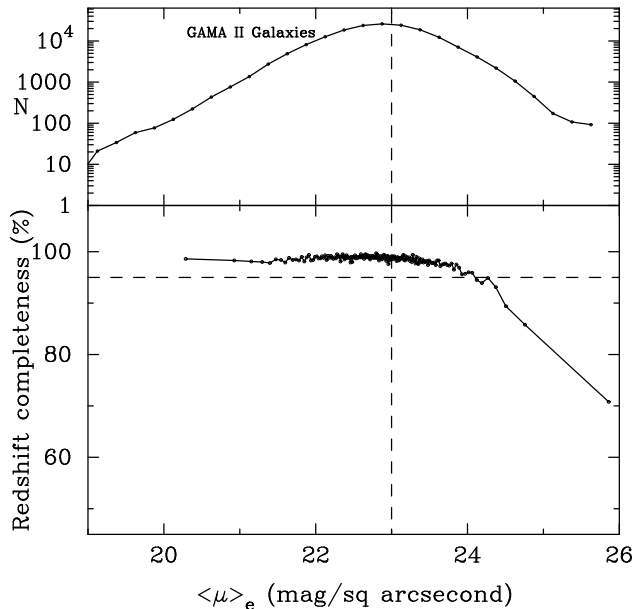


Figure 1. (Top) The number distribution of all GAMA II sources with $r \leq 19.8$ and reliable redshift. (Bottom) Redshift success rate for all GAMA II galaxies as a function of surface brightness in bins of 1000 galaxies. The horizontal and vertical dashed lines represent 95% redshift completeness and surface brightness completeness limit of $\mu_r = 23$ magnitude per square arcsecond adopted here (also for SDSS; see Blanton et al. 2005), respectively.

in this paper, we adopt a conservative surface brightness completeness limit of $\mu_e = 23.0$ magnitude per square arcsecond, consistent within the limit advocated in Blanton et al. (2005).

The sample used in this work is selected from the GAMA II catalogue such that $nQ \geq 3$ and $0.002 < z_{TONRY} \leq 0.02$ (see Baldry et al. 2012, for details on the local flow correction). This redshift range is optimally chosen to exclude galactic stars and enable visual classification of low-luminosity and low surface brightness galaxies in the shallow imaging data from the SDSS. All images were visually inspected for any remaining artefacts, stars and duplicates resulting in the final catalogue of 432 galaxies (including those fainter than $\mu_e = 23.0$ magnitude per square arcsecond) which is used throughout this paper unless stated otherwise.

All magnitudes used in this work are corrected for galactic extinction and K-corrected to $z = 0$ using `kcorr_z00v03` (Loveday et al. 2012) and `GalacticExtinctionv02` data management units for GAMA II. While GAMA II is 99% complete for galaxies down to $r = 19.8$, Figure 1 shows that the surface brightness incompleteness starts becoming significant at $\langle \mu \rangle_e > 24.2$ magnitude per square arcsecond. This limit thus restricts the overall angular size of the detected galaxies such that $\mu_{limit} = m + 2.5 \log(2\pi R_{limit}^2)$, where μ_{limit} , m and R_{limit} are the limiting effective surface brightness, magnitude and limiting size for the sample, respectively (see Graham & Driver 2005, for a detailed derivation of the formula).

2.2 The *Viking* data

The VISTA Kilo-degree Infrared Galaxy survey (*Viking*) will cover 1,500 square degrees of sky in five broadband filters (Z, Y, J, H, K_s) with the 4.1 metre Visible and Infrared Survey Telescope for Astronomy (VISTA), located at the Paranal Obser-

vatory in Chile. The *Viking* data has $2\times$ better resolution ($0.6''$), and depth approximately two magnitudes deeper than the SDSS.

Since the GAMA parent sample was selected in the r -band, while here we choose to present the analysis in the *Viking* Z -band, we need to evaluate the limiting values for all parameters in the Z -band. The difference between the SDSS r -band and the *Viking* Z -band magnitudes for our sample galaxies is in the range $-1.69 < m_r - m_Z \leq 1.96$, with a mean difference of 0.31 magnitude and standard deviation of 0.36 magnitude, respectively. For $\sim 95\%$ of the galaxies with data in both wavebands, the mean difference in the effective surface brightness within the effective radius is 0.35 magnitude per square arcsecond. Based on the above mentioned differences, in the following we assume a limiting effective surface brightness of $\mu_{limit} = 22.65$ magnitude per square arcsecond. The limiting magnitude $m_{limit} = 19.5$ mag at our $z_{max} = 0.02$ in the Z -band translates to $M_{Zlimit} = -15$.

2.3 Environment of galaxies

As we discuss in Section 1, most previous studies of low-luminosity galaxies focus on the dense environment of rich galaxy clusters. There are however, no massive clusters or groups in any of the three equatorial GAMA regions at $z \leq 0.02$ as demonstrated by the redshift distribution of GAMA galaxies in figure 5 of Baldry et al. (2012). In order to quantify environment for our sample, we also made use of the GAMA group catalogue (Robotham et al. 2011). However, due to the high lower redshift limit of this catalogue, the clustering information based on the friends-of-friends method is only available for a subsample of 263 of our galaxies at $z > 0.01$. Of these 263, 11 galaxies are found in groups with multiplicity ≥ 3 , and another 25 are in pairs.

Together, these properties demonstrate that the galaxies in our sample avoid the very dense cluster environments typical of most previous dwarf galaxy samples. This does not however imply that they are entirely isolated, but do satisfy our aim to determine the properties of galaxies evolving away from the influence of dense environments.

3 METHODOLOGY

We adopt three independent methods for the classification of galaxies: visual classification based on inspecting the five-colour SDSS images, quantitative structural analysis based on the modelling of the light profile of galaxies, and star formation properties obtained by fitting the spectral energy distribution of galaxies. We describe each of these methods below.

3.1 Visual classification of galaxy morphology

The visual appearance of galaxies has historically been used to understand their evolutionary sequence and formation mechanisms by separating them into classes. The Hubble tuning fork diagram (Hubble 1926), for instance is in use to the present day. The Hubble sequence, which was designed to be an evolutionary sequence in morphological types, can also be interpreted as a sequence in gas content, mass, and bar structure of galaxies, all of which are in a symbiotic relationship with their global star formation rate (SFR) and dynamic environment.

One of our key motivations for this project is to test if galaxies that can visually be classified into separate classes, are statistically distinguishable in other properties, indicative of an evolutionary

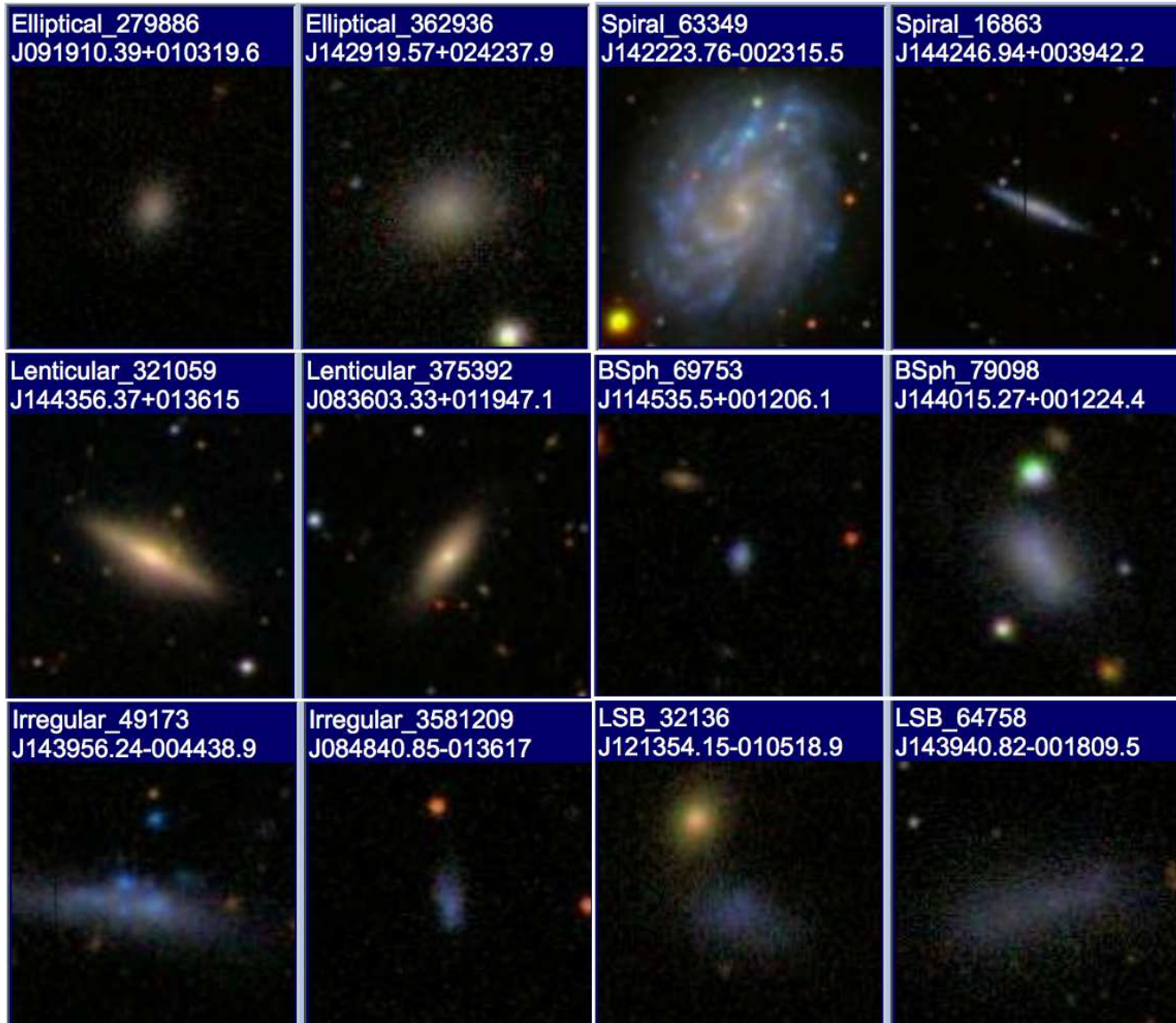


Figure 2. Montage of representative galaxies from each of the six visually identified morphology classes. The GAMA II ID and class are shown at the top of each image followed by its objID from SDSS(DR7).

link between them. To accomplish this, Mahajan visually classified the five colour SDSS images of all galaxies ($0.002 \leq z \leq 0.02$) five times, classifying them into six categories as detailed below. In this redshift range the physical resolution of the SDSS images is around $0.05 - 0.48$ kpc. The classification scheme was designed without any presumptions about how many categories were required. In other words, a new category was created based on the imaging data if there appeared to be galaxies which shared common traits in their appearance, but did not fit into any of the existing classes. Mahajan, Driver and Drinkwater also visually classified the three-colour images in the giH wavebands. On average, the various classifiers agreed 80% of the time with the low-surface brightness galaxies and blue spheroids with red centres appearing as the most doubtful cases. Amongst different types, different classifiers disagreed with themselves and with each other for galaxies mostly classified as irregulars; a non-negligible fraction of the irregular galaxies were equally likely to be classified as an LSB or a blue spheroid.

The visual classification scheme was designed to categorise

all galaxies using their broadband colour and morphology. Where possible however, broadband colour was not accounted for in order to simplify the morphology classification scheme (Figure 2). The different classes and total number of galaxies in each class are:

- *Elliptical (E; 21)*: Galaxies which are morphologically elliptical in shape. They are mostly red in colour.
- *Spirals (Sp; 47)*: Galaxies showing well-defined spiral arms or clearly identifiable edge-on disks. These galaxies often show conspicuous signs of ongoing star formation, such as HII regions, and stellar associations forming spiral arms.
- *Lenticulars (L; 26)*: Red, disk galaxies with a resolved nucleus. These galaxies are mostly big and bright, occasionally showing signs of some ongoing star formation in rings around nucleus, or low surface brightness disks without spiral arms.
- *Blue spheroidals (BSph; 73)*: Colour plays a key role in successfully identifying these galaxies. They are very blue and gener-

ally compact spheroids, morphologically similar to small elliptical galaxies or bulges of spiral galaxies.

