Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **462,** 1470–1500 (2016) doi:10.1093/mnras/stw1495
Advance Access publication 2016 July 6

Galaxy And Mass Assembly (GAMA): $\mathcal{M}_\star$–R_e relations of $z = 0$ bulges, discs and spheroids

Rebecca Lange,[1★] Amanda J. Moffett,[1] Simon P. Driver,[1,2★] Aaron S. G. Robotham,[1] Claudia del P. Lagos,[1,3] Lee S. Kelvin,[4] Christopher Conselice,[5] Berta Margalef-Bentabol,[5] Mehmet Alpaslan,[6] Ivan Baldry,[4] Joss Bland-Hawthorn,[7] Malcolm Bremer,[8] Sarah Brough,[9] Michelle Cluver,[10] Matthew Colless,[11] Luke J. M. Davies,[1] Boris Häußler,[12] Benne W. Holwerda,[13] Andrew M. Hopkins,[7] Prajwal R. Kafle,[1] Rebecca Kennedy,[5] Jochen Liske,[14] Steven Phillipps,[8] Cristina C. Popescu,[15,16] Edward N. Taylor,[17] Richard Tuffs,[18] Eelco van Kampen[19] and Angus H. Wright[1]

Affiliations are listed at the end of the paper

Accepted 2016 June 20. Received 2016 May 25; in original form 2016 February 14

ABSTRACT

We perform automated bulge + disc decomposition on a sample of ∼7500 galaxies from the Galaxy And Mass Assembly (GAMA) survey in the redshift range of $0.002 < z < 0.06$ using Structural Investigation of Galaxies via Model Analysis, a wrapper around GALFIT3. To achieve robust profile measurements, we use a novel approach of repeatedly fitting the galaxies, varying the input parameters to sample a large fraction of the input parameter space. Using this method, we reduce the catastrophic failure rate significantly and verify the confidence in the fit independently of χ^2. Additionally, using the median of the final fitting values and the 16th and 84th percentile produces more realistic error estimates than those provided by GALFIT, which are known to be underestimated. We use the results of our decompositions to analyse the stellar mass – half-light radius relations of bulges, discs and spheroids. We further investigate the association of components with a parent disc or elliptical relation to provide definite $z = 0$ disc and spheroid $\mathcal{M}_\star$–R_e relations. We conclude by comparing our local disc and spheroid $\mathcal{M}_\star$–R_e to simulated data from EAGLE and high-redshift data from Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey-Ultra Deep Survey. We show the potential of using the $\mathcal{M}_\star$–R_e relation to study galaxy evolution in both cases but caution that for a fair comparison, all data sets need to be processed and analysed in the same manner.

Key words: galaxies: elliptical and lenticular, cD – galaxies: formation – galaxies: fundamental parameters – galaxies: spiral – galaxies: statistics.

1 INTRODUCTION

At the fundamental level, galaxies are multicomponent systems (see for example Buta et al. 2010), consisting of at least a spheroid and/or disc. This is most obvious in the Sérsic index–colour plane where the single-component Sd and elliptical galaxies occupy distinct peaks with composite galaxies (S0abc) scattered between and around these peaks (see e.g. Driver et al. 2006; Cameron et al. 2009; Kelvin et al. 2012; Lange et al. 2015). These components have very different characteristics with spheroids typically having a featureless appearance and being pressure-supported. Discs, on the other hand, have features such as spiral arms and are rotationally supported. Furthermore, bulges are made up of redder stars with moderate to high metallicities and a high α-element abundance, while discs are made of younger, bluer stars with lower metallicities and typically are dust- and gas-rich. Spheroids are older, showing little to no star formation and are typically dust- and gas-depleted (see for example the review by Roberts & Haynes 1994). The simplest explanation for these stark differences is that spheroids and discs form via two distinct mechanisms over two distinct eras (Cook, Lapi & Granato 2009; Driver et al. 2013), i.e. a dynamically

---

★E-mail: rebecca.lange@icrar.org (RL); simon.driver@icrar.org (SPD)

Published by Oxford University Press on behalf of the Royal Astronomical Society

'hot mode' (spheroid formation) and 'cold mode' (disc formation) evolution.

Traditionally, the relative prominence of a bulge component is taken into account when classifying galaxies on to the *Hubble* sequence (see Hubble 1926, and later revisions by e.g. van den Bergh 1976; Kormendy & Bender 2012), however, studying global properties of galaxies by *Hubble* type could be misleading. For example, numerous evolution mechanisms have been proposed to explain the morphological diversity seen at $z = 0$, such as a (initial) major dissipative event, gas accretion, adiabatic contraction, major and minor mergers and secular processes (see e.g. Hopkins et al. 2010; Trujillo, Ferreras & Rosa 2011; L'Huillier, Combes & Semelin 2012; Cheung et al. 2013; Sachdeva et al. 2015). Each of these processes potentially acts to modify the prominence of the bulge, disc or other components. This indicates that galaxy components likely follow distinct formation pathways and structure effectively encodes the formation history. Therefore, to study galaxy evolution, bulge + disc decomposition is critical.

While the number of studies of large samples which employ bulge + disc decomposition to explore the nature of galaxies and their components is growing, the analysis is challenging (see e.g. Allen et al. 2006; Gadotti 2009; Simard et al. 2011; Bruce et al. 2012, 2014; Lang et al. 2014; Tasca et al. 2014; Meert, Vikram & Bernardi 2015; Salo et al. 2015). This is because multicomponent fitting is notoriously difficult, especially when trying to automate it for large samples. Nevertheless, a number of publicly available codes have now been created to allow bulge + disc decomposition, such as GIM2D (Simard 1998), BUDDA (de Souza, Gadotti & dos Anjos 2004), GALFIT3 (Peng et al. 2010) and IMFIT (Erwin 2015). Each code has advantages and disadvantages (see Erwin 2015, for example, for further discussion), here we elect to use GALFIT3 because of its ability to manage nearby objects, its computational reliability, and its speed.

Many studies that fit two-component Sérsic light profiles restrict the Sérsic index to $n = 1$ for the disc and in some cases, $n = 4$ for the bulge (e.g. Simard et al. 2011; Bruce et al. 2012; Lackner & Gunn 2012; Meert et al. 2015). This reduces the number of free parameters and ensures the fitting process is more robust but it restricts the possible interpretations of the fitting outcomes, e.g. classical and pseudo-bulges cannot be differentiated this way. A number of studies now show that the Sérsic index of discs and spheroids (be they pure or component) vary smoothly with mass and luminosity or due to dust or galaxy type (see e.g. Graham & Guzman 2003; Gadotti 2009; Kelvin et al. 2012; Graham 2013; Pastrav et al. 2013a,b). Hence, studies where the Sérsic index of the bulge or disc components are fixed may be overly restrictive. Furthermore, to correctly trace a galaxy's formation history, a full decomposition of all of its components would be ideal (e.g. the Spitzer Survey of Stellar Structure in Galaxies, S^4G; Salo et al. 2015). However, this is only viable for very nearby galaxies where all the components can be clearly resolved and hence for relatively small samples (S^4G is the largest study to date extending to 2352 galaxies for which a number have been fit with more than two components). To compare galaxies at different epochs going beyond a simple bulge and disc decomposition is difficult (Gadotti 2008). There are two reasons, however, why two components might be sufficient, (i) the majority of stellar mass resides in the bulge and disc components for most galaxies, and (ii) some components may simply represent minor perturbations to the underlying disc (e.g. bars, pseudo-bulges). Such perturbations should arguably be considered secondary rather than primary evolutionary markers.

Here we adopt the stance that bulge and disc components arise from two primary formation pathways (i.e. hot and cold mode evolution, respectively), and that additional components form in secondary formation pathways (i.e. tidal interactions, disc instabilities and perturbations). The likely primary pathways are: monolithic collapse followed by major mergers, which can produce elliptical galaxies by destroying and rearranging any structure previously present in a galaxy, resulting in a smooth light profile (Toomre 1977); and minor mergers and continued gas inflow, which can form or re-grow a disc around a pre-existing spheroid, resulting in a galaxy with two distinct components (see e.g. Steinmetz & Navarro 2002; Kannappan, Guie & Baker 2009; Wei et al. 2010). A key question worth asking is whether two generic components (spheroids and discs) really can explain the diversity seen, i.e. how many fundamental building blocks and structures are required to adequately reproduce the observed galaxy population? As most of the stellar mass is contained within the bulge and disc, how important are tertiary features like bars? Furthermore, how many different physical origins do the various spheroids and discs have? Are elliptical galaxies simply naked bulges and are bulges related to high-redshift compact galaxies (e.g. Berg et al. 2014; Graham, Dullo & Savorgnan 2015)? Are the discs of early types, late types and irregulars indistinguishable?

We believe that the stellar mass – half-light size (hereafter $\mathcal{M}_\star$–R_e) relation is a key scaling relation allowing us to address these questions for the following reasons.

(i) The size of a galaxy is related to its specific angular momentum making the mass and size of a galaxy fundamental observables of conserved quantities (e.g. Romanowsky & Fall 2012).

(ii) The simple assumption that angular momentum is conserved during the initial collapse of the dark matter halo links the angular momentum and mass of a galaxy with its dark matter halo (Fall & Efstathiou 1980; Dalcanton, Spergel & Summers 1997; Mo, Mao & White 1998).

(iii) Hydrodynamical simulations now produce galaxies with realistic sizes and direct comparisons (at different epochs) are possible to study formation and evolution histories of galaxies (see for example, the Evolution and Assembly of Galaxies and their Environments simulation suite, EAGLE; Crain et al. 2015; Schaye et al. 2015).

(iv) We can empirically measure and trace the masses and sizes of galaxies and their components over a range of redshifts and in different environments [e.g. with *Hubble Space Telescope* (*HST*) as well as high-redshift ground-based surveys and soon with *Euclid* and Wide-Field Infrared Survey Telescope (*WFIRST*)].

The $\mathcal{M}_\star$–R_e relation therefore represents the next critical diagnostic for galaxy evolution studies beyond simple mass functions (see e.g. Bouwens et al. 2004; van der Wel et al. 2014; Holwerda et al. 2015; Shibuya, Ouchi & Harikane 2015), enabling us to trace angular momentum build-up and the emergence of the component nature of galaxies while connecting observations to simulations.

Recent studies comparing the $\mathcal{M}_\star$–R_e relation of low and high redshift are already yielding interesting results. For example, at high redshift, galaxies might look disc-like or elliptical/spheroidal, but their physical properties are unlike any discs or ellipticals in the local Universe (see e.g. Bruce et al. 2012; Buitrago et al. 2013; Mortlock et al. 2013). Galaxies at high redshifts are typically more irregular with thick slab-like disc structures and clumpy star-forming regions (Wisnioski et al. 2012). In addition, they can be very compact but massive. In some cases, at redshift ∼2, they are a factor of up to 6 times smaller in size than galaxies of the same mass today (Daddi et al. 2005; Trujillo et al. 2007; Buitrago et al. 2008; van Dokkum et al. 2008, 2010; Weinzirl et al. 2011).

In this paper, we aim to provide a reliable low-redshift benchmark of the $\mathcal{M}_\star$–R_e relation for bulges, discs and spheroids. The bulge + disc decomposition sample is derived from a set of galaxies for which detailed morphological information is available (see Moffett et al. 2016). Section 2 describes the data and sample selection, Sections 3 and 4 describe the set up of our bulge + disc decomposition catalogue and component mass estimates. In Section 5, we present the $\mathcal{M}_\star$–R_e relations for bulges, spheroidal and disc galaxies and discuss the association of components with their possible parent populations. We then compare our distributions to the EAGLE simulation in Section 6 followed by a comparison of our low-redshift $\mathcal{M}_\star$–R_e relation with recent high-redshift data from Ultra Deep Survey (UDS) region within the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011) in Section 7. Finally, in Section 8, we present our summary and conclusions.

Throughout this paper, we use data derived from the Galaxy And Mass Assembly (GAMA) survey (Driver et al. 2011, 2016; Liske et al. 2015) with stellar masses derived from Taylor et al. (2011), sizes derived from Sérsic profile fitting using Structural Investigation of Galaxies via Model Analysis (SIGMA; Kelvin et al. 2012), and for a cosmology given by: Λ cold dark matter (ΛCDM) universe with $\Omega_\mathrm{m} = 0.3, \Omega_\Lambda = 0.7, H_0 = 70$ km s^{-1} Mpc^{-1}.

2 DATA

The GAMA survey is an optical spectroscopic and multiwavelength imaging survey of ∼300 000 galaxies, combining the data of several ground- and space-based telescopes (Driver et al. 2011, 2016). It is an intermediate survey in respect to depth and survey area (Baldry et al. 2010) and thus fits in between low-redshift, wide-field surveys such as SDSS (York et al. 2000) or 2dFGRS (Colless et al. 2003) and narrow deep field surveys like zCOSMOS (see Lilly et al. 2007 and Davies et al. 2015) or DEEP-2 (Davis et al. 2003).

In this paper, we use data from the GAMA II (Liske et al. 2015) equatorial regions, which are centred on 9h (G09), 12h (G12) and 14.5h (G15). The three regions are 12×5 deg^2 in extent and have an r band Petrosian magnitude limit of r <19.8 mag. The spectroscopic target selection is derived from SDSS DR 7 (Abazajian et al. 2009; see Baldry et al. 2010 for details) input catalogue and we reach a spectroscopic completeness of ≥98 per cent for the main survey targets. The redshifts (Baldry et al. 2014; Liske et al. 2015) are based on spectra taken with the AAOmega spectrograph at the 3.9m Anglo–Australian Telescope (Hopkins et al. 2013) located at Siding Spring Observatory. The supporting panchromatic imaging data extend from the FUV to the far-IR via *GALEX*, SDSS, VISTA, *WISE* and *Herschel* (for full details, see Driver et al. 2016). All optical and near-IR imaging data has matched aperture photometry (Hill et al. 2011; Liske et al. 2015) Here we focus on the redshifts (*SpecCatv27*), morphologies (*VisMorphv03*), optical imaging (*ApMatchedv06*) and stellar masses (*StellarMassCatv18*) data products.

2.1 Sample selection

We select galaxies with $0.002 < z < 0.06$, $r_\mathrm{petro} < 19.8$ mag and spectra quality NQ>2 (see Liske et al. 2015) for which visual morphologies have been established following Moffett et al. (2016). To briefly summarize the visual classification procedure: we use three-colour (*Hig*) postage stamps of the objects to visually inspect them. A simple classification tree is used to sort galaxies in the first instance into bulge- and disc-dominated, little blue spheroid (LBS) and star/artefact. In the following step, the bulge- and disc-dominated objects are further split into single- and multicomponent. Finally, the multicomponent galaxies are sorted into barred and unbarred. For each galaxy, the result was then translated to a *Hubble* type: E, S0-Sa, SB0-SBa, Sab-Scd, SBab-SBcd and Sd-Irr, plus the additional LBS, Star and Artefact classifications (see Moffett et al. 2016, for further details). Note that throughout this paper, the terms early and late type refer to our visual classification of objects being bulge- or disc-dominated and we do not impose any other parameter cuts (e.g. Sérsic index) to classify early or late type.

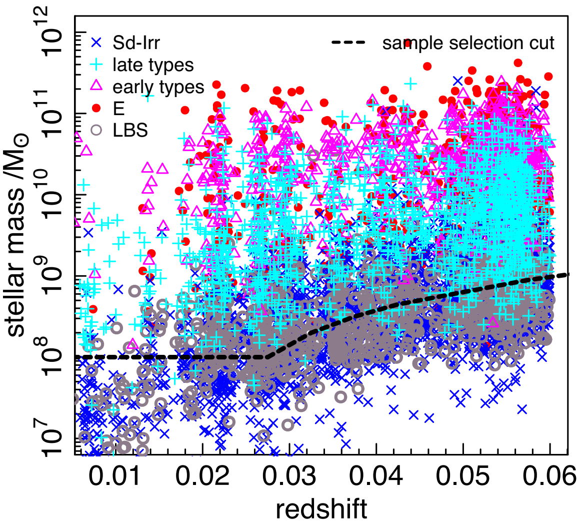


Figure 1. Shown is the redshift – stellar mass distribution for the GAMAnear sample. The points are coded according to the *Hubble* type established in the visual morphology classification. There appears to be no bias towards a particular *Hubble* type at the higher redshift boundary of our sample.

We perform bulge + disc decomposition in the r band only as our fitting approach is computationally expensive. For the analysis, we do not consider objects classified as stars or artefacts and we excluded one additional galaxy which was too large to be fit robustly. The resulting sample is hereafter called GAMAnear and comprises 7506 galaxies of which 2247 were visually classified as two-component (S0-Sa, SB0-SBa, Sab-Scd, SBab-SBcd) and 5259 as single-component (E, Sd-Irr, LBS) galaxies. Due to the low-redshift range, this sample extends well below 10^9 $\mathcal{M}_\odot$ allowing us to study the low-mass end of the $\mathcal{M}_\star$–R_e relation. However, because of the limited volume of our survey, this also means we do not have many very high-mass galaxies to study the curvature of the $\mathcal{M}_\star$–R_e relation at higher masses. This is important as the curvature in the (elliptical) $\mathcal{M}_\star$–R_e relation is likely indicative of the assembly history of the galaxy population (see e.g. Bernardi et al. 2011 and references therein).

As shown by Moffett et al. 2016, their fig. 5, the fraction of galaxy type by mass behave as expected, e.g. the fraction of early-type galaxies increases with increasing mass and the fraction of late-type galaxies decreases. To check whether the sample selection in this paper is biased, we show the sample distribution in the redshift – stellar mass plane in Fig. 1. The dashed line shows the final mass limit used in Section 5 to derive the $\mathcal{M}_\star$–R_e relation. The figure illustrates the mass segregation of the sample with the early-type galaxies being more massive than the late types. However, there is no clear bias with morphological type or our redshift range, especially at the upper redshift boundary.

3 BULGE + DISC DECOMPOSITION

Obtaining reliable bulge + disc fits is notoriously difficult, particularly in an automated fashion for large samples where the signal-to-noise ratio and resolution varies. Typically, 20–30 per cent of automatic fits are, in some way, non-physical. Previous studies have made use of a logical filter (e.g. Allen et al. 2006; see also Simard et al. 2011; Meert et al. 2015) to identify unphysical fits (e.g. component profiles which cross twice, the switching of the bulge and disc components etc.; see Allen et al. 2006 for more details) and manage these failures by replacement with a single Sérsic fit. As a first step, this reduces the catastrophic failure rate significantly but introduces a bias by removing the subset of two-component systems with poor fits. Following extensive exploration of our data using GALFIT3 (Peng et al. 2010) embedded in SIGMA (Kelvin et al. 2012), we identify five commonly occurring key factors which lead to poor and often catastrophic fitting outcomes. These are summarized below along with our adopted solution.

(1) *Becoming trapped in local minima and/or the limited movement of the final converged solution away from the initial conditions.*
The Levenberg–Marquart (LM) χ^2 minimization algorithm used by GALFIT3 can get stuck in a local rather than the global χ^2 minimum, especially when fitting multiple components. One way to overcome this is to vary the initial conditions (i.e. starting points) and repeat the fitting process. Convergence to a common solution, regardless of the starting point, provides confidence that the true minimum has been found. In due course, a full Markov Chain Monte Carlo (MCMC) approach, that appropriately samples the prior distribution should be developed but that is beyond the scope of our current investigation at this stage.

(2) *Unphysical solutions, e.g. a scalelength of 0.1 arcsec or a Sérsic index of 20.*
In some cases., the bulge or disc fits can migrate to the fit limits imposed, and these results are often not physical. While it is tempting to reduce the limits to plausible values, this causes a non-physical build-up of the solutions at the limits. Moreover, during the path towards convergence, it can sometimes be seen that solutions migrate into extreme values and then back again. To minimize the impact of our boundaries, we imposed no limits on the parameters, bar the constraint on the centre position, which is set to ± 5 pixels to account for the oversampling of the point spread function [PSF, i.e. GAMA pixel size is 0.339 arcsec and SDSS full width at half-maximum (FWHM) = 1.5 arcsec]. Instead, we elect to remove final solutions which settle on extreme values. We can afford to do this since we have multiple fits for each galaxy, i.e. some starting points lead to extreme outcomes but on the whole most converge to plausible values. Note that GALFIT does have some inbuilt constraints, such as a maximum Sérsic index of 20.

(3) *Decision on single or multiple components.*
A key problem in galaxy decomposition is to decide how many components are required. Ideally, this should be derivable from the independent 1-, 2- or multicomponent fits. Experimentation with the Akaike and Bayesian Information Criteria (AIC and BIC, respectively) was explored but no obvious automated process for determining the number of components, which agreed with our visual assessments, was identified. This is in part due to the limited information available in single-band fitting. Hence, we adopt our visual classifications as priors, i.e. E, Sd-Irr and LBS galaxies are taken as single-component systems and S0-Sa, SB0-SBa, Sab-Scd and SBab-SBcd as two-component. For completeness, however, we do derive and provide both one- and two-component fits for all systems.

(4) *Reversal of the bulge and disc components.*
On occasion, the initially assigned bulge component migrates to fit the disc component and the disc to the bulge. This effect was first noted in Allen et al. (2006) and can be rectified by switching the components if necessary. Here, regardless of the initial parameters, we assign the component with the lowest half-light radius as the bulge (i.e. inner, more compact component) and the other as the disc (i.e. outer, more extended component). This can, however, lead to cases where the bulge has a lower Sérsic index than the disc (see Appendix A for our treatment of these cases), which, in the majority of cases, is an unphysical solution.

(5) *Default GALFIT errors do not reflect the full complexity and uncertainty in the final fits.*
It is known that GALFIT (like other fitting codes) often underestimates the error on the returned parameters (see e.g. Häussler et al. 2007), possibly due to the poor treatment of correlated noise in real images. Essentially, the final errors do not provide any indication of fit confidence. By running GALFIT multiple times from a grid of initial conditions, we can assess the level of convergence which can be used to provide more realistic error estimates. This reassessment of the errors is probably the most important outcome of our adoption of a grid of initial conditions, providing some certainty for each galaxy as to the robustness of the fit.

The five strategies above proved critical in reducing the catastrophic error rate (as assessed from visual inspection) from ~20 per cent to ~5 per cent enabling us to dispense with the need for a logical filter, and most importantly obtain realistic errors.

We recognize that many of the above could also be addressed by improving the minimization algorithm and implementing an MCMC approach which fully samples the prior distribution. At the present time, however, in the absence of a known prior distribution and limited computing time, we believe our strategies minimize the obvious systematic issues which arise when using the GALFIT3 engine.

