Using morphology to understand galaxy evolution across the dwarf and massive-galaxy regimes

Author:
Ilin LAZAR

Supervised by:
Prof. Sugata KAVIRAJ
Dr. Garreth MARTIN
Prof. James GEACH

Centre for Astrophysics Research
School of Physics, Engineering and Computer Science
University of Hertfordshire

Submitted to the University of Hertfordshire in partial fulfilment of the requirements of the degree of Doctor of Philosophy.

May 2024

Abstract

The morphological properties of galaxies are important tracers of the physical processes, e.g. minor/major mergers, gas accretion and tidal interactions, that have shaped their evolution. This thesis first studies the morphological and structural properties of dwarf galaxies (which remain relatively unexplored compared to massive galaxies). It then explores the measurement of morphology using unsupervised machine learning (UML) algorithms, which will become important in the forthcoming era of peta and exascale surveys that will dominate the astrophysical landscape in the future.

I use a complete, unbiased sample of 257 dwarf (10^8 M$_\odot$ $< M_\star <$ $10^{9.5}$ M$_\odot$) galaxies at $z < 0.08$, in the COSMOS field, to study the morphological mix of the dwarf population in low-density environments. Visual inspection of extremely deep optical images reveals three principal morphological classes: 43 and 45 per cent of dwarfs exhibit the traditional 'early-type' (elliptical/S0) and 'late-type' (spiral) morphologies respectively, while 10 per cent populate a 'featureless' class, that lacks both the central light concentration seen in early-types and any spiral structure. This class is missing in the massive-galaxy regime. Compared to their massive counterparts, dwarf early-types show a much lower incidence of interactions, are significantly less concentrated and share similar rest-frame colours as dwarf late-types. This suggests that the formation histories of early-types is different in the dwarf and massive regimes. While massive early-types are likely shaped by interactions, their dwarf counterparts are shaped more by secular processes. This study suggests that featureless dwarfs in low-density environments are created via internal baryonic feedback, rather than by environmental processes. Finally, I show that while interacting dwarfs can be identified using the asymmetry parameter, it is challenging to cleanly separate early and late-type dwarfs using traditional morphological parameters, such as 'CAS', M_{20} and the Gini coefficient (unlike in the massive-galaxy regime).

I proceed by using 211 out of the 257 dwarf described above to study their structural properties: effective radii (R_e), effective surface brightnesses (μ_e) and colour profiles and gradients. I explore these properties as a function of stellar mass and the three principal dwarf morphological types (early-type galaxies (ETGs), late-type galaxies (LTGs) and featureless systems). The median effective radii of LTGs and featureless galaxies are factors of $\sim$2 and $\sim$1.2 larger than the ETGs. While the median effective brightness of the ETGs and LTGs is similar, the featureless class is much fainter. Dwarfs residing within the 'UDG' region in μ_e vs R_e space are a continuous extension of the galaxy population towards lower values of effective brightness, indicating that UDGs are not a distinct class of objects. While massive ETGs typically exhibit negative or flat colour gradients, dwarf ETGs show positive colour gradients (bluer

centres). The growth of ETGs changes from being 'outside-in' to 'inside out' as I move from the dwarf to the massive regime. The colour gradients of dwarf and massive LTGs are, however, similar. I show that, compared to their non-interacting counterparts, interacting dwarfs are larger, bluer at all radii and exhibit similar median effective surface brightness, suggesting that interactions enhance star formation across the entire galaxy.

Forthcoming 'Big data' surveys (e.g. LSST/SKA), which will produce peta and exabyte volumes of data will become the new 'normal' in the coming decades. These volumes will make morphological classification using traditional methods (e.g. direct visual inspection, as used in the first two chapters) impractical. Even semi-automated techniques, e.g. supervised machine learning with training sets built via visual inspection, may be difficult, because of the time-consuming nature of creating these sets. However, UML, does not require training on labelled data, making it ideal for galaxy morphological classification for large surveys.

Here I present an UML algorithm, that utilizes hierarchical clustering and growing neural gas networks to group together survey image patches with similar visual properties, followed by a clustering of objects (e.g. galaxies) that are reconstructed from these patches. I implement the algorithm on the Deep layer (27 deg^2) of the Hyper Suprime-Cam Subaru-Strategic-Program, to reduce a population of hundreds of thousands of galaxies to a small number ($\sim$150) of morphological clusters, which exhibit high purity. These clusters, rather than individual galaxies, can then be rapidly benchmarked via visual inspection and classified into typical morphological types. Using these morphological clusters, I successfully reproduce known trends of galaxy properties as a function of morphological type, which demonstrates the efficacy of the method.

Finally, I study 108 blue, low-mass ellipticals (which have a median stellar mass of $10^{8.7}$ $\mathrm{M}_{\odot}$) at $z <$ 0.3 in the COSMOS field, which have been autonomously detected by the algorithm described above. How elliptical galaxies form is a key question in observational cosmology. While the formation of massive ellipticals is strongly linked to mergers, the low mass ($M_{\star}/\mathrm{M}_{\odot} < 10^{9.5}$) regime remains less well explored. In particular, studying elliptical populations when they are blue, and therefore rapidly building stellar mass, offers strong constraints on their formation. Visual inspection of the deep optical HSC images indicates that less than 3 per cent of these systems have visible tidal features, a factor of 2 less than the incidence of tidal features in a control sample of galaxies with the same distribution of stellar mass and redshift. This suggests that the star formation activity in these objects is not driven by mergers or interactions but by secular gas accretion (i.e. steady gas accretion over Gyr timescales). I show that these blue ellipticals reside in low-density environments, further away from nodes and large-scale

filaments than other galaxies. At similar stellar masses and environments, blue ellipticals outnumber their normal (red) counterparts by a factor of 2. Thus, these systems are likely progenitors of not only normal ellipticals at similar stellar mass but, given their high star formation rates, also of ellipticals at higher stellar masses. Secular gas accretion, therefore, likely plays a significant (and possibly dominant) role in the stellar assembly of elliptical galaxies in the low mass regime.

Declaration

The following chapters of this thesis have been published previously:

1. Chapter 5: this was published as Lazar, I.; Kaviraj, S.; Martin, G.; Laigle, C.; Watkins, A.; Jackson, R. A., 2023, *Monthly Notices of the Royal Astronomical Society*, **520**, 2109.

2. Chapter 2: this was published as Lazar, I.; Kaviraj, S.; Watkins, A. E.; Martin, G.; Bichang'a, B.; Jackson, R. A., 2024, *Monthly Notices of the Royal Astronomical Society*, **529**, 499.

3. Chapter 3: this is accepted in MNRAS as Lazar, I.; Kaviraj, S.; Watkins, A. E.; Martin, G.; Bichang'a, B.; Jackson, R. A., 2024.

Acknowledgements

Firstly, I would like to thank my supervisor Sugata Kaviraj for an amazing guidance and mentorship throughout my studies. Thank you for the invaluable time and effort you put into this mentorship and for the exciting research conversations we had which fuelled my inspiration and ambition. I would like to thank my research group members and mentors Garreth Martin, Ryan Jackson and Aaron Watkins for the great advice and interesting discussions with regards to my work. I would like to thank other academics at University of Hertfordshire for great insight and advice with regards to my research projects such as Chiaki Kodayashi, Daniel Smith, James Geach and Emma Curtis Lake but also other researchers at other institutions such as Clotilde Laigle, Pierre Allain-Duc and Liza Sazanova. I would like to thank Vito Graffagnino for his IT and moral support. I would like to thank Victor Debattista, my masters supervisor and mentor, for helping me with guidance throughout my academic journey and for sharing his invaluable research expertise with me. I would like to thank previous project supervisors Eva Schinnerer and Daizhong Liu for the great knowledge and expertise they shared with me. I would like to also thank professor Cristina Popescu for her guidance and encouragement in my academic endeavour.

Secondly, I would like to thank all my PhD fellow colleagues and friends at University of Hertfordshire for a great moral support and friendship. In particular, my office colleagues Brian Bichang'a, James McGarry, Hariharan Mutu and Ben Law.

Finally, I would like to thank my family and especially my parents for making this achievement possible. Thank you for the moral, emotional and financial support, encouragement and for helping in any way possible in this journey. I would like to thank my highschool friends for moral support and a great friendship throughout this time. I would also like to thank my landlord for moral support and insightful discussions.

Contents

Abstract i

Acknowledgements v

Contents vi

1 Introduction 1
- 1.1 The morphological diversity of galaxies . . . 1
- 1.2 Stellar mass assembly over cosmic time . . . 6
  - 1.2.1 Gas accretion . . . 6
  - 1.2.2 Intrinsic star formation . . . 8
  - 1.2.3 Minor and major mergers . . . 9
  - 1.2.4 The regulation and quenching of star formation in galaxies . . . 13
    - 1.2.4.1 Feedback . . . 13
    - 1.2.4.2 Environment . . . 14
- 1.3 Galaxy morphology metrics: from classical to modern methods . . . 15
- 1.4 Overview . . . 16

2 The morphological mix of dwarf galaxies in the nearby Universe 17
- 2.1 Introduction . . . 17
- 2.2 Data . . . 21
  - 2.2.1 A complete sample of nearby dwarf galaxies . . . 21
- 2.3 Morphological classifications via visual inspection . . . 23
  - 2.3.1 Principal morphological classes in dwarf galaxies . . . 24
  - 2.3.2 Local densities of different dwarf morphological classes . . . 30
  - 2.3.3 The role of interactions . . . 31
  - 2.3.4 Rest-frame colours . . . 34
  - 2.3.5 Bars . . . 35
- 2.4 Morphological parameters . . . 38
  - 2.4.1 Concentration . . . 38
  - 2.4.2 Asymmetry . . . 39
  - 2.4.3 Clumpiness . . . 39
  - 2.4.4 Gini coefficient . . . 40
  - 2.4.5 M_{20} . . . 40
  - 2.4.6 Sérsic index . . . 42
  - 2.4.7 Comparison to dwarf and massive galaxies in the very local Universe (within 50 Mpc) . . . 42

2.4.8 Comparison to massive galaxies in the nearby Universe ($z < 0.1$) . . . 45
2.4.9 Can interacting and non-interacting dwarfs be separated using morphological parameters? . . . 49
2.5 Implications for the evolutionary histories of dwarf morphological classes . . . 51
2.5.1 Dwarf ETGs have different evolutionary histories to massive ETGs . . . 51
2.5.2 Featureless dwarfs in low-density environments are formed via baryonic feedback 54
2.5.3 Shallower potential wells cause late-type dwarfs to be structurally more asymmetric than massive late-types . . . 54
2.6 Summary . . . 55

3 The structure of nearby dwarf galaxies as a function of stellar mass and morphology – size, effective surface brightness and colour gradients 58
3.1 Introduction . . . 58
3.2 A sample of nearby dwarf galaxies . . . 61
3.2.1 Morphological classification via visual inspection . . . 62
3.2.2 Masking, PSF correction and construction of surface-brightness profiles . . . 64
3.3 Size . . . 65
3.4 Effective surface brightness . . . 72
3.5 Colour profiles and gradients . . . 74
3.6 Summary . . . 79

4 Autonomous classification of galaxy morphologies in the HSC-SSP using unsupervised machine learning 81
4.1 Introduction . . . 81
4.2 Data: the HSC-SSP . . . 84
4.3 Methodology: unsupervised machine learning algorithms for morphological analysis . . 84
4.3.1 Feature space . . . 85
4.3.2 Growing Neural Gas . . . 86
4.3.3 Hierarchical clustering . . . 87
4.3.4 Final object clustering . . . 88
4.3.5 Improvements made to the previous versions of the technique . . . 89
4.4 Galaxy morphologies in the HSC-SSP DR3 DEEP . . . 90
4.4.1 An example of the autonomous identification of specific morphological types - blue elliptical galaxy . . . 96
4.5 Summary . . . 99

5 Relaxed blue ellipticals: accretion-driven stellar growth is a key evolutionary channel for low mass elliptical galaxies 101
5.1 Introduction . . . 101
5.2 Galaxy morphologies via unsupervised machine learning . . . 104
5.3 COSMOS2020: accurate physical parameters for galaxies . . . 105
5.4 Measurement of local density and topological structure . . . 106
5.5 Blue ellipticals: the importance of secular gas accretion and the insignificance of mergers 109
5.5.1 Identification . . . 109
5.5.2 Physical properties . . . 112
5.5.3 Absence of tidal features - blue ellipticals typically do not evolve via galaxy interactions . . . 115
5.5.4 Blue ellipticals reside further away from nodes and large-scale filaments . . . 116

5.5.5 Blue ellipticals as progenitors of the general low mass elliptical population . . . 117
5.6 Summary . . . 118

6 Conclusions **121**
6.1 The morphological mix and structural properties of dwarf galaxies in the nearby Universe 121
6.1.1 Differences in stellar mass assembly in dwarf and high mass early type galaxies . 122
6.2 Galaxy morphological classification in the HSC-SSP survey using unsupervised machine learning . . . 124

7 Future work **125**

A **127**
A.1 Dwarfs in the DES comparison sample are biased towards blue, late-type galaxies 127

B **130**
B.1 Colour bimodality in the nearby dwarf population . . . 130

C **131**
C.1 Density maps . . . 131
C.2 Examples of sersic fits . . . 133

Bibliography **137**

Chapter 1

Introduction

1.1 The morphological diversity of galaxies

For over 100 years the study of galaxy morphology has been a cornerstone of extra-galactic research. The detailed analysis of galactic structure and morphology began in the late 19^{th} and early 20^{th} centuries when photographic plates were used to record the fluxes of astronomical objects (e.g. Wolf, 1908; Lundmark, 1926, 1927; Holmberg, 1937). A first general galaxy classification scheme was introduced by Hubble et al. (1926) and revised by Hubble (1936) under the terminology 'The Hubble Tuning Fork' (or the 'Hubble sequence'), which can be seen in Figure 1.1. This classification system generally segregates two main morphological types, early and late type galaxies, where late type systems are further sub-divided according to the sizes of their bulges, the tightness of their spirals arms and the presence of bars in the galactic centres.