- *Low surface brightness (LSB; 69) galaxies*: These extended objects show very poor contrast with the background in the five-band SDSS imaging. We note that a substantial fraction of these galaxies may have been misclassified due to the very shallow imaging data used here. Many of these galaxies may also be classified as irregular, and as we will show below, these two classes overlap in most of the parameter space explored here.

- *Irregulars (Irr; 196)*: All confirmed extended sources that do not belong to any of the above categories.

Some representative examples from each of the six classes are shown in Figure 2.

We did not use luminosity in our classifications and avoided labelling galaxies as dwarfs or giants based on appearance alone because the angular size of galaxies is a function of redshift. However, the resulting classifications were a strong function of luminosity with the last three classes listed above (BSph, LSB and Irr) dominating the low luminosity galaxies (as shown in Figure 7). The mean fraction of these galaxies in the whole sample is 78 per cent, with the fraction increasing from zero at $M_Z \sim -20$ to 100 per cent at $M_Z \sim -16$; it reaches the mean fraction at a luminosity of $M_Z = -18.5$. This is equivalent to the $M_r \sim -18$ mag limit for dwarf galaxies discussed above. We therefore refer to galaxies with $M_Z > -18.5$ collectively as dwarfs in order to make the following discussions more concise, unless stated otherwise.

3.2 Structural investigation of galaxy morphology with model analysis (SIGMA)

SIGMA (Kelvin et al. 2012) employs a range of image analysis software and logical filters to perform structural analysis on an input catalogue of galaxies. At the heart of SIGMA lies SOURCE EXTRACTOR (Bertin & Arnouts 1996), PSF EXTRACTOR (PSFEX; Bertin 2011) and GALFIT 3 (Peng et al. 2010), which are aided by additional packages in the fitting process. Detailed explanations for various packages and their respective roles in the fitting process are given by Kelvin et al. (2012). SIGMA only requires the image and the position of the primary galaxy therein as an input. It then outputs magnitude, fitted sky background and the fit parameters described below.

GALFIT fits a single Sérsic function to each primary galaxy with seven free parameters: object centres x_0 and y_0 , total integrated magnitude m_{tot} , effective radius along the semi-major axis R_{eff} , Sérsic index n , ellipticity and position angle. Secondary objects in the image are modelled by either a single Sérsic function or a scaled point spread function (PSF) for stars, as appropriate. The PSF comprises three free parameters x_0 , y_0 and m_{tot} . See Peng et al. (2010) for more information on the fitting process. A single Sérsic function which describes the light profile of a galaxy as a function of its radius is given by the Sérsic equation

$$I(r) = I_e \exp \left[-b_n \left(\left(\frac{r}{R_{eff}} \right)^{1/n} - 1 \right) \right], \quad (1)$$

where $I(r)$ is the intensity at radius r , I_e is the intensity at the effective radius R_{eff} , the approximate radius containing half of the total light, and n is the Sérsic index which determines the shape of the light profile of a galaxy. The value of b_n is a function of the Sérsic index. A large number of profile shapes can be fitted by varying n ; $n = 0.5$ gives a Gaussian profile, $n = 1$ an exponential profile suitable for galaxy discs, and $n = 4$ a de Vaucouleurs profile

mostly associated with spheroids such as giant elliptical galaxies (see Graham & Driver 2005, for a detailed description of the Sérsic model). The quality of fits can be judged from Figure 3, which shows a typical example of fitted light profile for each morphology class discussed above. Although $> 91\%$ of our sample galaxies are fitted well by SIGMA, the single component Sérsic model does not fit the centre of the nucleated larger galaxies, spirals and lenticulars, well. The number of galaxies in each morphology class whose light profile could not be adequately fitted by SIGMA are presented in Table 1. We note that some of the galaxies may have more than one issue and hence are counted more than once for this table. We discuss some limitations of the single component Sérsic model and its impact on the presented analysis further in §6.4.

It is however notable that only 2/21 elliptical galaxies in this GAMA sample may require a nuclear component, suggesting that in contrast to the (dwarf) elliptical galaxies in clusters (Thomas, Drinkwater, & Evstigneeva 2008), most of the low-luminosity elliptical galaxies are not nucleated (also see Oh & Lin 2000).

No explicit constraints are placed on the parameters fitted by SIGMA. However, GALFIT has an internal limitation for Sérsic index such that $0.05 < n < 20$, where the lower limit is a ‘soft’ limit and the upper bound is a hard limit (Kelvin et al. 2012).

3.3 Multi-wavelength Analysis of Galaxy Physical Properties (MAGPHYS)

The spectral energy distribution of all the galaxies in our sample was fitted using the Multi-wavelength Analysis of Galaxy Physical Properties (da Cunha, Charlot, & Elbaz 2008, MAGPHYS). The 18 band photometry was derived from cross-matching a number of distinct catalogues which include: the far ultraviolet (*FUV*) and near ultraviolet (*NUV*) *GALEX* data (GalaxyMainv02; best_mag_nuv, best_magerr_nuv, best_mag_fuv, best_magerr_fuv; see Seibert et al. in prep); the aperture matched $u - K$ photometry (Ap-Matchedv04; Kron aperture matched photometry using SExtractor, see Hill et al. (2011) for details); publicly available WISE data (Wright et al. 2010, <http://irsadist.ipac.caltech.edu/wise-allsky/>; w(1,2,3,4)mpro and w(1,2,3,4)sigmp), and bespoke measurements from the Herschel-Atlas SPIRE observations (18BandPhotometryv01) using the r -band defined apertures convolved to the relevant SPIRE band point-spread function, and taking care to apportion flux from overlapping targets as described in Appendix A1 of Bourne et al. (2012). A full description of the analysis of the assembly of these data will appear in Driver et al. (2014, in prep).

To overcome the background confusion level and the bright high redshift interlopers in the SPIRE bands, measurements were also made in apertures of identical area to the target galaxies but placed at random locations across the SPIRE data (Driver et al. 2014, in prep). The mean flux in the random apertures were then subtracted from the flux measured at the location of the target galaxy, and a flux error assigned based on the quadrature combination of the remaining measured target flux and the variance of the random measurements.

The 18 band measurements were then converted from magnitudes to Janskys and fed into MAGPHYS (da Cunha, Charlot, & Elbaz 2008). MAGPHYS then compares the data to an extensive library of stellar population and dust templates to derive fundamental measurements given flux measurements, the redshift and the known filter-set. For each galaxy, MAGPHYS provides both best-fit values for derived parameters, and marginalised median and quartile values. Here we adopt the

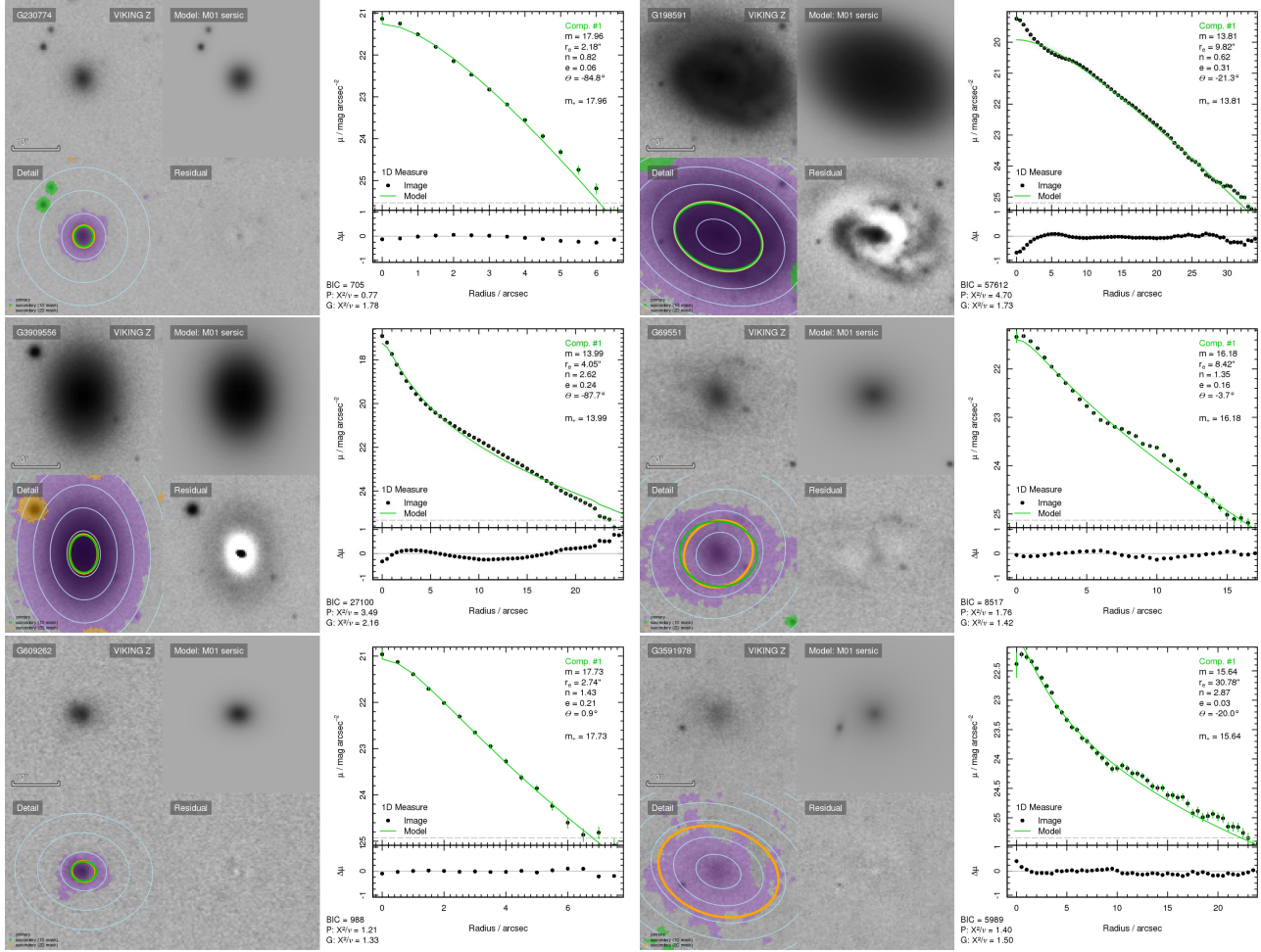


Figure 3. The surface brightness fits in the VIKING Z-band for some of the galaxies in our sample. (left to right from top:) An elliptical, spiral, lenticular, irregular, blue spheroid and low surface brightness galaxy, respectively. Each panel shows (clockwise from top left:) the original Viking Z-band image, Sérsic model, 1D light profile (with residuals: image-model at the bottom), residual image and, ellipses centred on the primary galaxy used for estimating the light profile along with masked objects, respectively.

Table 1. Success rate for SIGMA and MAGPHYS fitting for different morphological classes.

Class →	Elliptical	Lenticular	Spiral	BSph	Irr	LSB
Number (Percentage) of galaxies in each class →	21 (4.86)	26 (6.01)	47 (10.88)	73 (16.89)	196 (45.37)	69 (15.97)
Issues ↓						
SIGMA fails to fit the centre of the galaxy	2	8	21		15	7
SIGMA fails to fit the outskirts of the galaxy		4	4	6	16	3
SIGMA fits a profile with $\Delta\mu > 0.2 $ in at least one bin	2	5	7	2	18	6
SIGMA fails to fit a surface brightness profile		1	8		13	5
Companion object(s) interfering with the light profile fitting	2	1		2	12	3
MAGPHYS fit not good	1	1	5		5	1
MAGPHYS fails to fit the SED				1	1	

median and quartile values (which on occasion may lie significantly adrift from the best-fit values), for the following parameters: stellar mass M^* , dust mass M_{dust} , star formation rate and the specific star formation rate.