3.1 Construction of a robust decomposition catalogue

3.1.1 The initial grid and convergence

As stated, for completeness, we perform both single and double (bulge + disc decomposition) component fits in the r band on all 7506 galaxies in our sample using one or two Sérsic functions, respectively. We do not constrain any fitting parameters, except for the inbuilt limits within GALFIT. Hence, in our two-component fits, the bulge and disc Sérsic indices are not set to any particular value (e.g. 1 and 4) as is often done in other studies. We use the SIGMA (Kelvin et al. 2012) wrapper code for GALFIT (Peng et al. 2010). As a front-end wrapper, SIGMA creates cutouts from the GAMA regions, does a local background subtraction and detects objects and stars using SEXTRACTOR (Bertin & Arnouts 1996). To obtain reliable Sérsic fits, it is important that local background sky variations are accounted for, yet it is also important to not oversubtract light from the galaxy itself, as this will lead to systematic errors in the galaxy flux measurements. Our local background subtraction is in addition to the background subtraction applied during mosaicking of the GAMA data. The grid size used during this additional sky estimation depends on the size of the galaxy and varies from 32 × 32 to 128 × 128 pixels. Using this variable mesh approach was found to be the most robust method to remove small-scale sky variations without removing light from the galaxy (for further details, see Kelvin et al. 2012). After the sky subtraction, SIGMA constructs

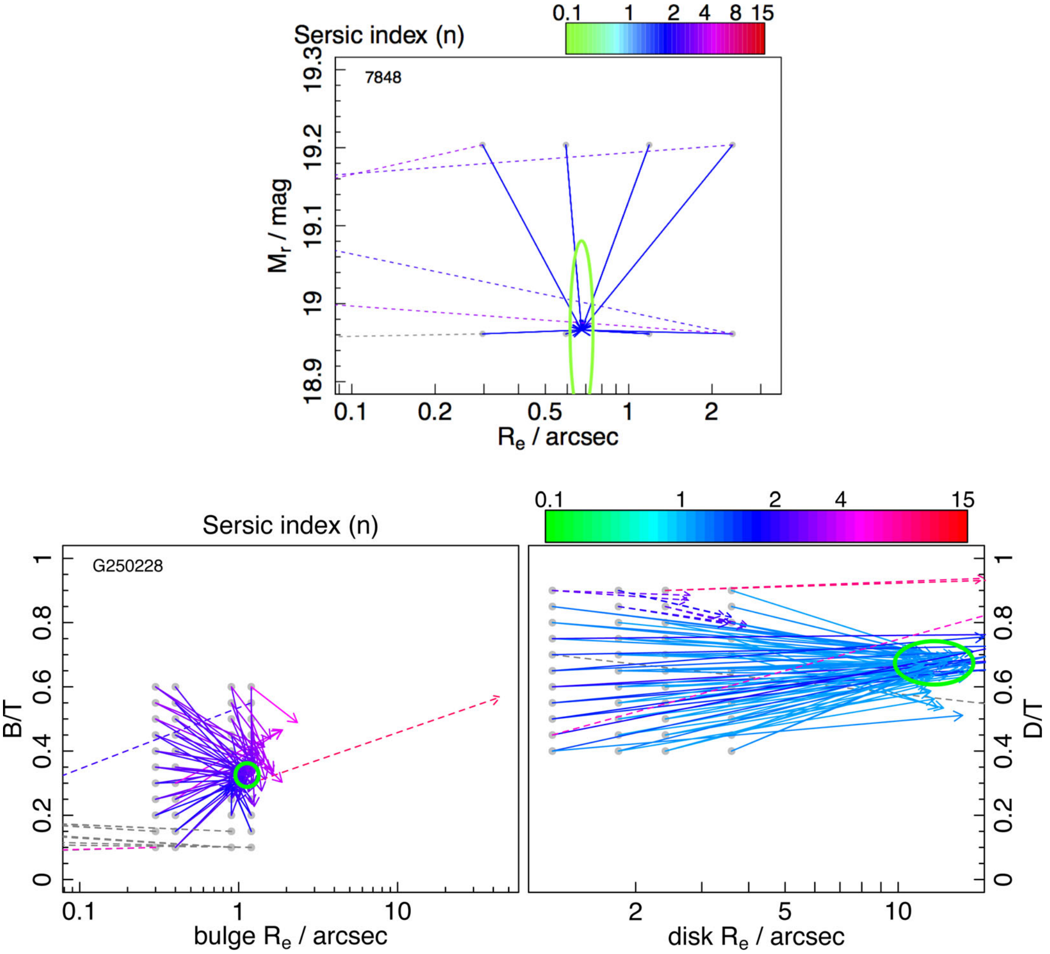


Figure 2. Convergence plot examples for a single-component galaxy in the top panel and a double component galaxy in the bottom panel. The plots show the starting (grey points) and end values (arrow head) for several fit parameters making it easy to evaluate how well the galaxy was fit. For the single-component fits, we show the galaxy's total magnitude, size and Sérsic index. For the two-component fit, we show the bulge-to-total (B/T) and disc-to-total (D/T) flux ratios of the components instead of the magnitude. The green ellipse is centred at the median output values and its size corresponds to the adopted error on the median. Note that for the single-component fit, the error on the magnitude corresponds to our error floor of 0.11 mag. The dashed lines indicate the fitting outcomes we consider to have failed. Note that the arrow colours correspond to the final Sérsic index values only (if it is grey, then the Sérsic index is outside the range of the values we considered physical, see Section 3.1.4) and each grey point has several arrows associated with it due to the combination of starting values of our initial grid. See the text for a detailed description.

a PSF using PSFEXTRACTOR (Bertin 2013) which is later used to convolve the GALFIT models. The SEXTRACTOR outputs are also used to inform the fitting of neighbouring objects as well as provide initial starting values for the GALFIT run (for full details on SIGMA, see Kelvin et al. 2012). During the actual GALFIT routine the primary and all secondary objects are simultaneously modelled using a Sérsic function. In the case of a two-component fit, GALFIT minimizes the χ^2 over two Sérsic functions centred on the primary object while also fitting the secondary objects with a single Sérsic profile.

To identify convergence to the global minimum, we use a grid of initial starting points (as previously discussed) for both the bulge and disc components as described below.

(i) Two-component fitting: a total of 88 starting combinations varying input parameters as follows,

(a) ratio of bulge to disc size (bulge size/disc size, R_{SE}= SExtractorradius):
1:1 ($R_{\mathrm{SE}}/R_{\mathrm{SE}}$),
1:2 ($0.75 \times R_{\mathrm{SE}}/1.5 \times R_{\mathrm{SE}}$),
1:4 ($0.5 \times R_{\mathrm{SE}}/2 \times R_{\mathrm{SE}}$),
1:9 ($0.33 \times R_{\mathrm{SE}}/3 \times R_{\mathrm{SE}}$),
(b) two sets of component starting Sérsic index:
$n = 4+1$ (bulge + disc)
$n = 2.5+0.7$ (bulge + disc),
(c) component bulge and disc flux ratio:
60 per cent : 40 per cent (bulge : disc) to 10 per cent : 90 per cent in steps of 5 per cent,

(ii) Single-component fitting: a total of 33 combinations of the input parameters

(a) $R = 2, 1, 0.5, 0.25 \times R_{\mathrm{SE}}$
(b) Sérsic index $n = 1, 2, 3, 4$
(c) total magnitude mag $= 1, 0.8 \times M_{\mathrm{SE}}$
(d) one additional model starting with $R = R_{\mathrm{SE}}$ and mag $= M_{\mathrm{SE}}$ and $n = 2.5$.

R_{SE} and M_{SE} denote the initial size and magnitude values taken from the SEXTRACTOR outputs for the entire galaxy.

Fig. 2 shows an example convergence plot for a single-component fit (top, G7848) and a double-component fit (bottom, G250228). The plots show the grid of initial conditions (grey points) and vectors pointing to the final solution for each parameter combination. We plot the size versus magnitude plane for the single fits and the size versus component light fraction plane for the double fits (bulge component, left, disc component, right). The colour bar at the top shows the Sérsic indices considered and spans the same range for all convergence plots. The arrows pointing to the final output parameters are coloured according to the final Sérsic index. In practice (e.g. Fig. 2, bottom), not all fits converge to a plausible solution and hence screening is required to remove obvious bad fits. Dashed lines indicate fitting outcomes which were excluded due to bad values (see the screening descriptions below) or a large reduced χ^2. If the (dashed) lines are grey, the final Sérsic index was outside the range displayed in the colour bar. The green ellipse shows the median solution and its size corresponds to the adopted error on the median, i.e. the error is symmetrical and taken to be the average of the 16th and 84th percentile range. We produce convergence plots for all one- and two-component fits of our 7506 galaxies. As mentioned previously convergence towards a tight median value is by no means assured and a number of situations need to be managed, including component flipping, unphysical solutions and poor-quality fits. We refer to this management as screening and define the various steps below.

3.1.2 Screening via profile switching

Each of our 2247 two-component systems will have 88 model outputs from our grid of initial parameters. When fitting the galaxies, component 1 has been assigned the bulge initial parameters and component 2 the disc initial values. Since GALFIT components can migrate significantly, we ensure, after the fitting has finished, that the more compact component is taken as the bulge and the more extended as the disc. However, we find some cases where, even though the bulge is smaller in size, the disc has the higher Sérsic index. Visually inspecting a number of the resulting profiles, we find that the more extreme cases typically are bad fits and flag these (see Appendix A). Additionally, we relax the criterion for switching the components and allow bulges with lower Sérsic index than the disc if they are no more than 10 per cent larger than the disc. For 1447 galaxies, at least one of the 88 parameter combinations required switching the profiles output by GALFIT.

3.1.3 Screening via rejection of poor-quality fits

We also reject fits with poor reduced χ^2 values. To decide on an appropriate reduced χ^2 cut, we randomly inspected 20 fitting outcomes for each of 100 two-component galaxies. For each fitting outcome, we decided (by eye) whether it was acceptable or not based on the light profile of the model and the resulting residuals. Fig. 3 shows the distribution of reduced χ^2 for these galaxies split into 'good' and 'bad' fitting outcomes, shown as green and red histograms, respectively. The vertical dashed line indicates the final cut of reduced $\chi^2 = 4$ which is left deliberately high to ensure that even for galaxies with a lot of structure, we exclude none of the acceptable fit outcomes and have enough outputs to evaluate the 'best' fit. This cut is also implemented for the single Sérsic fits. In total, 20 505 (∼10 per cent) of the 197 736 fitting results were removed from the two-component sample and 5686 (∼7 per cent) from the 173 547 fitting results of the single-component sample.

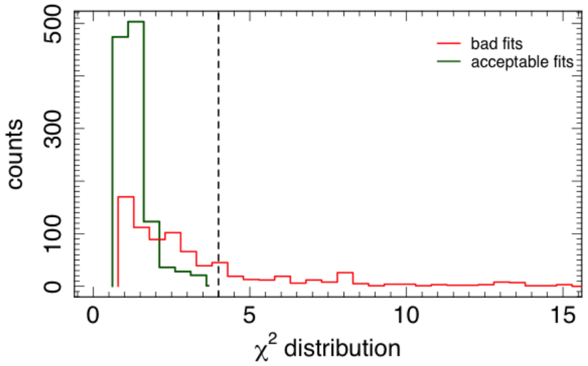


Figure 3. Shown is the reduced χ^2 distribution of a test sample of 100 two-component galaxies for which 20 fitting outcomes each were visually inspected. The green histogram shows the fits classified as 'good' and the red histogram shows the bad fits. The dashed vertical line is the implemented reduced χ^2 cut.

3.1.4 Screening via rejection of unphysical fits

Fig. 4 summarizes the derived GALFIT-fitted values of all combinations for our bulges and discs showing in the upper panels the bulge (left) and disc (right) sizes, and in the lower panels the bulge (left) and disc (right) Sérsic indices. Taking the top-left panel, we see that the bulge sizes follow two distinctive bands, one at plausible sizes (i.e. 0.1–10 arcsec scales), and one at unphysical sizes (0.001–0.01 arcsec) given the data resolution of ∼1.5 arcsec. We reject the fitting results which result in overly compact 'bulges' and remove these from further considerations (red dashed line). Overly large bulges or discs are not a prominent problem but we remove obvious outliers based on the distribution of all solutions. Similarly, in Fig. 4 (bottom), we show the Sérsic index distribution. Once again, the vertical red dashed lines indicate the division between the fitting results we consider physical and those we consider unphysical and that should therefore be rejected. The limits adopted leading to rejection (red dashed lines) are:

(i) for bulge sizes: $R_\mathrm{e} < 0.01$ arcsec or $R_\mathrm{e} > 20$ arcsec
(ii) for disc sizes: $R_\mathrm{e} < 0.1$ arcsec or $R_\mathrm{e} > 200$ arcsec
(iii) for bulge and disc Sérsic index: $n < 0.1$ or $n > 15$.

These cuts are deliberately permissive and should cut out only the most unrealistic fitting outcomes. For our two-component galaxy sample, of the 197 736 combinations fitted, we reject 17 343 (∼18 per cent) based on bulge size, 616 based on disc size (<1 per cent), 18 357 (∼19 per cent) based on bulge Sérisc index and 11 462 (∼12 per cent) based on disc Sérsic index. For the single-component fits, we reject fitting outcomes based on the same limits as the disc size and Sérsic cuts of the two-component fits. Of the 173 547 combinations fit to the single-component systems, we reject 2753 (∼3 per cent) based on their size and 2355 based on their Sérsic index (∼3 per cent). In total, 49 688 (∼25 per cent) fitting results are rejected from our two-component sample fits and 7757 (∼4 per cent) from our single-component fits. Note that, in many cases, fitting results are rejected by more than one criterion (i.e. reduced χ^2 and/or size and/or Sérsic index).

We also screen our galaxies for various flags, described in Appendix A. However, we only consider two flags important during the component $\mathcal{M}_\star$–R_e relation fits, namely the very high (or low) B/T galaxies and reversed Sérsic index galaxies. We deem the high (and low) B/T galaxies single-component systems and move them from our two-component sample to our single-component sample.

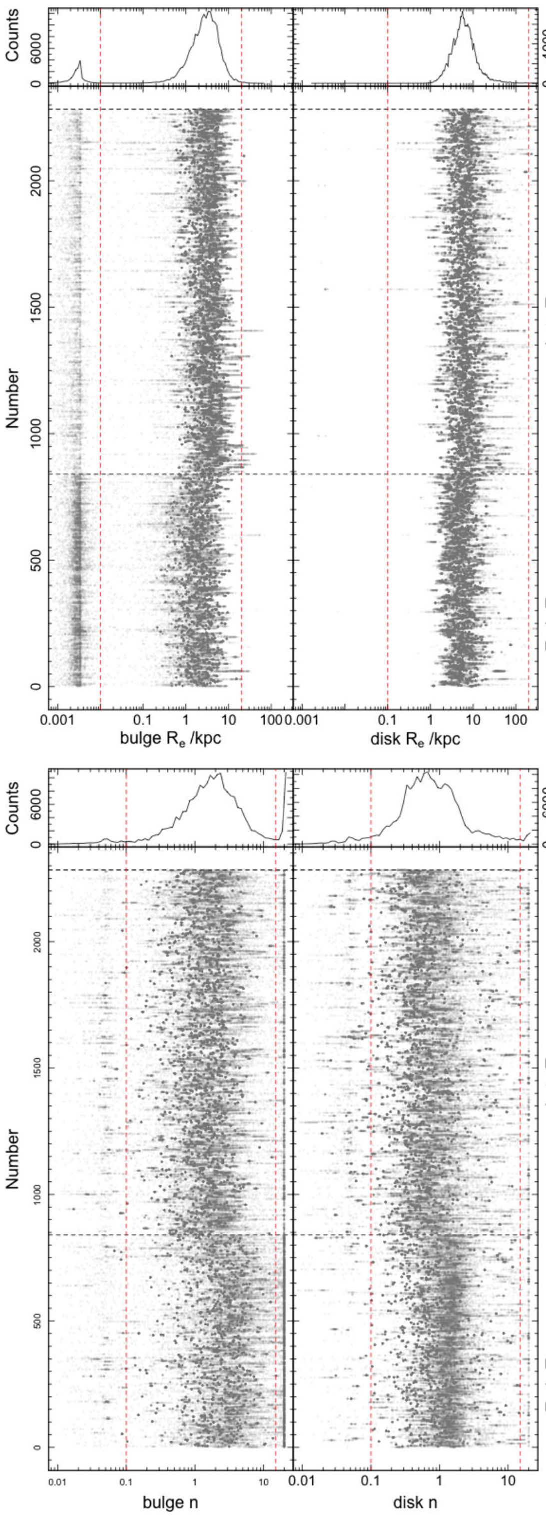


Figure 4. The top panel shows the distribution of the output size for all bulges (left) and discs (right) for the fitted two-component models (no reduced χ^2 cut has been imposed) of each galaxy (indicated as 'Number'). The dashed vertical lines show the implemented cuts on the size before the median is established. The bottom panel shows the corresponding distribution for the output Sérsic index for the bulges (left) and discs (right). In all panels, the galaxies are sorted by *Hubble* type with the late-type two-component systems at the top and early-type two-component systems at the bottom. The horizontal dashed black line shows where the *Hubble* type changes.

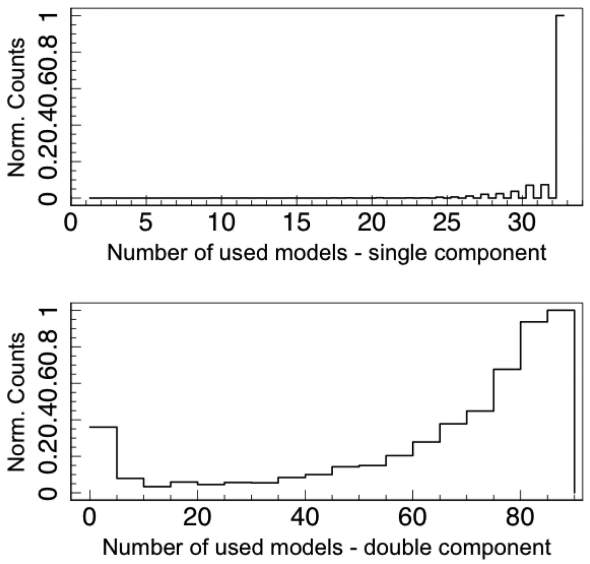


Figure 5. Here we show the distribution of the number of fitting outcomes used to calculate the median. The top panel shows the distribution for single-component fits based on single-component galaxies only. The bottom plot shows the distribution of the number of two-component fits used for two-component galaxies only. It can be seen that the single Sérsic fits generally converge nicely and the two-component fits have a broader distribution with a spike at very low numbers.

Galaxies with inverted Sérsic index have bulges with lower n than discs. Visually inspecting several of the profiles, we find that in most cases, these are bad fits, i.e. we find the disc Sérsic index $n > 2$. This itself would not be a problem if the errors reflect our confidence in the fit. Many of these profiles, however, have converged to this unphysical solution. We find 182 late-type two-component systems and 87 early-type two-component systems have inverted Sérsic index and converged profiles. We remove these galaxies from our component consideration, but use their single-component profile fits to establish their global $\mathcal{M}_\star$–R_e relation.

3.1.5 Final parameter selection

For each galaxy, we consider two possible profile fit solutions taken from the remaining fitting results:

(i) the minimum χ^2 model with the associated GALFIT parameters and errors, and

(ii) the median fit values of the remaining fitting results and the 16th and 84th percentiles (i.e. the 1σ deviation of a normal distribution) as an uncertainty indicator.

While the minimum χ^2 solution should represent the best formal fit from our grid, the median model is our preferred solution, as the errors on the median reflect the level of convergence and robust errors are critical. Note that the median values are calculated for each output parameter individually and do not directly represent any single solution. In cases where the fitting converged, the median and minimum χ^2 solution will be almost identical and the 16th and 84th percentile range often is smaller than the GALFIT errors. We therefore adopt an error floor of 10 per cent of the median value, which assures that in almost all cases, the median solution is consistent with the minimum χ^2 solution within the estimated errors.

Fig. 5 shows the histogram of the number of the remaining fitting results used to calculate the median for all single-component systems (top) and all two-component systems (bottom). The single-component fits often converge and the histogram peaks at ∼33 fitting results with only a small tail towards lower numbers. The two-component fits on the other hand do not converge as often. It is encouraging that the peak is ≥85 fitting results, however, there is a large fraction of galaxies for which very few solutions remain for the median calculations. In addition, there is a rise towards very low numbers indicating that some galaxies are likely too complex to be fit with two components only. We find that, while the galaxies with low model counts span the whole mass range, most of them lie close to our upper redshift boundary and were classified as late-type double-component systems. This shows the inherent difficulty of fitting multiple component systems in poorer image quality regimes. In addition to the tightness of the median errors, we can also use the number of fitting results left for the calculation of the median to help establish our confidence in the fitting results.

3.2 Convergence examples

We finish this section by presenting five examples which highlight some of the issues encountered and show that the median values present a robust alternative to the minimum χ^2 solutions. Fig. 6 (upper panels) shows the convergence plots (as introduced in Section 3.1.1 and Fig. 2), and diagnostic plots for the median and minimum-reduced χ^2 solutions (middle and lower panels, respectively). The four images which make up the diagnostic plots (middle and lower left panels) show, from the top left in a clockwise direction, the SDSS r band image stamp, the model produced by GALFIT, the residual and the SEXTRACTOR segmentation map overplotted on the SDSS image stamp (the primary object is shown in purple and the secondaries in green). The red and blue ellipse show the R_e of the bulge and disc, respectively. The yellow ellipse is the original SEXTRACTOR radius and the cyan ellipses show the radii for which the surface brightness was evaluated. Also shown (middle and lower-right panels) are the 1D light-profile comparisons. The black points are the values extracted from within the blue ellipses, the red and green lines are the 1D light profile for the bulge and disc, evaluated from the GALFIT model and the green line is the total light profile. The lower inset panel shows the residual between the model and data. Below we discuss five examples of various fit and convergence outcomes. We only discuss examples for two-component galaxies here, since we find that the single Sérsic fits generally converge well (e.g. example a):

(i) *Example a: full convergence*
Fig. 6(a) shows the ideal case of full convergence from all combinations of initial parameters. The median and minimum χ^2 solution diagnostic plots also show that both reached the same answer. The residual images show little structure and the final errors of the median fit are small as indicated by the green ellipse on the convergence plot. For two-component fits, we consider them fully converged when they have more than 80 fitting results remaining after rejection of spurious fits and the error on the median is set to the 10 per cent error floor. This is the case for 423 (∼19 per cent) galaxies. Similarly, for single components, over 30 fitting results must remain for the median calculation and all errors are set to the 10 per cent error floor. This is true for 4566 (∼87 per cent) single-component galaxies.