Early type galaxies consist of two morphological sub-classes: elliptical and lenticular (S0). Elliptical galaxies are spheroidal, dispersion-dominated systems with a prominent central light concentration (or 'bulge') and a diffuse and smooth outer structure (e.g. Almeida et al., 2007). The lenticulars, on the other hand, present some structural characteristics resembling a smooth disk component alongwith a bulge (e.g. Laurikainen et al., 2010; D'Onofrio et al., 2015). Late-type galaxies are typically disk-like, rotation-dominated systems generally exhibiting low central light concentrations and containing non-axisymmetric features such as spiral arms and galactic bars (e.g. Pohlen & Trujillo, 2006) (see examples in Figure 1.2).

Although the visual morphological differences between early and late type galaxies already point to different formation and evolution channels, there is a rich literature which presents evidence supporting a

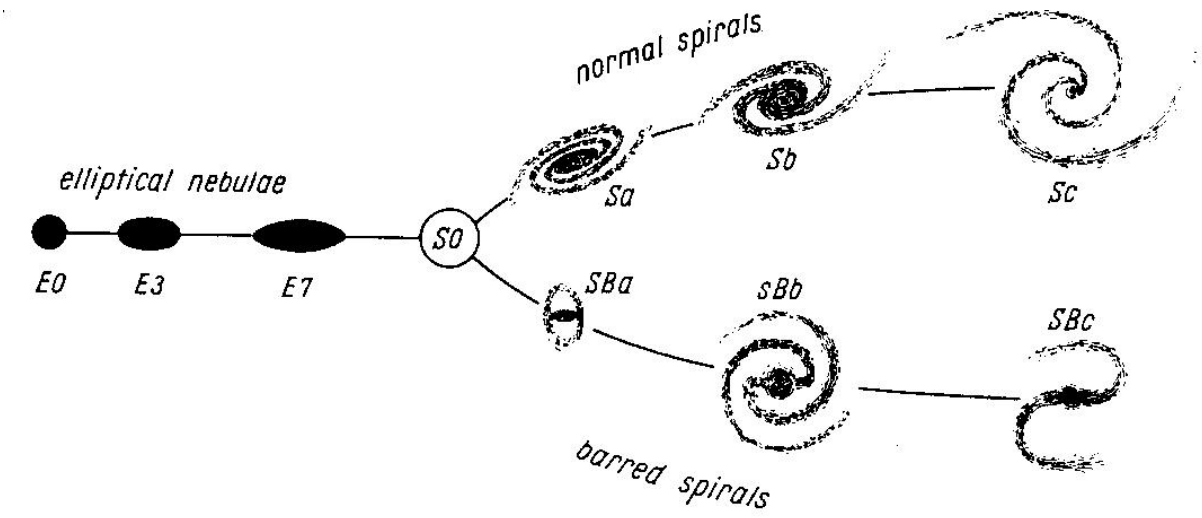

FIGURE 1.1: The classification scheme from Hubble (1936).

strong interdependence of galaxy structure with galaxy properties like stellar mass, star formation rate or environmental density (e.g. Dressler et al., 1997; Hogg et al., 2002; Conselice, 2006a; Bamford et al., 2009; Lee et al., 2013; Hashemizadeh et al., 2021; Dimauro et al., 2022). Since the physical properties of galaxies are strongly linked to their morphology, morphological information provides important insights into the physical mechanisms that shaped their intrinsic structure. For example, early type galaxies generally dominate the high mass regime, have old stellar populations and are in a quiescent state (presenting red optical colours) (e.g. Kelvin et al., 2014; Domínguez Sánchez et al., 2023) as opposed to late type galaxies, which are still star-forming (exhibiting blue optical colours), have generally younger stellar populations and are less massive (e.g. Bell & de Jong, 2000; Domínguez Sánchez et al., 2023) (see examples in Figures 1.3 and 1.4).

Many studies, such as Buitrago et al. (2013), Conselice (2014) and Huertas-Company et al. (2024), show that the galaxy morphological fractions change as we approach the present day with the number density of massive ($M_\star > 10^{10.5}$ $M_\odot$) early type galaxies increasing as the number density of the late types decreases. In addition, a significant number of high redshift ($z > 2$) galaxies are highly star forming and have been found to have disturbed and peculiar morphologies as opposed to present day galaxies which are more quiescent and have a more organized intrinsic structure (e.g. Huertas-Company et al., 2024; Conselice et al., 2024). Therefore, the morphological diversity of galaxies changes as a function of time and the major questions of extra-galactic astrophysics revolve around understanding the dominant evolutionary channels of galaxies as a function of morphology and redshift.

The advent of galaxy morphological analysis at a large scale begins in the last decade of the 20$^{\text{th}}$ century with HST and continues in the first two decades of the 21$^{\text{st}}$ century with surveys such as Sloan Digital Sky

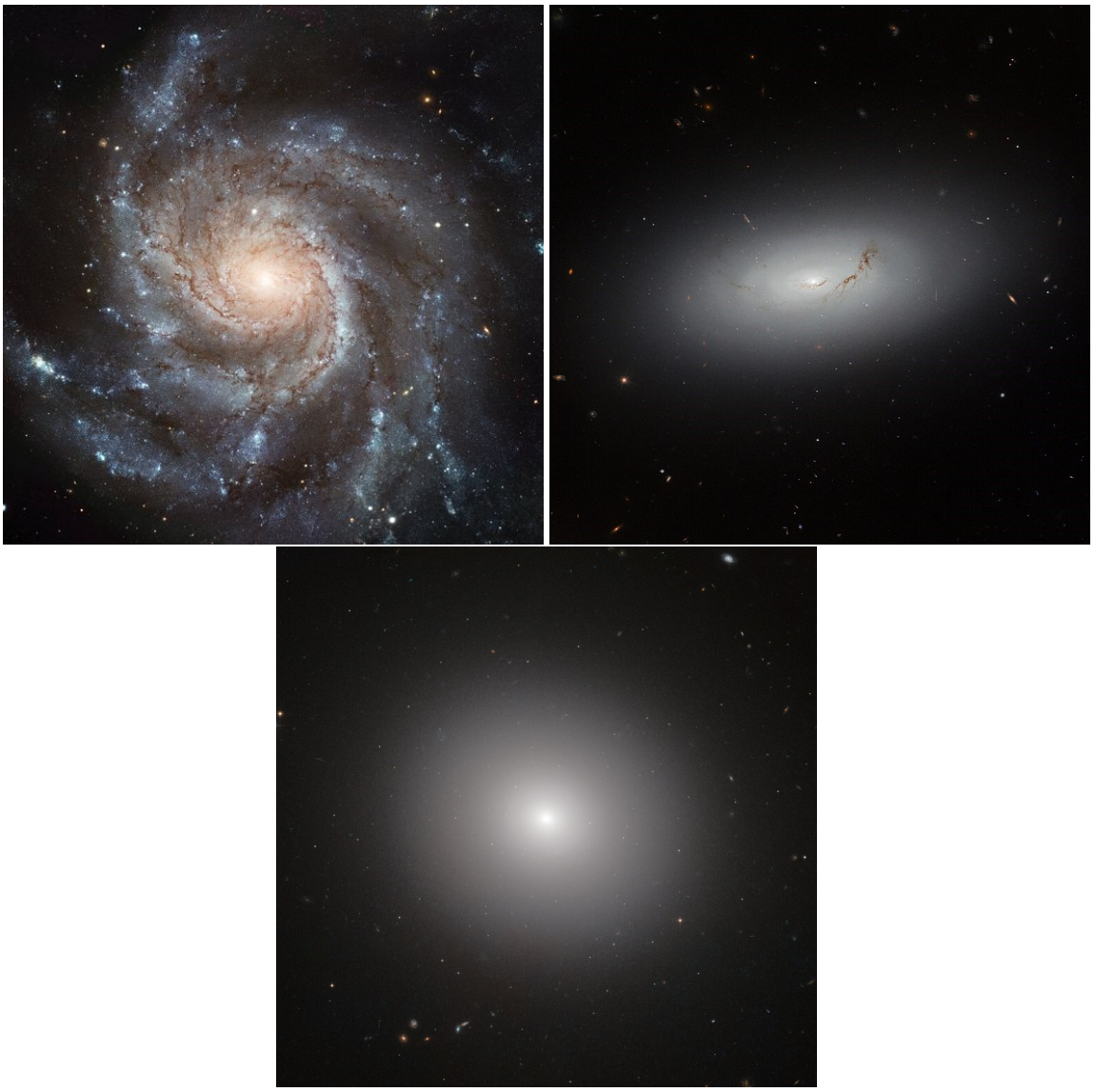

FIGURE 1.2: Examples of galaxies from different morphological classes. Late type galaxy M101 (top-left), lenticular galaxy (or S0 type) NGC 3156 (top-right) and elliptical galaxy IC 2006 (bottom). Credit: ESA/NASA.

Survey (SDSS) (e.g. Yamauchi et al., 2005; Fukugita et al., 2007; Ball et al., 2008) or the Spitzer Survey of Stellar Structure in Galaxies (e.g. Buta et al., 2010; Holwerda et al., 2014; Buta et al., 2015) which offer complete samples of thousands of galaxies in the high mass regime (i.e. high surface brightness objects), particularly at low redshift. More recently, deep-wide surveys such as those from the Hyper Suprime-Cam (Aihara et al., 2018) have been able to resolve millions of objects out to higher redshift and down to much fainter surface brightnesses than shallower surveys such as SDSS. This is now allowing for large scale unbiased morphological studies not only in massive galaxies (e.g. Kawinwanichakij et al., 2021; Ghosh et al., 2023) but also in the dwarf regime (e.g. Greco et al., 2018; Lazar et al., 2024). New and forthcoming surveys such as LSST, Euclid, Roman and SKA will shed more light into the low