4 THE STRUCTURAL PARAMETERS OF MORPHOLOGICALLY DISTINCT GALAXIES

In this section we analyse the relation between various structural properties of galaxies estimated by fitting their light profile in the *Viking Z*-band with a single component Sérsic model.

In order to test if the visual morphology of a galaxy is associated with its light profile, in Figure 4 we show the distribution of different morphological classes in surface brightness and Sérsic

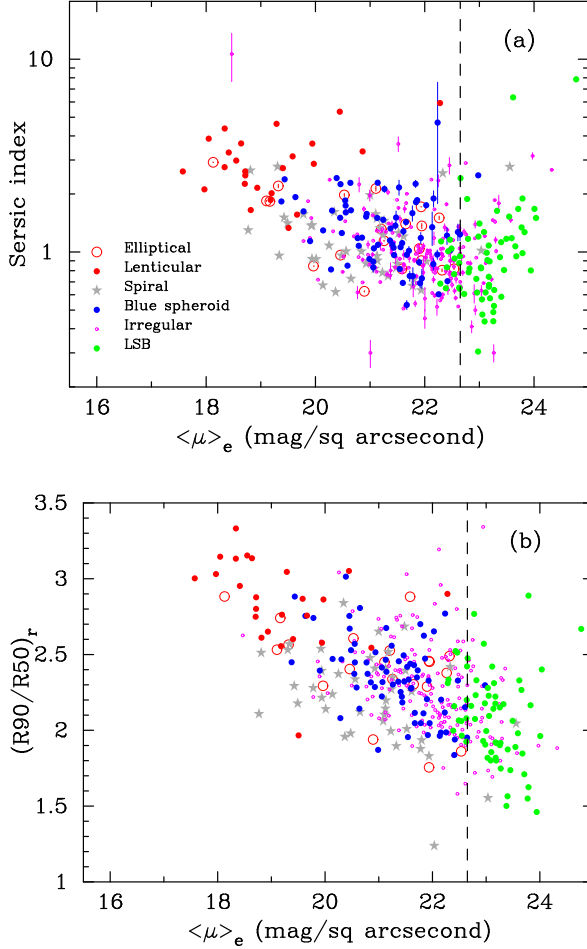


Figure 4. (a) The Sérsic index n is shown as a function of the effective surface brightness $\langle\mu\rangle_e$ for morphologically distinct galaxies in our GAMA II sample. Different morphological classes are represented by different colours and symbols as per the legend, and measurement uncertainties are shown for all classes except the LSBs to maintain clarity. The vertical dashed line represents the surface brightness completeness limit for our sample in the Z -band. (b) The same, but shown as a function of the concentration index $(R_{90}/R_{50})_r$. These figures show that albeit with a huge scatter, the effective surface brightness is correlated with the light concentration of galaxies, and the correlation between n and $\langle\mu\rangle_e$ seen in the upper panel is not a result of parameter coupling in the Sérsic model.

index. Although there is almost an order of magnitude of scatter in n at any $\langle\mu\rangle_e$, a trend is clearly visible, such that higher surface brightness galaxies have higher stellar concentration towards the centre. In the bottom panel of Figure 4, we show the same replacing n with the concentration index $(R_{90}/R_{50})_r$, where R_x is the Petrosian radius containing $x\%$ of the galaxy’s light in the r -band. Together, these figures show that the correlation between n and $\langle\mu\rangle_e$ does not result from parameter coupling in the Sérsic model.

Several authors have suggested a correlation between the physical radius and n for luminous elliptical and lenticular galaxies, as well as dEs in clusters (e.g. Caon, Capaccioli, & D’Onofrio 1993; Caon & Einasto 1995; Young & Currie 1995). Even though our sample does not include classical $n = 4$ ellipticals, in Figure 5 we find a similar correlation for the bulge-dominated galaxies in this sample. It is particularly interesting that the star-forming blue spheroids form a continuous distribution with the passively evol-

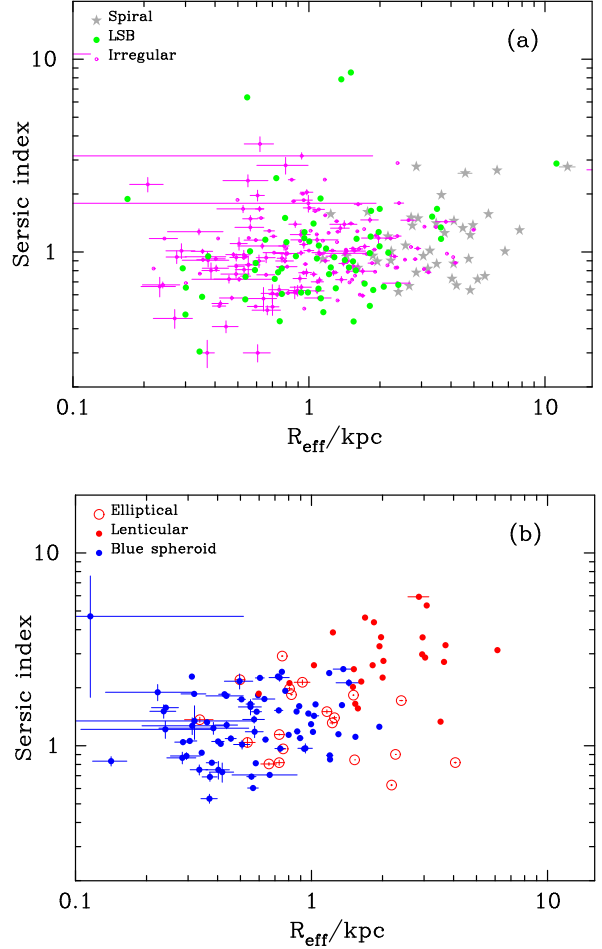


Figure 5. The Sérsic index n as a function of the effective radius R_{eff} for (a) disk-dominated, and (b) bulge-dominated galaxies, respectively. Unlike the n - $\log R_{eff}$ correlation presented in the literature (see text), and seen here for the bulge-dominated galaxies, we do not find any apparent trend between $\log R_{eff}$ and n for non-spheroidal galaxies.

ing lenticulars in the n - R_{eff} plane such that the product moment correlation $r = 0.414$, while including ellipticals slightly worsens the correlation with $r = 0.365$ (Figure 5(b)). Figure 5(a) shows that although spiral galaxies are somewhat disjoint from the rest of the ensemble, none of the disk-dominated categories (Spirals, Irregulars and LSBs) show any apparent trend in this parameter space ($r = 0.134$). Some of the intrinsic scatter within each morphological class may be due to the measurement uncertainties, or the incapability of the single component Sérsic model to fully represent the stellar distribution in low-mass, faint galaxies represented by the irregulars, blue spheroids and LSBs. However, since the fraction of galaxies in each class where the light profile is not estimated properly is small (Table 1), we claim that the lack of a correlation between R_{eff} and $\langle\mu\rangle_e$ for non-spheroidal galaxies is real. The limitations of our data and methodology are discussed further in §6.4.

Figure 6 shows the distribution of morphologically distinct galaxies in the plane spanned by effective surface brightness and the effective radius. As in Figure 4, the different classes split into distinct populations, but with considerable overlap. We quantify this observation using the Kolmogorov-Smirnov (K-S) statistical probability for various parameters for all pairs of visually distinct galaxies in Appendix A. Table A5 for R_{eff} shows that ellipticals

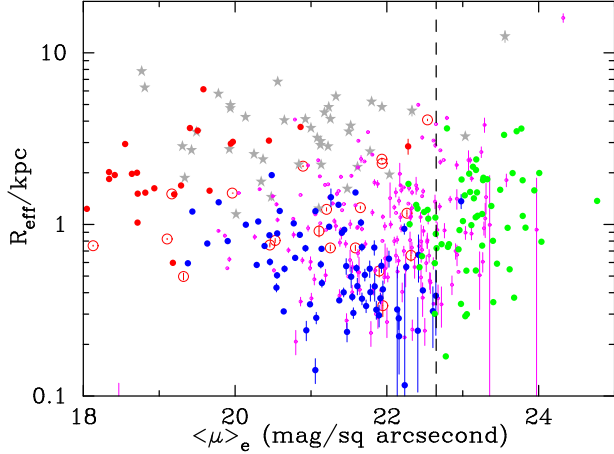


Figure 6. Effective radius, R_{eff} as a function of the effective surface brightness within R_{eff} , $\langle\mu\rangle_e$ for all galaxies. The symbols and colours are same as in Figure 4, and the vertical dashed line is the surface brightness limit for our sample. The different locus of various morphological types shows the unconscious bias affecting visual classification.

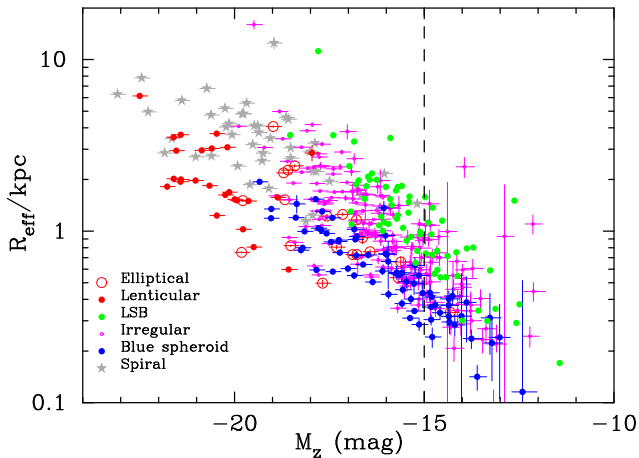


Figure 7. Effective radius, R_{eff} as a function of the Z -band magnitude for galaxies in our sample. The symbols and colours are same as in Figure 4. The vertical dashed lines represents the limiting magnitude for our sample. This figure shows that albeit with some scatter, all galaxies show a monotonic relation between R_{eff} and luminosity, such that the more luminous galaxies tend to be larger. At lower luminosities, LSBs tend to be larger than the blue spheroids and the irregulars, while at the bright end, spirals have larger R_{eff} relative to the lenticulars.

are likely to have a statistically similar range in R_{eff} to all types of galaxies except spirals and lenticulars. The LSBs and irregulars are also likely to have similar R_{eff} distributions. Table A8 on the other hand shows that the $\langle\mu\rangle_e$ distributions are statistically different for all types of galaxies except ellipticals, which show statistically similar distribution to spirals as well as blue spheroids. Without the visual classification however, the sample fails to segregate into distinct population. Such a lack of correlation between $\langle\mu\rangle_e$ and R_{eff} has previously been observed among luminous galaxies (e.g. Boyce & Phillipps 1995).

Figure 7 shows the distribution of the galaxies in the Z -band magnitude and R_{eff} plane. This figure adds strength to the above argument, by showing that the morphologically distinct galaxies have a distinct locus in the 3D space mapped by absolute magnitude, surface brightness and effective radius, but *they all form a*

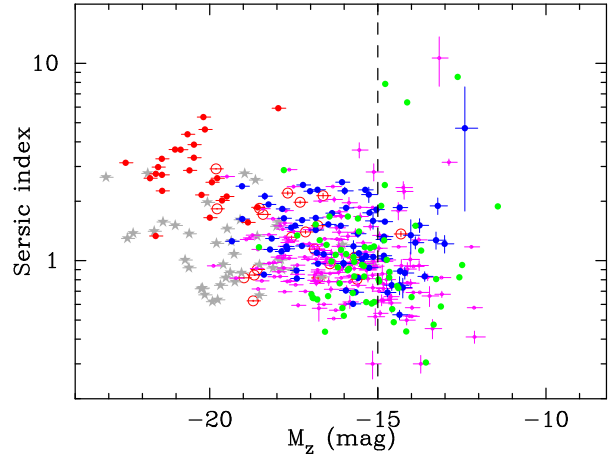


Figure 8. Sérsic index of galaxies as a function of M_Z . The vertical line marks the magnitude limit for our sample, and symbols and colours are same as in Figure 4. While giant galaxies seem to segregate by morphological type, the low-luminosity galaxies form a unimodal population. It is noteworthy that the unimodality of the low-luminosity galaxies is not entirely a consequence of the measurement uncertainties in the estimated parameters, shown for all galaxies except the LSBs (see Table 2).

continuous distribution in this parameter space representing mass and angular momentum.