(ii) *Example b: partial convergence*
The median and minimum-reduced χ^2 model diagnostic plots in Fig. 6(b) show good agreement. From the convergence plot, it is obvious that many of the solutions found by GALFIT were rejected during the screening process, due to an unphysical Sérsic index or high reduced χ^2. The remaining models after screening show convergence resulting in a good solution with tight errors. For two-component systems, we consider good convergence to be reached when we have 60–80 solutions remaining, with the errors set to the 10 per cent error floor. This is the case for 297 (∼13 per cent) galaxies. Equivalently, for single components, 25–30 solutions must remain for the median calculation with the errors set to the error floor. We find this true for 220 (∼4 per cent) single-component galaxies.

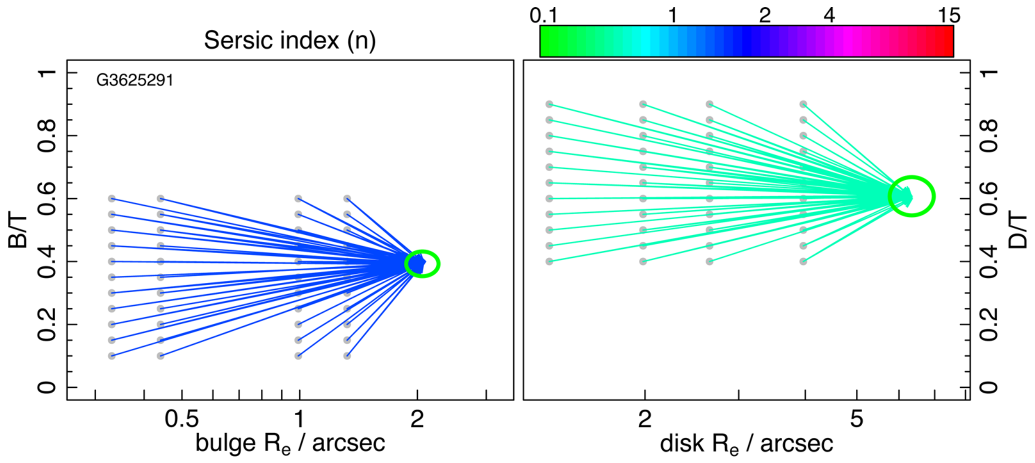
(iii) *Example c: two plausible solutions*
The diagnostic plots in Fig. 6(c) suggest that the median model gives a physically more meaningful two-component solution than the minimum-reduced χ^2 solution which is converging towards a single-component solution. The convergence plot, however, highlights that the median solution is not one of the actual solutions found by GALFIT. Nevertheless, the errors on the median (green error ellipse) enclose both solutions. While the median fit cannot be considered as robust, this uncertainty is fairly reflected in the final errors. To establish whether several plausible solutions have been found, we test how many solutions are near the median. If less than 10 per cent of the solutions of the median, for at least one of the size, Sérsic index or B/T values of either the bulge or disc, lie within 10 per cent (i.e. the error floor), then we consider the fits to have converged to several plausible solutions which are distinct from the median. This is the case for 205 (∼5 per cent) double-component systems. For the single-component systems, we consider the size, Sérsic index and magnitude and find six (<1 per cent) single-component galaxies have converged to several distinct solutions.

(iv) *Example d: no convergence*
Fig. 6(d) shows a case where no obvious single converged solution is found, but the median model returns acceptable parameters with an appropriately broad error distribution. The diagnostic plots also show that the median model returns a physically possible solution with good residuals. The minimum χ^2 solution, however, returns a fit where the bulge, even though it has a smaller R_e is the dominant component in the outer parts of the galaxy. Since the errors associated with the median model are large, this particular galaxy will not have much influence on the $\mathcal{M}_\star$–R_e relation we fit in Section 5.2, but using the median parameters and large error bars means that the galaxy will not be discarded from the sample. To test non-convergence, we use the same metric as in example 3, i.e. the percentage of solutions found within 10 per cent of the median. We consider galaxies not clearly converged if more than 10 per cent but less than 50 per cent of the solutions lie close to the median. We test the size, Sérsic index and B/T measurements for the bulge and disc and find for 931 (∼41 per cent) two-component systems that at least one of them is not converged. For the single-component systems, this is the case for 141 (∼3 per cent) galaxies.

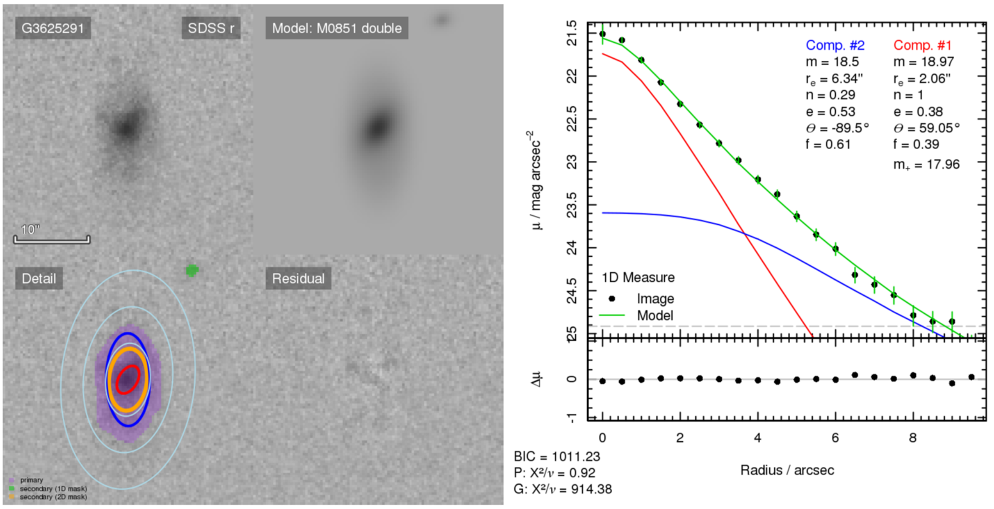
(v) *Example e: no solution*
Fig. 6(e) shows a case where convergence is found, however, all fits are excluded from the final catalogue due to the screening process. No median model diagnostic plot is shown due to all fit parameters being unrealistic. Only 120 (∼5 per cent) of our two-component galaxies and 129 (∼2.5 per cent) of our single-component systems fall into this category.

Convergence plots for all systems are available from the GAMA data base.

(a) Convergence plots



Median fit model



Minimum χ^2 model

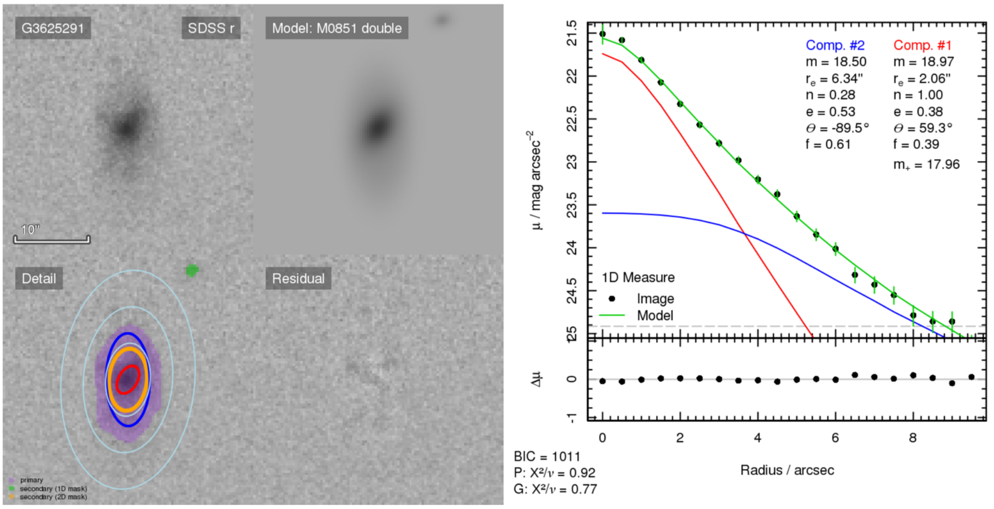
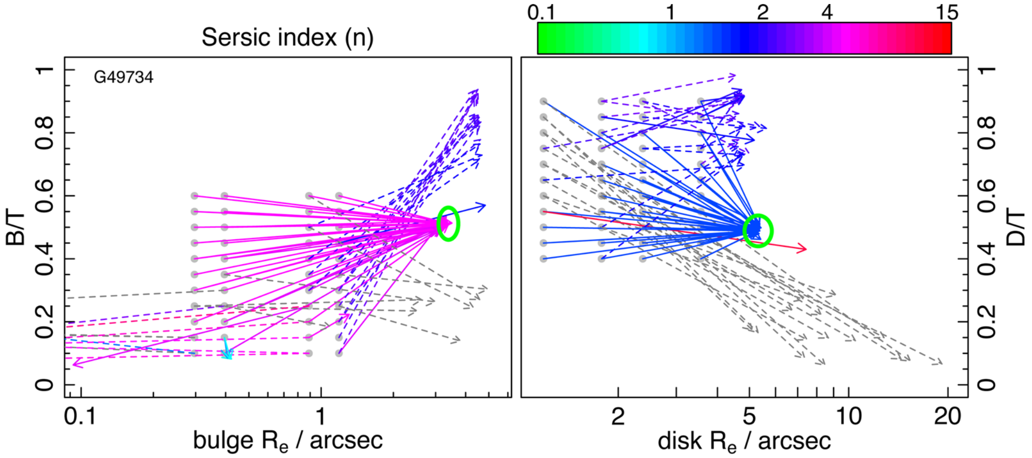
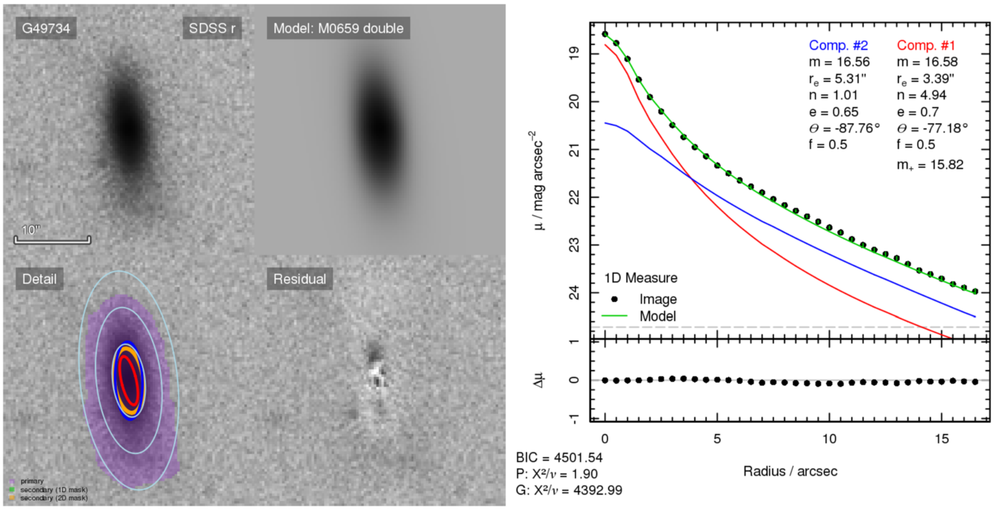


Figure 6. Presented are the convergence plots (top) and corresponding diagnostic fit plots for the median fit model (middle) and minimum-reduced χ^2 solution (bottom). We show four examples ranging from full convergence to no convergence (panels a, b, c, d) and one example where all fits are unrealistic (panel e). A detailed discussion can be found in the text.

(b) Convergence plots



Median fit model



Minimum χ^2 model

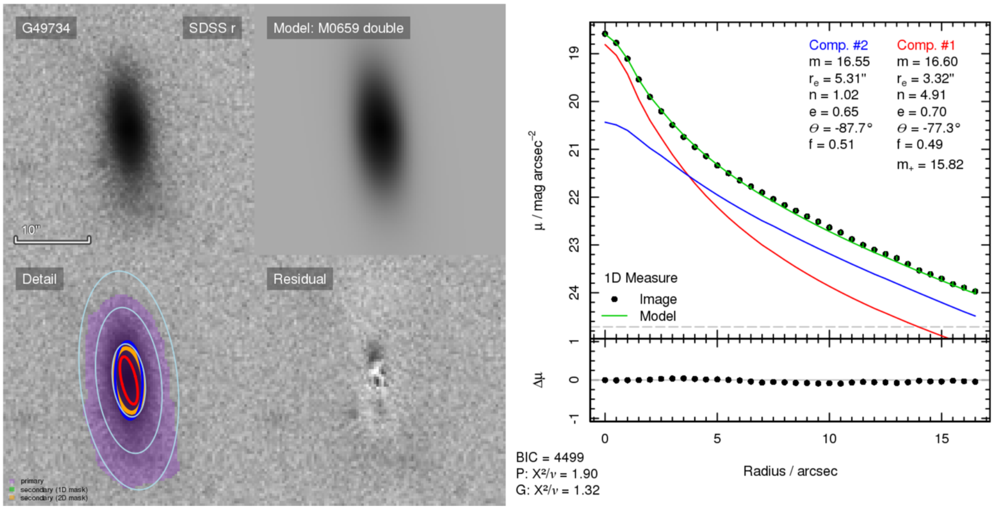
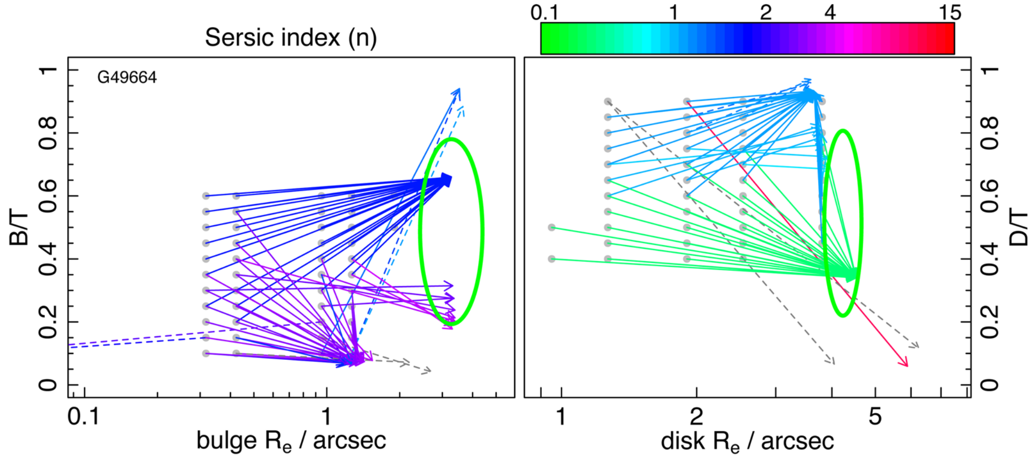
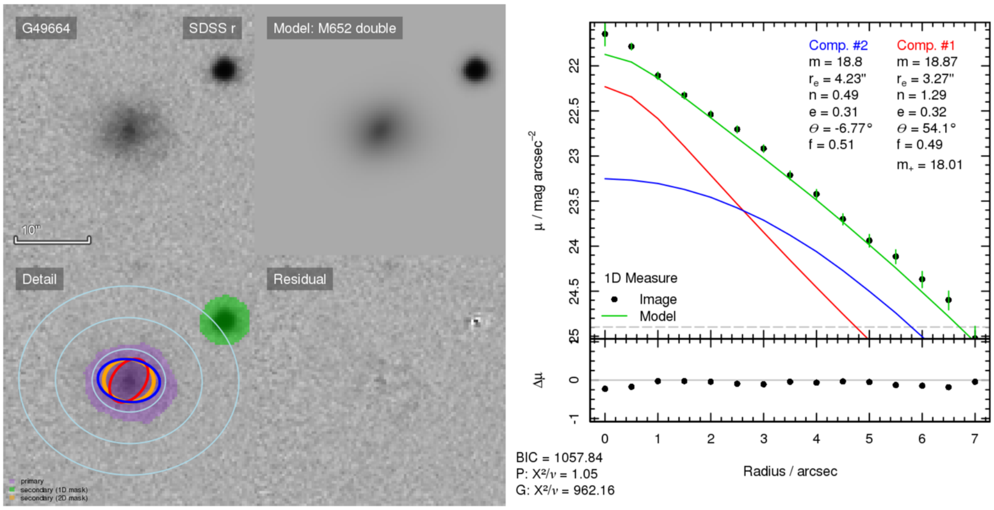


Figure 6 – *continued*

(c) Convergence plots



Median fit model



Minimum χ^2 model

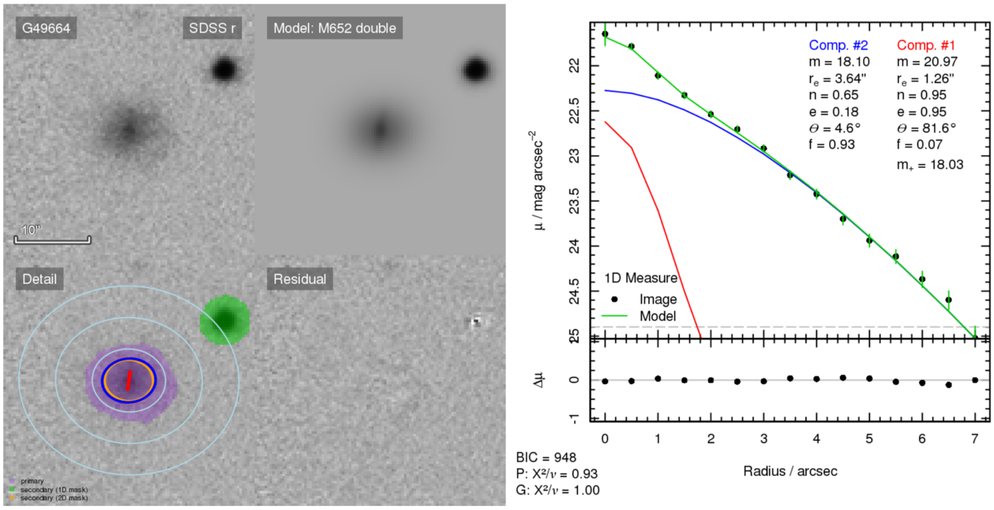
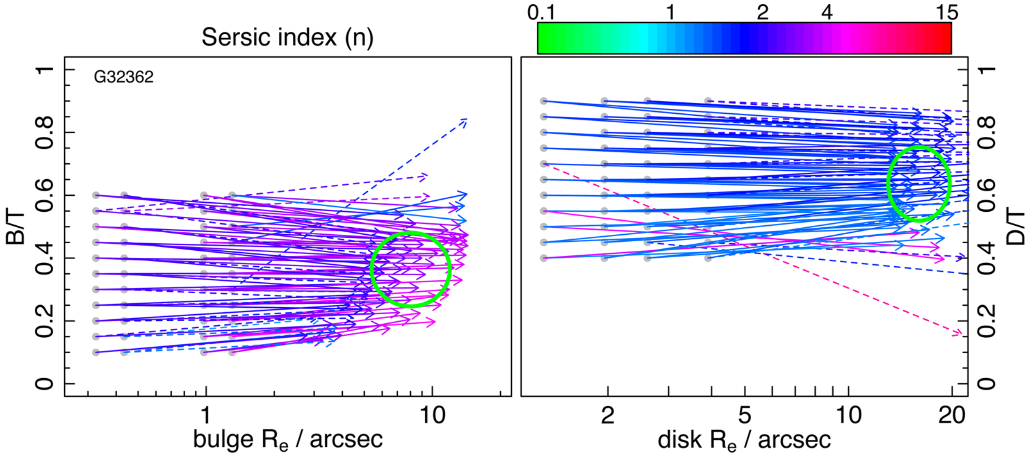
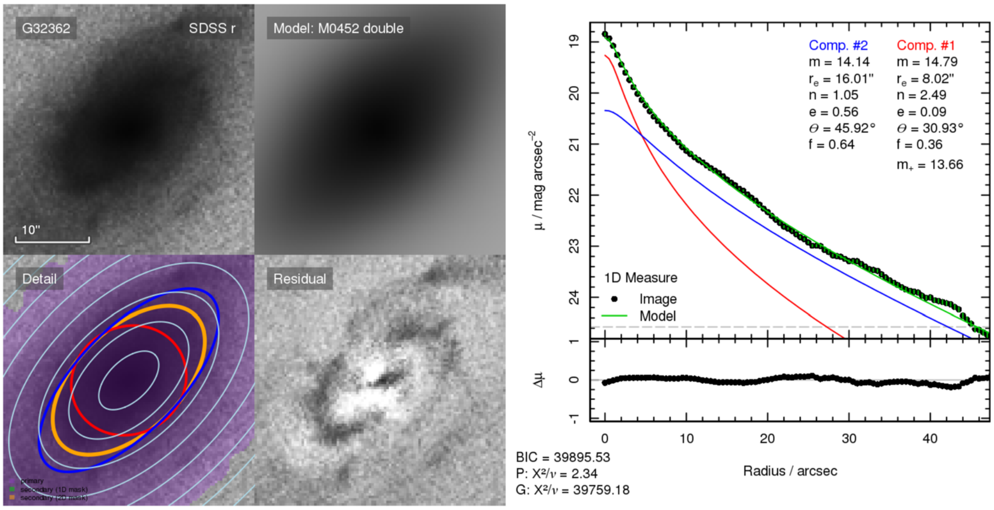