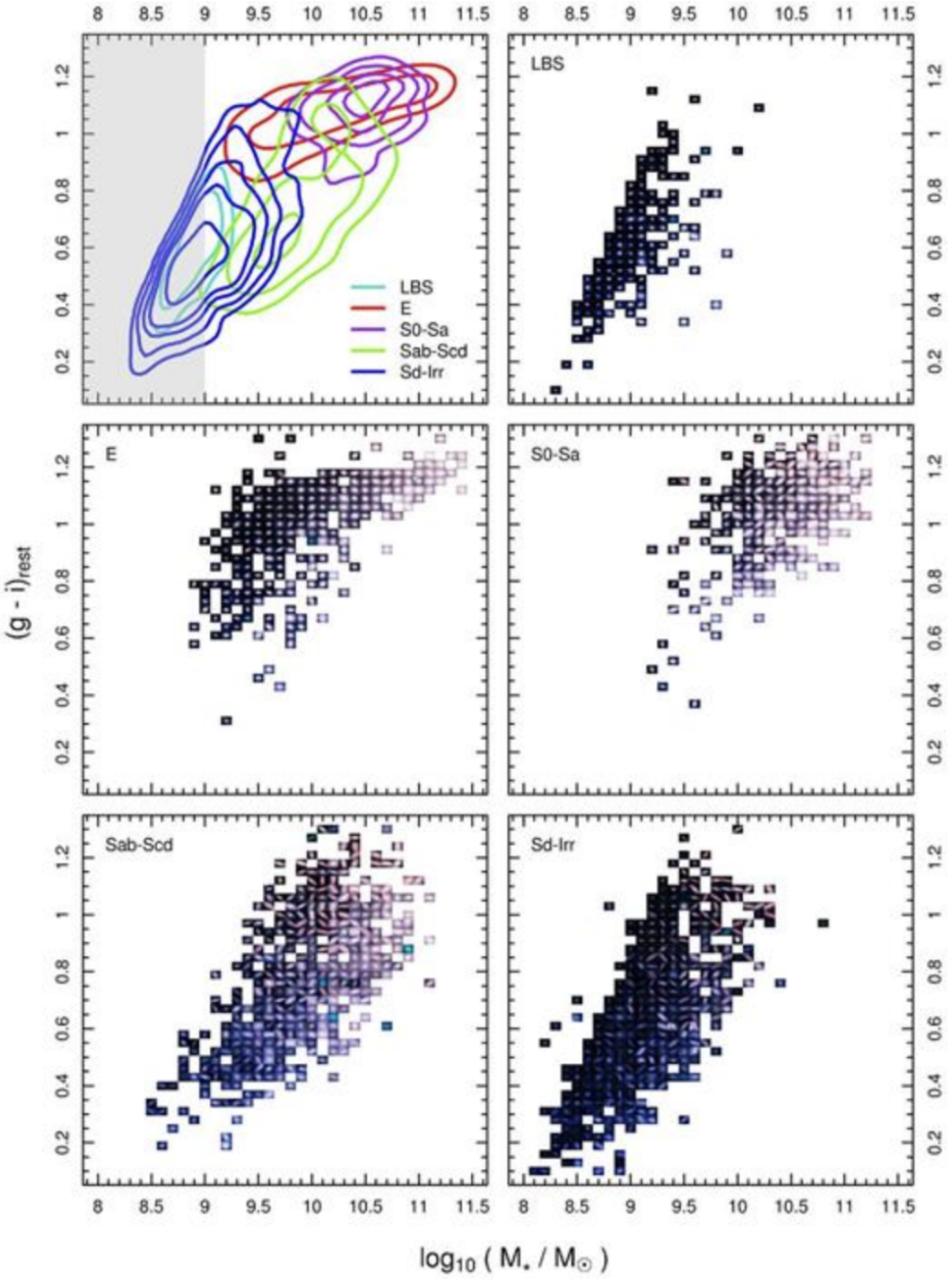

FIGURE 1.3: Figure showing examples of galaxies from different morphological classes present in different locations on the color-stellar mass diagram. Top-left panel shows all the different morphologies while the other panels present galaxy postage stamps in the locations of the different galaxy morphologies. Early type galaxies are given by E and S0-Sa and late type galaxies are Sab, Scd, Sd and Irr (as shown in Figure 1.1). LBS are a class of blue spheroids. Credit: Kelvin et al. (2014).

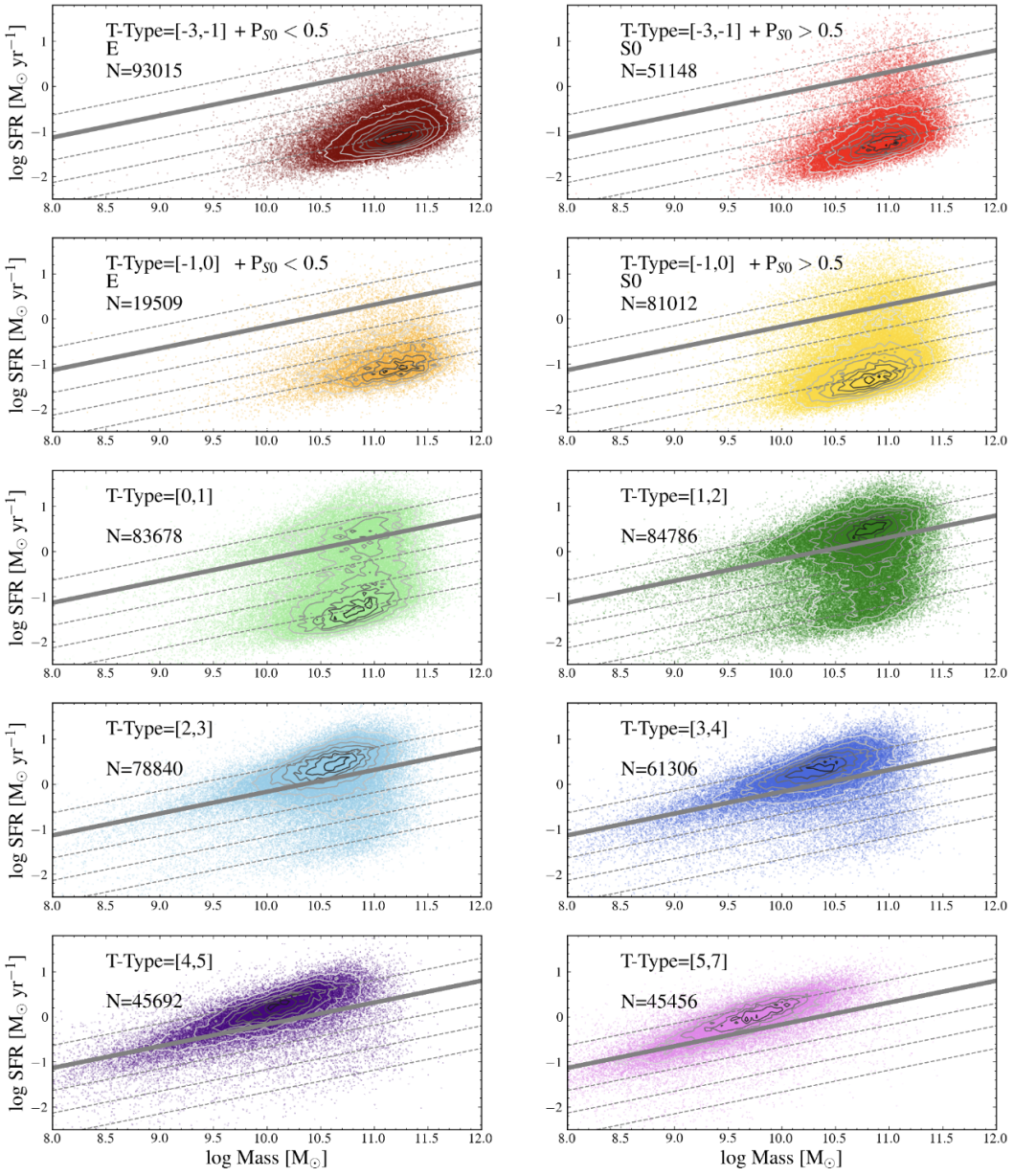

FIGURE 1.4: Figure showing examples of galaxies from different morphological classes present in different locations on the SFR-stellar mass diagram. Here the morphological classes are shown as the Hubble T-Type number which is based on the Hubble sequence and corresponds to early type galaxies for T-Type=[-3,0] and late type galaxies for T-Type=[1,7]. Credit: Domínguez Sánchez et al. (2023).

surface brightness regime from a more statistical perspective including objects at higher redshifts.

1.2 Stellar mass assembly over cosmic time

The first galaxies emerging after the Big Bang were most likely gas disks embedded within dark matter (DM) haloes (e.g. Pawlik et al., 2011). The general formation model for a disk galaxy assumes a spherical overdense region of space consisting of both dark and baryonic matter which are spun up by the tidal torques caused by the large scale structure (Peebles, 1969; Doroshkevich, 1970). Over time the dark and baryonic matter (which is initially relatively pristine hydrogen and helium) contained within the overdense region collapses and some regions increase in density. This causes baryons to cool in these regions by decoupling from the dark matter particles which have halted their collapse once the system reached virialization, since they are dissipationless particles. The baryons condense into the centre of the dark matter halo and since they possess non-zero angular momentum they settle onto a rotating disk (e.g. Dalcanton et al., 1997; Mo et al., 1998). The initial angular momentum of the baryons is thought to be created as a direct consequence of the bottom-up assembly (White & Rees, 1978) or cold flows (Kimm et al., 2011; Stewart et al., 2013). Even though the exact physical process is less known, as soon as the gas accumulates into dense clumps within the disk and is sufficiently cool, the optimal conditions for star formation are reached and stars are able to form (Tonini et al., 2016). However, it has been shown by studies such as Sommer-Larsen et al. (1999), Thacker & Couchman (2001) and Zavala et al. (2008) that baryonic feedback regulates the gas cooling in the early times an plays an important role in the formation of the final disk structure.

Even by considering stellar mass loss due to stellar evolutionary processes (e.g. AGB winds or SNe events) galaxies need access to gas in order to continue their star formation activity at the same rate otherwise the star formation would decrease after the first few Gyrs of evolution (Bigiel et al., 2008; Leroy et al., 2008; Rahman et al., 2012). Once these primordial disk structures are formed, the assembly of stellar mass in these systems can continue via two main channels: gas accretion from the intergalactic medium and mergers with other galaxies. The major mechanisms that play an important role in the formation and evolution of galaxies are described by these two evolutionary channels.

1.2.1 Gas accretion

One of the mechanisms that allows galaxies to access fresh gas is via direct accretion from ambient inter-galactic medium on the halo of the galaxy. Once the gas crosses the viral radius, it experiences

subsequent shocks with the halo gas which increase the temperature of the in-flowing gas. If the temperature of the gas maintains below the virial temperature (i.e. it does not escape the potential of the halo), it cools and collapses onto the galaxy. The cooling of the in-falling gas proceeds at a slower pace if the accreting gas has a high temperature (i.e. comparable to virial temperature) and star formation is delayed as a consequence of shock heating at the edges of the halo (Rees & Ostriker, 1977; White & Rees, 1978). This results in 'hot' mode accretion (e.g. Rees & Ostriker, 1977; Cole et al., 1994; Stevens et al., 2017; Stern et al., 2024), where the halo is dominated by new shocked gas which slowly cools giving rise to isotropic accretion. However, cold streams of gas can spiral in on the dark matter halo at high angular momentum. As a result, shock heating takes place at smaller distances from the galactic centre where the energy injected by the shock on the baryons is easily dissipated, since the temperature of the in-falling gas is significantly below the virial temperature. This 'cold' mode accretion is an efficient mechanism of the inside-out formation of a stellar disk (Dekel & Birnboim, 2006; Dekel et al., 2009; Danovich et al., 2012; Stewart et al., 2017).

Studies which use hydrodynamical simulations highlight the importance of accretion into the central regions of galaxies via cold flows from the cosmic web (e.g. Kimm et al., 2011; Danovich et al., 2012, 2015) at high angular momentum (Powell et al., 2011). Supported also by observations, cold mode accretion is thought to represent an important element for galaxy mass assembly, where evidence of this type of events is found at low redshifts in the form of regular or warped gas disks with significant spins (e.g. Diamond-Stanic et al., 2016; Ho et al., 2017; Bouché et al., 2016) or at high redshifts in the form of gas disks or rings (e.g. Prescott et al., 2015). Additionally, signatures of cold mode accretion can also be seen in the scale lengths of disks (Brooks et al., 2009) and the galaxy geometry with respect to cosmic web filaments (e.g. Chen et al., 2019; Barsanti et al., 2022).

Some studies such as Kereš et al. (2005) and van de Voort et al. (2011) suggest that cold mode accretion has a significant contribution in galaxy mass assembly generally in the early Universe and for low mass dark matter haloes due to denser intergalactic filamentary gas and due to the lower galactic halo temperatures which cause gas accretion do develop in a rapid manner and over shorter timescales since heating caused by shocks is not as important. At low redshifts hot mode accretion becomes important since haloes start to heat up and the gas density of the cosmic web drops. As a result, gas accretion is slowed down due to shock heating and takes a longer time to collapse. Nevertheless, while gas cold flows become less dominant, they may still contribute to the stellar mass assembly of low mass galaxies (e.g. van de Voort et al., 2011).

The importance of gas accretion is also provided by the observed star formation rates of galaxies where

there is observational evidence that mergers alone could not have accounted for all the stellar assembly in galaxies. High redshift studies such as Conselice et al. (2013) find that in order to explain the high star formation rates observed in galaxies at $z \sim 2$ with $M_{\star} < 10^{11}$ $M_{\odot}$, gas accretion must happen at a rate of $\sim$90 $M_{\odot}$ yr^{-1}. Therefore, accumulating gas through accretion can play an equally important role as mergers in driving star formation and stellar mass growth.

Lastly, high gas in-fall rates from the intergalactic medium or from minor mergers causing chaotic accretion can also drive mass growth. This chaotic accretion can give rise to a gravitational instability within a galactic disk and the gas must lose angular momentum in order for the disk to re-enter a state of hydrodynamical equilibrium (Tonini et al., 2016). The gas excess fails to maintain a high angular momentum and falls towards the center of the galaxy generating high levels of star formation and increasing the central light concentration of the system.

1.2.2 Intrinsic star formation

The accreted supply of gas has to be converted into stars for the accretion events to have effects on stellar mass assembly. Since stellar mass growth in galaxies can be the result of in-situ star formation from accreted gas but can also originate from ex-situ mass from mergers, the major fraction of stellar mass assembly must have occurred in-situ (i.e. within the galaxy's halo) for galaxies below $M_{\star} \sim 10^{11}$ $M_{\odot}$ (e.g. Rodriguez-Gomez et al., 2016a).