A similar trend continues in the plane spanned by absolute magnitude (M_Z) and n . As shown in Figure 8, giant galaxies split into two branches, the passively evolving lenticulars at the high- n end ($n > 2$), and spirals at low- n ($n \lesssim 1$). The dwarfs on the other hand do not show any trend in this space with LSBs, irregulars and blue spheroids forming a single ‘clump’ in the dwarf regime ($M_Z > -18$). The product moment correlation for the entire sample in this plane is $r = -0.210$, which reduces to a negligible -0.065 when only dwarfs are considered. The K-S statistic suggests that among dwarfs, blue spheroids and ellipticals, and irregulars and LSBs show statistically similar distributions in n (Appendix A). The n distribution of spiral galaxies is also statistically similar to irregulars and LSBs.

Together, Figures 4 to 8 show that while Sérsic index is a good quantitative representation of the morphology of luminous galaxies, it is not an effective measure for dwarfs, even at the low redshift ($z \leq 0.02$) considered here.

Figure 9 shows the bivariate brightness distribution (BBD) for our sample in the M_Z - $\langle\mu\rangle_e$ (μ_0) planes. It is remarkable how different morphological classes segregate into distinct spaces in the plane mapped by the (central/effective) surface brightness and absolute magnitude, but link to form a continuum across the entire range. This is reflected in the K-S statistical probabilities for different types of galaxies showing the distribution of $\langle\mu\rangle_e$, μ_0 and magnitude to be drawn from the same parent population. Appendix A shows that the distributions of $\langle\mu\rangle_e$ and μ_0 are different for all combinations of visually distinct galaxies except the ellipticals and blue spheroids, and, ellipticals and spirals. It is interesting that this trend is seen in the more commonly used μ_0 and not just in $\langle\mu\rangle_e$, because $\langle\mu\rangle_e$ can be measured more robustly relative to μ_0 whose measure is strongly affected by the quality of seeing (Graham & Guzmán 2003; Boselli et al. 2008).

The giant galaxies in Figure 9 split into two sequences, one of star-forming spirals and the other formed by lenticulars. The

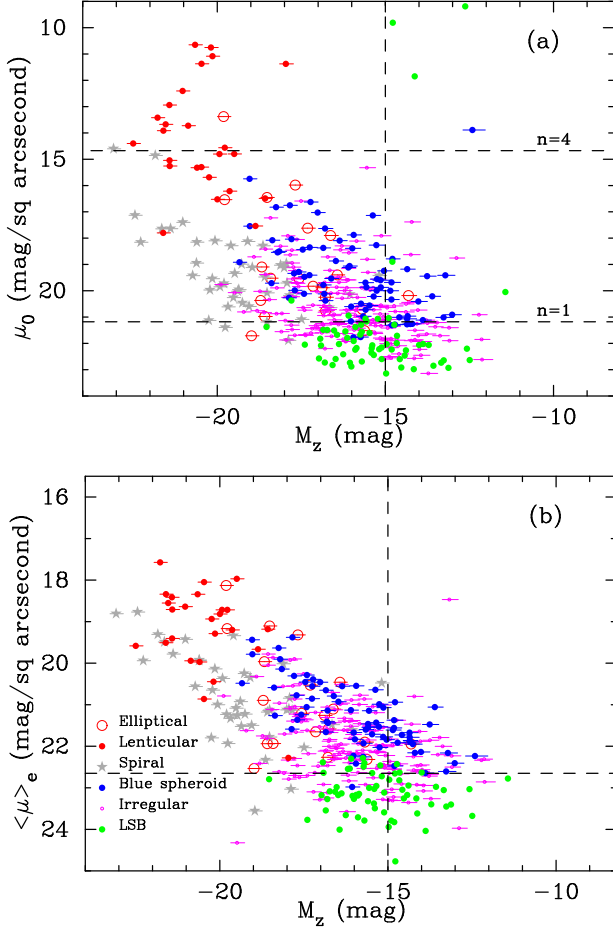


Figure 9. (a) Central surface brightness and (b) effective surface brightness, respectively as a function of M_Z for our sample. The symbols and colours are as in the legend. The vertical dashed line marks the limiting magnitude for our GAMA II sample. For the central surface brightness (*top*), the horizontal dashed lines represent the expected values corresponding to the Sérsic index, $n=1$ and 4 using the formulae derived by Graham & Driver (2005). In the *bottom* panel the horizontal line represents the surface brightness completeness limit for our sample.

lenticulars in particular have statistically distinctive distribution of $\langle\mu\rangle_e$ relative to all other types of galaxies (Appendix A). At a given magnitude, spirals on an average have $\langle\mu\rangle_e$ fainter by around two magnitudes per square arcsecond than the lenticulars, and μ_0 fainter by around four magnitudes per square arcsecond, respectively.

To summarise, in this section we have shown that although morphologically distinct galaxies have different centroids and statistically different distributions in many of the 2D parameter spaces mapped by luminosity and the structural parameters, the different classes especially among the dwarf galaxies overlap significantly. Table 2 shows the mean, standard deviation and the standard error on mean for $\langle\mu\rangle_e$, μ_0 , R_{eff} and n , respectively for different morphological types. Table 2 shows that despite being empirically distinctive, the distributions of structural parameters of visually distinct galaxies are very broad and overlap extensively, even though the measurement uncertainties are negligible relative to the scatter within each class. The broad width of the distributions, specifically of R_{eff} and μ_0 (i) may be an indication of different evolution histories of morphologically similar galaxies, or, (ii) a result of an artificial categorisation of a continuous galaxy morphology distribution

into discontinuous classes. The same however may also indicate (i) the inefficient modelling of the light profile for some structural parameters such as μ_0 (Table 1; also see §6.4), and, (ii) uncertainty in our visual classification of the low-luminosity galaxies. Better imaging data and multi-component modelling of light profile are required to confirm this.

5 STAR FORMATION AND DUST IN MORPHOLOGICALLY DISTINCT GALAXIES

With the aim of getting further insight into the differences between morphologically distinct galaxies, we fitted the multi-band photometry available for most of our sample, with MAGPHYS (§3.3). Here we describe how morphologically distinct galaxies behave in the parameter space formed by the star formation and dust properties of galaxies.

Figure 10 shows the well-known trend in the star formation rate (SFR) with M^* , viz., SFR increases, while specific SFR, SFR/M^* (sSFR henceforth) decreases with increasing stellar mass of galaxies, respectively. In this sample spanning almost four orders of magnitude in M^* , once again the star-forming and passive giant galaxies occupy very different ranges in both, M^* and sSFR. Since our sample does not include clusters (§2), it is unsurprising that the visually identified elliptical galaxies are mostly low-mass ellipticals ((d)ellipticals, henceforth) and not the classic giant ellipticals dominating the cores of rich clusters.

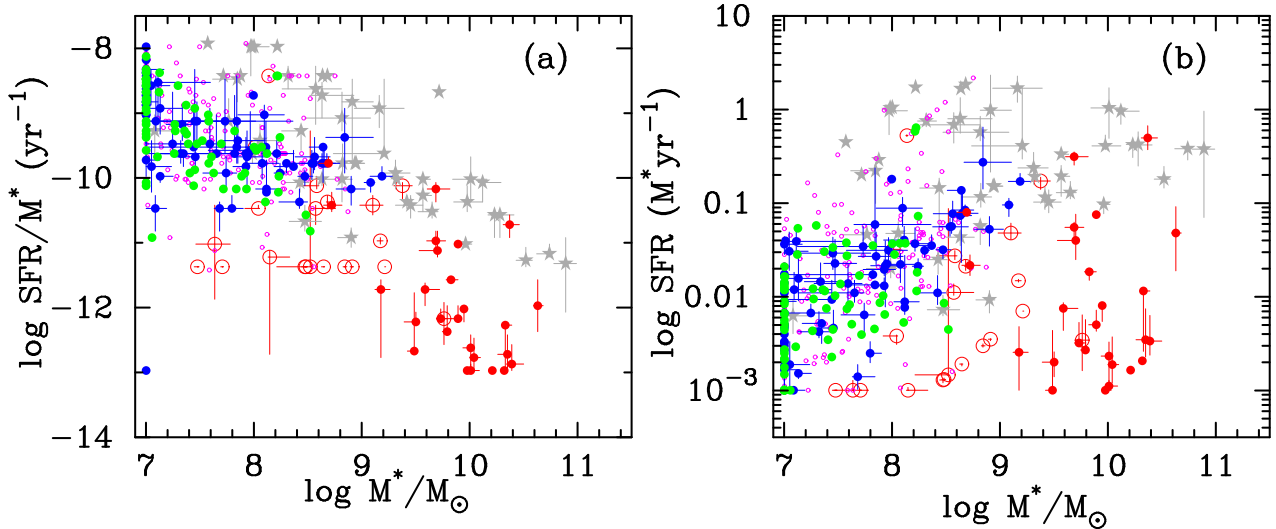
As an ensemble all dwarf galaxies ($M^*/M_\odot \lesssim 10^9$) show steadily increasing SFR with M^* . The K-S statistic suggests that the distribution of M^* for all except the irregulars and blue spheroids, and SFR for all except the lenticular and (d)ellipticals, and lenticulars and LSBs, respectively are statistically different (Appendix A). The sSFR distribution on the other hand is similar for the irregulars, blue spheroids and the LSBs, thereby showing a unimodality among the star-forming dwarf galaxies. Some of this overlap in the star formation properties of dwarf galaxies may however be a result of the scatter within each morphological class as suggested by the mean and standard deviations for the distributions of SFR, sSFR, M^* and M_{dust} for visually distinct galaxies (Table 3).

If a galaxy has spent most of its life away from the influence of cluster and large groups, it may be assumed that the phase of its passive evolution begins once it runs out of all the cold gas which would have fuelled star formation. Under these assumptions, Figure 10 qualitatively indicates that large spiral galaxies such as those probed here, turn into passively evolving lenticulars, while the low-mass galaxies may turn into small red (d)ellipticals. This observation is based on the distribution of galaxies in both panels of Figure 10, where the (d)ellipticals appear in the same mass range as star-forming dwarf galaxies, but with extremely low star formation.

Figure 11 shows the distribution of our galaxies with UV data on the $(NUV - r)$ versus M_z colour magnitude diagram. The UV-optical colour of (d)elliptical galaxies shows that at fixed luminosity, they are on average redder than their counterparts, while all the other dwarf galaxies form a unimodal population. In the high luminosity regime on the other hand, the lenticulars and spiral galaxies clearly segregate, such that the lenticulars are on average around 2.5 magnitude redder than the spiral galaxies. This figure further strengthens the above hypothesis that a star-forming dwarf galaxy (LSB, irregular or blue spheroid) may evolve into red (d)elliptical, while a spiral galaxy will turn into a lenticular.

Table 2. Mean (standard deviation) and the standard error on mean of the distribution of structural parameters for different morphological classes.