Figure 6 – *continued*

(d) Convergence plots



Median fit model



Minimum χ² model

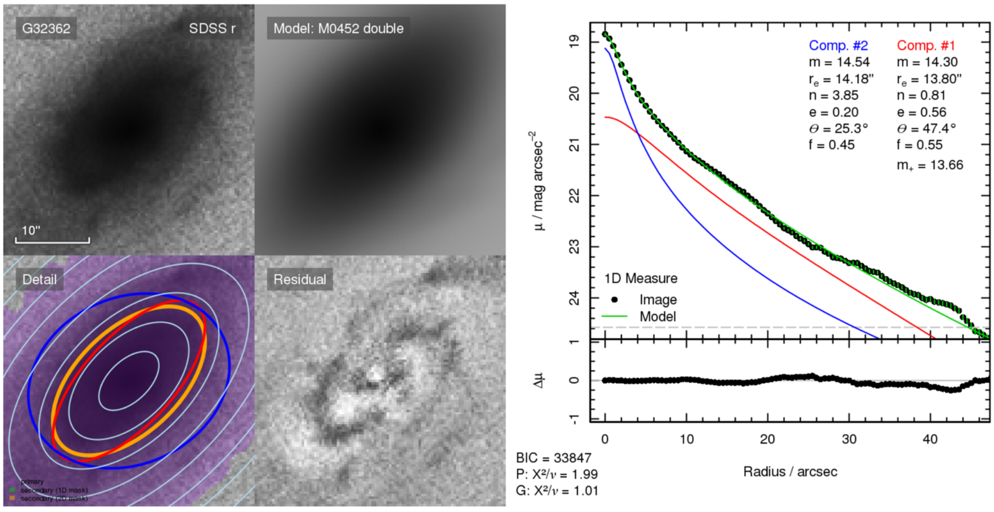
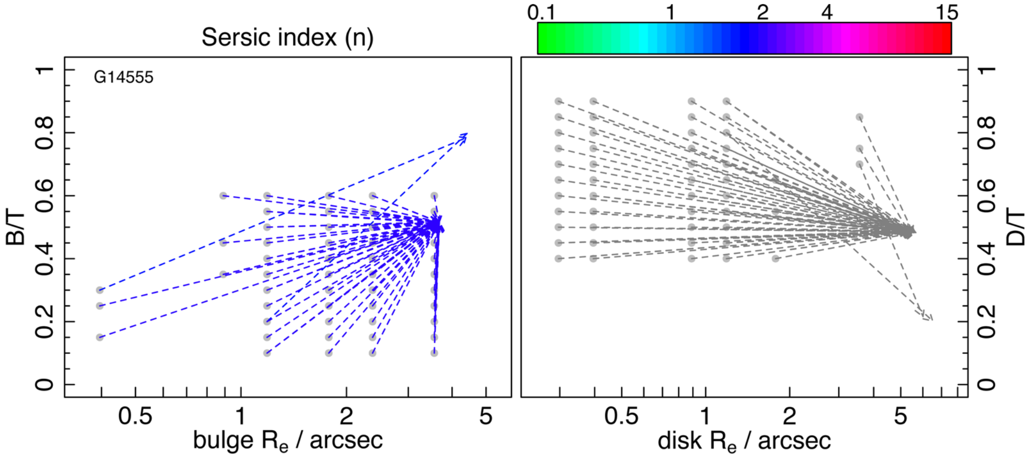


Figure 6 – *continued*

(e) Convergence plots



Median fit model

NA

Minimum χ^2 model

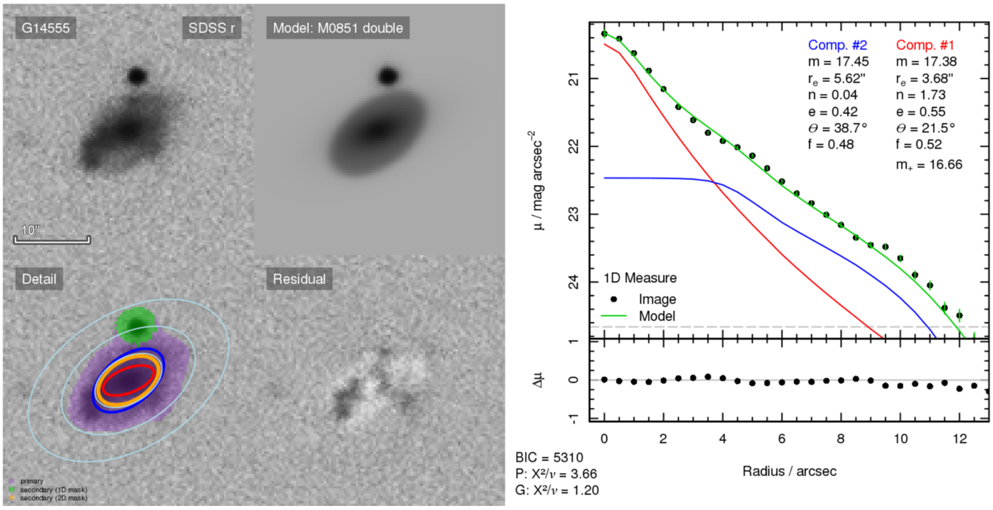


Figure 6 – *continued*

4 COMPONENT MASSES

To derive the $\mathcal{M}_\star$–R_e relations for our galaxy components, we now need component mass estimates. For the single Sérsic fits, we can directly use the GAMA stellar mass estimates from the StellarMassesv18 catalogue and apply the flux-scale correction (for a detailed description, see Taylor et al. 2011, essentially the flux-scale correction accounts for the differences between aperture matched and Sérsic photometry). These masses are based on synthetic stellar population models from the BC03 library (Bruzual & Charlot 2003) with a Chabrier (2003) initial mass function and the Calzetti et al. (2000) dust obscuration law. We find that, for our sample, the typical error, which has been derived in a Bayesian way (sections 3.2–3.4 of Taylor et al. 2011), is of the order of ~0.12 dex.

For the double-component galaxies, we calculate the component mass from the component colours (Driver et al. 2006) using the relationship between optical colour $(g - i)$ and mass-to-light ratio as calibrated by Taylor et al. (2011):

$$\log M_*/\mathrm{M}_\odot = -0.68 + 0.7(g - i) - 0.4(M_i - 4.58), \qquad (1)$$

where M_i is the absolute magnitude in the i band and we use the $(g{-}i)$ colour of either the bulge or disc to calculate the component mass. The stellar masses derived via equation (1) are estimated to be accurate within a factor of 2. Note that this equation is sensitive to the evolution of colour and magnitude, however, as our sample has a low-redshift range, the effects will be negligible.

Ideally, we would use bulge and disc colours derived from the bulge + disc decompositions in the *g* and *i* band, however, this is beyond the scope of this paper. Hence, we have to estimate the colours of the components. For this, we measure the PSF and total magnitudes of our galaxies in the *g, r,* and *i* band and we then use the GALFIT measured *r* band component magnitudes and B/T to estimate the bulge and disc colours.

We measure the core (i.e. PSF) and total magnitudes, which we correct for foreground extinction, in the *g, r,* and *i* band using LAMBDAR, a code developed to measure PSF+weighted aperture photometry (for more details, see Wright et al. 2016). We then equate the colours measured using the PSF magnitudes to bulge colour measurements (i.e. assuming the bulge has no colour gradient) and combine these colours with our *r* band bulge magnitude from GALFIT to obtain bulge flux measurements in both *g* and *i* (i.e. $m_{i,\mathrm{bulge}} = m_{r,\mathrm{bulge}} - (r - i)_{\mathrm{PSF,bulge}}$). In cases where the bulge colours could not be measured, we use the median bulge colour of the entire population as there is no significant trend between bulge colour and mass. We derive *g* and *i* disc fluxes by assuming that the disc flux in each band is equal to the LAMBDAR total flux minus the previously derived bulge flux.

We examine the disc $(g - i)$ colour distribution and find a small number of extreme outliers ($>3\sigma$) from the colour distribution, whose disc $(g - i)$ colours we subsequently replace with the running median disc $(g - i)$ colour.

Finally, we use the derived bulge and disc $(g - i)$ colours and total *i* band magnitudes to derive stellar mass estimates for each component according to equation (1). The component colour versus Sérsic index distribution is shown in Fig. 7 for bulges (top) and discs (bottom). Fig. 7 shows that for late-type galaxies, bulges are generally redder than discs but for early-type galaxies, the colours are very similar (for a detailed study of the wavelength dependence of bulge + disc decompositions in GAMA, see Kennedy et al. 2016). But it also highlights the problem galaxies for which the component colour had to be set to the median colour (vertical band in the bulge plot, top) in order to be able to calculate a stellar mass estimate.

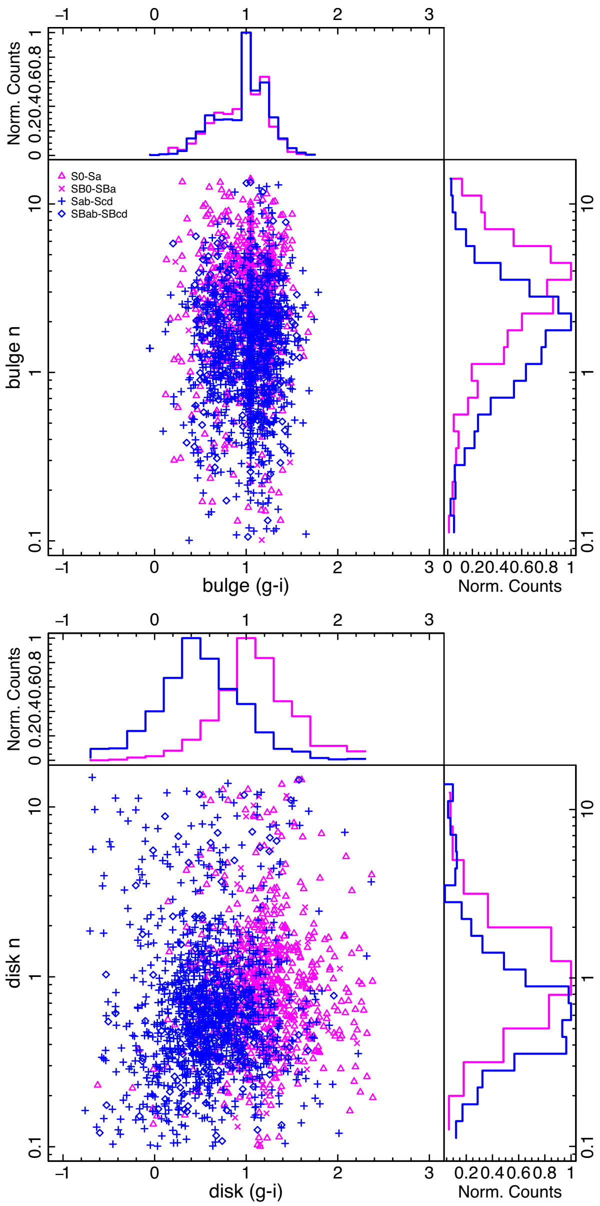


Figure 7. Shown are the $(g-i)$ colour versus Sérsic index distributions for the bulges (top) and discs (bottom) in our two-component galaxy sample. The colour coding in both plots is the same with late types in blue and early types in magenta. The bulges of late-type galaxies have smaller Sérsic indices than the early-type bulges. The Sérsic index distribution of the late-type discs also peaks slightly lower than the early-type discs. Late-type discs are also bluer than early-type discs.

5 $\mathcal{M}_\star$–R_e RELATIONS

We now present the $\mathcal{M}_\star$–R_e relations, first, for the different *Hubble* types and secondly, by structural components. We conclude our analysis by presenting a combined $\mathcal{M}_\star$–R_e relation for disc (i.e. Sd-Irr galaxies and disc components) as well as spheroids (i.e. ellipticals and classical bulge components).

Fig. 8 shows the selection of the final sample, based on total stellar mass and single Sérsic profile fits, used to fit the $\mathcal{M}_\star$–R_e relation. Note that we do not consider any galaxies below $\mathcal{M}_* = 10^8\,\mathcal{M}_\odot$, since number counts are too low to establish a robust weight (after flux-scale correction, this reduces the sample to 6788 galaxies). First, we find and exclude outliers from the general mass–size distribution (top panel of Fig. 8). For this, we fit the entire sample with a simple power law and remove all galaxies which are more than 3σ sigma offset from the best-fitting linear relation (28 galaxies in total). We then establish the lower mass limit for a volume-limited sample at $z = 0.06$. In the middle panel of Fig. 8, we plot the maximum redshift at which each galaxy can be seen versus its stellar mass. To establish the lower mass limit of a volume-limited sample, we find the point at which more than 95 per cent of our galaxies could be seen at a redshift of 0.06 (i.e. their maximum redshift is $z_\mathrm{max} \geq 0.06$, indicated as the dashed line). We find a lower mass limit of $\mathcal{M}_* = 10^9\,\mathcal{M}_\odot$ (solid blue line), which would reduce our sample size to 3679 galaxies. To include lower mass galaxies, we implement a smooth volume- and mass-limited sample for galaxies below $\mathcal{M}_* = 10^9\,\mathcal{M}_\odot$. For each galaxy, we evaluate if their measured redshift is larger than their expected maximum redshift and remove them (1624 galaxies removed). The bottom panel of Fig. 8 shows the resulting sample distribution. All galaxies in red are included in our final sample (5136 total) and all grey points are excluded from the volume-limited sample. For all galaxies below

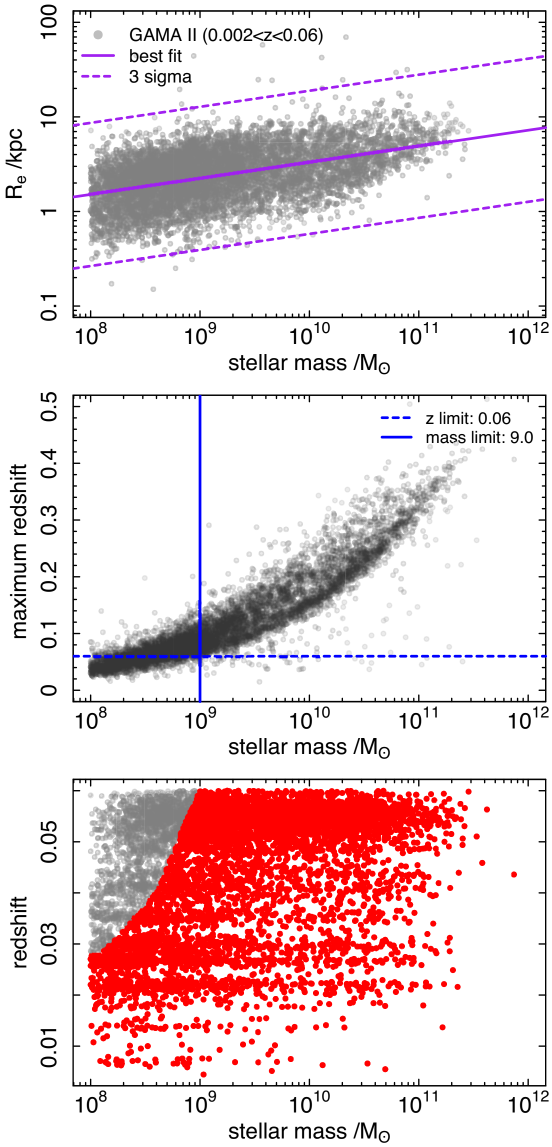


Figure 8. The top panel shows the total stellar mass – half-light size distribution (derived from single Sérsic fits) of the GAMAnear sample. All galaxies more than 3σ offset from the line of best fit are removed as outliers from our $\mathcal{M}_\star$–R_e relation fits. The middle panel shows the total stellar mass – maximum redshift distribution of the sample. The blue dashed line shows our redshift limit and the solid blue line the lower mass limit for a volume-limited sample. All galaxies to the right of the mass limit and above the redshift limit are included in the volume-limited sample. The bottom panel shows the total stellar mass – redshift distribution of our sample. For all galaxies below $10^9\,\mathcal{M}_\odot$, we have implemented a smooth volume-limited sample selection. Our final sample is highlighted in red.

$\mathcal{M}_* = 10^9\,\mathcal{M}_\odot$, we also calculate a V/V$_{\rm max}$ weighting based on their redshift and our sample redshift limits of $0.002 < z < 0.06$. For galaxies with $\mathcal{M}_* > 10^9\,\mathcal{M}_\odot$, the V/V$_{\rm max}$ is set to 1. To ensure that we only include galaxies with good, physical fits, we require:

(i) the flux-scale correction is within 0.5 and 1.5;

(ii) the components have Sérsic indices within the range $0.3 < n < 10$;

(iii) the components are resolved, i.e. $R_e > 0.5\ \times$FWHM of the PSF which is determined from each galaxy's fit image individually;

(iv) at least five solutions remained to calculate the median fit parameters.

This reduces the sample size to 2669 single-component and 1470 double-component systems.

We adopt a simple power law, following Shen et al. (2003) and Lange et al. (2015), to fit the $\mathcal{M}_\star$–R_e relation:

$$R_e = a\left(\frac{\mathcal{M}_*}{10^{10}\,\mathcal{M}_\odot}\right)^b, \tag{2}$$

where R_e is the effective half-light radius in kpc and $\mathcal{M}_*$ is the mass of the galaxy. To perform the actual fitting, we utilize the HYPERFIT package (Robotham & Obreschkow 2015) which estimates the $\mathcal{M}_\star$–R_e relation via Bayesian inference for each morphological group and component. During fitting, we assume uniform priors and each galaxy is weighted by its V/V$_{\rm max}$ and the (convergence) errors for each individual galaxy are fully taken into account during the fitting process.

5.1 Global $\mathcal{M}_\star$–R_e relations by *Hubble* type

To establish the global (i.e. single-component Sérsic fit) $\mathcal{M}_\star$–R_e relation by *Hubble* type, we have grouped the GAMAnear sample into five populations:

(i) 1564 late-type single-component galaxies (including 40 high/low B/T galaxies),

(ii) 890 late-type multicomponent systems (comprised of Sab-Scd and SBab-SBcd galaxies) of which 708 have also good two-component fits,

(iii) 580 early-type multicomponent systems (which include S0-Sa and SB0-SBa galaxies) of which 493 have also good two-component fits,

(iv) 806 early-type single-component (including 33 high B/T galaxies), and

(v) 372 LBS.

The resulting global $\mathcal{M}_\star$–R_e relations are shown in panel (i) of Fig. 9, from left to right, the plots are (a) Sd-Irr, (b) visually late-type multicomponent systems, (c) visually early-type multicomponent systems, (d) ellipticals and (e) LBS. The fit parameters can be found in Table 1 (i).

We find that the single Sérsic $\mathcal{M}_\star$–R_e relation fits to the different morphological types lie on almost parallel lines (i.e. comparable gradients but offset in normalization). Most two-component systems are more massive than Sd-Irr galaxies, but compared at the same mass, we find that two-component systems are smaller than Sd-Irr galaxies. Compared to the ellipticals, however, we find that two-component systems are larger at a given mass. This corroborates the composite nature of these galaxies, i.e. the disc surrounding the bulge makes their global R_e appear larger than ellipticals but smaller than Sd-Irr galaxies at a given mass, with the offset between the Sd-Irr and elliptical relation depending on the relative dominance of the disc or bulge component. The LBS galaxies, on the other hand, are our smallest and least massive population. The slope of their $\mathcal{M}_\star$–R_e relation is flatter than that of any of the other morphological types. Nevertheless, their sizes are mostly consistent with an extension of the elliptical population. In fact, within the errors, the LBS relation is consistent with the low-mass elliptical relation ($\mathcal{M}_* < 10^{10}\,\mathcal{M}_\odot$), see Table 1 (ii).

Here our morphological subdivisions are finer than in our previous work (L15) but broadly agree. In detail, our Sd-Irr class has a steeper $\mathcal{M}_\star$–R_e relation than the late-type relation in L15. This is

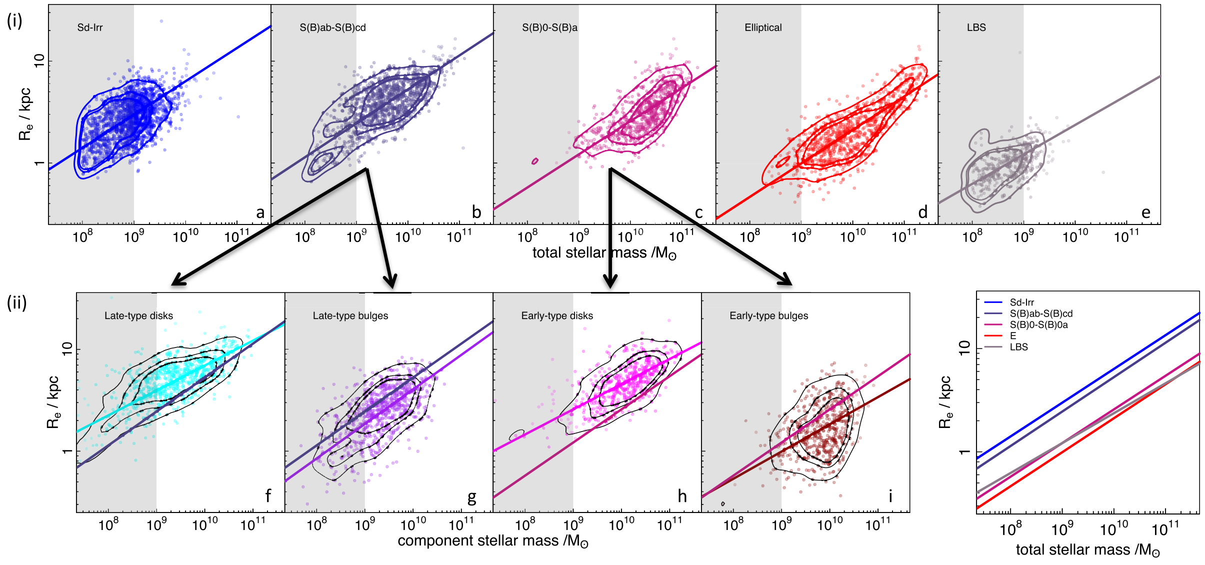


Figure 9. The top panel (i) shows the global $\mathcal{M}_\star$–$R_{\rm e}$ relation for: (a) Sd-Irr (blue), (b) late-type multicomponent (dark purple), (c) early-type multicomponent (rose), (d) ellipticals (red), and (e) LBS (grey) galaxies. Panel (ii) shows, from left to right, the component $\mathcal{M}_\star$–$R_{\rm e}$ relation for late-type discs (cyan) and bulges (purple), plots (f) and (g), respectively. We also plot the global S(B)ab–S(B)cd relation from panel (b) in comparison. The early-type discs (magenta) and bulges (dark red) are shown in plots (h) and (i). Again we plot the global S(B)0–S(B)a (shown in panel c) in comparison. The grey shaded areas indicate where our smooth volume-limited sample selection starts and galaxies are upweighted by their V/V$_{\rm max}$. The black lines are the 90th, 68th and 50th percentiles of the respective mass–size distributions. The arrows show from which population the components were derived. Finally, we show all global $\mathcal{M}_\star$–$R_{\rm e}$ relations in comparison in the far right plot in panel (ii). This highlights the similarities between the different populations, i.e. the relations are parallel but offset from each other depending on their bulge fraction.

Table 1. The regression fit parameters to equation (2) for the different *Hubble* types and structural components as well as a combined early- and late-type relation (see Fig. 9).