The exact underlying physics of star formation is poorly understood mostly due to the large dynamical range that must be taken into account. Nevertheless, the star formation law also know as the Kennicutt-Schmidtt law (Kennicutt, 1998), provides a reliable picture of the relationship between star formation rate density and surface gas density for a wide range of spatial scales down to $\sim$100 pc (e.g. Onodera et al., 2010; Khoperskov & Vasiliev, 2017). Moreover, galaxy observations also point towards a tight correlation between the star formation rate and stellar mass (i.e. the "main-sequence") (e.g. Noeske et al., 2007; Rodighiero et al., 2010; Salmon et al., 2015) which may suggest that the majority of star formation carry on at a constant rate.

The correlation between the galaxy gas fractions and star formation rate density maintains across cosmic time (e.g. Genzel et al., 2010) where the star formation activity is at its maximum at around $z \sim 2$ (Madau & Dickinson, 2014). Gas-rich galaxies exhibit significant start formation rates (greater than 150 $M_{\odot}$ yr^{-1}) at high redshifts ($z \sim 2$) (e.g. Daddi et al., 2005). However, star formation is limited at low

FIGURE 1.5: Two merging galaxies in the constellation Come Berenices. Credit: NASA/ESA.

redshifts due to feedback processes and low gas densities. As a result, in-situ star formation allows for gradual growth of stellar mass and is an important channel for the stellar mass assembly of galaxies.

1.2.3 Minor and major mergers

The in-situ formation of stars is not the only way that stellar mass can be assembled within galaxies. An alternative channel that allows for mass growth is through mergers of two or more galaxies. This is viewed as a significant evolutionary channel via which galaxies accumulate stellar mass and change morphologies, particularly at the highest stellar masses (e.g. Robotham et al., 2014; Ferreras et al., 2017; Qu et al., 2017) in the hierarchical paradigm of galaxy formation and evolution (White & Rees, 1978).

Mergers are broadly split into two categories: minor and major mergers. Major mergers are defined as mergers between galaxies of similar stellar masses where the stellar mass ratios are greater than 1:4. In minor mergers one galaxy is considerably larger than its companion and the system has a stellar mass ratio less than 1:4 (e.g. Somerville & Davé, 2015).

Mergers are identified observationally by the disturbed morphology of the remnant or, in the case where both of galaxies are seen interacting or in close proximity (Conselice et al., 2014), through the tidal features the merger produces around them (see examples in Figures 1.5 and 1.6).

Mergers create gravitational torques which produce tidal forces that strip away stellar mass from the host galaxies. Depending on the merger ratio the tidal forces inflicted on the stellar content can vary in intensity, where major mergers are usually the most violent, displacing the most amount of mass from the interacting galaxies, creating tidal features. Fly-by events, galaxy interactions that do not end with coalescence, can also create tidal features usually of less intensity and which are less extended than in the case of major mergers but which can also have important effects on the morphology of the host system, such as disk thickening (e.g. Kumar et al., 2021).

Observationally, the timescale of the tidal features for which they remain visible to the observer depends on the depth of the galaxy images, where shallower instruments produce images with more amount of noise, putting constraints on the clarity of the image. For example, in an empirical study done by Conselice et al. (2009), the author uses Hubble Space Telescope imaging data from the Extended Groth Strip and COSMOS surveys at $z<1$ to find that, based on the merger rates observed in this redshift range, the timescale of the tidal features for which they remain visible to the observer (until they are dispersed or pulled back towards the galaxy) is in the range between couple hundred Myrs to 1 Gyr which is also confirmed in N-body simulations (e.g. Conselice, 2006b; Conselice et al., 2007).

Major and minor mergers can contribute to stellar mass assembly with the aid of several underlying physical mechanisms. Firstly, mergers simply increase the overall stellar mass of the systems in question. However, they can also enhance the density of the existing gas, in addition to increasing the galaxy's gas supply (Barnes, 2004; Renaud et al., 2014; Buck et al., 2023). Nevertheless, mergers do not always enhance star formation (Robotham et al., 2014) even though, generally, galaxy populations at any given redshift exhibiting very high star formation rates tend to have a high fraction of mergers. The requirement for a merger to result in a significant star formation activity is that one or more of the interacting galaxies has to be rich in gas (also known as 'wet' mergers) (e.g. Lin et al., 2008; Athanassoula et al., 2016). Wet mergers generally increase the star formation activity of the host galaxy considerably (an enhancement by 3 or 4 times). Very significant sustained enhancement of star formation is not common since the duration of the starburst period is very short (Conselice et al., 2014).

Mergers are also a direct cause for changes in the kinematics and morphology of a galaxy by randomizing stellar orbits which results in an increase of the system's velocity dispersion and an enhancement of its bulge component (e.g. Toomre & Toomre, 1972; White & Rees, 1978; Bournaud et al., 2007; Kormendy

et al., 2009; Martig et al., 2009; Bezanson et al., 2009). Some studies suggest that single major mergers could directly produce elliptical galaxies (e.g. Farouki & Shapiro, 1982; Negroponte & White, 1983; Zeng et al., 2021). Such mergers are typically violent and short lived events lasting only a couple of orbits before the galaxies coalesce (a few hundred Myrs) (e.g. Conselice et al., 2009; Bournaud, 2010). Major merger events act on shorter timescales on the gravitational potential of a galaxy than its dynamical timescale. This causes stellar orbits to be randomised, while their angular momentum and energy is altered via 'violent relaxation' (Bournaud, 2010), resulting a dispersion-dominated system. Due to violent relaxation, stars that gain energy progress towards the edges of the system or are removed entirely from it and stars which lose energy migrate towards the centre of the merger remnant. The end result is an increase in the remnant's central light concentration and also an increase in the extent of the remnant's outer envelope (Hernquist & Mihos, 1995). Using the TNG 100 simulation Zeng et al. (2021) find that the morphology of the major merger remnant is strongly dependent on the orbit type where spiral-in collisions lead to a disk remnant and head-on collisions result in a bulge-dominated remnant. However, there is some debate about whether the dominant channel by which elliptical galaxies develop their morphologies is uniquely due to major mergers (e.g. Brennan et al., 2015) followed by subsequent quenching of star formation via AGN feedback (e.g. Hopkins et al., 2008).

Minor mergers are statistically more frequent since the luminosity function increases towards the faint end (e.g. Man et al., 2012). Although, minor mergers have little impact in altering the structural properties of the galaxy, they can still produce conditions for starting star formation events by triggering the gas supply of the host galaxy (Tonini et al., 2016), resulting in stellar mass growth. The mass growth due to a minor merger is more gradual since the orbital decay rate of the interacting galaxies is much smaller than in the case of a violent galaxy coalescence with a 1:1 or 1:2 stellar mass ratio (Toomre, 1977). This mechanism produces a framework where the larger galaxy progressively strips matter from its smaller companion. If the satellite is large enough it can affect the host's structure by randomizing some of its stellar orbits and causing warps and thickening of its disk (Qu et al., 2011).

When considering a minor merger system, the stellar mass growth of the host can depend on its morphology. If the host is a bulge-dominated galaxy the minor merger may result in an inside-out mass growth where the galaxy accretes matter in the outskirts of the system (Wellons et al., 2016). For a disk-dominated host (i.e. rotationally supported system) radial migration of stars towards the galactic center can be triggered via a minor merger therefore increasing the central mass density of the system (Croton et al., 2016).

Minor mergers are thought to play an important role in the assembly of late type galaxies throughout

FIGURE 1.6: Galaxy showing shell-like tidal features. Credit: NASA/ESA and Liverpool John Moores University.

cosmic time mainly because they can provide a natural explanation for which late type galaxies have maintained their structured morphology while at the same time acquiring significant amounts of mass (Schweizer, 2000; Somerville & Davé, 2015; Martin et al., 2018a). Martin et al. (2018a) also show that the survival of the galactic disk after a merger is highly dependent on the geometry of the orbiting system and support a scenario where pairs of disk mergers can maintain a rotationally supported system as long as the galaxies undergo a wet merger (which maintains star formation) and the galaxies collide while their spins are aligned.

In addition, studies such as Martin et al. (2018a) or Zeng et al. (2021) use the Horizon-AGN and TNG

100 cosmological simulations to show that minor mergers represent a dominant channel for the stellar mass assembly of both late and early type galaxies where early type galaxies experience more major mergers or mergers with higher mass ratios in their histories. The authors also find that the majority of the interactions experienced by early type galaxies are minor mergers at $z < 1$. This claim is also supported by Kaviraj et al. (2009) who demonstrate by means of numerical simulations that the UV light excess seen in early type galaxies in the local Universe likely originates from minor mergers.

1.2.4 The regulation and quenching of star formation in galaxies

1.2.4.1 Feedback

Feedback plays an important role in regulating the star formation activity in a galaxy and can either reduce the efficiency of star formation or quench it entirely, halting the stellar mass growth of the galaxy.

The epoch of reionization, know from observations to have taken place at around $z \sim 6$ (e.g. Fan et al., 2006), is the era when the gas was largely ionized to temperatures of around 10^4 K. Although the source of this ionization is not well known, it is thought to originate from the UV radiation produced by young stars and QSOs. The result of this process was to infuse the gas within low mass haloes with enough energy that would prevent its cooling and collapse, since once heated it is easier for baryons in shallow potential wells to escape the collapse. Therefore, low mass haloes that could have hosted galaxies with stellar masses less than $M_\star \sim 10^7$ $M_\odot$ are thought to be prevented from doing so (Silk, 2011). This can make the dwarf part of the galaxy mass function very dependent on the UV radiation from the reionization epoch, which regulates the abundance of galaxies in that mass range. A better understanding of the source of the ionization is necessary in order to constrain the masses of the haloes that are being prevented from forming galaxies.

In the low mass regime ($M_\star < 10^{9.5}$ $M_\odot$) stellar feedback is a significant driver of baryonic outflows and heating of the interstellar medium via supernovae (SN) events and stellar winds (Collins & Read, 2022). As a result, such mechanisms regulate the star formation efficiency of the system through the heating and removal of its gas (e.g. Hayward & Hopkins, 2017). The material in the outflows can then be recycled by cooling and re-entering the galaxy, a process which gradually forms new generations of stars over long timescales. Type II SN events and stellar winds from young stars have a significant role in heating the interstellar medium in the initial stages of the star formation episode (Tenorio-Tagle, 1996). Type Ia SN events and stellar winds from new generations of stars act over longer timescales and they become an important source for heating the interstellar medium over Gyr timescales.

AGN feedback is thought to have significant effects on galaxies at the massive end of the mass function ($M_\star > 10^{9.5}\ M_\odot$) (e.g. Beckmann et al., 2017). The energy released by an AGN originates from the underlying accretion events onto the supermassive black hole located in the galactic centre. The effects of stellar feedback are negligible in massive galaxies due to the deep potential wells of these systems.

Generally, AGN are of two types: quasars (also called radiative or wind mode AGN) and jet (or radio) mode AGN (e.g. Fabian, 2012). Depending on the type of AGN, its released energy can affect both the interstellar and the intergalactic medium. The jet mode outputs lower levels of energy due to lower accretion rates of matter onto the supermassive black hole and acts to continuously keep the halo at high temperatures, resulting in the halo gas not being able to cool and be accreted onto the galaxy. This type of AGN has been shown to also have effects on the inter-galactic medium where it can prevent gas in galaxy clusters and groups to cool (e.g. Fabian, 2012; Hardcastle et al., 2019). The quasar mode AGN feedback results from periods of rapid accretion of matter onto the supermassive black hole, generally occurring in the most massive halos at high redshifs (e.g. Combes, 2017). It typically affects the interstellar medium around the black hole at smaller spatial scales (pc to kpc) by injecting energy into this environment and therefore, preventing it from forming stars (e.g. Veilleux et al., 2017).

1.2.4.2 Environment

Environment has been shown to have a significant role in galaxy evolution, where the observed distribution of galaxy morphologies at low and intermediate redshift is strongly correlated with environmental density (Dressler, 1980; Dressler et al., 1999). These observations indicate that early type galaxies typically inhabit denser environments compared to late type galaxies. Dense environments such as clusters or groups can have a strong influence on galaxies due to strong tidal forces which can inject energy into the stars and the gas and processes like ram pressure which can strip gas from the system (e.g. Hester, 2006). Both these processes can result in the quenching of star formation.

The influence of environment on galaxies at high redshift ($z > 1$) is generally not as clear. While some observations show a tendency for massive galaxies to be quiescent in clusters (e.g. Papovich et al., 2012), some authors find elevated star formation rates in cluster galaxies at high redshift (Alberts et al., 2016; Laishram et al., 2024).

While much progress has been made in recent years, the principal drivers of morphological transformation and stellar assembly for different galaxy types are still under debate. This is particularly true in the low mass regime. However, the arrival of data from surveys such as LSST, Euclid, Roman and

SKA is set to change our understanding of this regime. Accurate and efficient galaxy morphological classification analysis methods will be crucial for this endeavour in the coming years.