Class → Parameter ↓	Elliptical	Lenticular	Spiral	BSph	Irr	LSB
μ_0 (mag/sq arcsecond)	17.95(5.45) \pm 1.19	14.20(2.02) \pm 0.40	19.25(1.51) \pm 0.22	19.10(3.20) \pm 0.37	20.18(3.95) \pm 0.28	21.15(3.45) \pm 0.41
$\langle\mu\rangle_e$ (mag/ sq arcsecond)	22.06(1.64) \pm 0.36	20.37(1.07) \pm 0.21	21.59(1.02) \pm 0.15	22.19(0.96) \pm 0.11	22.63(1.01) \pm 0.07	23.79(1.31) \pm 0.16
R_{eff} (kpc)	5.21(4.11) \pm 0.90	9.43(6.56) \pm 1.29	19.91(12.24) \pm 1.78	3.36(2.68) \pm 0.31	6.03(5.05) \pm 0.36	5.74(4.30) \pm 0.52
n	2.05(2.95) \pm 0.64	3.01(1.13) \pm 0.22	1.24(0.56) \pm 0.08	1.59(1.70) \pm 0.20	1.29(1.68) \pm 0.12	1.37(1.57) \pm 0.19

**Figure 10.** (a) sSFR and (b) SFR of different types of galaxies as a function of $\log M^*$. The symbols and colours are the same as in Figure 4. All parameters shown here are the medians of the PDFs obtained for the corresponding parameter from the spectral energy distribution fitting using MAGPHYS (see text). The lower and upper uncertainties represent the 16th and the 84th percentiles of the PDFs, respectively.

In Figure 12 we analyse the dust mass of galaxies in our sample as a function of M^* and the total SFR, respectively. Although with considerable scatter, the sample shows a trend of increasing dust mass M_{dust} , with M^* , and SFR. As expected, spiral galaxies have the highest SFRs and dust masses at fixed M^* (also see Table 3). The lenticulars on the other hand, have at least an order of magnitude lower M_{dust} at fixed M^* , and several orders of magnitude lower SFRs on average. K-S test probabilities (Appendix A) suggest that the distribution of M_{dust} for (d)ellipticals is statistically similar to irregulars and blue spheroids, while that of the latter two is similar to LSBs as well.

The star-forming dwarf galaxies; irregulars, blue spheroids and LSBs, show a monotonic increase in M_{dust} with M^* and SFR, but no apparent trends segregating different classes. At fixed M^* , (d)ellipticals have lowest M_{dust} . This trend however reverses with SFR, such that at fixed SFR (d)ellipticals have the highest dust masses relative to other dwarf classes. In the SFR- M_{dust} space, (d)ellipticals mark the lower limit of passively evolving galaxies in M_{dust} .

If we assume that dust is not easily destroyed in galaxies, and that galaxies can not produce dust once their star formation turns off, the trend seen in Figure 12(a) suggests that the dwarf galaxies will turn into (d)ellipticals as they age. On the other hand if M_{dust} in the giant spiral galaxies is fixed when star formation turns off, they will fade into the lenticular class identified here. The fact that M_{dust} of lenticulars is on average a magnitude lower than the spirals does not contradict this hypothesis because star-forming galaxies tend to loose dust via outflows such as supernovae explo-

sions during their “active” phase. Therefore the total dust mass of a galaxy towards the end of its star-forming phase is likely to be less than at the peak of its star formation activity. These observations thus imply that star-forming dwarf galaxies may have evolved into (d)ellipticals, while lenticulars succeed star-forming spirals.

To summarise, the star formation and dust properties derived from MAGPHYS efficiently separate giant galaxies into spirals and passively evolving lenticulars, but fail to distinguish the morphologically distinct star-forming dwarf classes. This is likely a consequence of intrinsic scatter within each class and not the measurement uncertainties as indicated by Table 3. The (d)ellipticals on the other hand, span a distinguishable range of SFR and sSFR relative to other galaxies of similar luminosity.

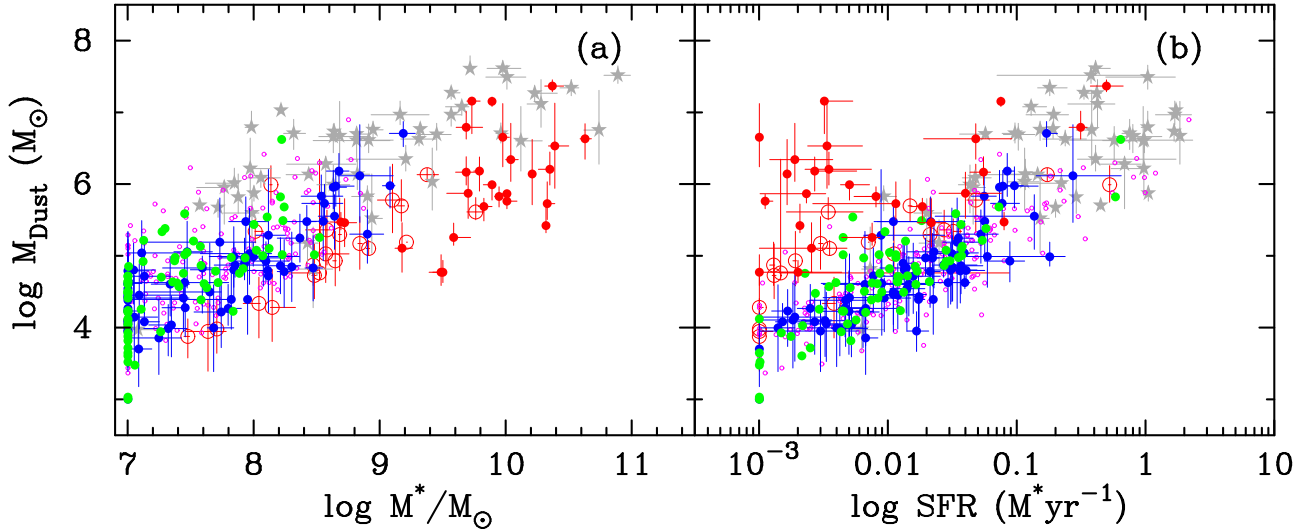
6 DISCUSSION

6.1 Stellar mass distribution for different morphological types and a comparison with the cluster environment

In Figure 13 we show the stellar mass function for all the galaxies in our sample, function corresponding to different morphological classes and a representative function for cluster galaxies. The cluster mass function was presented by Baldry, Glazebrook, & Driver (2008, see their figure 7) for clusters in the RASS-SDSS galaxy cluster survey (Popesso et al. 2006). Using the relation given by Taylor et al. (2011) for our sample complete to $M_Z = -15$ at $z = 0.02$, the stellar mass completeness ranges from $7.15 \leq \log M^* \leq 7.99$, corresponding to $0 \leq (g-i) \leq 1.2$. The upturn at $\log M^* <$

Table 3. Mean (standard deviation) and the standard error on mean of the distribution of star formation and dust properties for different morphological classes.

Class → Parameter ↓	Elliptical	Lenticular	Spiral	BSph	Irr	LSB
$\log \text{SFR} (M_{\odot} \text{yr}^{-1})$	$-2.20(0.77) \pm 0.17$	$-2.03(0.75) \pm 0.15$	$-0.65(0.64) \pm 0.09$	$-1.83(0.57) \pm 0.07$	$-1.74(0.69) \pm 0.05$	$-2.05(0.56) \pm 0.07$
$\log \text{SFR}/M^* (\text{yr}^{-1})$	$-10.84(0.81) \pm 0.18$	$-11.89(0.93) \pm 0.18$	$-9.60(1.00) \pm 0.14$	$-9.42(0.73) \pm 0.08$	$-9.28(0.71) \pm 0.05$	$-9.30(0.66) \pm 0.08$
$\log M^* (M_{\odot})$	$8.55(0.58) \pm 0.13$	$9.86(0.47) \pm 0.09$	$8.95(0.92) \pm 0.13$	$7.68(0.64) \pm 0.07$	$7.63(0.54) \pm 0.04$	$7.41(0.45) \pm 0.05$
$\log M_{\text{dust}} (M_{\odot})$	$5.02(0.64) \pm 0.14$	$6.06(0.65) \pm 0.13$	$6.41(0.75) \pm 0.01$	$4.74(0.66) \pm 0.08$	$4.85(0.65) \pm 0.05$	$4.64(0.64) \pm 0.08$


Figure 12. (a) Dust mass in different types of galaxies as a function of M^* and (b) the total SFR, respectively. All quantities plotted are determined from spectral energy distribution fitting using MAGPHYS (see text). The symbols and colours are the same as in Figure 4. The plotted values are the median, while uncertainties are the 16th and the 84th percentile values obtained from the probability distribution function for all the parameters for each galaxy, respectively.

9 seen in clusters (also see Jenkins et al. 2007; Yamanoi et al. 2012, and references therein) is apparent in our GAMA sample as well, consistent with other comparisons of galaxy stellar mass function in different environments (Baldry, Glazebrook, & Driver 2008). The low-mass upturn in clusters is evidently steeper than for our GAMA mass function, suggesting that the slope of the mass function at the low-mass end is environment-dependant. Note, however, that the cluster distribution is derived by statistical subtraction of background galaxy counts.

There is also a marked difference in the galaxy populations which contribute to the low-mass branch of the mass function. It is noteworthy that in the environments probed here, LSBs, blue spheroids and irregulars, all of which are forming stars (Figure 10), dominate at $\log M^* < 9$, while in clusters this branch is mostly populated by passively evolving dEs. Deep, spectroscopic data homogeneously sampling different environments are thus required to confirm the differences seen in the faint-end slope of the mass function for different environments.

6.2 The bivariate brightness distribution: A comparison with literature

Generically speaking, galaxies should be describable using two independent parameters (Brosche 1973; Kodaira, Okamura, & Watanabe 1983). While several combinations of such parameters have been explored in literature (e.g. Caldwell 1983; Kodaira, Okamura, & Watanabe 1983; Binggeli, Sandage, & Tarengi 1984; Kodaira 1989;

Young & Currie 1995; Prugniel & Simien 1997), the bivariate brightness distribution (BBD) appears to be the most enlightening in the context of structural evolution of galaxies during their life cycle. The BBD showing the natural morphological segregation in the space mapped by central surface brightness and absolute magnitude has been shown by several authors (Caldwell 1983; Binggeli, Sandage, & Tarengi 1984; Kormendy 1985; Binggeli 1994). From these seminal works, in Binggeli’s words, “The $M - \mu$ diagram for stellar systems might well become the equivalent of the HR diagram for stars”.

In this work, the BBD in its two forms comparable to Binggeli’s figures is shown in Figure 9. The first conclusion that can be drawn from these figures is that the effective surface brightness within the effective radius is directly correlated with luminosity. This is an intuitive, yet non-obvious inference. In other words, low-luminosity, high surface brightness galaxies (e.g. M60-UCD1; Strader et al. 2013), and luminous, low surface brightness galaxies like Malin I (Bothun et al. 1987) are very rare.

In the environments probed here, the giant galaxies split into two categories such that the high surface brightness branch comprises red and passively evolving lenticular galaxies (Figures 10 and 11). The low surface brightness branch on the other hand is made up of star-forming spiral galaxies. The dwarfs do not show a trend with luminosity; rather each class of dwarf galaxies spans the same range of luminosity but the effective surface brightness increases between classes, in increasing order: LSBs, the irregulars and the blue spheroids (also see Table 2). This trend is visible with

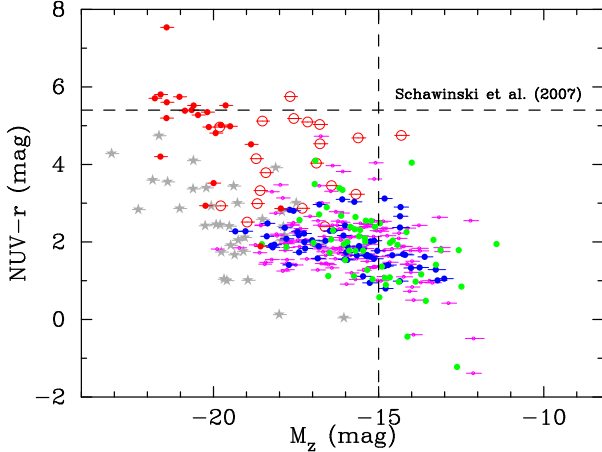


Figure 11. $NUV-r$ colour as a function of the z -band absolute magnitude for all galaxies in our sample. The symbols and colours are the same as in Figure 4. The dashed horizontal line at $NUV-r=5.4$ is the lower limit for the passively evolving early-type galaxies in the SDSS (Schawinski et al. 2007), while the vertical dashed line marks the magnitude completeness limit of our sample. This figure shows that the passively-evolving lenticulars are on average around 2.5 magnitude redder than the spiral galaxies of similar luminosity. In the low-luminosity regime however, all dwarf galaxies except the ellipticals form a unimodal population. The ellipticals on average are redder than all other dwarfs of similar luminosity.