Case	a/kpc	b
(i) *Hubble* type		
Sd-Irr	6.347 ± 0.174	0.327 ± 0.008
S(B)ab-S(B)cd	5.285 ± 0.098	0.333 ± 0.009
S(B)0-S(B)a	2.574 ± 0.051	0.326 ± 0.015
E	2.114 ± 0.035	0.329 ± 0.01
LBS	2.366 ± 0.166	0.289 ± 0.019
(ii) Additional mass constraints		
E ($M_* \geq 10^{10}$ M$_\odot$)	1.382 ± 0.065	0.643 ± 0.032
E ($M_* \geq 2 \times 10^{10}$ M$_\odot$)	0.999 ± 0.089	0.786 ± 0.048
E ($M_* < 10^{10}$ M$_\odot$)	1.978 ± 0.077	0.265 ± 0.022
E ($M_* < 2 \times 10^{10}$ M$_\odot$)	2.108 ± 0.041	0.326 ± 0.012
(iii) Structural components		
late-type disc	6.939 ± 0.17	0.245 ± 0.008
late-type bulge	4.041 ± 0.129	0.339 ± 0.014
early-type disc	4.55 ± 0.097	0.247 ± 0.015
early-type bulge	1.836 ± 0.054	0.267 ± 0.026
(iv) Combined case		
all discs	5.56 ± 0.075	0.274 ± 0.004
all discs + LTB	5.125 ± 0.065	0.263 ± 0.004
final $z = 0$ discs	5.141 ± 0.063	0.274 ± 0.004
E + ETB	2.033 ± 0.028	0.318 ± 0.009
final $z = 0$ spheroids	2.063 ± 0.029	0.263 ± 0.005
global late types	4.104 ± 0.044	0.208 ± 0.004

Table 2. Fitting parameters taken from L15 (Table 2, 3 and B2) for the r band morphological late- and early-type $\mathcal{M}_\star$–$R_{\rm e}$ relation.

Case	a	b
Late type	3.971 ± 1.745	0.204 ± 0.018
Early type	1.819 ± 1.186	0.46 ± 0.023
Early type $M_* > 2 \times 10^{10}$ M$_\odot$	1.390 ± 1.557	0.624 ± 0.033

an effect of our sample selection. In fact, fitting an $\mathcal{M}_\star$–$R_{\rm e}$ relation to a combined sample of Sd-Irr and all two-component systems [see Table 1 (iv)] results in a relation fully consistent with the late-type relation in L15 (reproduced in Table 2). Our elliptical class has a shallower $\mathcal{M}_\star$–$R_{\rm e}$ relation compared to the early-type relation in L15. This is also largely a sample selection effect, caused by the relative increase in the number of low-mass to high-mass ellipticals within the sample. Fitting a high-mass elliptical relation (see Table 1 (ii), similar to L15, with $\mathcal{M}_* > 10^{10}\,\mathcal{M}_\odot$), we again find good agreement with our earlier work.

5.2 $\mathcal{M}_\star$–$R_{\rm e}$ relations by galaxy component

The distributions and fits to the structural component $\mathcal{M}_\star$–$R_{\rm e}$ relations are shown in panel (ii) of Fig. 9 and the fitting parameters can be found in Table 1 (iii).

From left to right, panel (ii) shows the late-type discs and bulges (LTD and LTB, plots f and g, respectively) followed by early-type discs and bulges (ETD and ETB, plots h and i). In each panel, we also show the global $\mathcal{M}_\star$–$R_{\rm e}$ relation of the population they were derived from, which is indicated by the large black arrows. Not surprisingly, we find that generally discs are larger and bulges are smaller than the global single Sérsic fits of the population they were derived from. Additionally, our data show that the component $\mathcal{M}_\star$–$R_{\rm e}$ relations

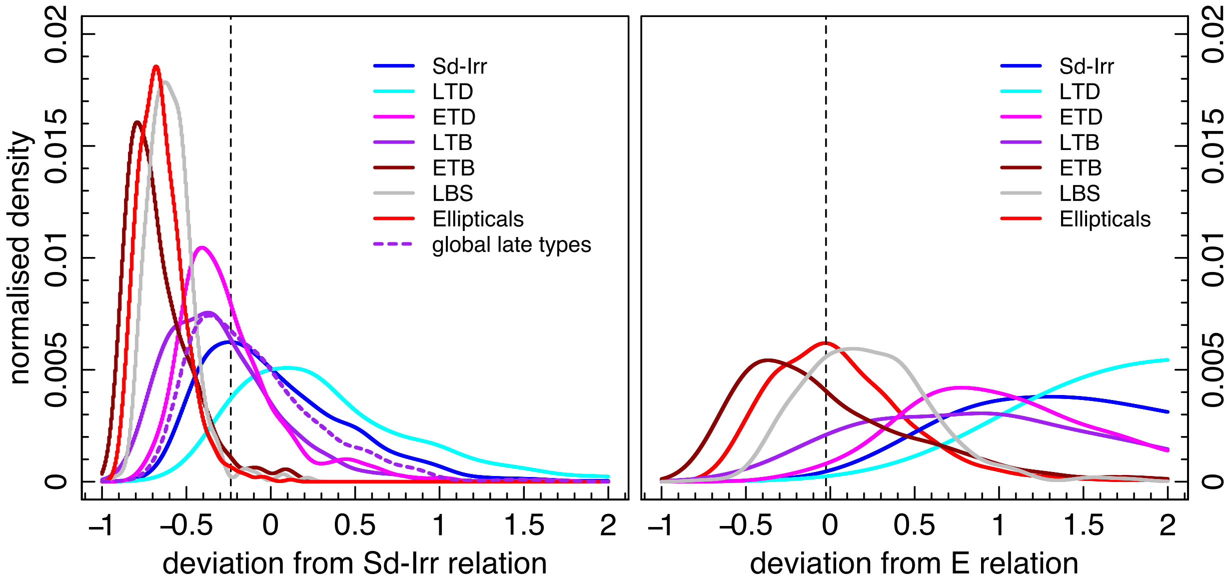


Figure 10. Shown are the relative fractional deviations of the different components from the best-fitting $\mathcal{M}_\star$–R_e relation of the tested parent population. The left-hand plot shows the deviation of the data from the Sd-Irr $\mathcal{M}_\star$–R_e relation and the right-hand plot shows the deviation from the elliptical relation. In both plots, the area under the curve has been normalized over the range of deviations shown in the plot.

are typically less curved than the global relation they were derived from. However, the late-type discs are an exception to this and show a slight upward curvature at high masses. This is in qualitative agreement with results from Bernardi et al. (2014) who found that the discs of late-type galaxies cannot be fit with a single power law, whereas the bulges of early-type galaxies follow a pure power law which is not exhibited by their parent population.

We now wish to establish whether these component fits are consistent with the Sd-Irr or E $\mathcal{M}_\star$–R_e relations. In particular, we wish to test the following hypotheses:

(a) ETD and Sd-Irr galaxies are associated,
(b) ETB and ellipticals are associated,
(c) LTD and Sd-Irr galaxies are associated,
(d) LTB and ellipticals are associated,
(e) LTB and Sd-Irr galaxies are associated,
(f) LBS and ellipticals are associated,
(g) LBS and Sd-Irr galaxies are associated,
(h) S(B)ab–S(B)cd systems fit with a single component (see Section 5.3) and Sd-Irr galaxies are associated.

In Fig. 10, we visualize the affiliation of various components with either the Sd-Irr (left) or ellipticals (right) by plotting their relative deviations, defined as

$$(R_\mathrm{observed} - R_\mathrm{predicted})/R_\mathrm{predicted}, \tag{3}$$

where R_observed represents the sizes of the tested populations, and $R_\mathrm{predicted}$ represents the predicted size using either the Sd-Irr or elliptical $\mathcal{M}_\star$–R_e relation. In both panels, the area under the curve has been normalized and each population is colour coded as shown in the legend with the dashed, vertical black line showing the peak of the deviations for the Sd-Irr galaxies and ellipticals, i.e. the populations to which we compare. The disc components, late-type bulges and the global late-type populations all broadly align with the Sd-Irr relation. The late-type discs tend to higher deviations indicating that they are larger than the Sd-Irr population, however, their peak deviation is close to 0. The good agreement of the late-type bulges with the Sd-Irr relation hints at their possible ‘pseudo’-bulge nature which is corroborated by their Sérsic index distribution. On the other hand, dust in latet-type galaxies can artificially lower the Sérsic index of LTB, making them appear larger and thus fall within the parameter space of the Sd-Irr relation. Ellipticals, early-type bulges and LBS on the other hand do not align with the Sd-Irr relation and their distributions are completely offset. Conversely, we find that the LBS and early-type bulges visually align with the elliptical $\mathcal{M}_\star$–R_e relation. They have similar peaks, however, the early-type bulges have a broader distribution, which is in good agreement with the larger scatter observed in Fig. 9. Whether this is intrinsic or a by-product of the decomposition is unclear. Late-type bulges, discs and Sd-Irr galaxies are offset from the elliptical distribution and do not follow their relation.

In addition to the qualitative nature of Fig. 10, we also perform a two sample Kolmogorov–Smirnoff-test (KS-test) for each hypothesis. For this, we compare the different samples to the $\mathcal{M}_\star$–R_e relation of either Sd-Irr or elliptical galaxies. Since the KS-test, in essence, compares the cumulative distributions of two populations in one dimension, it does not take into account the spread of our data around the $\mathcal{M}_\star$–R_e relation. Hence, we decided to bin our data in $\Delta\log(M_\star) = 0.2$ steps to establish the median mass and size of the bin and we use the median bin mass to calculate the expected size based on either the Sd-Irr or elliptical $\mathcal{M}_\star$–R_e relation. This also allows us to test whether the expected size distribution from the $\mathcal{M}_\star$–R_e relation fit agrees with the observed median size distribution of the sample, even for cases which have only little or no overlap in the mass–size plane (e.g. LBS galaxies and the elliptical $\mathcal{M}_\star$–R_e relation). The resulting KS statistics are shown in Table 3. For our tested assumptions, combining the results of Fig. 10 and the KS-test, we conclude:

(i) a,b,c,e,f,h = True,
(ii) d,g = False.

Table 3. Shown are the *D*- and *p*-values for a two-tailed KS-test on the hypothesis stated in Section 5.2.

Case	*D*	*p*-values
Components versus Sd-Irr and E relation		
(a) ETD versus Sd-Irr	0.4	0.418
(b) ETB versus E	0.38	0.66
(c) LTD versus Sd-Irr	0.29	0.635
(d) LTB versus E	0.5	0.1
(e) LTB versus Sd-Irr	0.42	0.256
(f) LBS versus E	0.43	0.575
(g) LBS versus Sd-Irr	1	0.001
Global fits versus Sd-Irr and E relation		
Sd-Irr versus Sd-Irr	0.33	0.73
Global late types versus Sd-Irr	0.23	0.898
Global early types versus Sd-Irr	0.57	0.019
E versus E	0.2	0.994
Global late types versus E	0.62	0.013
Global early types versus E	0.29	0.635

5.3 Should S(B)ab–S(B)cd systems be described as single or multicomponent?

In the previous section, we found that both late-type discs and late-type bulges are associated with the Sd-Irr relation. Additionally, the Sérsic index distribution of these components (Fig. 7, bulge, top and disc, bottom) show that late-type discs extend to lower Sérsic indices than typically expected. In fact, the LTD in our sample have a median Sérsic index of $n \sim 0.6$ (i.e. a more Gaussian-like light profile). The Sérsic index distribution of our late-type bulges peaks at $n \sim 2$ which is much lower than would be expected for an intermediate- to high-mass classical bulge where $n \sim 4$. There are two possible reasons, (i) we are seeing the effects of dust affecting both bulge and disc, and/or (ii) late-type systems are composed of pseudo-bulges and discs.

Seeing a bulge or disc through dust has the effect of lowering the Sérsic index as well as making them appear larger (Pastrav et al. 2013a,b). On the other hand, if the late-type galaxies do indeed contain pseudo-bulges, which are arguably perturbations of the disc, this begs the question as to whether the late-type galaxies should or should not be decomposed into two components.

Fig. 10 (left-hand panel) and a KS-test show that the global (i.e. single Sérsic fit) late-type $\mathcal{M}_\star$–R_e relation can also be associated with the Sd-Irr relation and a decomposition of the late-type two-component systems is not strictly necessary. Another issue to consider here is that if S(B)ab–S(B)cd galaxies are truly two-component systems than fitting their light profile with a single component only would bias our size estimation to larger sizes (Bernardi et al. 2014), especially for brighter and larger galaxies. However, the majority of our S(B)ab–S(B)cd systems are comparatively small when considering this effect found by Bernardi et al. (2014). Furthermore, fitting a final disc relation using either global or component fits for the S(B)ab–S(B)cd galaxies, we find that the resulting $\mathcal{M}_\star$–R_e relations are nearly identical (these relations are also given in Table 1 for reference). Hence, as considering either single- or two-component sizes has little effect on the $\mathcal{M}_\star$–R_e relation and as we cannot conclusively tell the difference between a pseudo-bulge and a classical bulge without kinematic data, we opt to use the global (single component) S(B)ab–S(B)cd galaxies for our final $z = 0$ disc relation to avoid overinterpreting our results. This is also in concordance with other recent studies where late-type two-component galaxies are considered 'bulgeless' discs (for example Sachdeva et al. 2015).

5.4 Combined disc and spheroid $\mathcal{M}_\star$–R_e relations

In this section, we now aim to establish the definitive $z = 0$ disc and spheroid $\mathcal{M}_\star$–R_e relation composed of associated global and component populations as identified in the previous section. In summary, we consider the following populations to be associated.

(i) Sd-Irr and late-type galaxies and the discs of early-type galaxies and,

(ii) Ellipticals, early-type bulges and LBS.

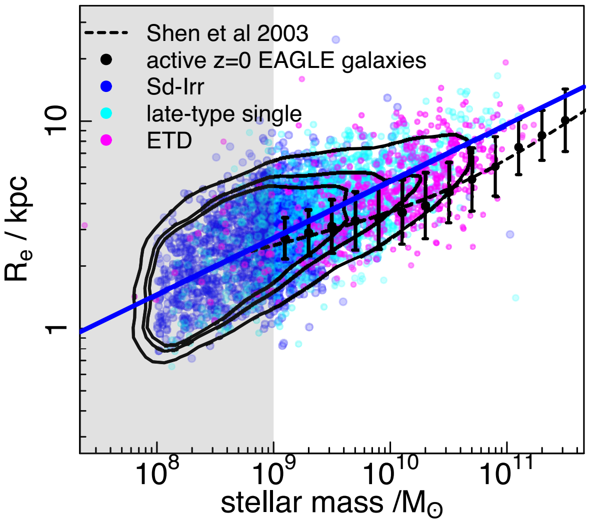
We show our final combined $\mathcal{M}_\star$–R_e relations for discs (top) and spheroids (bottom) in Fig. 11. The data points are colour coded by the population they belong to. The solid lines show our final disc and spheroid $\mathcal{M}_\star$–R_e relation fits. The dashed red line in the lower panel shows our high-mass elliptical relation, which would be more appropriate to use for comparisons with high-redshift data (see Section 7). In comparison, we also show simulated $z = 0$ EAGLE galaxies which are actively star forming (top) and passive (bottom) as black points with error bars. We discuss the selection and comparison of the EAGLE data in the next section. Additionally, we show the $\mathcal{M}_\star$–R_e relation for late- and early-type galaxies by Shen et al. (2003) as a dashed black line. We show these relations as they were used to calibrate the simulated data from EAGLE. Note that we plot the Shen et al. (2003) relations only over the mass range for which they were established and we have corrected them for the size–wavelength dependence using the equations given in L15 (their table 4), assuming $\mathcal{M}_\star = 10^{10}\,\mathcal{M}_\odot$. Additionally, we also correct for the fact that Shen et al. (2003) use a circularized radius.[1] The $\mathcal{M}_\star$–R_e relation parameters can be found in Table 1 (iv).

For the combined disc populations (Fig. 11, top),we find that the relation flattens considerably from the Sd-Irr-only relation, an effect of including the high-mass early-type discs and late-type galaxies. If our assumption is true, that the Sd-Irr galaxies, early-type discs, and 'reconstituted' late-type galaxies are related, then this could hint at a possible change in the slope of the late-type $\mathcal{M}_\star$–R_e at high masses. Comparing our $\mathcal{M}_\star$–R_e relation to the Shen et al. (2003) relation, it is obvious that our data follow an opposite trend at higher masses and does not turn up but down. This effect arises because we compare a component relation to a global relation, we see a similar steepening of our data if only global sizes are considered in the $\mathcal{M}_\star$–R_e distribution instead. In addition, since our data extend to lower masses than the Shen et al. (2003) analysis, we can also see that the $\mathcal{M}_\star$–R_e relation does not actually flatten out at lower masses. Overall, this shows that a linear fit (in log–log space) to the data is sufficient to describe the disc $\mathcal{M}_\star$–R_e relation over the observed stellar mass range.

Compared to the elliptical-only relation, the slope of the combined elliptical, LBS and ETB $\mathcal{M}_\star$–R_e relation also flattens and essentially lies parallel to the final disc relation. We also see a turn-off in the data, albeit in the opposite sense to the combined disc relation, i.e. the data flatten at the low-mass end. For the spheroid $\mathcal{M}_\star$–R_e relation, this flattening is caused by the inclusion of more low-mass components (mainly LBS), this is corroborated by a comparison of our data to the Shen et al. (2003) relation. We see a very good agreement over the mass range in which the Shen et al. (2003)

---

[1] $R_c = \sqrt{b/a}R_\mathrm{e}$, where b and a are the semiminor and semimajor axes. Using the average ellipticity of our sample, we get $R_\mathrm{e} = 1.33 \times R_c$.



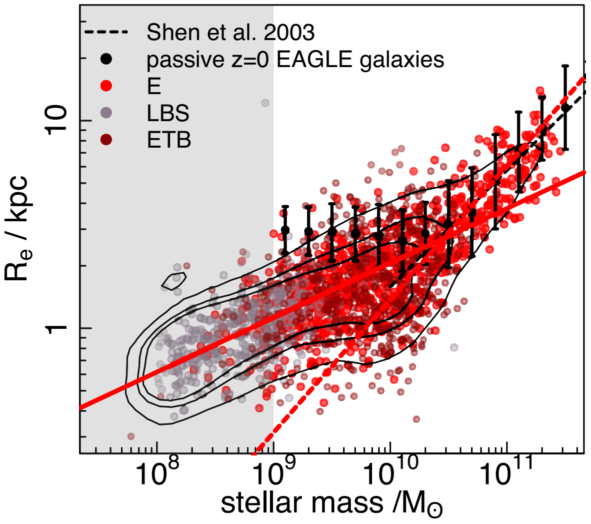


Figure 11. Shown are the final $z = 0$ disc (top) and spheroid (bottom) $\mathcal{M}_{\star}$–$R_{\rm e}$ relations. The disc relation consists of Sd-Irr galaxies, early-type disc components and reconstituted late-type two-component systems. The spheroid relation is composed of ellipticals, early-type bulges and LBS. The solid coloured lines are the final disc and spheroid $\mathcal{M}_{\star}$–$R_{\rm e}$ relation fits to the data, the dashed red line shows the high-mass elliptical-only relation, the grey shaded areas indicate where our smooth volume-limited sample selection starts and galaxies are upweighted by their $V/V_{\rm max}$. The black (solid) lines are the 90th, 68th and 50th percentiles of the respective mass–size distributions. Additionally, we show the active (top) and passive (bottom) $z = 0$ galaxies from the EAGLE simulation (black points + error bars) and the $\mathcal{M}_{\star}$–$R_{\rm e}$ relation from Shen et al. (2003) corrected for waveband and circularized radius (dashed black line, only shown over the mass range in which it was established).

$\mathcal{M}_{\star}$–$R_{\rm e}$ relation was established. However, a clear turn-off, or flattening can be seen at low masses. That the flattening is caused by low-mass galaxies is further supported by comparing the final spheroid relation to the low-mass elliptical-only relation, see Table 1 (ii), which are within the errors identical. Additionally, we fit the $\mathcal{M}_{\star}$–$R_{\rm e}$ relation to ellipticals and early-type bulges only and find that this relation is, within the errors, identical to the elliptical only relation, supporting the notion that early-type bulges are indeed classical bulges and 'elliptical-like'. This begs the question whether the LBS are all indeed early-type galaxies. As shown in Fig. 10, the distribution of the relative deviations of LBS from the elliptical $\mathcal{M}_{\star}$–$R_{\rm e}$ relation is consistent with them being associated. Their broad distribution, however, which looks similar to a top-hat function, extends to high deviations which suggests that not all LBS galaxies are the same. Without higher resolution imaging data, we cannot yet conclusively say if LBS are indeed all early-type galaxies. Preliminary visual inspection of LBSs available in the higher resolution imaging from the VISTA Kilo-degree Infrared Galaxy (VIKING) survey shows that a significant fraction of them are actually two-component systems.

If the LBS and ETBs, as well as the high- and low-mass ellipticals, are all indeed the same population, then a curved fit, i.e. a double power law, is necessary to describe the $\mathcal{M}_{\star}$–$R_{\rm e}$ relation. In lieu of a definitive answer, we recommend considering the high- and low-mass early-type populations separately. In Table 1 (ii), we provide a $\mathcal{M}_{\star}$–$R_{\rm e}$ relation for high- and low-mass ellipticals considering two different mass separators at $M_{\star} = 10^{10}\,{\rm M}_{\odot}$ and $2 \times 10^{10}\,{\rm M}_{\odot}$. The combined $\mathcal{M}_{\star}$–$R_{\rm e}$ relation to all ellipticals, ETBs and LBS in Table 1 (iv) should only be considered for samples that primarily contain galaxies with $M_{\star} \leq 10^{10}\,{\rm M}_{\odot}$. Furthermore, it should be noted that previous studies have reported a second deviation of the early-type $\mathcal{M}_{\star}$–$R_{\rm e}$ relation from a simple power law at very high masses ($\mathcal{M}_{\star} \sim 2 \times 10^{11}\,\mathcal{M}_{\odot}$; e.g. Bernardi et al. 2007; Hyde & Bernardi 2009; Bernardi et al. 2011, 2014). Due to the limited volume we survey, we do not see this deviation. However, as this second change in slope is likely linked to the formation history of galaxies, it needs to be taken into account when studying the high-mass end of the $\mathcal{M}_{\star}$–$R_{\rm e}$ relation.

Finally, we include the caveat that the flattening of both the disc and spheroid population could be real or due to misclassification in our visual morphology sample or a significant bias in our bulge + disc decomposition. Ultimately deeper data, such as that provided by VST KiDS and a fit in all available wavelengths to establish robust masses, should clarify whether the flattening is real or an artefact of our methodology.