1.3 Galaxy morphology metrics: from classical to modern methods

Galaxy morphological classification via visual inspection is a common traditional method of quantifying the structural properties of galaxies using morphological classes defined by studies like Hubble et al. (1926). While individual astronomers can classify small datasets, citizen science platforms such as Galaxy Zoo have performed millions of galaxy classifications by crowd sourcing the process using several thousand citizen scientists (Lintott et al., 2011) and making use of galaxy images from large scale surveys (e.g. Willett et al., 2013; Ahn et al., 2014).

However, more automated ways of analysing galaxy structure exist. A common parametric way of performing this task is via Sérsic fitting of galaxy light profiles (Sérsic, 1963), where the shape of the profile is characterized by the Sérsic index (n). The standard classification threshold for late and early type galaxies is at $n = 2.5$, where $n > 2.5$ corresponds to early type galaxies and $n < 2.5$ corresponds to late type galaxies (e.g. Shen et al., 2003). Other popular non-parametric methods are the CAS system (Conselice, 2003), M_{20} and the Gini coefficient (Lotz et al., 2004) which do not assume any underlying shape for the galaxy light profile and are able to measure galaxy structure out to high redshift.

However, the arrival of big data surveys such as SDSS, DES and DECaLS has prompted a move towards automated methods that exploit machine-learning techniques. A successful classification technique is using supervised machine learning via convolutional neural networks, where labelled data is trained to obtain a model which is then used for large scale galaxy classification (e.g. Domínguez Sánchez et al., 2018; Cheng et al., 2021; Walmsley et al., 2022). This method is however often dependent on human classifiers to construct the training sets. The peta and exa-scale volumes of upcoming surveys such as LSST and Euclid may present challenges in creating these training sets. This is particularly true for LSST data, where the data itself will become deeper and therefore change over small periods of time. Unsupervised machine learning does not rely on labelled training datasets and may become important for classifying such surveys (as we explore in this thesis).

1.4 Overview

The broad aim of this work is two fold. First, we aim to further our understanding of the physical processes that drive galaxy evolution in the dwarf regime and compare our results to what is known in massive galaxies. We do this by studying morphology and structure in complete, unbiased samples of dwarf galaxies in the nearby Universe using deep photometric data in the COSMOS field. Second, we develop algorithms that employ unsupervised machine learning on datasets that are similar to those from new and forthcoming surveys (e.g. LSST) and use them to classify galaxies across the dwarf and massive galaxy regimes.

In Chapter 2, we quantify the morphological mix of dwarf galaxies in the local Universe ($z < 0.08$) as a function of stellar mass, colour, interaction rate and environment. We use both visual inspection and morphological parameters such as CAS, M_{20} and the Gini coefficient to quantify dwarf morphology and highlight differences between the physical processes that drive galaxy evolution in dwrafs and massive galaxies using our results. Finally, we investigate how reliably traditional morphological classification techniques such as CAS, M_{20} and Gini can discriminate between visually classified early and late type galaxies in the dwarf regime.

In Chapter 3, we continue our study from Chapter 2 by investigating the structural properties of dwarfs, as a function of stellar mass and morphology: effective radii, effective surface brightnesses and colour profiles and gradients. We combine our results with those from Chapter 2 to further investigate the differences in the evolutionary histories of dwarfs and massive galaxies.

In Chapter 4, we present an unsupervised machine learning technique designed for large scale galaxy morphological classification. We demonstrate how the algorithm classifies galaxies in the Hyper Suprime-Cam surveys, estimate the accuracy of these classifications and investigate the capability of the algorithm to detect rare or peculiar objects.

In Chapter 5, we continue the analysis of a population of low mass blue elliptical galaxies at $z \sim 0.2$ found using the algorithm presented in Chapter 4. By studying their interaction rates and environments we explore the relative role of interactions and secular processes in their evolution.

Finally, in Chapters 6 and 7 we end with a summary of this thesis and a plan for future work.

Chapter 2

The morphological mix of dwarf galaxies in the nearby Universe

2.1 Introduction

Morphology is a fundamental parameter in observational astrophysics, which is strongly correlated with the evolution of galaxies. In the massive-galaxy ($M_\star > 10^{9.5}$ $M_\odot$) regime, the principal morphological classes are well described by the classical 'Hubble sequence' (Hubble, 1936): 'early-type' galaxies (ETGs), which include S0 and elliptical systems, and 'late-type' galaxies (LTGs), which include different types of disc (spiral) galaxies. Many studies have shown that, in massive galaxies, these morphological types differ not only in their structural characteristics but also in their physical properties, such as colour, star formation rate (SFR) and stellar mass (e.g. Conselice et al., 2014; Martin et al., 2018a; Sampaio et al., 2022). ETGs, which are more pressure-supported than LTGs, typically exhibit low SFRs, red colours and high stellar masses (e.g. Bernardi et al., 2003; Kaviraj et al., 2007; Cappellari et al., 2013). LTGs, which are largely supported by rotation (and exhibit spiral arms as a result), typically exhibit higher SFRs, blue colours and lower stellar masses.

Morphology encodes key information about the dominant physical processes that have shaped the formation of galaxies. The general consensus, in the massive-galaxy regime, is that LTGs largely form via secular cold gas accretion (White & Rees, 1978; Fall & Efstathiou, 1980) and experience either a relatively quiet history (e.g. Jackson et al., 2022) or a larger fraction of gas-rich (and mainly prograde) mergers that helps maintain their discy morphologies and relatively high SFRs over their lifetimes (e.g. Martin et al., 2018a). ETGs, on the other hand, are thought to be the end products of multiple minor or

major mergers (e.g. Martin et al., 2018a) and, in high density environments, also experience processes that promote gas depletion, such as tidal interactions (Moore et al., 1998; Jackson et al., 2021a) and ram pressure stripping (Walker et al., 1996; Martin et al., 2019; Jackson et al., 2021a).

A wide array of methods have traditionally been used to determine the morphology of galaxies (where some notable innovative studies include the work done in e.g. de Vaucouleurs, 1959; Sandage, 1961; Buta & Combes, 1996). Direct visual inspection of galaxy images, whilst the most time-consuming, produces accurate classifications (e.g. Lintott et al., 2011; Kaviraj, 2014a), against which automated methods are often calibrated. The classification of galaxies in large surveys typically requires automation, which is often achieved using both parametric and non-parametric methods. A commonly used parametric method for distinguishing between early and late-type galaxies is fitting Sérsic indices (Sérsic, 1963) using galaxy surface brightness profiles (e.g. Bottrell et al., 2019). In massive galaxies, the steeper profiles of ETGs results in them having larger values of the Sérsic index ($n > 2.5$), while LTGs typically exhibit indices lower than this value (e.g. Shen et al., 2003).

While parametric techniques require the assumption of a fitting function (such as a Sérsic profile), many non-parametric methods have been successfully used to measure morphology from survey data. The pioneering studies of Conselice (2003, C03 hereafter), Abraham et al. (2003) and Lotz et al. (2004, L04 hereafter) introduced five such morphological parameters which have been widely used in galaxy morphological classification in the literature: concentration, asymmetry and clumpiness (known collectively as the 'CAS' system), M_{20} and the Gini coefficient (e.g. Pearson et al., 2021). Many studies that probe optical and near-infrared images, particularly in the nearby Universe where galaxies are well resolved, show that massive ETGs exhibit higher concentration, lower asymmetry and lower clumpiness than their LTG counterparts (e.g. Conselice, 2003; Holwerda et al., 2014; Cheng et al., 2021).

Mergers and tidally-disturbed systems typically show the highest values of asymmetry, while galaxies undergoing starbursts typically show the highest values of clumpiness (e.g. Conselice, 2003). The Gini – M_{20} space in the optical wavelengths offers a reliable discriminant, at least in the massive-galaxy regime, both between early and late type galaxies and between interacting systems and their non-interacting counterparts (e.g. Lotz et al., 2004, 2008). Similar results have been found in the near-infrared (Holwerda et al., 2014), which is less sensitive to dust obscuration (and therefore a more reliable indicator of the underlying structure of the system), using galaxies from the local Universe in the Spitzer Survey of Stellar Structure in Galaxies (S^4G) (Sheth et al., 2010). Forthcoming 'Big Data' surveys, from instruments like the Rubin Observatory's LSST (Ivezić et al., 2019) and Euclid (Laureijs et al., 2010), will present unique challenges for morphological classification work due to their unprecedented data

volumes. Machine learning techniques (e.g. Hocking et al., 2018; Martin et al., 2020; Walmsley et al., 2022; Pearson et al., 2022; Domínguez Sánchez et al., 2023), used in conjunction with visual inspection and morphological parameters, are therefore likely to become important for the next generation of surveys.

While a rich literature already exists on galaxy morphology in the massive-galaxy regime (e.g. Bílek et al., 2020; Fraser-McKelvie & Cortese, 2022; Nersesian et al., 2023), much less is known about the dwarf regime ($M_\star < 10^{9.5}$ $M_\odot$) outside the very local Universe. For example, are the principal visual morphological classes and their formation histories in dwarf galaxies similar to what is seen in their massive counterparts? How frequent are morphological details like bars and interactions? How do commonly-used morphological parameters perform in the dwarf regime in separating morphological classes and identifying interacting systems?

The current lack of statistical knowledge about dwarf morphology is partly driven by the fact that, while dwarfs can be studied in detail in the very local Universe, up to distances of $\sim$50 Mpc (e.g. Mateo, 1998; Tolstoy et al., 2009; Besla et al., 2016), typical dwarfs are too faint to be detected, at cosmological distances, in past wide-area surveys like the SDSS (e.g. Jackson et al., 2021a; Davis et al., 2022). This is because, while these surveys provide large sky areas, they have relatively shallow detection limits. The dwarfs that are visible in large shallow surveys of the past tend to be those which have relatively high SFRs. The high star formation boosts the luminosity of the dwarfs above the detection thresholds of shallow surveys, making them detectable (e.g. Jackson et al., 2021a). However, this also biases these dwarf samples towards systems that are blue and therefore more likely to have late-type morphology (Kaviraj et al. in prep). Unbiased morphological studies in the dwarf regime have so far been possible, either in our local neighbourhood (e.g. Tolstoy et al., 2009), in nearby groups and clusters (e.g. Boselli et al., 2008; Venhola et al., 2017; Eigenthaler et al., 2018) or around nearby massive galaxies (e.g. Duc et al., 2015; Geha et al., 2017; Carlsten et al., 2021; Mao et al., 2021; Trujillo et al., 2021; Holwerda et al., 2023), regions where relatively complete dwarf samples can be assembled. Exploring the morphological mix of dwarfs in the general Universe, in low-density environments, requires surveys which are both deep and wide, like the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP; Aihara et al., 2018), in which dwarf galaxy populations are likely to be complete, down to $M_\star \sim 10^8$ $M_\odot$ out to at least $z \sim 0.3$.

It is worth noting here that recent work has shown that typical morphological classes, such as ETGs and LTGs, behave differently in the dwarf regime compared to the massive-galaxy regime. For example,

Lazar et al. (2023) use ultra-deep HSC-SSP imaging, which has a point source depth of ∼28 magnitudes[1] (more than 5 magnitudes deeper than standard-depth SDSS imaging) of over 100 blue dwarf ellipticals to show that less than 3 per cent of these systems show signs of tidal interactions, lower than the interaction fraction seen in the general galaxy population at similar stellar masses. This is in contrast to the massive-galaxy regime, in which more than 70 per cent of ellipticals show tidal features in deep images (e.g. van Dokkum, 2005).

The past literature shows that, in the high mass regime, ETGs typically exhibit Sérsic indices (n) around 4, while $n \sim 1$ corresponds to LTGs (de Vaucouleurs, 1959, 1977; De Propris et al., 2016). However, some studies indicate that these trends may not be preserved in the dwarf regime. For example, Conselice (2003), who study galaxy morphology in the local Universe (out to ∼50 Mpc) using optical data, show that concentration cannot be used to separate the ETG and LTG populations in the dwarf regime. Such results suggest that, rather than being strongly shaped by interactions, which would increase the central concentration and result in tidal features, ETGs in the dwarf regime may evolve preferentially via secular processes like steady gas accretion over large (i.e. Gyr) timescales. More broadly, the evolutionary histories of traditional morphological types (e.g. ETGs and LTGs) in the dwarf regime could potentially be different from those in the massive-galaxy regime.

A detailed census of the morphological mix of dwarf galaxies in the general Universe, outside the local neighbourhood and in low-density environments, is largely missing and clearly desirable for establishing the morphological trends in the dwarf population. Such a census, which requires an exploration using surveys that are both deep and wide, is the purpose of this study. This chapter is organized as follows. In Section 2.2, we describe the datasets used in this chapter and the selection of a sample of nearby dwarf galaxies that underpins this study. In Section 2.3, we explore the morphological mix of our dwarfs via visual inspection of their HSC *gri* composite images and their unsharp-masked counterparts, identify the principal morphological classes in the dwarf regime, explore the role of morphological details like bars and interactions and study their recent star formation histories via their rest-frame colours. We also compare our findings in the dwarf regime to what is known in massive galaxies. In Section 2.4, we explore dwarf morphology using commonly-used morphological parameters, such as the Sérsic index, concentration, asymmetry and clumpiness (CAS), M_{20} and the Gini coefficient, and compare the distributions of these parameters in dwarfs to those in their massive counterparts in the nearby Universe. In Section 2.5, we bring our results together and discuss the implications of our findings for the evolution of different dwarf morphological classes. We summarise our findings in Section 2.6.