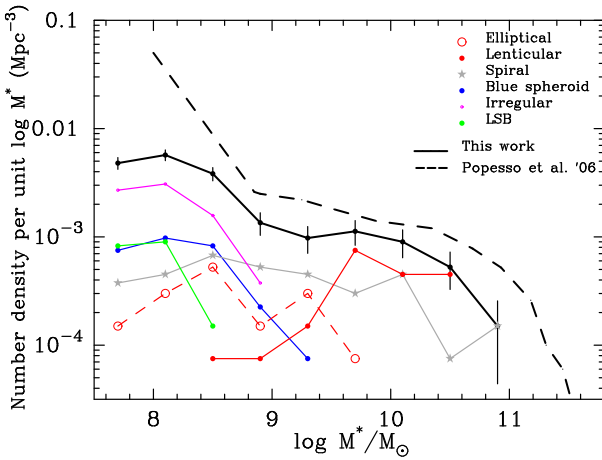


Figure 13. The stellar mass function of galaxies derived for individual morphological classes, the composite for the entire sample, and the composite of cluster galaxies based on the luminosity function derived by Popesso et al. (2006) and converted to M^* by Baldry, Glazebrook, & Driver (2008). The mass function for cluster galaxies was scaled down by a factor of five to aid in comparison. Poissonian uncertainties are shown for the composite mass function for GAMA II sample. While an upturn at $\log M^* < 9$ is apparent in both samples, the low-mass end in clusters rises more sharply compared to the environments probed by our GAMA sample. It is also evident that the low-mass branch in the GAMA sample is mainly populated by the LSBs, Irregulars and blue spheroids, while in clusters dEs are more prominent at this end.

μ_0 and $\langle \mu \rangle_e$, although the segregation and mild trend with luminosity is more obviously seen with the latter.

Most of the literature following the papers published in the early 1980s either concentrated on resolving the dichotomy between giant and dwarf elliptical galaxies (e.g. Graham & Guzmán 2003; Kormendy & Bender 2012; Graham 2013, and references therein), or the evolutionary sequence from bulges of spiral galaxies

to dwarf ellipticals (e.g. Boselli et al. 2008; Toloba et al. 2011, and references therein). But since dwarf ellipticals have predominantly been discovered in nearby rich clusters such as Virgo and Coma, all such papers are somewhat related to the impact of environmental processes in clusters on the structural properties of galaxies. Our work in this paper, on the other hand, gives a complimentary view of galaxy properties, particularly dwarfs, in environments outside rich clusters.

6.3 On the issue of statistically distinct populations of galaxies in the local Universe

One of our primary aims was to objectively test how many distinct classes are required to classify the nearby galaxy population, as opposed to the number of classes we inferred from our visual classifications. As shown by the 2-dimensional parameter plots in §4 and §5, galaxies with different visual classifications have overlapping distributions at fainter luminosities, but appear to split into star-forming and passively-evolving classes at brighter luminosities. It is still possible that the dwarf galaxies separate into different clusters in the full multi-dimensional parameter space that may have been missed by the two-dimensional projections. We therefore applied a quantitative statistical analysis to test our data for evidence of distinct groups in the multi-dimensional parameter space.

6.3.1 Application of the Clustering Algorithm

We used the “k-means” algorithm (MacQueen 1967) to decompose the data into 2-20 clusters. For a given number of clusters, this finds the cluster positions that minimise the sum of the squares of the distances from each data point to its cluster centre. We determined the best number of clusters to choose by using the Calinski & Harabasz (1974) variance ratio criterion (as implemented in the NbClust package Charrad et al. 2012). This involves choosing the partition that gives the highest ratio of the variance of the distances between objects in different clusters to the variance of distances between objects within clusters. Calinski & Harabasz (1974) plot the variance ratio as a function of the number of clusters and use the first local maximum to define the best number of clusters. We modified this criterion by requiring that the selected peak was significantly (3 standard deviations) above the neighbouring points¹. On rare occasions a double peak was formed by two variance ratios with the same value (to within 1 standard deviation) that was still significantly higher than the neighbouring points. In this case the first of the two (always $n = 2$ clusters) was chosen.

Before starting the clustering analysis we removed any objects with missing data and then scaled the remaining galaxies to have a mean of zero and standard deviation of unity in each parameter. All the parameters we analysed were logarithmic measurements (or magnitudes): the algorithm failed to find a preferred number of clusters when applied to linear units, apparently because the distribution in each parameter consisted of a group plus one or two extreme outliers.

¹ The NbClust package (Charrad et al. 2012) does not take into account uncertainties in the parameters and does not calculate uncertainties in the variance ratios, so we used a Monte Carlo approach to estimate uncertainties. We recalculated each variance ratio 30 times using different initial random seeds for the k-means algorithm and used the resulting distributions to estimate the uncertainties: see Appendix B.

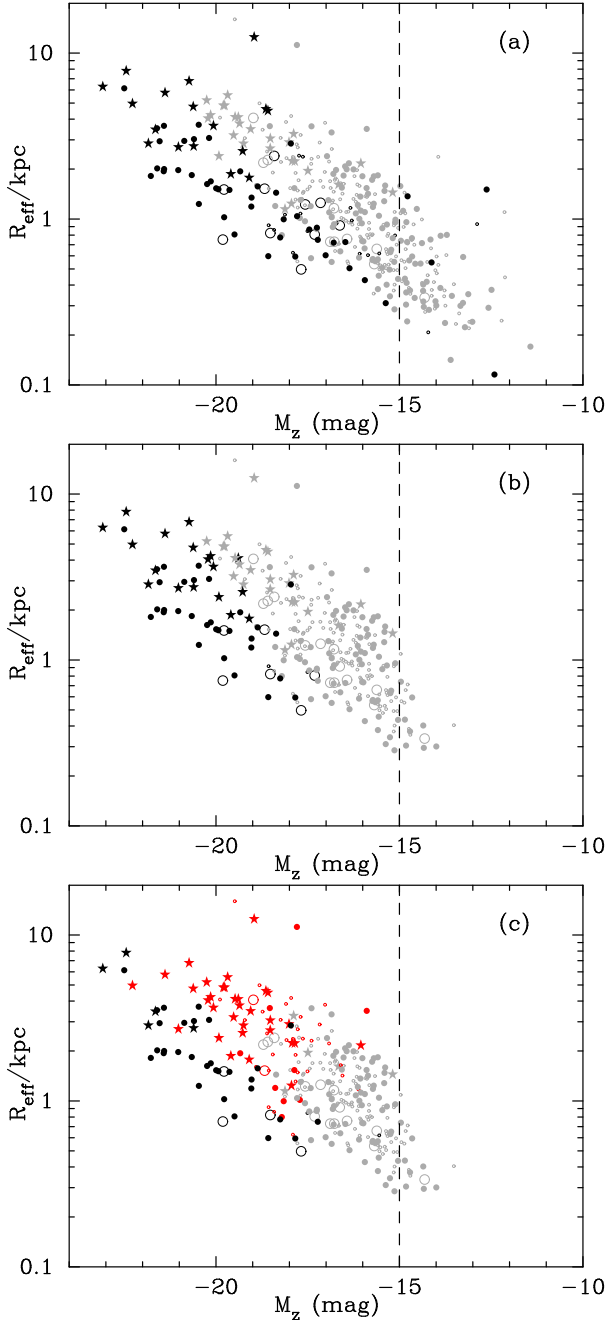


Figure 14. Automated classification of galaxies in our sample. This is the same as Figure 7, but colour-coded for “clusters” identified by the k-means algorithm using different combinations of parameters: (a) Structural+SF 1, and (c) Structural+SF 3, respectively as discussed in the text (also see Table 4). The symbol types are same as in the above figures. Two clusters are preferred statistically for this sample for all but one combination of parameters. There is however no statistical evidence for any partition among the low-luminosity galaxies.

6.3.2 Clustering with Structural Parameters

We first applied the analysis to the four structural parameters from SIGMA (μ_0 , $\langle\mu\rangle_e$, $\log n$ and R_{eff}) and the optical ($g-i$) colour giving a sample of 405 galaxies. We define this as our reference data set as it is closest to the information we used for our visual classifications. For these parameters, two clusters were preferred over any other partition of the data. This partition is summarised (as “Struc-

tural”) in Table 4 and illustrated in Figure 14(a). The figure is same as the M_Z - R_{eff} distribution shown in Figure 7 above, but colour-coded by the clusters identified by the k-means algorithm. The first cluster (91 galaxies) tends to higher luminosity and the second cluster (314) has lower luminosities. Where the two clusters overlap in luminosity, objects in the first cluster have smaller radii as shown in Figure 14(a). In terms of our morphological classification, the first cluster contains all the “lenticulars” as well all the more luminous spirals. It also contains some of the more luminous “elliptical” and “blue spheroid” galaxies.

We investigated the dependence of the partitions on the parameter choice by removing different parameters from the clustering calculation. Removing the colour had very little effect (2 clusters of 97 and 308 members), but when we removed the colour and any one of the structural parameters, the k-means algorithm could not find any preferred partition. We then retained the colour but removed each of the structural parameters in turn. Again, these gave very similar results, except when either of the surface brightness terms was removed. The size of the first cluster increased to 129 when the central surface brightness was removed (μ_0) and it decreased to 78 when the effective surface brightness ($\langle\mu\rangle_e$) was removed. These changes reflect the properties of the respective surface brightness measures shown in Figure 9. The central surface brightness very strongly separates the different luminous galaxy classes, so removing it causes the clustering algorithm to tend to put all the luminous galaxies together. The converse applies to the effective surface brightness.

We observed two common features of all the partitions based on the structural plus colour parameters. First, the lenticular galaxies were always all included in the first (high-luminosity) cluster, as were the luminous ($M_Z < -20.5$) spirals. Second, there was no separation of the low-luminosity galaxies into separate classes. Instead, the majority (88%) of the objects we visually classified as star-forming dwarfs (BSph, Irr, LSB) were placed in the second cluster. The fraction of *all* low luminosity ($M_Z > -18.5$) galaxies in the second cluster is 87 per cent.

6.3.3 Clustering with Structural and MAGPHYS Parameters

We then added the four MAGPHYS parameters (\log of sSFR, SFR, M^* and M_{dust}), creating a nine-dimensional space (but with degeneracy between some of the parameters). This sample was limited to 313 galaxies due to incomplete MAGPHYS data. This also gave a preference for two clusters but with fewer galaxies (57) in the high-luminosity group than with just the structural parameters. This is shown in Figure 14(b) and listed as “Structural+SF 1” in Table 4. The main difference from the previous partition is that very few low luminosity galaxies now appear in the first cluster. The fraction of galaxies in the second cluster is now 97 per cent at luminosities fainter than $M_Z = -18.5$. The first cluster consists of just the visually-classified lenticulars and the brighter spirals. The second cluster contains 97 per cent of all the visually-classified star-forming dwarf galaxies.