6 DOES THE SIZE DISTRIBUTION POSE A PROBLEM FOR SIMULATIONS?

In this section, we have a first look at comparing our final disc and spheroid relations with data from the EAGLE simulation (Crain et al. 2015; Schaye et al. 2015). EAGLE is a suite of cosmological hydrodynamical simulations performed at two numerical resolutions, in periodic volumes with a range of sizes, and using a variety of subgrid implementations to model physical processes below the resolution limit. The subgrid parameters governing energetic feedback mechanisms of the EAGLE reference model were calibrated to the $z = 0$ galaxy stellar mass function, galaxy stellar mass–black hole mass relation, and galaxy stellar $\mathcal{M}_{\star}$–$R_{\rm e}$ relations (see Crain et al. 2015 for details and motivation). The EAGLE reference model reproduces many observed galaxy relations that were not part of the calibration set, such as the evolution of the galaxy stellar mass function (Furlong et al. 2015b), of galaxy sizes (Furlong et al. 2015a), of their optical colours (Trayford et al. 2015) and of their atomic (Bahé et al. 2016) and molecular gas content (Lagos et al. 2015), among others, and thus is an excellent testbed to compare with our observations.

We use the public data base of EAGLE described in McAlpine et al. (2016). In particular, we focus our attention on the reference model of EAGLE run in a cubic volume of length 100 comoving Mpc on a side with 2×1504^3 dark matter and gas particles (particle masses are 9.7×10^6 and $1.81 \times 10^6\,\mathcal{M}_{\odot}$, respectively), which help to reach a physical resolution of 0.7 kpc. One of the notable aspects of EAGLE is the plethora of subgrid baryonic physics included in the model: (i) radiative cooling and photo-heating rates, (ii) star formation, (iii) stellar evolution and metal enrichment, (iv) stellar feedback and (v) black hole growth and active galactic nuclei feedback. These physical models are the key ingredient to reproduce a large set of properties of the observed galaxy population in the local Universe. For more details of the simulation, we refer the reader to Schaye et al. (2015).

To compare EAGLE data with our observations, we use the masses and sizes of $z = 0$ active and passive galaxies, calculated from data in the public data base, following the method described in Furlong et al. (2015b). Briefly, the size is the mean half-mass radius taken within a 100 kpc aperture projected along the x-y, x-z and y-z axes. Note that this definition of size differs from the size used to calibrate the simulation against the Shen et al. (2003) late-type galaxies. The galaxy sizes used in the calibration are based on the Sérsic scale-length obtained from fitting Sérsic profiles to the surface density profiles of the simulated galaxies. To separate active and passive galaxies, Furlong et al. (2015b,a) use a specific star formation rate cut of 0.01 Gyr^{-1}, which is approximately one decade below the observed main sequence of star formation. For the following comparison, there are two notable caveats:

(i) we compare our measurements of the disc and spheroid $\mathcal{M}_\star$–R_e relation for the half-light radius to the half-mass radius of the simulated galaxies, and

(ii) we separate active and passive simulated galaxies to compare to our discs and bulges, respectively. We caution the reader that although a correlation between being bulge (disc)-dominated and being passive (active) is expected, they are not necessarily the same populations.

Fig. 11 shows our final disc and spheroidal $\mathcal{M}_\star$–R_e relations. We plot the EAGLE data as black points and the error bars indicate the 16th and 84th percentile of their distribution. We also show the Shen et al. (2003) relation for $n < 2.5$ and $n > 2.5$ galaxies as the dashed line. The difference in the shape of the disc $\mathcal{M}_\star$–R_e relation (Fig. 11, top) is, in part, caused by comparing disc only components (our data) with global sizes (EAGLE). This is because also considering a bulge component in a global profile fit has the effect of lowering the half-mass radius due to their typically smaller size and higher concentration compared to discs. For the spheroid relation (Fig. 11, bottom), we agree well with the EAGLE data down to $\mathcal{M}_* \sim 10^{10}\,\mathcal{M}_\odot$. Below this mass limit, we see a much less marked change of slope compared to the EAGLE data. This could be due to comparing passive galaxies in EAGLE with our spheroid sample that contains LBS galaxies. The latter are arguably active systems. On the other hand, at least in part, this could be explained by the known limitations of the simulation, e.g. the high fraction of passive low-mass galaxies at $z = 0$ due to the finite sampling of the star formation in low-mass galaxies. A one-to-one comparison, in which EAGLE galaxies are analysed with the same pipeline applied to GAMA galaxies, is needed to shed light on this issue. A further point to consider is the shape of the $\mathcal{M}_\star$–R_e relation at very high masses ($\mathcal{M}_* \geq 10^{11.3}\,\mathcal{M}_\odot$) which has been shown to deviate from a pure power law (e.g. Bernardi et al. 2007; Hyde & Bernardi 2009; Bernardi et al. 2011, 2014). As mentioned previously, due to our small survey volume, we do not sample the high-mass end of the $\mathcal{M}_\star$–R_e relation well and cannot confirm the curvature of the relation. However, to ensure that simulations return realistic galaxy sizes at all masses, not only the mass range over which they were matched, a detailed comparison with the relevant studies is necessary.

Fig. 12 summarizes the findings of Fig. 11. Here we compare the distribution of the vertical scatter for the GAMA and EAGLE data from our disc (left) and spheroid (right) $\mathcal{M}_\star$–R_e relation. We divide our sample in bins of 0.3 dex in stellar mass and fit a Gaussian to the distributions (see Figs B1 and B2) to study the scatter of the data from the $\mathcal{M}_\star$–R_e relation at fixed stellar mass. The top panel of Fig. 12 shows the offset (μ of the Gaussian fit) of the $\log(R_\mathrm{e})$ scatter from the $\mathcal{M}_\star$–R_e relation. Note that for the high-mass spheroid/passive galaxy sample, we actually compare to the high-mass ($\mathcal{M}_* > 10^{10}\,\mathcal{M}_\odot$) elliptical $\mathcal{M}_\star$–R_e relation. This does not change the scatter of the data but has the effect of moving the offset from the $\mathcal{M}_\star$–R_e relation to ∼0. The bottom panel shows the $\sigma_{\log(R_\mathrm{e})}$ versus stellar mass distribution. We show the sigma (standard deviation, SD) of the best-fitting Gaussian to the underlying distribution as well as the 16th and 84th percentile scatter of the GAMA and EAGLE data. In cases where the underlying distributions are somewhat skewed, the SD and 16/84th percentile sigma do not agree well and could even hint at a possible bimodality in the underlying distribution. We also include the scatter versus mass relation from Shen et al. (2003), corrected from $\log_\mathrm{e}$ to $\log_{10}$, for comparison. Note that the Shen et al. (2003) scatter–mass relation is based on the combined scatter of their early- and late-type $\mathcal{M}_\star$–R_e relations. In Appendix B, we show the histograms of the scatter of the data from the $\mathcal{M}_\star$–R_e relation for each mass bin with the best-fitting distribution overplotted.

Examining Fig. 12, the offset in the modes of the GAMA and EAGLE data is expected, as seen in Fig. 11. This is due to the EAGLE simulation being calibrated using the Shen et al. (2003) relations. To some extent, the variance (or sigma values as indicated on the figure) is of more interest as these have not been explicitly tailored in the simulation to match the data distributions.

We focus our analysis on the sigma derived from the Gaussian fits, i.e. the points labelled SD in Fig. 12. We find that for the disc/active population, the variance in the simulated data is almost always smaller than the observations. For the spheroid/passive populations, there is a divide at $\mathcal{M}_* \sim 2 \times 10^{10}\,\mathcal{M}_\odot$ with the simulated data having a smaller variance for less massive systems compared to the data and a comparable variance for more massive systems. This is somewhat surprising as the dark matter spin distributions are known to be quite broad. If coupling is strong, one would expect the distribution of specific angular momentum and disc sizes to be comparably broad. Because we are comparing light against mass and components against classes, we should be careful with interpreting any deviations. Clearly, an improved comparison can be made from bulge + disc decompositions of the EAGLE images which would place the observational and simulated data on to the same footing. Although *gri* images of EAGLE galaxies with $\mathcal{M}_* > 10^{10}\,\mathcal{M}_\odot$ have been made publicly available from their data base, it is still not sufficient to perform an analysis like the one done here for GAMA. Images of individual bands, preferably to lower stellar masses, would be required for this. For the moment, we consider Figs 11 and 12 to provide a good demonstration of the potential of the mass–size plane for comparing observational and simulated data.

7 COMPARISON WITH HIGH-REDSHIFT DATA

There is a well-known discrepancy between the $\mathcal{M}_\star$–R_e relation of high-redshift galaxies ($z > 1$) and the local $\mathcal{M}_\star$–R_e relation, with the high-redshift galaxies at the same mass being smaller than their low-redshift counterparts (e.g. Daddi et al. 2005; Longhetti et al. 2007; Trujillo et al. 2006, 2007; van Dokkum et al. 2008). This

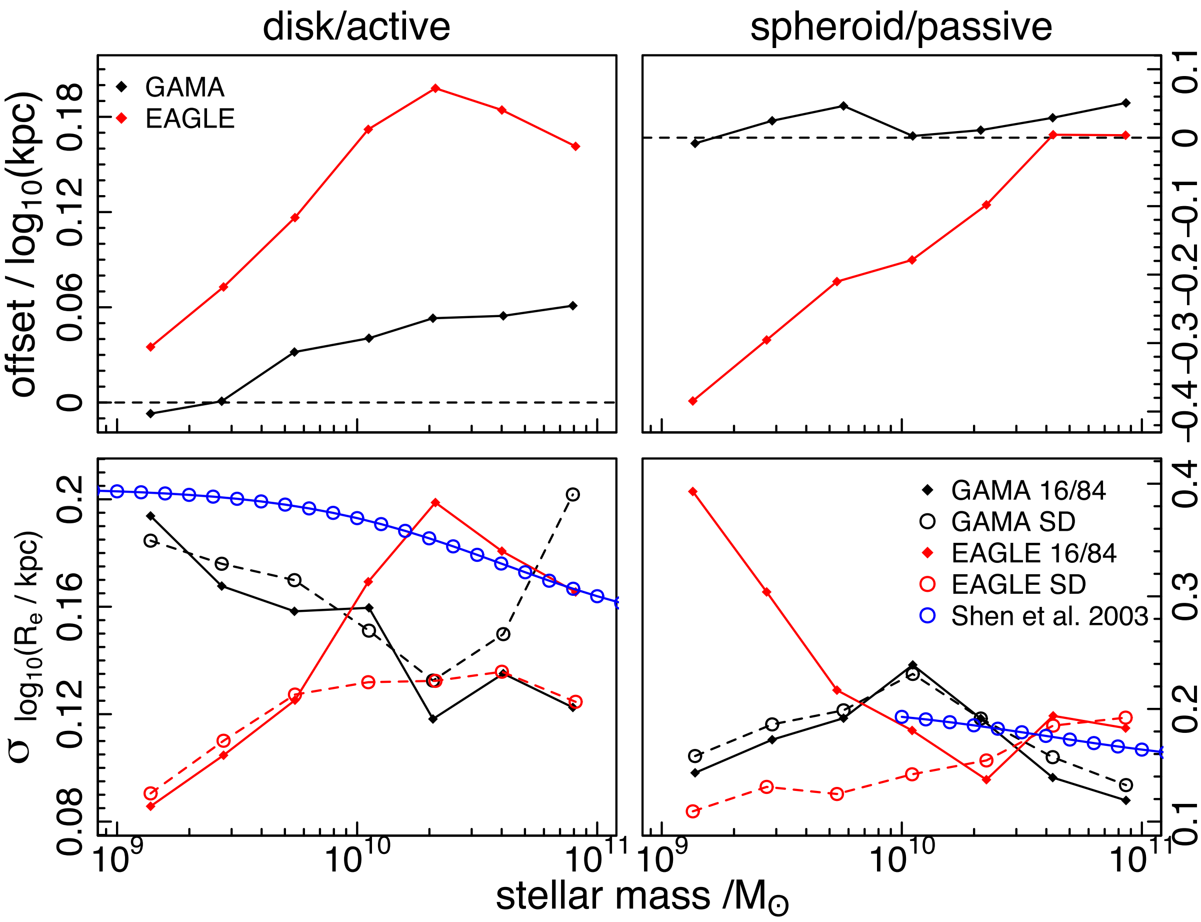


Figure 12. Shown is the offset (top) of the Gaussian distribution fit of the observational and simulated data from the disc (left) and spheroid (right) $\mathcal{M}_{\star}$–R_{e} relation. The bottom panel shows the corresponding scatter $\sigma_{\log(R_{\mathrm{e}})}$ (SD, circle plus dashed lines) as well as the 16th/84th percentile scatter (16/84, filled symbols plus solid line) of the distribution. Additionally, we show the σ–mass relation from Shen et al. 2003 for comparison.

itself is not a problem since an evolution in the $\mathcal{M}_{\star}$–R_{e} relation is expected due to galaxies at later times being formed through less dissipative events (i.e. low-redshift progenitors are gas poorer than high-redshift progenitors, see e.g. Hopkins et al. 2009). Hence, at lower redshift, newly formed galaxies are expected to be bigger than their high-redshift counterparts. However, the lack of old massive compact systems at low redshift means that the galaxies observed at high redshift must have grown by a factor of up to ~5 to 6 to end up on the local $\mathcal{M}_{\star}$–R_{e} relation. One might argue that the measurements of high-redshift galaxies are inherently difficult and the observed size growth is biased by systematics. However, several studies have shown that, even considering all the uncertainties in the mass and size measurements of the high-z galaxies, the size growth is real. Even for the most unfavourable cases, the high-redshift galaxies lie well below the present-day $\mathcal{M}_{\star}$–R_{e} relation (see e.g. Weinzirl et al. 2011; Buitrago et al. 2013).

In this section, we briefly compare published measurements of high-redshift galaxies from the CANDELS-UDS field (Mortlock et al. 2013, 2015; Margalef-Bentabol et al. 2016), against our local disc and spheroid $\mathcal{M}_{\star}$–R_{e} relations. To briefly summarize, the CANDLES-UDS data contain 1132 galaxies with $\mathcal{M}_{*} > 10^{10}\,\mathcal{M}_{\odot}$ and $1 < z < 3$. Of these, 683 are fit with a single-component Sérsic profile while 449 with a bulge + disc profile. For the two-component fits, the disc is set to $n = 1$ and the bulge n is free unless the fit failed in which case the bulge was reset to either $n = 1$ or 4 (Margalef-Bentabol et al. 2016).

The CANDELS-UDS data were obtained using the *HST*'s Wide Field Camera 3 (WFC3/IR) F160W (*H*) band, which is well matched in physical resolution and rest wavelength to the GAMA low-redshift SDSS data. Here we restrict ourselves to the redshift range $1 < z < 1.5$, where the *H* band equates to a comparable rest wavelength of 640–800 nm. For the *HST* data, stellar masses were derived by the CANDELS team using BC03 stellar populations, a Chabrier IMF and with the same ΛCDM cosmology as our analysis. One minor difference worth highlighting between the *HST* and GAMA analysis, is the derivation of component masses. For the *HST* data, this was based on the single band bulge-to-total ratio rather than component colours (see Section 4).

The high-redshift CANDELS-UDS sample naturally divides into three distinct populations: single-component systems, bulges of two-component systems and discs of two-component systems. We now explore whether any of these three populations follow a similar $\mathcal{M}_{\star}$–R_{e} relation to the local benchmarks.

In Fig. 13, we show the $\mathcal{M}_{\star}$–R_{e} plane with the various high-redshift samples overlaid on the local benchmark data (left-hand panel against discs and right-hand panel against spheroids). Note that we show both the local spheroid $\mathcal{M}_{\star}$–R_{e} (solid red line) and the high-mass elliptical $\mathcal{M}_{\star}$–R_{e} relation ($\mathcal{M}_{*} > 10^{10}\,\mathcal{M}_{\odot}$, dashed red line). For our comparison, we concentrate on the high-mass elliptical relation since this describes our low-redshift data better in the mass range observed for high-redshift galaxies. It is immediately apparent that none of the populations provide a good association with the

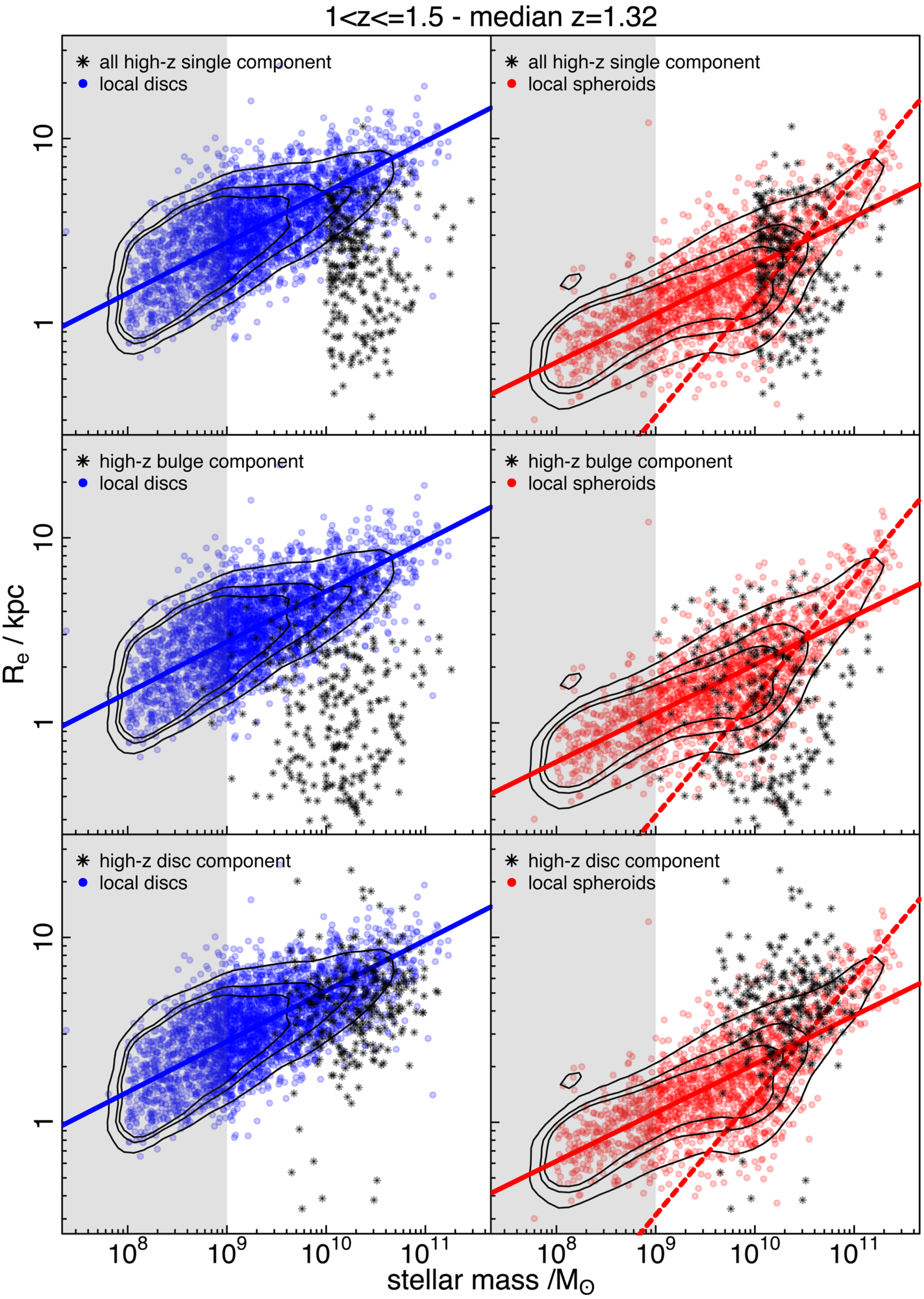


Figure 13. We show the local disc (blue, left side) and spheroid (red, right side) $\mathcal{M}_\star$–R_e distributions in comparison to the high-redshift CANDELS data ($1 < z < 1.5$) for our three high-redshift populations. The solid blue and red lines are our final disc and spheroid $\mathcal{M}_\star$–R_e relation as presented in Section 5.4. The dashed red line is the local high-mass ($\mathcal{M}_* > 10^{10}\,\mathcal{M}_\odot$) elliptical $\mathcal{M}_\star$–R_e relation as given in Table 1 (ii). It is immediately obvious that most data do not agree with the disc $\mathcal{M}_\star$–R_e relation. On the other hand, there is good agreement between our local spheroids and the high-redshift single-component systems and the high-redshift bulge components (with a compact extension of the high-redshift bulge components).

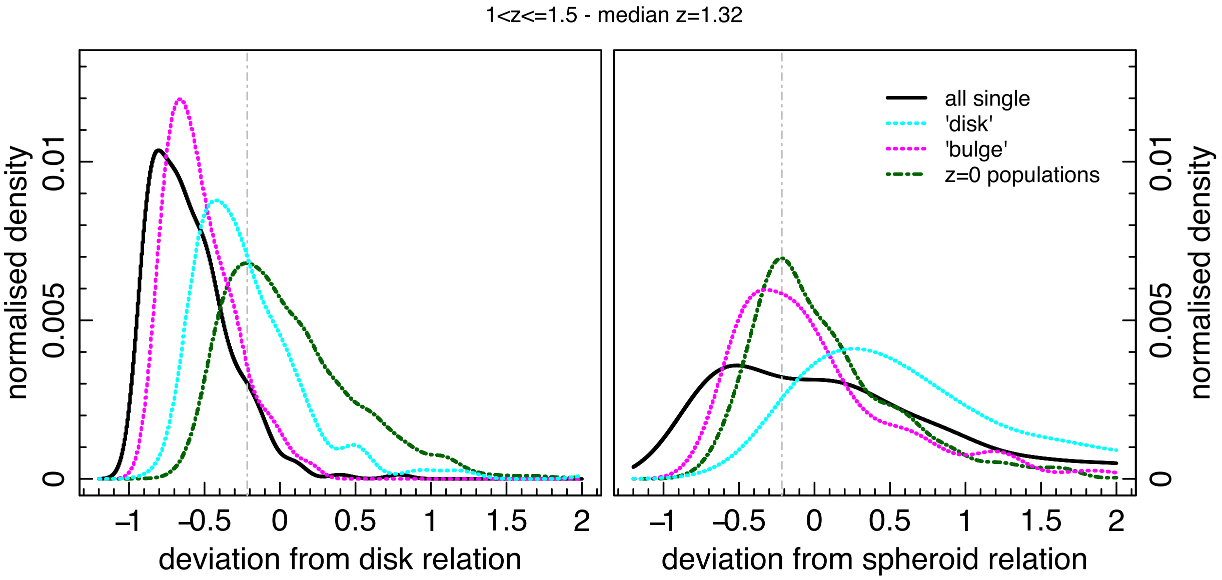


Figure 14. Shown is the same comparison as Fig. 10 but for the CANDELS high-redshift data ($1 < z \leq 1.5$) compared to our final disc (left) and high-mass elliptical (right, $\mathcal{M}_* > 10^{10}\,\mathcal{M}_\odot$) $\mathcal{M}_\star$–R_e relations.

low-redshift disc benchmark (left-hand panel), with all systems being significantly more compact. However, the high-redshift discs lie the closest and fall just below the low-redshift relation (bottom-left panel). On the other hand, we see what appears to be a fairly close association between the high-redshift single-component systems and the low-redshift spheroid benchmark (top-right panel). The high-redshift bulges also overlap with the low-redshift spheroid relation (middle-right panel), but extend to significantly lower sizes as well.