[1]The i-band 3σ limiting surface-brightness (measured in a 10×10 pixel patch) is around 31 mag arcsec^{-2}.

2.2 Data

Our study is underpinned by the Classic version of the COSMOS2020 catalogue (Weaver et al., 2022), which provides physical parameters, such as photometric redshifts, stellar masses and SFRs for $\sim$1.7 million sources in the $\sim$2 deg^2 COSMOS field (Scoville et al., 2007). These parameters are calculated using deep photometry from 40 broadband and narrowband filters, across the UV through to the mid-infrared, from the following instruments: GALEX (Zamojski et al., 2007), MegaCam/CFHT (Sawicki et al., 2019), ACS/HST (Leauthaud et al., 2007), Hyper Suprime-Cam (Aihara et al., 2019), Subaru/Suprime-Cam (Taniguchi et al., 2007, 2015), VIRCAM/VISTA (McCracken et al., 2012) and IRAC/Spitzer (Ashby et al., 2013; Steinhardt et al., 2014; Ashby et al., 2015, 2018). The optical and infrared aperture photometry is extracted using the SExtractor and IRACLEAN codes respectively and physical parameters are calculated using the LePhare SED-fitting algorithm (Arnouts et al., 2002; Ilbert et al., 2006). The 40-filter photometry results in accurate parameters, with photometric redshift accuracies better than $\sim$1 and $\sim$4 per cent for bright ($i < 22.5$ mag) and faint ($25 < i < 27$ mag) galaxies respectively. Our visual inspection uses optical *gri* colour composite images from the HSC-SSP Ultra-deep layer, which have a 5σ point source depth of $\sim$28 magnitudes. This is around 5 magnitudes deeper than standard depth SDSS imaging and almost 10 magnitudes deeper than the detection limit of the SDSS spectroscopic main galaxy sample. The median seeing of the HSC images is $\sim$0.6 arcseconds (around a factor of 2 better than the SDSS).

2.2.1 A complete sample of nearby dwarf galaxies

To construct our dwarf galaxy sample, we first select objects which are classified as galaxies by LePhare ('type' = 0 in the COSMOS2020 catalogue). We then restrict our study to galaxies which have stellar masses in the range 10^8 M$_\odot$ $< M_\star <$ $10^{9.5}$ M$_\odot$ and redshifts in the range $z < 0.08$ and which lie within the HSC-SSP footprint and outside the HSC-SSP bright-star masks defined by Coupon et al. (2018). We restrict ourselves to $z < 0.08$ because visual inspection of these dwarfs via their HSC images suggests that morphological classification becomes more difficult beyond this redshift. The sample obtained after applying these constraints is comprised of 283 dwarf galaxies.

We then further exclude 26 objects that either have an 'extendedness' of 0 in the HSC *griz* filters (i.e. are classified as stars via this parameter, even though they are classified as galaxies by LePhare[2]) or, after visual inspection, either appear to have low signal to noise or are objects where a large foreground

[2]These objects clearly look like stars when visually inspected (see Figure 2.6).

or background galaxy is contaminating the photometry of the object in question. The final sample we use for our analysis contains 257 dwarf galaxies. The median redshift and stellar mass errors in our dwarf sample are 0.02 and 0.08 dex, respectively. Note that $z \sim 0.08$ is much lower than the redshift out to which COSMOS2020 is expected to be complete ($z \sim 0.3$), for galaxies that have stellar masses of $M_\star > 10^8$ $M_\odot$ (see e.g. Figure 1 in Jackson et al., 2021a). This completeness range is also confirmed by Kaviraj et al. (2024; submitted) who investigate the completeness of dwarf galaxies in the HSC survey by considering a hypothetical object with a purely old simple stellar population (SSP) at low redshift that forms in a star formation burst at a redshift of 2. Even though such object does not exist in reality, the authors use this object as the faintest limiting case which if it is brighter than the detection limit of the survey then all galaxies brighter than this hypothetical object will be detectable. Figure 2.1 (reproduced with permission from Kaviraj et al. (2024), submitted to MNRAS) shows redshift estimates for 2 different metallicity assumptions (0.5 $Z_\odot$ and 0.05 $Z_\odot$ shown by the dotted and solid lines, respectively) for the SSPs built using the Bruzual & Charlot (2003) stellar models, at which a survey with a given detection limit would be complete as a function of stellar mass. For the survey used in this work, which is HSC-SSP UDEEP the completeness limit is shown by the black solid and dotted lines, where all galaxy samples with lower redshifts and higher masses than the black lines are complete.

The sample of dwarfs used here therefore offers an unbiased statistical sample of galaxies which can be used to study the morphological properties of the general dwarf population in the nearby Universe.

It is instructive to consider the types of large-scale structures that are actually present in the COSMOS2020 footprint at the redshifts probed in this study. Using the X-ray group catalogues compiled by Finoguenov et al. (2007), George et al. (2011) and Gozaliasl et al. (2019), which incorporate HST-ACS data from Leauthaud et al. (2007) and photometric redshifts from Ilbert et al. (2009), we check the M_{200} values of groups that reside within the COSMOS2020 footprint at $z < 0.1$. The virial masses of the three groups that fit this description lie in the range $10^{12.9}$ $M_\odot < M_{200} < 10^{13.2}$ $M_\odot$. In comparison, a small cluster like Fornax has a virial mass of $\sim 10^{13.9}$ $M_\odot$ (Drinkwater et al., 2001), while larger clusters like Virgo and Coma have virial masses of $\sim 10^{15}$ $M_\odot$ (e.g. Fouqué et al., 2001; Gavazzi et al., 2009). The COSMOS2020 galaxy population, in our redshift range of interest, therefore resides preferentially in low-density environments.

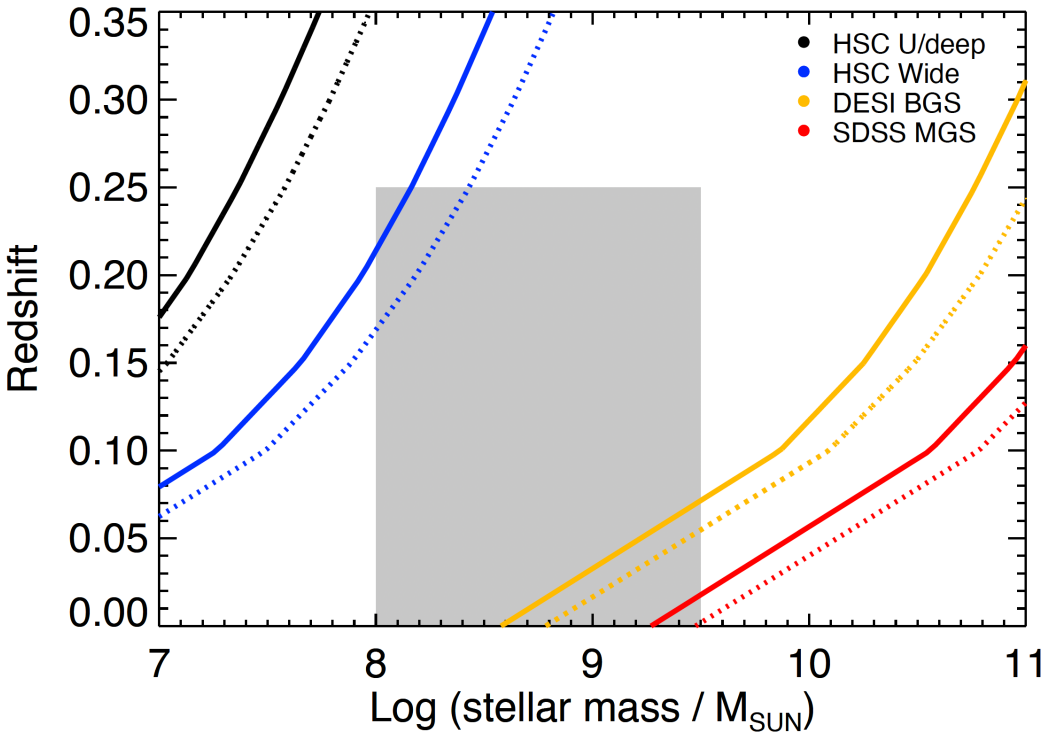

FIGURE 2.1: Redshift completeness estimates as a function of stellar mass. The solid and dashed lines represent redshift estimates for the purely old SSPs with metallicity assumptions of 0.05 $Z_{\odot}$ and 0.5 $Z_{\odot}$, respectively. The black lines correspond to the detection limit of HSC UDEEP (used in this work). The blue, yellow and red lines are associated with the detection limit of HSC WIDE, DESI BGS and SDSS MGS, where SDSS presents the brightest detection limit among the surveys presented in this plot. This Figure has been reproduced with permission from Kaviraj et al. (2024), submitted to MNRAS.

2.3 Morphological classifications via visual inspection

We use visual inspection of HSC *gri* colour composite images to classify our dwarf galaxies into their principal morphological classes. In order to generate the red-green-blue composite images we make use of the Python function `make_lupton_rgb` (described in Lupton et al., 2004) from the python library `astropy`. Unlike massive galaxies, dwarfs are intrinsically faint and their internal structures (disks, bars, tidal features etc.) may have low contrast in our images. Therefore, for each galaxy, we also create and visually inspect an 'unsharp masked' version of each colour composite image. In unsharp masking (e.g. Malin, 1977), a blurred image is created by convolving the original with a kernel. The blurred version is subtracted from the original and the difference image is then multiplied by a factor which represents the strength of the sharpening. The kernel size and sharpening strength are the free parameters in this process. The technique has the effect of sharpening the edges of structures in a galaxy. Given that the sizes of structures may vary in different galaxies, and since our dwarfs span a range (albeit small) in redshift, we explore multiple values for both the blurring kernel and the sharpening strength in each

galaxy[3].

Unsharp masking has previously been used in astronomy to detect faint, low-contrast features like shells and tidal features inside and around nearby massive galaxies (e.g. Malin & Carter, 1983). Since they are intrinsically bright, the internal constituents of massive galaxies (spiral arms, bars etc.) are readily visible in contemporary surveys without unsharp masking. However, the same may not be true for dwarfs, which are relatively fainter and unsharp masking could increase the efficacy of our morphological classifications. The visual classifications were performed by one expert classifier (Sugata Kaviraj). The images were randomised, both the original and unsharp-masked images of each galaxy were classified at the same time and physical parameters (e.g. stellar mass and redshift) were kept hidden during the classification process to avoid introducing any biases.

Finally, visual inspection is performed after applying unsharp masking on HSC *gri* colour composite images. We classify galaxies with a disk-like spiral structure as late type galaxies and objects with a diffuse structure and high central light concentration as early type galaxies. Galaxies which present a diffuse structure but with a lack of central light concentration are classified as featureless galaxies and objects which present disturbed/irregular morphology are classified as irregular. We also assign a flag for each galaxy if it presents a bar. A separate flag (signalling an interaction) is used if the objects show signs of ongoing or recent interactions such as tidal features, visible large-scale internal asymmetry, shells and dust lanes.

2.3.1 Principal morphological classes in dwarf galaxies

We begin by presenting, in Figures 2.2, 2.3 and 2.4, the three principal morphological classes yielded by our visual inspection: ETGs, i.e. ellipticals and S0s (Figure 2.2), LTGs (Figure 2.3) and featureless dwarfs (Figure 2.4). ETGs and LTGs in the dwarf regime are similar to the classical morphological types known from studies of massive galaxies. While we do not attempt to divide our LTGs into more granular subclasses, our dwarf LTGs are typically akin to the 'Sc' or 'Sd' morphologies seen in the massive-galaxy population. The featureless dwarfs, on the other hand, have flat, smooth profiles that lack either the central light concentration that typifies ETGs or any disk structure that is found in LTGs. Not only that the featureless type lacks a central light concentration but, as we show in Chapter 3, compared to the ETG class, it has similar effective radii but significantly fainter effective surface brightness values for a fixed stellar mass, which signals that the distribution of stellar mass within these galaxy types is

[3]We use the interactive unsharp mask filter in the GIMP software package to interactively vary the kernel and the strength in individual dwarf galaxy images.

different. A very small number of objects (less than 2 per cent) have a mostly irregular morphology that does not fit into these classes. Examples of these objects are shown in Figure 2.5. Given their small contribution to the overall number of dwarfs, these galaxies are omitted from the analysis below for clarity.