As previously, we removed the ($g-i$) colour and found that this made no significant difference to the result (59 in the first cluster; “Structural+SF 2” in Table 4). This is not surprising as the MAGPHYS parameters are based on extensive colour information; indeed $g-i$ correlates strongly with specific star formation ($R = -0.72$). For the remaining tests we therefore excluded the $g-i$ colour. We then removed each of the four MAGPHYS parameters in turn. Removing any one of specific SFR, SFR and dust mass still resulted in two clusters, but with varying numbers of spirals in

Table 4. Automated classification results with different choices of parameters

partition	parameters	N	n_1	n_2	n_3	p_{Vis}	$p_{-18.5}$
Structural	12345----	405	91	314	-	88%	87%
Structural+SF 1	123456789	313	57	256	-	97%	97%
Structural+SF 2	1234-6789	313	59	254	-	97%	97%
Structural+SF 3	1234-67-9	313	41	198	74	79%	83%

Note: the parameters used in each choice are numbered as follows. 1: μ_0 ; 2: $\langle\mu\rangle_e$; 3: $\log n$; 4: $\log R_{eff}$; 5: $g - i$; 6: $\log sSFR$; 7: $\log SFR$; 8: $\log M^*$; 9: $\log M_{dust}$. N is the number of galaxies and n_i are the number of galaxies assigned to cluster i as colour-coded in Figures 14(a)-(c). p_{Vis} is the percentage of visually-classified star-forming dwarf galaxies assigned to the second cluster and $p_{-18.5}$ are the percentages of galaxies with absolute magnitude $M_Z > -18.5$ in the second cluster.

the first cluster (cluster 1 sizes 93, 56, and 47 respectively). Removing the stellar mass, by contrast, resulted in a significant preference for a 3-cluster partition which we show in Figure 14(c) (“Structural+SF 3” in Table 4). The first cluster (high luminosity and low radius) is now smaller and corresponds very closely to the visual lenticular classification. The remaining two clusters contain all the other galaxies split into high- and low-luminosity groups (74 and 198).

6.3.4 Summary of clustering analysis

Clustering algorithms like k-means provide an objective basis to determine how many distinct groups are defined by our measured galaxy parameters. Our analysis has shown that the results are somewhat sensitive to the choice of parameters, notably at higher luminosities where the numbers of galaxies in the first cluster can vary by up to 30 per cent. However at lower luminosities ($M_Z > -18.5$) there was much less variation: in the two-cluster partitions the number of galaxies in the second cluster varied by only 5 per cent.

In all but one of the partitions we have discussed, only 2 clusters were identified in the data. This is markedly different from our visual classifications which identified 6 different galaxy types. The most striking result is that the three visual classes of dwarf star-forming galaxies (blue spheroid, irregular and low surface brightness) are assigned to the same cluster. Using just the structural parameters 88 per cent of the star-forming dwarfs are in cluster 2; this rises to 97 per cent when the MAGPHYS parameters are added. There is no evidence that these dwarf galaxies can be separated into distinct types, despite our visual impression that this was the case. Instead they form a single population with a continuous distribution that can be parameterised by surface brightness or size (as in Figure 2). *This implies that the different morphological classifications given to low luminosity star-forming galaxies in literature only reflect variations in surface brightness and do not correspond to any intrinsic physical differences between them.*

6.4 Limitations of the analysis and future prospects

The GAMA II data analysed here suggests that the dwarf galaxies ($-18 \leq M_r \leq -15$) in the local Universe (8-87 Mpc distance) form a unimodal population in the parameter space spanned by M_Z , R_{eff} , n , $\langle\mu\rangle_e$, μ_0 , $(g - i)$, $NUV - r$, M^* , M_{dust} , SFR and sSFR. This is most likely due to the intrinsic scatter resulting from stochastic evolution within different morphological classes of galaxies. Other factors may however also have affected the conclusions drawn here, and we discuss them in brief below.

- *Wavelength dependency:* The light profile of galaxies and the structural parameters derived from it are extremely wavelength-dependant. As shown by Kelvin et al. (2012) for a sample of more than 138,000 galaxies in GAMA I, the Sérsic index for galaxies does not vary significantly with bandpass but minor variations are expected in individual galaxies because different optical wavebands trace different stellar components. On the other hand, the effective radius is a strong function of waveband in which the stellar light is modelled, such that the size of galaxy increases with wavelength (Kelvin et al. 2012, their figure 22). Zhang et al. (2012) reached similar conclusions when modelling ultraviolet to infrared data for the nearby dwarf irregular galaxies.

Although we employ the high-resolution *Viking* data in the Z -band for low surface brightness, low-luminosity galaxies here, the trends seen for the SDSS data used by Kelvin et al. (2012) still apply:

- *Imaging data:* In this work we made use of the SDSS (Data Release 7) imaging to visually classify our sample into different classes. The SDSS images are taken with an exposure time of $\lesssim 54$ seconds per field, and have an angular resolution of $1.2''$. Given that the bulk of our sample comprises low-luminosity dwarfs, the low-resolution imaging is a major limiting factor in the visual classification, and may have significantly contributed to the scatter within individual classes. This may also have caused morphologically similar classes such as the irregulars and the LSBs, or irregulars and the blue spheroids to overlap artificially in the parameter spaces explored in §4 and §5. However, at this point we are unable to quantify this effect.

The morphological classification of the sample presented here is 80% reliable, meaning that among the visual classification by Mahajan, Driver and Drinkwater, and repetitive classification by Mahajan, the visual class assigned to galaxies agreed $\sim 80\%$ of the times. The disagreement is mostly caused in the case of low surface brightness galaxies, and small, blue spheroids (the latter are as likely to be classified as irregulars).

The light profiles for our galaxy sample are modelled in the *Viking* Z -band data to automatically quantify structural parameters such as the Sérsic index and effective radius using SIGMA. The *Viking* data are ~ 2 mag deeper and have a twofold improvement in angular resolution compared to the SDSS. It is the best available dataset in the near infrared waveband for such a spectroscopically complete sample, yet still shallow for dim or low-luminosity galaxies. Deeper datasets with higher resolution and depth are thus required for resolving the issues raised in this work due to measurement uncertainties, and confirm the observed trends in the structural and star formation properties of galaxies.

- *Single component Sérsic fit:* While a single component Sérsic

profile fits the giant elliptical galaxies (Tal & van Dokkum 2011, and references therein) and spirals (Pohlen & Trujillo 2006) successfully, little is known of its behaviour at the faint end of the luminosity function. Recently though, Herrmann, Hunter, & Elmegreen (2013) have fitted single, double and triple exponential profiles to the stellar disk of 141 nearby dwarf galaxies in multi-wavelength data from ultraviolet to mid infrared. They find that blue compact dwarfs are over-represented by profiles where the light falls off less steeply, and Magellanic-like spirals by profiles where the light falls more steeply after the break in the first exponential, relative to dwarf irregulars. This observation is in agreement with the results presented by Pohlen & Trujillo (2006), who showed that $> 90\%$ of their sample deviated from a classical $n = 1$ exponential profile, and is instead better represented by a broken exponential profile.

Caon, Capaccioli, & D’Onofrio (1993) showed the correlation between Sérsic index n and (model-independent) size of early-type galaxies does not result from the parameter coupling in the Sérsic model. This implies that if a model fails to capture the range of stellar distribution in a galaxy, it will result in over- or under-estimated magnitude, surface brightnesses and size as a function of the stellar concentration (n), and consequently M^* (see section 2.1 of the excellent review by Graham 2013).

As we have shown in Figure 5 (also see Table 1), while most bulge-dominated galaxies (blue spheroids and lenticulars) follow the correlation shown by Caon, Capaccioli, & D’Onofrio (1993) for early-type and S0 cluster galaxies, spirals and other non-spheroidal dwarf galaxies do not. Table 1 suggests that almost half of the luminous spirals in our sample could be better fit by additional components to fit their light profile, especially in the centre. Some dwarf galaxies may also benefit by the inclusion of a second component. However, given the complex geometry of the low-luminosity irregular galaxies it is hard to believe that measurement uncertainties could be greatly improved by multiple component fits. Moreover, the measurement uncertainties obtained from these data are much smaller than the intrinsic scatter within each visually identifiable class. Hence, it is reasonable to assume that better models for the light profile may further tighten the correlations seen for the luminous galaxies.

7 SUMMARY

We have shown that morphologically distinct star-forming dwarf galaxies are not distinguishable in the parameter space comprising $\langle \mu \rangle_e$, μ_0 , n , R_{eff} , SFR, sSFR, M^* , M_{dust} , $(NUV - r)$ and $(g - i)$. The (d)elliptical galaxies remain indistinguishable from the other dwarf classes in structural parameters, but their low sSFR and SFR makes them easily identifiable in star formation properties. In various 2D-parameter spaces formed by structural parameters, morphologically distinct dwarf galaxies occupy overlapping, yet different regions of the parameter space. However, all except the low-mass ellipticals also show similar star formation and dust properties. The more luminous galaxies on the other hand, clearly separate into star-forming spirals and passively evolving lenticulars, respectively. For the ensemble of galaxies in our sample, the “k-means” algorithm prefers a bimodal distribution independent of the number of parameters used to partition the data.

We have shown for the first time the distribution of morphologically distinct dwarf galaxies in environments outside rich clusters in $M - \mu_0$ ($\langle \mu \rangle_e$). Our analysis shows that although morphologically distinct galaxies occupy different regions of these spaces,

in the dwarf regime the sub-populations overlap extensively. The giants on the other hand split into two as mentioned above.

We showed that the model-independent correlation between stellar concentration, parametrized by the Sérsic index n and size (R_{eff}), shown to exist for elliptical and S0 cluster galaxies, is also followed by the blue spheroids and passively evolving ellipticals. The non-spheroidal galaxies on the other hand do not show such a correlation.

The SFR, sSFR, M_{dust} and M^* derived by fitting the spectral energy distribution of all galaxies using MAGPHYS qualitatively suggest that dwarf galaxies in these environments may evolve into red, passive low-mass ellipticals, while the luminous spiral galaxies turn into lenticulars. This hypothesis is based on the assumption that galaxies can not acquire much dust after their star formation turns off. A detailed analysis of the neutral hydrogen content of galaxies in these environments is required to confirm this speculation.

To conclude, we showed that galaxies across a wide range in magnitudes have statistically distinct, yet overlapping distributions in all their structural parameters. In the low-luminosity regime, the star-forming dwarf galaxies: LSBs, irregulars and blue spheroids also overlap in their star formation and dust properties, while the ellipticals have distinguishably low sSFR relative to other galaxies of similar luminosity. The giant galaxies on the other hand show clear separation such that the spirals are characterised by high SFR, sSFR and M_{dust} , while vice-versa is true for the lenticulars. Hence, “classification” (visual or otherwise) of galaxies forming a continuum in one, two, or multiple parameter space into discrete categories although useful in understanding the evolution mechanism(s) to first order, must be undertaken with caution, and consequent implications from such a classification must be used with caution.

8 ACKNOWLEDGEMENTS

GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programs including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is <http://www.gama-survey.org/>.

This project was partly funded by the UWA-UQ Bilateral Research Collaboration Awards (University of Queensland, 2013; CI: Mahajan).

We thank Michael Brown, Alister Graham, Maritza Lara-Lopez, Trevor Ponman and Stephen Wilkins for their comments and suggestions on an early draft of this paper. We also thank the anonymous referee for their constructive criticism which greatly helped in clarifying the presentation of this work.