To elucidate these issues further, we show in Fig. 14 the deviation distribution. We follow our methodology from before (see Section 5.2) and use equation (3) to calculate the deviation from the local benchmark relations. Here, R_observed represents the sizes of the CANDELS-UDS data, and $R_\mathrm{predicted}$ represents the predicted size using either the low-redshift disc relation or low-redshift high-mass ($\mathcal{M}_* > 10^{10}\,\mathcal{M}_\odot$) elliptical relation, $\mathcal{M}_\star$–R_e relation. On Fig. 14, each of the three populations are shown compared against either the local disc (left-hand panel) or local spheroid (right-hand panel) relations. The dark green dot–dashed lines show the scatter of the local data about the disc and spheroid relations. Note that the low-redshift spheroid deviation (dark green dashed line, right-hand panel) is only evaluated down to $\mathcal{M}_* = 10^{10}\,\mathcal{M}_\odot$ to give a fair comparison, since a high-mass-only $\mathcal{M}_\star$–R_e relation is used for evaluation. The solid black line shows the high-redshift single-component sample, while the dotted lines show the high-redshift discs (cyan) and bulges (magenta). Fig. 14 reiterates our findings from Fig. 13. Compared to the local spheroid benchmark, two populations show a plausible fit, the single-component systems, and the high-redshift bulges, albeit with a greater spread potentially indicative of the greater measurement error associated with fitting high-redshift data. Whereas none of the distributions agree with the disc $\mathcal{M}_\star$–R_e relation. However, the high-redshift disc components have the largest overlap with our local disc $\mathcal{M}_\star$–R_e relation.

In addition to the qualitative nature of Figs 13 and 14, we also perform a KS-test, as described in Section 3, to establish the association of the high-redshift data with our local disc and high-mass spheroid $\mathcal{M}_\star$–R_e relation. The resulting test statistics are shown in Table 4, and corroborate our visual inspections.

Table 4. Shown are the D- and p-values for a two-tailed KS-test on the hypothesis that high-redshift $1 \leq z \leq 1.5$ galaxies and components are associated with either the local disc (i) or spheroid (ii) $\mathcal{M}_\star$–R_e relation.

Case	D	p-values
(i) high redshift versus local disc $\mathcal{M}_\star$–R_e relation		
a) all high-redshift single	1	0.002
b) disc	0.62	0.087
c) bulge	1	0
(ii) high redshift versus local high-mass spheroid $\mathcal{M}_\star$–R_e relation		
a) all high-redshift single	0.5	0.474
b) disc	0.75	0.019
c) bulge	0.4	0.418

Putting aside external observations (e.g. visual morphology used as priors), our analysis suggests that the majority of high-redshift systems overlap satisfactorily with the low-redshift spheroid $\mathcal{M}_\star$–R_e relation, with the exception of the high-redshift disc components. The obvious and simplest conclusion is that we are essentially seeing bulge and spheroid formation/emergence at high redshift, with some two-component systems existing, which adhere reasonable closely to the $z = 0$ spheroid relations but with the high-redshift discs somewhat more compact than their $z = 0$ counterparts. This observation meshes well with the notion of rapid spheroid formation at high redshift ($z > 1.5$; Tacchella et al. 2015) followed by disc growth at intermediate to lower-z ($z < 1.5$; e.g. Sachdeva et al. 2015), i.e. two-phase evolution as described in Driver et al. (2013). The obvious objections, however, are that observations of high-redshift systems generally show them to be visually clumpy, vigorously star-forming, and exhibiting clear evidence of systemic rotation, and hence are often described as disc-like (see for example Förster Schreiber et al. 2006, 2009; Green et al. 2010, and the recent review by Glazebrook 2013). However, these observations also show strong vertical velocity dispersions (e.g. Barro et al. 2014), verging on or exceeding the Toomre stability criterion (Toomre 1964), i.e. if these are discs, they are highly unstable, and unlike any type of disc, seen locally.

At some level, there is a semantic issue worth raising: when exactly does a spheroid become a spheroid, or a disc become a disc? Even in the ideal scenario of an entirely isolated collapsing gas cloud, it is likely to go through several starburst phases, fragmentation and merging of these fragments before finally resembling what we consider a classical elliptical. Exactly at what point should we start calling such a system an elliptical, at the moment of first collapse or only after all star formation has ceased and the system becomes dynamically relaxed? If the critical criteria are along the lines of the *Hubble* classification, then clearly the high-redshift systems are not spheroids, however, it is also clear these are not conventional discs (smoothly rotating systems with minimal velocity dispersions and aspect ratios of 1:10) and the use of this terminology is equally misleading.

The $\mathcal{M}_\star$–R_e relation essentially maps fundamental (conservable) quantities of mass and angular momentum. In this sense, the $\mathcal{M}_\star$–R_e relation is quite powerful and appears to be arguing that the majority of systems at high redshift are, if not spheroids, proto-spheroids in the process of settling into spheroids. Those systems that do appear to have two-components exhibit bulges consistent or slightly smaller than low-redshift bulges (which could be due to increased nuclear activity making the bulges appear more compact), and discs which are offset to lower sizes, however, disc growth is expected to continue to lower redshifts.

While the above paints a consistent and tantalizing picture, the caveats at this stage are significant. The analysis of the CANDLES-UDS data has been conducted by an independent group using distinct methods and strategies which could introduce systematic offsets. Distances are also based on photometric redshifts for the vast majority of the high-redshift sample. The sample size is also relatively small (subject to cosmic variance effects), and spans a particularly narrow mass range. Nevertheless, as we push the depth and area boundaries with facilities such as *Euclid* and *WFIRST*, the $\mathcal{M}_\star$–R_e scaling relation shows great promise for providing not only a connection to the hydrodynamical simulations but also as a bridge between the low- and high-redshift Universe.

8 SUMMARY AND CONCLUSION

We have presented our bulge-disc decomposition catalogue for 7506 galaxies from the GAMA survey in the redshift range of $0.002 < z < 0.06$ (Section 3). To overcome the limitations of the LM minimization algorithm used in GALFIT, which can get trapped in local minima (especially for two-component fits), we repeatedly fit our galaxy sample with varying starting points to map out the parameter space. For the single-component galaxies, we use a set of 33 combinations of the starting parameters, and for the two-component fits, we use a set of 88 combinations. We implement a screening process to prune bad fits and determine the final fitting values and errors from the median of the acceptable fits. We use the 16th and 84th percentile of the remaining output parameter distribution to determine the error on the median model combined with a 10 per cent error floor. Through this strategy, we reduce our catastrophic failure rate from ∼20 per cent to ∼5 per cent.

We then presented the $\mathcal{M}_\star$–R_e relations of our sample by *Hubble* type and component with the component masses based on an estimation from the bulge and disc colours. Next, we explored the association of the bulge and disc components with either the Sd-Irr or elliptical $\mathcal{M}_\star$–R_e relation. We find that S(B)ab–S(B)cd galaxies likely consist of a disc plus a pseudo-bulge. Considering that a pseudo-bulge is a perturbation of the disc, we decide that our late-type two-component systems are best represented by a single-component Sérsic fit. Thus, we associate elliptical, early-type bulges and LBS for the spheroid $\mathcal{M}_\star$–R_e relation and Sd-Irr, single-component fit S(B)ab–S(B)cd galaxies and early-type discs for the final disc $\mathcal{M}_\star$–R_e relation, which we provide as a definitive low-redshift benchmark:

$$R_\mathrm{e} = 5.141\left(\frac{\mathcal{M}_*}{10^{10}\,\mathcal{M}_\odot}\right)^{0.274} \quad \text{for discs and,}$$

$$R_\mathrm{e} = 2.063\left(\frac{\mathcal{M}_*}{10^{10}\,\mathcal{M}_\odot}\right)^{0.263} \quad \text{for spheroids.}$$

However, we caution the reader that the spheroid relation is heavily dominated by low-mass galaxies. If a comparison to high-mass spheroids is needed, then the high-mass elliptical $\mathcal{M}_\star$–R_e relation (see Table 1) should be used in lieu of a curved spheroid relation.

Next, we used our local disc and spheroid $\mathcal{M}_\star$–R_e distributions to compare to data from the EAGLE simulation. We find a qualitatively good agreement between the sizes of the EAGLE data and our $\mathcal{M}_\star$–R_e relations. This is not surprising as the sizes in EAGLE were calibrated using the (Shen et al. 2003) $\mathcal{M}_\star$–R_e distribution. Hence, the variance is of more interest as this has not been explicitly matched between the observed and simulated data. Comparing the scatter of the observed and simulated data, we find that in almost all cases, the simulated data has a smaller scatter which is unexpected, considering that the dark matter spin distribution is known to be fairly broad and we would expect the sizes and angular momentum distributions to have similarly broad distributions. Since we are comparing half-light sizes to half-mass sizes and components versus active and passive galaxies, we caution the reader to not overinterpret this comparisons. Instead, we would like to highlight the potential of using the mass–size plane to compare observational and simulated data.

Finally, we compare our local $\mathcal{M}_\star$–R_e relations to high-redshift data from the CANDELS-UDS field. We concentrate on available data in a redshift range of $1 < z \leq 1.5$ with available single and two-component fits (Mortlock et al. 2013; Margalef-Bentabol et al. 2016). We generally find that low-mass high-redshift galaxies agree better with the local $\mathcal{M}_\star$–R_e distributions than high-mass high-redshift galaxies. Furthermore, high-redshift systems, with the exception of disc components, more closely follow that of our local spheroid relation. The high-redshift discs on the other hand follow the local disc $\mathcal{M}_\star$–R_e relation, albeit offset to slightly smaller size. We interpret this as evidence for spheroid formation at high redshift and propose that further disc formation and/or growth does not occur until later times.

ACKNOWLEDGEMENTS

RL would like to acknowledge funding from the International Centre for Radio Astronomy Research and the University of Western Australia. SB acknowledges the funding support from the Australian Research Council through a Future Fellowship (FT140101166). CL is funded by a Discovery Early Career Researcher Award (DE150100618).

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. The VISTA VIKING data used in this paper are based on observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 179.A-2004. 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/.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Allen P. D., Driver S. P., Graham A. W., Cameron E., Liske J., de Propris R., 2006, MNRAS, 371, 2
Bahé Y. M. et al., 2016, MNRAS, 456, 1115
Baldry I. K. et al., 2010, MNRAS, 404, 86
Baldry I. K. et al., 2014, MNRAS, 441, 2440
Barro G. et al., 2014, ApJ, 795, 145
Berg T. A. M., Simard L., Mendel Trevor J., Ellison S. L., 2014, MNRAS, 440, L66
Bernardi M., Hyde J. B., Sheth R. K., Miller C. J., Nichol R. C., 2007, AJ, 133, 1741
Bernardi M., Roche N., Shankar F., Sheth R. K., 2011, MNRAS, 412, L6
Bernardi M., Meert A., Vikram V., Huertas-Company M., Mei S., Shankar F., Sheth R. K., 2014, MNRAS, 443, 874
Bertin E., 2013, Astrophysics Source Code Library, record ascl:1301.001
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bouwens R. J., Illingworth G. D., Blakeslee J. P., Broadhurst T. J., Franx M., 2004, ApJ, 611, L1
Bruce V. A. et al., 2012, MNRAS, 427, 1666
Bruce V. A. et al., 2014, MNRAS, 444, 1660
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Buitrago F., Trujillo I., Conselice C. J., Bouwens R. J., Dickinson M., Yan H., 2008, ApJ, 687, L61
Buitrago F., Trujillo I., Conselice C. J., Häußler B., 2013, MNRAS, 428, 1460
Buta R. J. et al., 2010, ApJS, 190, 147
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cameron E., Driver S., Graham A., Liske J., 2009, ApJ, 699, 105
Chabrier G., 2003, ApJ, 586, L133
Cheung E. et al., 2013, ApJ, 779, 162
Colless M. et al., 2003, preprint (astro-ph/0306581)
Cook M., Lapi A., Granato G. L., 2009, MNRAS, 397, 534
Crain R. A. et al., 2015, MNRAS, 450, 1937
Daddi E. et al., 2005, ApJ, 626, 680
Dalcanton J. J., Spergel D. N., Summers F. J., 1997, ApJ, 482, 659
Davies L. J. M. et al., 2015, MNRAS, 447, 1014
Davis M. et al., 2003, in Puragra G., ed., Proc. SPIE Conf. Ser. Vol. 4834, Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II. SPIE, Bellingham, p. 161
de Souza R. E., Gadotti D. A., dos Anjos S., 2004, ApJS, 153, 411
Driver S. P. et al., 2006, MNRAS, 368, 414
Driver S. P. et al., 2011, MNRAS, 413, 971
Driver S. P., Robotham A. S. G., Bland-Hawthorn J., Brown M., Hopkins A., Liske J., Phillipps S., Wilkins S., 2013, MNRAS, 430, 2622
Driver S. P. et al., 2016, MNRAS, 455, 3911
Erwin P., 2015, ApJ, 799, 226
Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189
Förster Schreiber N. M. et al., 2006, ApJ, 645, 1062
Förster Schreiber N. M. et al., 2009, ApJ, 706, 1364
Furlong M. et al., 2015a, preprint (arXiv:1510.05645)
Furlong M. et al., 2015b, MNRAS, 450, 4486
Gadotti D. A., 2008, MNRAS, 384, 420
Gadotti D. A., 2009, MNRAS, 393, 1531
Glazebrook K., 2013, Publ. Astron. Soc. Aust., 30, e056
Graham A. W., 2013, in Oswalt T. D., Keel W. C., eds, A Review of Elliptical and Disc Galaxy Structure, and Modern Scaling Laws. Springer Publishing, Springer Science+Business Media Dordrecht, p. 91
Graham A. W., Guzman R., 2003, AJ, 125, 2936
Graham A. W., Dullo B. T., Savorgnan G. A. D., 2015, ApJ, 804, 32
Green A. W. et al., 2010, Nature, 467, 684
Grogin N. A. et al., 2011, ApJS, 197, 35
Häussler B. et al., 2007, ApJS, 172, 615
Hill D. T. et al., 2011, MNRAS, 412, 765
Holwerda B. W., Bouwens R., Oesch P., Smit R., Illingworth G., Labbe I., 2015, ApJ, 808, 6
Hopkins P. F., Bundy K., Murray N., Quataert E., Lauer T. R., Ma C.-P., 2009, MNRAS, 398, 898
Hopkins P. F., Bundy K., Hernquist L., Wuyts S., Cox T. J., 2010, MNRAS, 401, 1099
Hopkins A. M. et al., 2013, MNRAS, 430, 2047
Hubble E. P., 1926, ApJ, 64, 321
Hyde J. B., Bernardi M., 2009, MNRAS, 394, 1978
Kannappan S. J., Guie J. M., Baker A. J., 2009, AJ, 138, 579
Kelvin L. S. et al., 2012, MNRAS, 421, 1007
Kennedy R. et al., 2016, MNRAS, 460, 3458
Koekemoer A. M. et al., 2011, ApJS, 197, 36
Kormendy J., Bender R., 2012, ApJS, 198, 2
L'Huillier B., Combes F., Semelin B., 2012, A&A, 544, A68
Lackner C. N., Gunn J. E., 2012, MNRAS, 421, 2277
Lagos C. d. P. et al., 2015, MNRAS, 452, 3815
Lang P. et al., 2014, ApJ, 788, 11
Lange R. et al., 2015, MNRAS, 447, 2603
Lilly S. J. et al., 2007, ApJS, 172, 70
Liske J. et al., 2015, MNRAS, 452, 2087
Longhetti M. et al., 2007, MNRAS, 374, 614
Margalef-Bentabol B., Conselice C. J., Mortlock A., Hartley W., Duncan K., Ferguson H. C., Dekel A., Primack J. R., 2016, MNRAS, 461, 2728
McAlpine S. et al., 2016, Astron. Comput., 15, 72
Meert A., Vikram V., Bernardi M., 2015, MNRAS, 446, 3943
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Moffett A. J. et al., 2016, MNRAS, 457, 1308
Mortlock A. et al., 2013, MNRAS, 433, 1185
Mortlock A. et al., 2015, MNRAS, 447, 2
Pastrav B. A., Popescu C. C., Tuffs R. J., Sansom A. E., 2013a, A&A, 553, 23
Pastrav B. A., Popescu C. C., Tuffs R. J., Sansom A. E., 2013b, A&A, 557, A137
Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, AJ, 139, 2097
Roberts M., Haynes M., 1994, ARA&A, 32, 115
Robotham A. S. G., Obreschkow D., 2015, Publ. Astron. Soc. Aust., 32, 33
Romanowsky A. J., Fall S. M., 2012, ApJS, 203, 17
Sachdeva S., Gadotti D. A., Saha K., Singh H. P., 2015, MNRAS, 451, 2
Salo H. et al., 2015, ApJS, 219, 4
Schaye J. et al., 2015, MNRAS, 446, 521
Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978
Shibuya T., Ouchi M., Harikane Y., 2015, ApJS, 219, 15
Simard L., 1998, in Albrecht R., Hook R. N., Bushouse H. A., eds, ASP Conf. Ser., Vol. 145, Astronomical Data Analysis Software and Systems VII. Astron. Soc. Pac., San Francisco, p. 108
Simard L., Mendel J. T., Patton D. R., Ellison S. L., McConnachie A. W., 2011, ApJS, 196
Steinmetz M., Navarro J. F., 2002, New Astron., 7, 155
Tacchella S. et al., 2015, Science, 348, 314
Tasca L. A. M. et al., 2014, A&A, 564, L12
Taylor E. N. et al., 2011, MNRAS, 418, 1587
Toomre A., 1964, ApJ, 139, 1217
Toomre A., 1977, in Tinsley B. M., Larson D., Campbell R. B. G., eds, Evolution of Galaxies and Stellar Populations. Yale University Observatory, New Haven, p. 401
Trayford J. W. et al., 2015, MNRAS, 452, 2879
Trujillo I. et al., 2006, ApJ, 650, 18

Trujillo I., Conselice C. J., Bundy K., Cooper M. C., Eisenhardt P., Ellis R. S., 2007, MNRAS, 382, 109
Trujillo I., Ferreras I., Rosa I. G. D., 2011, MNRAS, 415, 3903
van den Bergh S., 1976, ApJ, 206, 883
van der Wel A. et al., 2014, ApJ, 788, 28
van Dokkum P. G. et al., 2008, ApJ, 677, L5
van Dokkum P. G. et al., 2010, ApJ, 709, 1018
Wei L. H., Kannappan S. J., Vogel S. N., Baker A. J., 2010, ApJ, 708, 841
Weinzirl T. et al., 2011, ApJ, 743, 87
Wisnioski E., Glazebrook K., Blake C., Poole G. B., Green A. W., Wyder T., Martin C., 2012, MNRAS, 422, 3339
Wright A. H. et al., 2016, MNRAS, 460, 765
York D. G. et al., 2000, AJ, 120, 1579

APPENDIX A: FLAGGING

As described in Section 2.1, our sample of 7506 $0.002 < z < 0.06$ galaxies has been classified on to the *Hubble*-type system as described in Moffett et al. (2016). The sample contains 860 E, 826 S(B)0-S(B)a, 1421 S(B)abc, 3531 Sd-Irr systems and 868 LBS. The E, LBS and Sd-Irrs we consider single-component systems, best fit by a single Sérsic profile, while we assume the remainder to be best described by a two-component, double Sérsic profile.

Following the decision on one or two components, we flag the resulting fits for various criteria:

(i) very high or low B/T (B/T > 0.8 and < 0.1, respectively),
(ii) disc $n >$ bulge n,
(iii) B/T reverses between minimum reduced χ^2 and median model,
(iv) bulge and disc position angle (PA) offset,
(v) the minimum reduced χ^2 solution is an outlier to the median value.

Flags (i)–(iv) are only evaluated on the sample of two-component galaxies. Flag (v) is evaluated for all galaxies (i.e. single and two-component).

For our two-component sample, we have a total of 962 galaxies with at least one flag (∼44 per cent). This drops drastically when checking for galaxies with several flags, and we find only 206 galaxies with more than one flag raised. Our single-component sample has a total of 164 flagged galaxies (∼3 per cent).

To check whether any of the flags are more likely to produce unsatisfactory fits, we visually inspect a random sample of 50 flagged two-component galaxies. We found that in many cases, the median fit is acceptable and only flags (i) and (ii) are more likely to yield potentially bad fits.

(i) *Galaxies with high (low) B/T*
We selected galaxies with a median B/T > 0.6 and visually inspected their minimum reduced χ^2 and median fit results as well as the convergence plots. For each galaxy, we decided whether it is better fit with a double component or single component or if it has a bad or uncertain fit (i.e. the fit has bad apertures and no solution can be found, or it is unclear whether a two-component fit is appropriate).

Fig. A1 shows the distribution of high B/T galaxies. We set all 46 galaxies with a B/T > 0.8 to a single-component fit. Additionally, since our mean B/T error is 0.1, we also consider all galaxies with a B/T < 0.1 to be a single-component galaxy. This adds another 42 galaxies for which the two-component fit is considered not appropriate.

(ii) *Disc Sérsic index > bulge Sérsic index*
As mentioned in Section 3.1.2 during the set up of our final B/T decomposition catalogue, we screen galaxies for bulge and disc component switching. To do this, we assume that the bulge R_e is smaller than the disc R_e and swap the assigned component for those galaxies where this is not the case. However, for galaxies where the bulge R_e is up to 10 per cent larger than the disc R_e, we also check whether the disc n is larger than the bulge n and only swap the assigned component if this is the case.

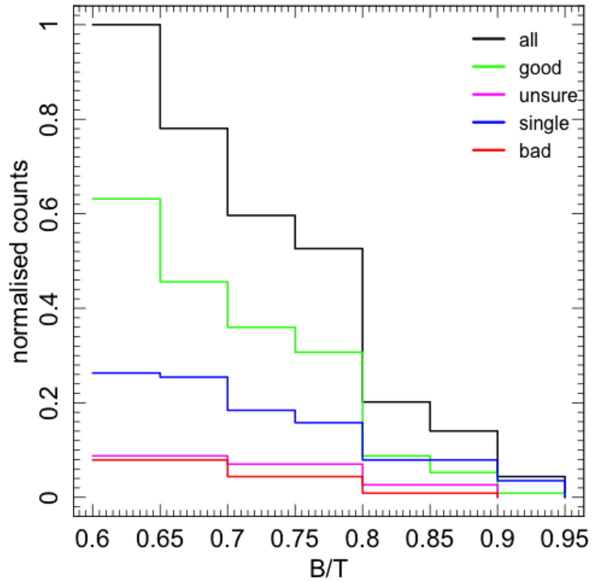


Figure A1. The distribution of galaxies with a median B/T > 0.6 and the visual decision whether they are two-component, single-component systems, or if the fit failed.