Finally, in Figure 2.6, we show examples of the 26 objects that are excluded from our analysis (as described in Section 2.2.1), either because they are classified as stars by the HSC extendedness parameter (even though they are classified as galaxies by LEPHARE) or because they are contaminated by nearby massive galaxies or are low signal to noise objects. Table 2.1 summarises both the number fractions of dwarfs in these different morphological classes (upper sub-table) and the fractions of galaxies in each morphological class which show evidence for ongoing or recent interactions (lower sub-table). The first column (presented in bold) indicates fractions for the entire dwarf sample, while the second and third columns show values for the lower and upper halves of our stellar mass range respectively. Throughout this work uncertainties are calculated following Cameron (2011).

We note that the featureless class strongly resembles what in the nearby Universe are commonly referred to as dwarf spheroidals (dSph). Indeed, similar galaxies were discovered near the Milky Way as early as the 1930s (Shapley, 1938), and in the Virgo Cluster as early as the 1950s (Reaves, 1956; van den Bergh, 1959). Reaves (1956) referred to those in Virgo as IC 3475-type objects, and described them as having "no trace of any small central nucleus or arm structure", with "relatively slight concentration of light to the center, if any", which is extremely similar to our description of the featureless galaxies found in this study. These featureless dwarfs thus may be dSph-type galaxies.

In this context, it is also instructive to explore the connection between the featureless class and the population of 'ultra-diffuse galaxies' (UDGs, e.g. van Dokkum et al., 2015; Koda et al., 2015), which are a subset of the dSph population (Conselice, 2018) and have been prominent in the recent literature. Typical UDGs have effective surface brightnesses ($\mu_{\rm eff}$) fainter than $\sim$25 mag arcsec^{-2} and effective radii ($R_{\rm eff}$) larger than 1.5 kpc (e.g. Conselice, 2018). However, while 19 out of 24 galaxies in our featureless class have $R_{\rm eff}$ larger than 1.5 kpc, virtually all of these galaxies have $\mu_{\rm e}$ brighter than 25 mag arcsec^{-2}. Only two of our featureless galaxies are consistent with the definition of a UDG (if either r or i band images are considered). Thus, while the featureless objects are likely related to dSphs, and could be brighter versions of UDG-like systems, we choose to call them featureless for now, until we can conduct more detailed comparisons of their structural properties with dSph galaxies using larger datasets.

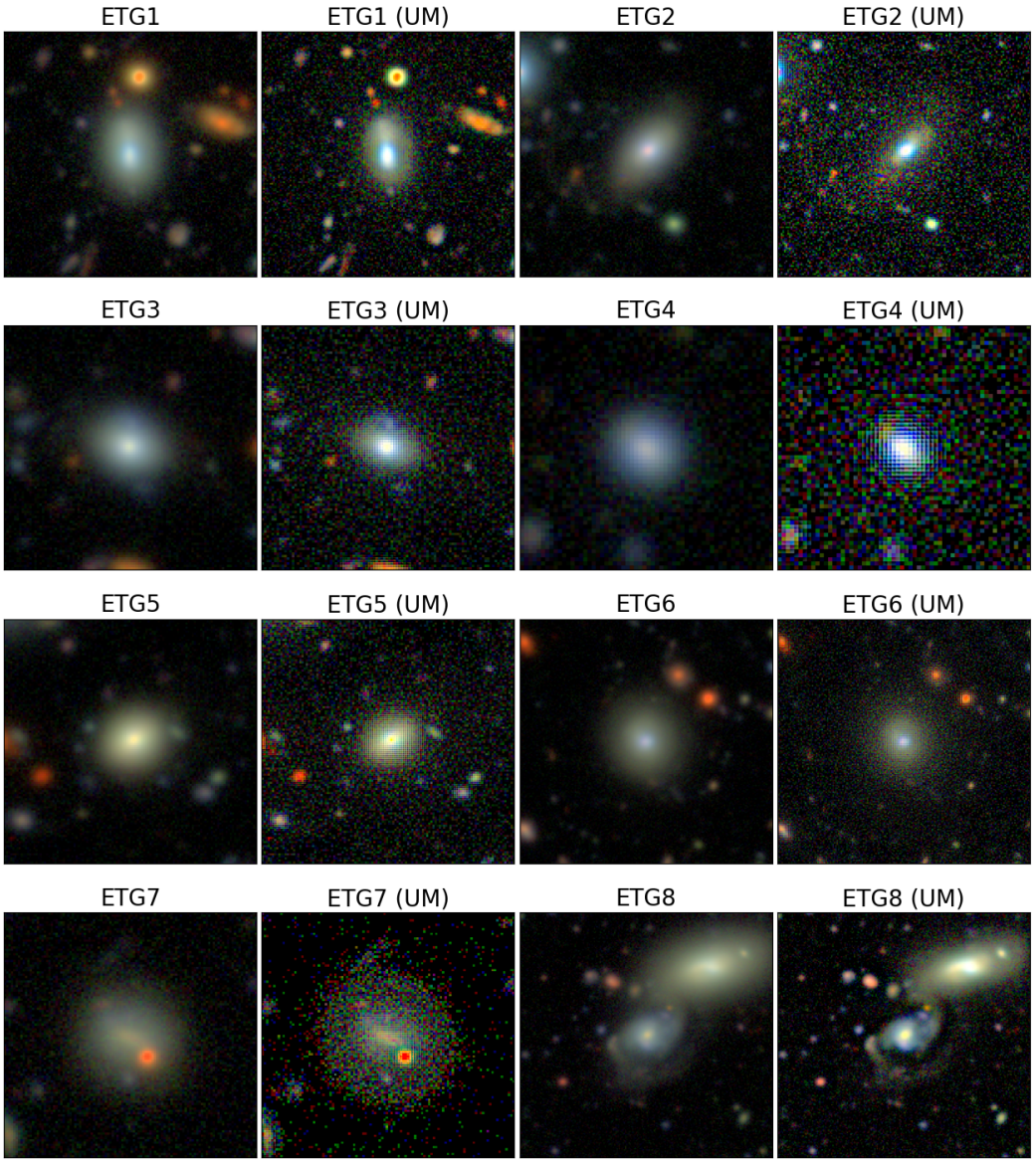

FIGURE 2.2: Examples of ETGs in our dwarf sample. For each galaxy the left-hand panel shows the HSC colour image, while the right-hand panel shows its unsharp-masked counterpart. ETG1 is an example of a system which shows an internal asymmetry. ETG2 shows a faint tidal feature to the north east of the galaxy. ETG7 is the only dwarf ETG in our sample that exhibits a dust lane which also contains a foreground star to south-west, while ETG8 shows an interacting system with two noticeable tidal tails to the north and south of the central galaxy. For ETG8 the tidal tail to the south seems more extended than the one to the north, containing more diffuse flux to a larger extent around the central galaxy in an arc-like pattern towards the larger companion. As the images suggest, blue cores and regions are common in dwarf ETGs whether they are interacting or not.

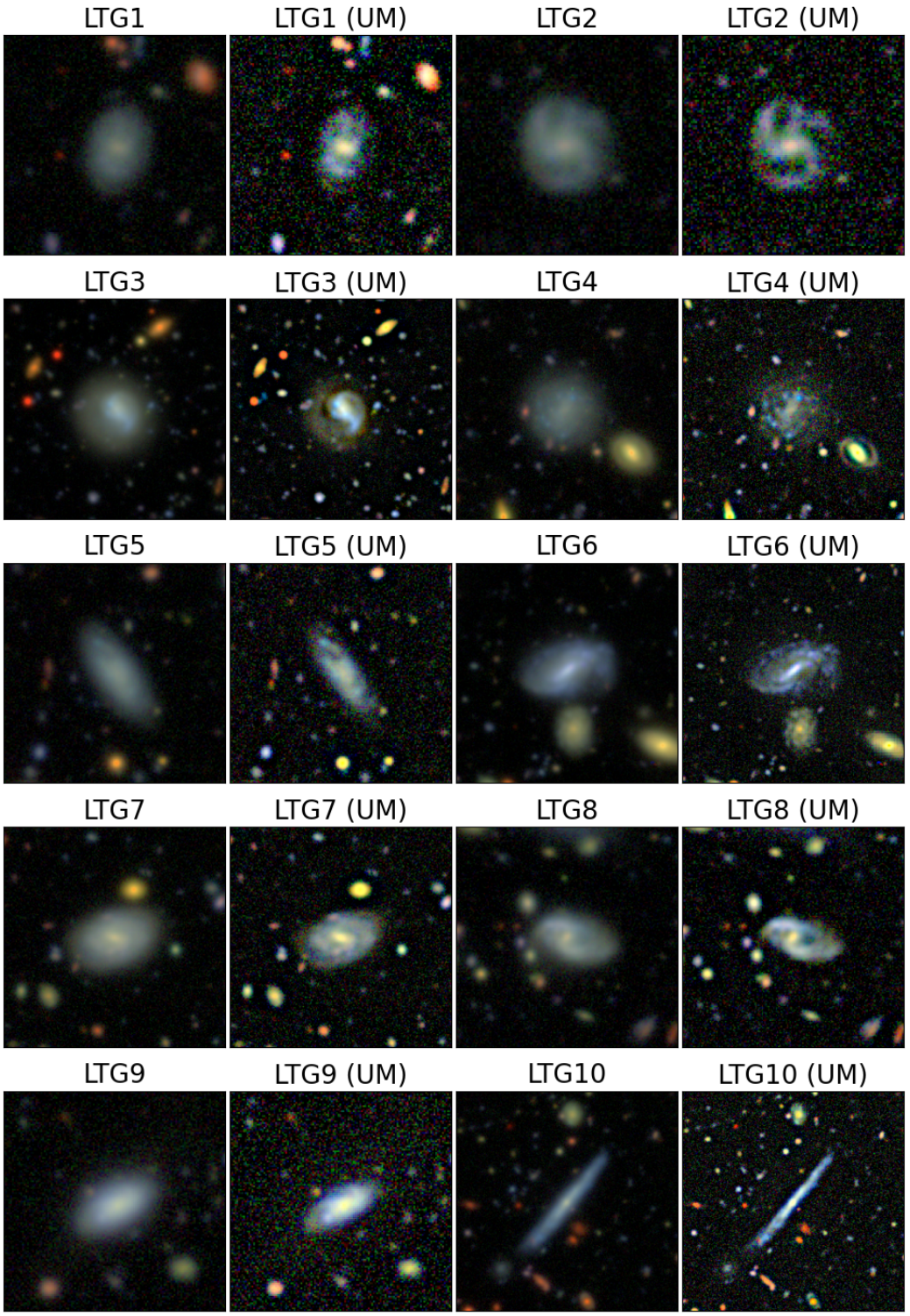

FIGURE 2.3: Examples of LTGs in our dwarf sample. For each galaxy the left-hand panel shows the HSC colour image, while the right-hand panel shows its unsharp-masked counterpart. LTG3 shows an example of a system which shows an internal asymmetry, while LTG6 shows an example of a system with a tidal feature.

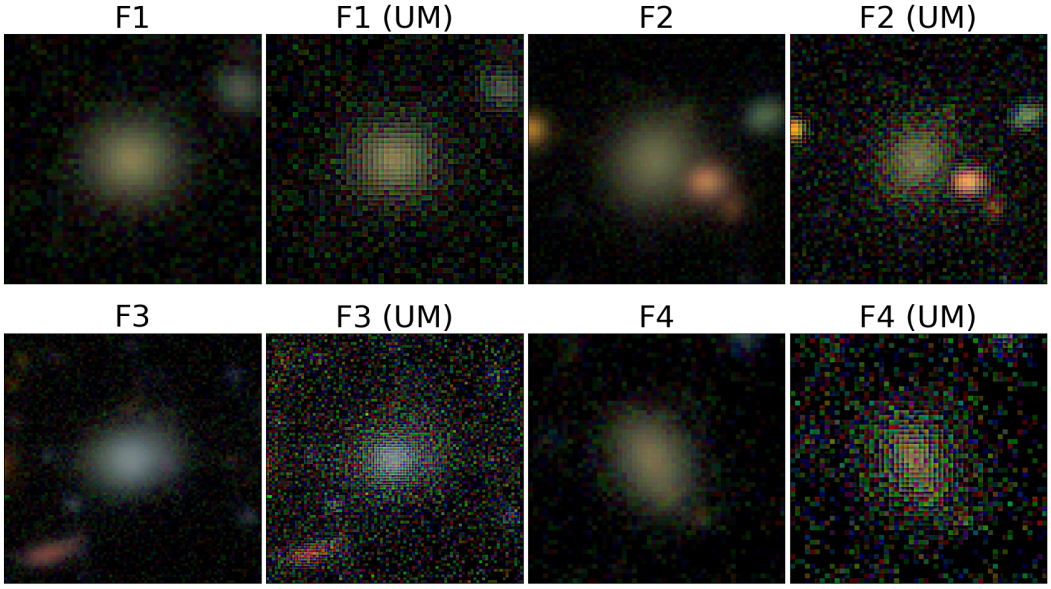

FIGURE 2.4: Examples of dwarf galaxies classified as featureless. For each galaxy the left-hand panel shows the HSC colour image, while the right-hand panel shows its unsharp-masked counterpart.