REFERENCES

- Baldry I. K., Glazebrook K., Driver S. P., 2008, MNRAS, 388, 945
- Baldry I. K., et al., 2010, MNRAS, 404, 86
- Baldry I. K., et al., 2012, MNRAS, 421, 621

Bertin E., Arnouts S., 1996, A&AS, 117, 393
 Bertin, E., 2011, Astronomical Data Analysis Software and Systems XX, 442, 435
 Binggeli B., Sandage A., Tarengi M., 1984, AJ, 89, 64
 Binggeli B., 1994, ESOC, 49, 13
 Blanton M. R., Lupton R. H., Schlegel D. J., Strauss M. A., Brinkmann J., Fukugita M., Loveday J., 2005, ApJ, 631, 208
 Boselli A., Boissier S., Cortese L., Gavazzi G., 2008, A&A, 489, 1015
 Bothun G. D., Impey C. D., Malin D. F., Mould J. R., 1987, AJ, 94, 23
 Bourne N., et al., 2012, MNRAS, 421, 3027
 Boyce P. J., Phillipps S., 1995, A&A, 296, 26
 Brosche P., 1973, A&A, 23, 259
 Brough S., et al., 2013, MNRAS, 435, 2903
 Caldwell N., 1983, AJ, 88, 804
 Calinski T., Harabasz J., 1974, Communications in Statistics - Theory and Methods, 3(1), 1-27. doi:10.1080/03610927408827101.
 Caon N., Capaccioli M., D’Onofrio M., 1993, MNRAS, 265, 1013
 Caon N., Einasto M., 1995, MNRAS, 273, 913
 Cortese L., Catinella B., Boissier S., Boselli A., Heinis S., 2011, MNRAS, 415, 1797
 Charrad M., Ghazzali N., Boiteau V., Niknafs A., 2012, UseR! 2012, CGB12a
 da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595
 Dressler A., 1980, ApJ, 236, 351
 Driver S. P., et al., 2011, MNRAS, 413, 971
 Driver S. P., et al., 2012, MNRAS, 427, 3244
 Fujita, Y., 1998, ApJ, 509, 587
 Graham A. W., 2013, pss6.book, 91
 Graham A. W., Guzmán R., 2003, AJ, 125, 2936
 Graham A. W., Driver S. P., 2005, PASA, 22, 118
 Herrmann K. A., Hunter D. A., Elmegreen B. G. 2013, AJ, 146, 104
 Hill D. T., et al., 2011, MNRAS, 412, 765
 Hubble E. P., 1926, ApJ, 64, 321
 Hunter D. A., Elmegreen B. G., Ludka B. C., 2010, AJ, 139, 447
 Janz J., et al., 2013, arXiv, arXiv:1308.6496
 Janz J., et al., 2012, ApJ, 745, L24
 Jenkins L. P., Hornschemeier A. E., Mobasher B., Alexander D. M., Bauer F. E., 2007, ApJ, 666, 846
 Lake G., Katz N., Moore B., 1998, ApJ, 495, 152
 Lee B., et al., 2013, ApJ, 774, 47
 Oh K. S., Lin D. N. C., 2000, ApJ, 543, 620
 Kaviraj S., et al., 2007, ApJS, 173, 619
 Kelvin L. S., et al., 2012, MNRAS, 421, 1007
 Kodaira K., Okamura S., Watanabe M., 1983, ApJ, 274, L49
 Kodaira K., 1989, ApJ, 342, 122
 Kormendy J., 1985, ApJ, 295, 73
 Kormendy J., Bender R., 2012, ApJS, 198, 2
 Loveday J., et al., 2012, MNRAS, 420, 1239
 Macqueen J., “Some methods for classification and analysis of multivariate observations”, Proc. Fifth Berkeley Symp. on Math. Statist. and Prob., Vol. 1 (Univ. of Calif. Press, 1967), 281, 297
 Mahajan S., Raychaudhury S., Pimblett K. A., 2012, MNRAS, 427, 1252
 McConnachie A. W., 2012, AJ, 144, 4
 Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, AJ, 139, 2097
 Pohlen M., Trujillo I., 2006, A&A, 454, 759

Table A1. K-S probabilities for $\log M^*$.

	E	Sp	Irr	BSph	LSB	Len
E	1	1.98E-01	1.17E-06	3.99E-05	9.63E-08	3.86E-08
Sp		1	6.21E-14	1.34E-08	1.55E-13	3.41E-06
Irr			1	1.05E-01	3.61E-03	1.66E-20
BSph				1	3.54E-02	2.64E-16
LSB					1	6.07E-18
Len						1

Table A2. K-S probabilities for SFR.

	E	Sp	Irr	BSph	LSB	Len
E	1	4.25E-08	3.27E-03	6.02E-03	2.61E-02	5.84E-01
Sp		1	1.26E-16	2.76E-15	1.11E-17	8.92E-09
Irr			1	4.66E-01	9.39E-04	8.00E-03
BSph				1	2.44E-02	1.99E-02
LSB					1	7.98E-02
Len						1

Popesso P., Biviano A., Böhringer H., Romaniello M., 2006, A&A, 445, 29
 Prugniel P., Simien F., 1997, A&A, 321, 111
 Reynolds J. H., 1920, MNRAS, 80, 746
 Robotham A. S. G., et al., 2011, MNRAS, 416, 2640
 Schawinski K., et al., 2007, ApJS, 173, 512
 Sharina M. E., et al., 2008, MNRAS, 384, 1544
 Strader J., et al., 2013, ApJ, 775, L6
 Tal T., van Dokkum P. G., 2011, ApJ, 731, 89
 Taylor E. N., et al., 2011, MNRAS, 418, 1587
 Thilker D. A., et al., 2007, ApJS, 173, 538
 Thomas P. A., Drinkwater M. J., Evstigneeva E., 2008, MNRAS, 389, 102
 Toloba E., Boselli A., Cenarro A. J., Peletier R. F., Gorgas J., Gil de Paz A., Muñoz-Mateos J. C., 2011, A&A, 526, A114
 Wright E. L., et al., 2010, AJ, 140, 1868
 Yamanoi H., et al., 2012, AJ, 144, 40
 Young C. K., Currie M. J., 1995, MNRAS, 273, 1141
 Zhang H.-X., Hunter D. A., Elmegreen B. G., Gao Y., Schruha A., 2012, AJ, 143, 47

APPENDIX

A Kolmogorov-Smirnov Statistics

In this appendix, we present tables of probabilities from the Kolmogorov-Smirnov static, showing that the physical and star formation properties of various visual morphological classes arise from the same parent distributions.

B Preferred number of clusters with k-means clustering algorithm

In this Appendix we illustrate the Monte Carlo approach we used to estimate uncertainties in the statistics used to select the best number of clusters from the k-means analysis. As noted above (Section 6.3.1), we used the Calinski & Harabasz (1974) variance ratio as a figure-of-merit. We used the NbClust package (Charrad et al. 2012) for all the calculations, but this did not provide uncertainties

Table A3. K-S probabilities for SFR/M^* .

	E	Sp	Irr	BSph	LSB	Len
E	1	8.51E-06	2.41E-12	4.78E-10	4.20E-09	1.45E-05
Sp		1	2.23E-03	8.16E-03	3.08E-02	7.10E-10
Irr			1	2.41E-01	7.24E-01	2.26E-17
BSph				1	2.27E-01	6.59E-14
LSB					1	5.59E-14
Len						1

Table A4. K-S probabilities for M_{dust} .

	E	Sp	Irr	BSph	LSB	Len
E	1	3.35E-08	3.58E-01	6.73E-02	3.70E-02	1.70E-04
Sp		1	3.63E-21	1.04E-17	7.35E-20	2.87E-02
Irr			1	1.51E-01	1.08E-01	1.05E-09
BSph				1	6.76E-01	3.12E-09
LSB					1	3.85E-10
Len						1

for the calculated variance ratios. We applied a simple Monte-Carlo approach to estimate the variance ratio uncertainties by repeating each calculation 30 times using a different random seed for the k-means partition algorithm each time. We used the mean value of the 30 calculations as the final value of the statistic and the standard deviation of the 30 values to estimate an uncertainty (the standard error of the mean) which we show using the red symbols in Figure B1 for each of the three partitions shown in Figure 14. As the standard errors on each point are very small we also show the standard deviations (in grey) for comparison. We also show the results of a single calculation, showing that quite strong systematic trends can be present (e.g. for 6-10 clusters in the first panel): this is an additional reason to take the mean of several calculations.

In panels (a) and (c) of Figure B1 there is a clear local maximum of the variance ratio (at 2,3 clusters respectively) that is separated by more than 3 standard errors (the smaller error bars) from the adjacent points. In panel (b) the first point is barely three standard errors higher than the second, but we take the preferred number of clusters as two using our rule to take the smaller number of clusters in the case of a double peak.

Table A5. K-S probabilities for R_{eff} .

	E	Sp	Irr	BSph	LSB	Len
E	1	1.14E-07	6.66E-01	5.66E-02	4.61E-01	5.09E-03
Sp		1	3.27E-15	3.40E-16	5.86E-12	1.51E-04
Irr			1	2.84E-06	5.19E-01	3.61E-04
BSph				1	5.79E-06	5.33E-10
LSB					1	2.21E-03
Len						1

Table A6. K-S probabilities for n .

	E	Sp	Irr	BSph	LSB	Len
E	1	1.88E-01	9.83E-03	9.37E-01	2.61E-02	1.28E-04
Sp		1	5.27E-01	5.23E-03	4.55E-01	7.10E-10
Irr			1	1.80E-06	4.32E-01	1.23E-13
BSph				1	2.43E-04	2.33E-08
LSB					1	1.22E-10
Len						1

Table A7. K-S probabilities for μ_0 .

	E	Sp	Irr	BSph	LSB	Len
E	1	7.67E-01	5.90E-03	5.05E-01	8.93E-09	2.46E-06
Sp		1	1.05E-07	4.54E-01	9.01E-15	8.92E-10
Irr			1	7.81E-07	8.57E-11	1.14E-17
BSph				1	3.66E-18	2.98E-14
LSB					1	1.44E-14
Len						1

Table A8. K-S probabilities for $\langle\mu\rangle_e$.

	E	Sp	Irr	BSph	LSB	Len
E	1	1.06E-01	2.78E-02	5.49E-01	1.77E-10	1.36E-05
Sp		1	4.72E-09	3.12E-03	3.65E-20	1.47E-05
Irr			1	2.76E-05	2.81E-18	1.67E-14
BSph				1	1.60E-23	2.90E-12
LSB					1	5.69E-16
Len						1

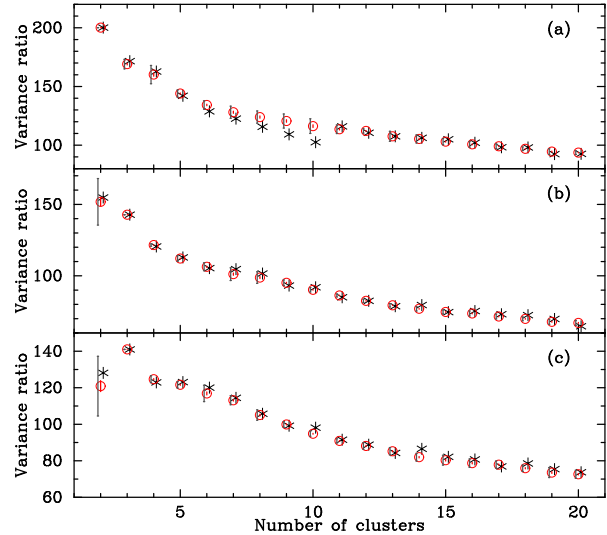


Figure B1. The k-means variance ratio as a function of the number of clusters analysed. The three plots are for the same parameter sets as in Figure 14: (a) Structural (b) Structural+SF 1, and (c) Structural+SF 3, respectively as discussed above and listed in Table 4. In each plot the black stars represent a single k-means calculation using a random seed of 12341. The red circles show the mean of 30 calculations using different random seeds. The red error bars give the standard error of the mean and the grey error bars show the 1-standard deviation range of the values. We select the best number of clusters by identifying a local maximum in the variance ratio: only in plot (c) is a 3-cluster partition preferred over the 2-cluster case.