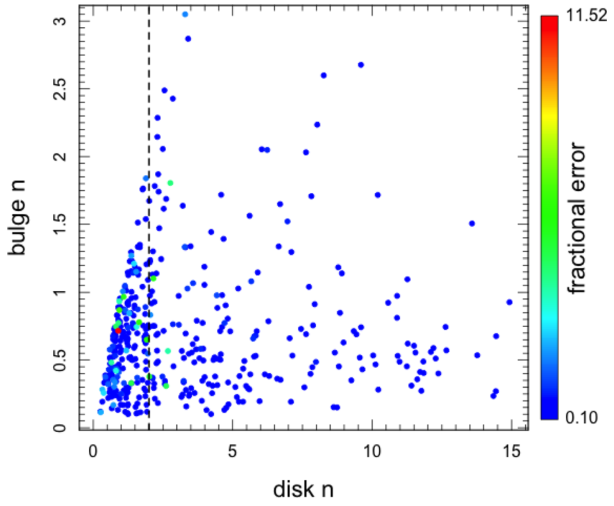


Figure A2. The disc versus bulge Sérsic index distribution for galaxies with bulge $n <$ disc n. The points are colour coded by the fractional error on the disc Sérsic index. It is obvious that many of the fits for these galaxies converged (i.e. 10 per cent error). We remove all converged and disc $n > 2$ galaxies from our final sample since we consider their fits unphysical.

However, even though the swapping ensures that our bulge is smaller than the disc, it is not guaranteed that the bulge n is larger than the disc n. We check our two-component fits and find 443 galaxies where the bulge n is smaller than the disc n. Fig. A2 shows the bulge versus disc n distribution coloured by the percentage error for the 443 flagged galaxies. Visually inspecting a number of the resulting fits, we find that galaxies with a disc $n > 2$ and/or

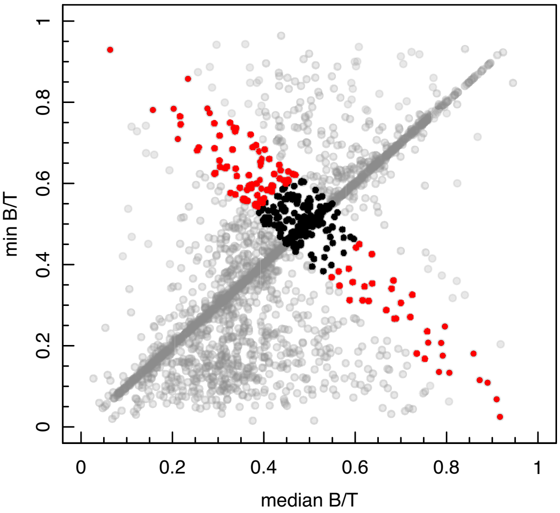


Figure A3. The median versus minimum reduced χ^2 distribution of the B/T in grey. The black points show the galaxies for which the sum of the min and median models lies between 0.9 and 1.1. The red points show the galaxies for which additionally the absolute difference between the min and median models B/T is larger than 0.15.

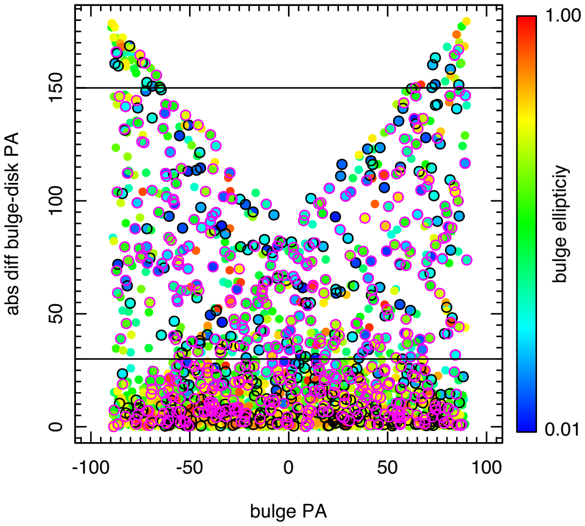


Figure A4. The absolute difference between bulge and disc position angle versus the bulge position angle. The points are coloured by the bulge ellipticity (e = 1-b/a). The two vertical lines at 30° and 150°show the range between which we flag the galaxies. The black circles show galaxies with bulge $e < 0.3$ and the pink circles show galaxies with a disc $e < 0.3$, which are excluded from the flagging.

converged fits are typically bad and we remove them from consideration of the component samples. In total, we remove 100 S(B)0-S(B)0a galaxies and 215 S(B)ab-S(B)cd galaxies.

(iii) *Bulge and disc fraction reverse between median and minimum reduced χ^2 model*

We have 116 galaxies for which the median- and minimum-reduced χ^2 models have opposite B/T values. To establish this sample, we select galaxies where the sum of the median and min B/T is between 0.9 and 1.1. The spread of this sum should include our inherent uncertainty in establishing the B/T, which is around 0.1–0.15. However, since this criterion alone also flags galaxies with a B/T = 0.5, we add a second criterion that the absolute difference between the median and min B/T has to be larger than 0.15 (i.e. larger than our average uncertainty in establishing a B/T). Fig. A3 shows the B/T distribution for the median versus minimum-reduced χ^2 model, and highlights the process of flagging the galaxies with a reversed B/T between the median and minimum-reduced χ^2 models.

(iv) *Bulge and disc PA more than 30°offset*

This flag is defined by identifying all galaxies for which the bulge and disc PA differs by more than 30°. However, for round bulges or discs, the PA is not very meaningful and even a large offset between bulge and disc PA is not indicative of a problem with the fitting results. To exclude these galaxies, we set a second condition for this flag, namely that the bulge or disc ellipticity has to be large enough to have a clearly identifiable preferred major axis, i.e. e = 1-b/a > 0.3. This reduces the flagged galaxies to 100 objects.

Fig. A4 shows the distribution of the absolute PA offset versus bulge PA, with the points coloured by their ellipticity. The horizontal lines show the angle offset between which we flag, and the objects circled show galaxies with $e < 0.3$ which are excluded from the flagging.

(v) *Minimum-reduced χ^2 solution represents an outlier to the median model*

Finally, we flag galaxies for which the minimum-reduced χ^2 solution represents and outlier, i.e. where the error on the median values does not include the minimum-reduced χ^2 solution. For minimum-reduced χ^2 solutions where the parameters are smaller than the median values, we consider the error on the lower end, and vice versa; where the median is smaller, we consider the error on the higher end.

We tested size, magnitude and Sérsic index distributions for outliers. We only consider the minimum-reduced χ^2 values to be a true outlier, and possibly better fit, if both the disc and bulge values fall outside the error range. Additionally, we check that the error on the median is larger than our assumed error floor, otherwise we consider the various fits of the galaxy to have converged and thus the minimum-reduced χ^2 solution likely represents a failed fit.

Fig. A5 shows the distribution of the absolute difference between the median and minimum-reduced χ^2 values versus the error for the median values. The left-hand panel shows the bulges and the right-hand panel the disc distribution. From top to bottom, we investigate size, magnitude and Sérsic index.

The red lines show the 1:1 correspondence, i.e. where the minimum-reduced χ^2 value lies on the edge of the median error distribution. The horizontal dashed lines shows our error floor, and we assume that most models converged to give rise to such a low error.

All galaxies to the right of the red lines and above the dashed black lines are potentially bad. We do, however, only flag the ones that are bad for both the disc and bulge values. This gives rise to 54 galaxies with flagged sizes, 393 galaxies with flagged magnitudes and 88 galaxies with flagged Sérsic indices. In total, this results in 449 galaxies for which the minimum-reduced χ^2 solution is an outlier to at least one of the three fitting parameters tested.

The equivalent test on our single-component fits finds 73 (42) spheroid (disc) sizes, 27 (30) spheroid (disc) magnitudes, and 62 (48) spheroid (disc) Sérsic indices flagged. In total, this equates to 90 (74) spheroid (disc) galaxies with at least one of the parameters flagged (≤5).

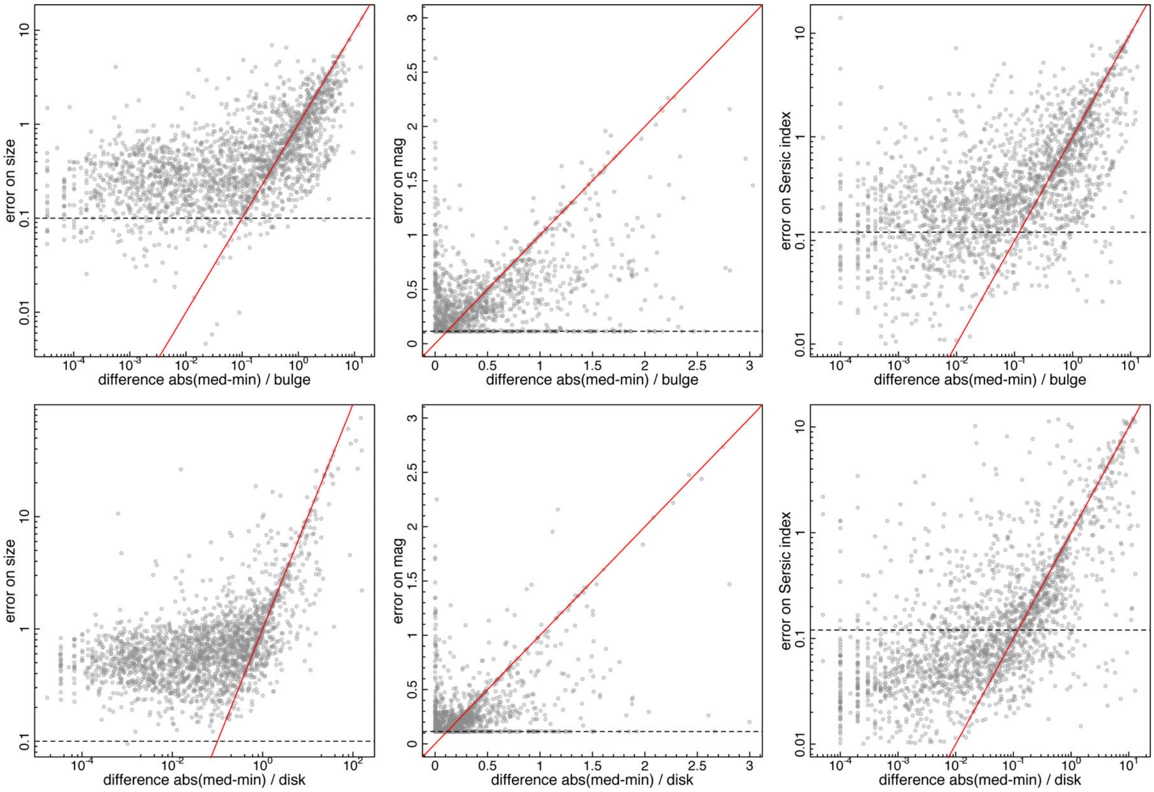


Figure A5. The top panels show bulge values, on the bottom panels the disc values. From left to right we show: difference in median and minimum-reduced χ^2 model sizes versus the median size error difference in median and minimum-reduced χ^2 model magnitudes versus the median magnitude error difference in median and minimum-reduced χ^2 model Sérsic indices versus the median Sérsic index error. The red lines show the 1:1 correspondence, the black dashed lines indicate our lower limits on the errors. Everything to the right of the red lines and above the dashed lines is flagged.

APPENDIX B: SIMULATION COMPARISONS

In Figs B1 and B2, we show the distribution of the vertical scatter from the $\mathcal{M}_\star$–R_e relation for discs and spheroids. The black histogram shows the distribution of our data and the blue line is a fit of a normal distribution to it. The red histograms show the distribution of the $z = 0$ EAGLE data with the red line being the normal distribution fit to it.

Fig. 12 is essentially derived from these plots where the SD points refer to the sigma of the Gaussian fit and the 16th and 84th percentiles are the width of the underlying distribution. Typically, these measurements of sigma should be the same, however, in some cases, the distributions have tails towards higher deviations which causes the 16th/84th percentile and SD measurements to be different (as seen in Fig. 12).

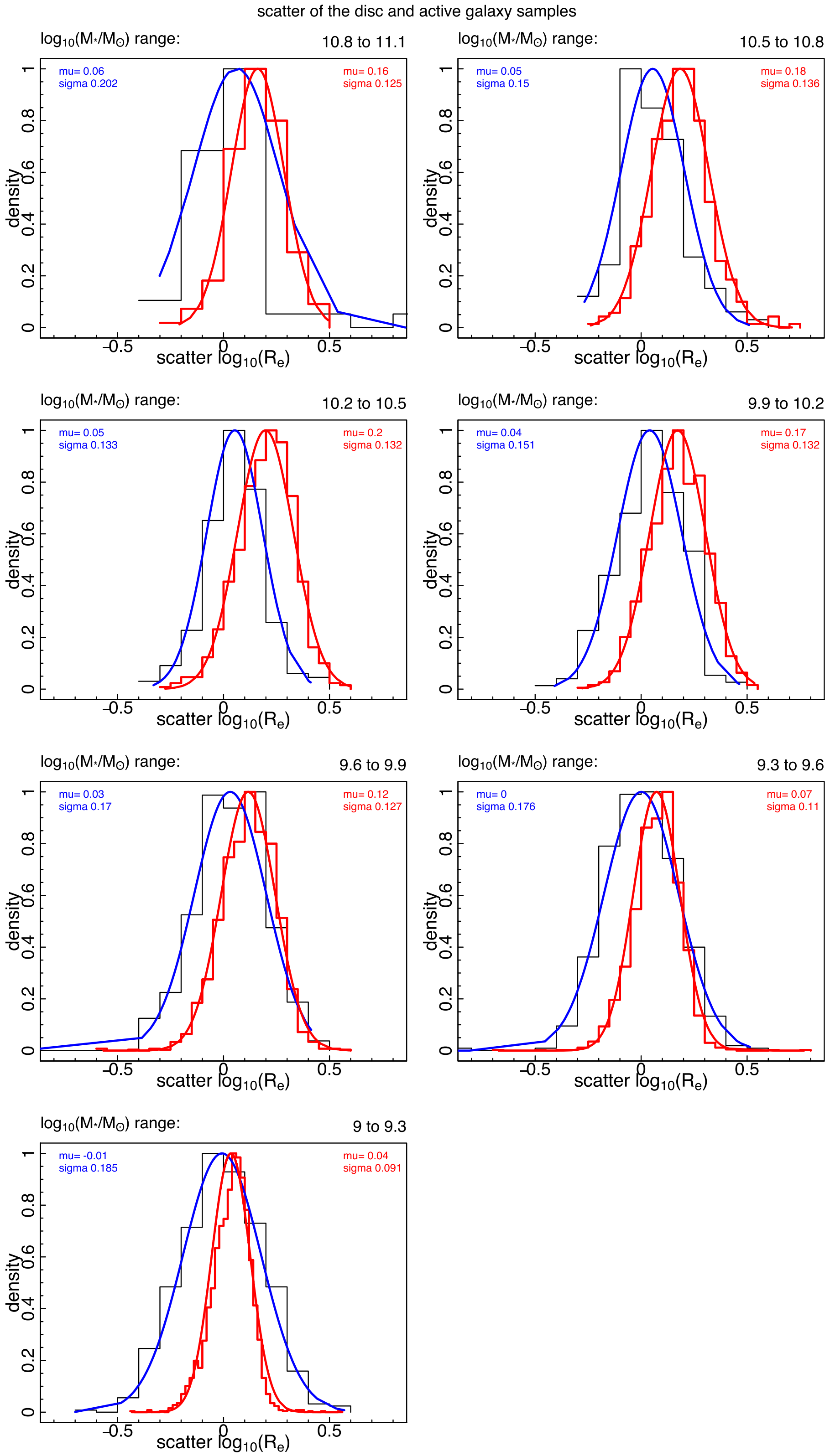


Figure B1. (a) Distribution of the half-light radius scatter of discs for bins in stellar mass (shown at the top of each panel) for our data. The blue line shows the normal distribution fit to the observed data (black histogram) and the red line is the normal distribution fit for the simulated data (red histogram), with the caveat that we use the half-mass radius for star-forming galaxies in the case of EAGLE. The legend on the left shows the best-fitting values for the observed data and on the right for the simulated data.

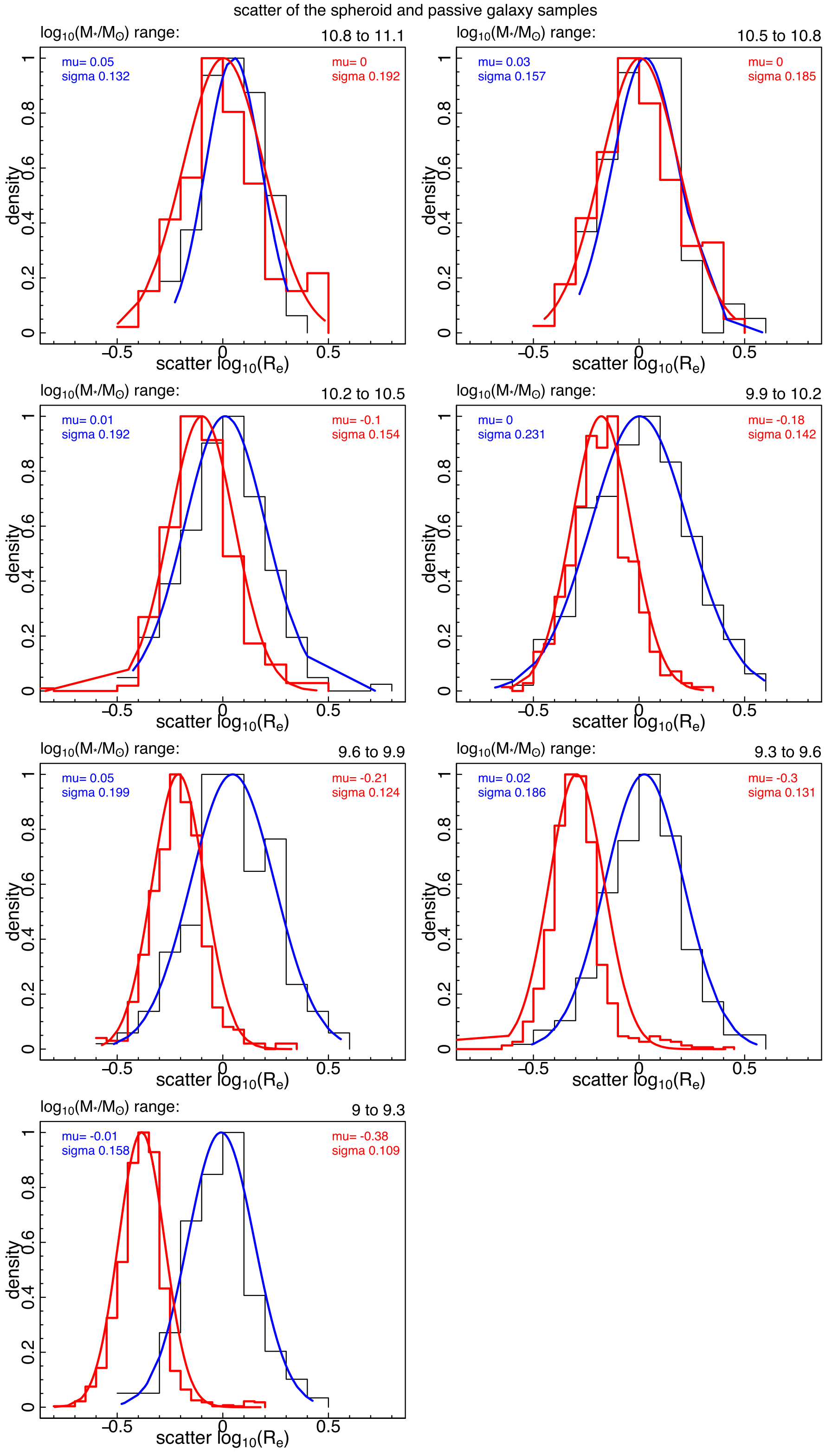


Figure B2. Same as Fig. B1 but shown is the $\log{(R_e)}$ scatter for the spheroid component distribution and in the case of EAGLE, we show the scatter of the half-mass radius of passive galaxies.

[1] *International Centre for Radio Astronomy Research, University of Western Australia, M468, 35 Stirling Highway, Crawley, WA 6009, Australia*
[2] *Scottish Universities' Physics Alliance (SUPA), School of Physics and Astronomy, University of St. Andrews, North Haugh, St. Andrews KY16 9SS, UK*
[3] *Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), 44 Rosehill Street, Redfern, NSW2016, Australia*
[4] *Astrophysics Research Institute, Liverpool John Moores University, IC2, Liverpool Science Park, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[5] *School of Physics and Astronomy, The University of Nottingham, University Park, Nottingham NG7 2RD, UK*
[6] *NASA Ames Research Center, N232, Moffett Field, Mountain View, CA 94035, USA*
[7] *Sydney Institute for Astronomy, School of Physics A28, University of Sydney, NSW 2006, Australia*
[8] *Astrophysics Group, School of Physics, University of Bristol, Bristol BS8 1TL, UK*
[9] *Australian Astronomical Observatory, PO Box 915, North Ryde, NSW 1670, Australia*
[10] *Department of Physics and Astronomy, University of the Western Cape, Robert Sobukwe Road, Bellville 7535, South Africa*
[11] *Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia*
[12] *European Southern Observatory, Alonso de Cordova 3107, Vitacura, Santiago, Chile*
[13] *University of Leiden, Sterrenwacht Leiden, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands*
[14] *Universität Hamburg, Hamburger Sternwarte, Gojenbergsweg 112, D-21029 Hamburg, Germany*
[15] *Jeremiah Horrocks Institute, University of Central Lancashire, Leighton Building, Preston PR1 2HE, UK*
[16] *The Astronomical Institute of the Romanian Academy, Str. Cutitul de Argint 5, Bucharest, Romania*
[17] *School of Physics, the University of Melbourne, VIC 3010, Australia*
[18] *Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, D-69117 Heidelberg, Germany*
[19] *European Southern Observatory, Karl-Schwarzschild-Str.2, D-85748 Garching, Germany*