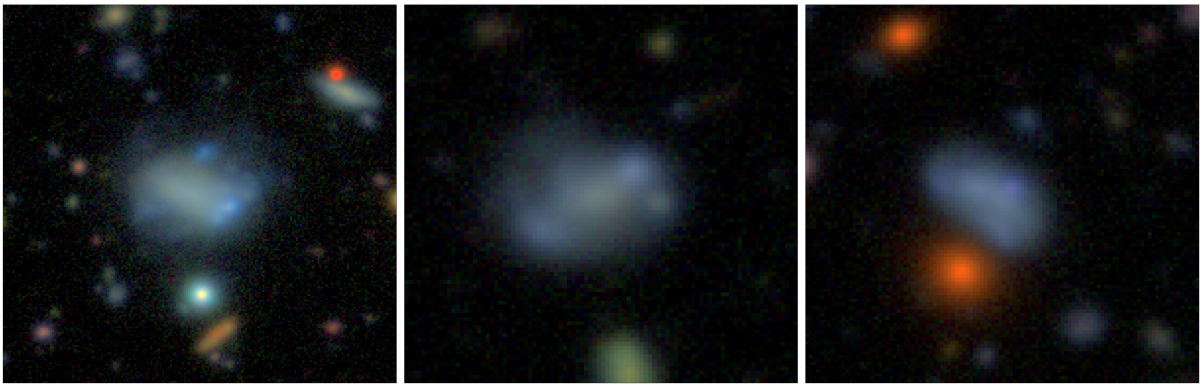

FIGURE 2.5: Examples of dwarf galaxies which have largely irregular morphology and are difficult to put into the ETG, LTG and featureless categories above. The fraction of such objects in our sample is negligible (less than 2 per cent).

The number fractions of ETGs and LTGs in dwarfs in the general Universe are reasonably similar, although there is an excess of late-types at higher stellar masses. This is different from the situation in high-density environments, reported in other work, where the fraction of dwarf ETGs tends to be higher. For example, the dwarf ETG fraction measured in the MATLAS survey, which samples dwarf satellites around nearby massive galaxies, is $\sim$ 70 per cent (Poulain et al., 2021). Similar results have been obtained by Carlsten et al. (2021), who performed a comparative study of ETGs and LTGs in the ELVES (Carlsten et al., 2022), NGVS (Ferrarese et al., 2012) and NGFS (e.g. Muñoz et al., 2015) surveys, which probe dwarfs in the local volume, Virgo and Fornax respectively. These trends suggest that the likelihood of dwarfs having early-type morphology increases in regions of higher density, akin to the morphology-density relation seen in the massive-galaxy regime (e.g. Oemler, 1974; Dressler, 1980;

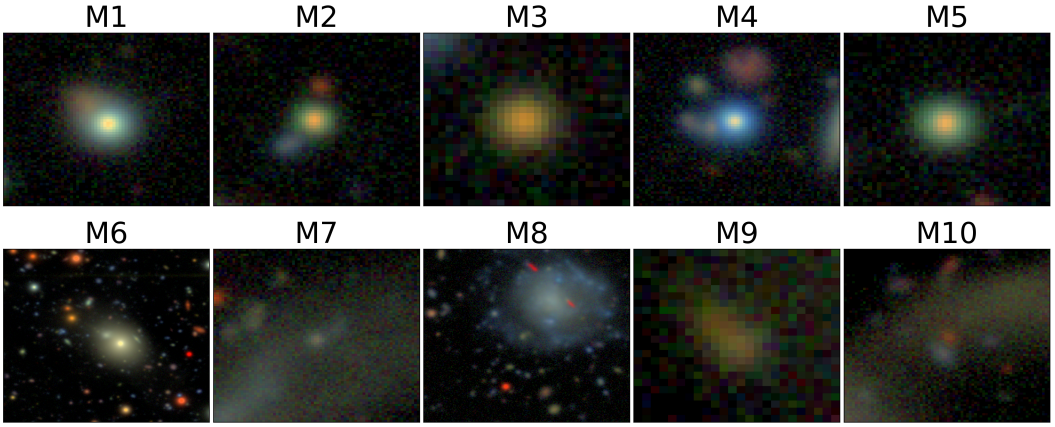

FIGURE 2.6: Examples of objects that are misclassified and removed from our analysis. These include stars (M1 - 5) which have HSC extendedness values of 0, i.e. are classified as stars by the HSC pipeline even though they are classified as galaxies by LEPHARE. Note that these objects clearly look like stars in the images. M6 and M7 are likely to be contaminated by the light from an overlapping massive galaxy. M8 appears to be an object embedded within the outskirts of a massive galaxy. M9 is a very low S/N object, while M10 is likely to be a high-redshift background galaxy residing behind a broad tidal feature from a nearby galaxy.

Number fractions in different morphological classes

	$\mathbf{10^8\ M_\odot < M_\star < 10^{9.5}\ M_\odot}$	$10^8\ M_\odot < M_\star < 10^{8.75}\ M_\odot$	$10^{8.75}\ M_\odot < M_\star < 10^{9.5}\ M_\odot$
dETG	$\mathbf{0.43^{0.03}}$	$0.45^{0.03}$	$0.37^{0.06}$
dLTG	$\mathbf{0.45^{0.03}}$	$0.40^{0.03}$	$0.63^{0.06}$
dF	$\mathbf{0.10^{0.02}}$	$0.12^{0.02}$	-
dIrr	$\mathbf{0.02^{0.01}}$	$0.03^{0.01}$	-

Interaction fractions in different morphological classes

	$\mathbf{10^8\ M_\odot < M_\star < 10^{9.5}\ M_\odot}$	$10^8\ M_\odot < M_\star < 10^{8.75}\ M_\odot$	$10^{8.75}\ M_\odot < M_\star < 10^{9.5}\ M_\odot$
dETG	$\mathbf{0.14^{0.03}}$	$0.15^{0.04}$	$0.10^{0.06}$
dLTG	$\mathbf{0.28^{0.04}}$	$0.24^{0.04}$	$0.35^{0.07}$
dF	$\mathbf{0.20^{0.07}}$	$0.20^{0.07}$	-
dIrr	$\mathbf{0.40^{0.16}}$	$0.40^{0.16}$	-

TABLE 2.1: **Upper sub-table:** Number fractions of dwarf galaxies in each morphological class (dETG = dwarf early-type galaxies, dLTG = dwarf late-type galaxies, dF = dwarf featureless galaxies, dIrr = dwarf irregular galaxies), in three mass ranges. The left-hand column shows the number fractions for the full mass range in our sample (10^8 $M_\odot$ < $M_\star$ < $10^{9.5}$ $M_\odot$). In the middle and right-hand columns we split the mass range into its lower and upper halves respectively. The errors, which are shown as superscripts, are calculated following Cameron (2011) and represent the lower and upper 68 per cent confidence levels of a Beta probability distribution (suitable to model the random behaviour of proportions and percentages). **Lower sub-table:** The fraction of galaxies in a given morphological class which shows evidence of interactions (see text in Section 2.3 for details). The dashes indicate instances where there are no galaxies corresponding to a specific morphological type.

Sazonova et al., 2020).

To compare the visual morphological mix of dwarfs to what is known in the massive-galaxy regime, we compare the results of our visual classification to those from Kaviraj (2010), who performed morphological classification, via visual inspection, of ∼1000 massive galaxies in the SDSS Stripe 82. The galaxies in this sample have upper redshift and lower mass limits of $z \sim 0.05$ and $M_\star \sim 10^{10}$ $M_\odot$ respectively. Around 60 per cent of the Kaviraj (2010) sample is composed of LTGs, slightly higher than the late-type fraction in the dwarf regime.

However, the largest difference, in terms of major morphological classes, between dwarfs and massive galaxies is the existence of the featureless galaxies, which do not have an equivalent in the massive-galaxy regime. It is worth noting that all featureless galaxies reside in the lower half of our stellar mass range (see Table 2.1). In other words, the fraction of featureless dwarfs increases towards lower stellar masses, which suggests that their formation becomes easier as the gravitational potential well becomes shallower. This, in turn, indicates that their formation likely depends on processes like baryonic feedback or tidal perturbations, whose effect will become stronger as the potential well becomes shallower (e.g. Jackson et al., 2021a).

2.3.2 Local densities of different dwarf morphological classes

We quantify the relative local densities of the various dwarf morphological classes by calculating projected distances of individual dwarfs from the first to the tenth nearest massive ($M_\star > 10^{10}$ $M_\odot$) neighbours, using a redshift tolerance of 0.02. While the distance to the first nearest neighbour provides a measure of very local density, distances to progressively farther neighbours provide estimates of the ambient densities averaged over larger distances (e.g. Dressler, 1980). Recall that the galaxy population in COSMOS, at the redshifts we are considering here ($z < 0.08$), reside in low-density environments (Section 2.2.1). Our aim here is to understand if there are any strong trends in *relative* density between the different dwarf morphological classes. In Figure 2.7, we show the distances to the nearest, third nearest and tenth nearest massive neighbour. We find no significant differences between the different dwarf morphological classes in terms of their local density, with the median values overlapping within the statistical errors. The same general trends are present in all other distances (not shown for brevity). It is worth noting, however, that there are more featureless dwarfs at distances less than 250 kpc from the nearest massive galaxies than in other morphological classes (top panel of Figure 2.7), although this trend has to be confirmed using larger samples of dwarfs in future work.

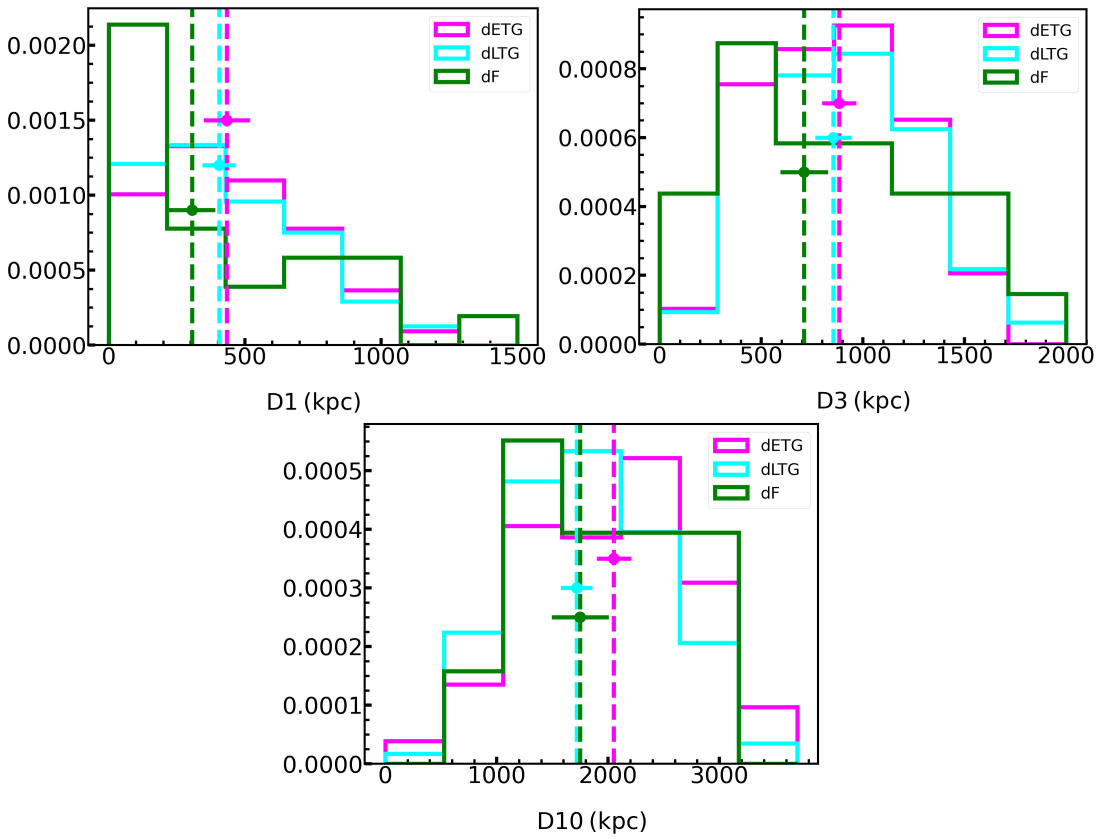

FIGURE 2.7: Distributions of projected distances to the first (top-left panel), third (top-right panel) and tenth (bottom panel) nearest massive ($M_\star > 10^{10}$ $M_\odot$) neighbour, within a redshift tolerance of 0.02. The histograms are normalized by their areas.

2.3.3 The role of interactions

The interaction fraction in dwarf LTGs (∼28 per cent, Table 2.1) is around a factor of 2 higher than that found in massive LTGs, which exhibit interaction fractions between 11 and 16 per cent, depending on the exact morphological subclass (Kaviraj, 2014a). We note that we define a galaxy to be interacting if it presents tidal features, large-scale internal asymmetries, shells or dust lanes. The higher interaction fraction in dwarf LTGs, compared to their massive counterparts, is likely to reflect the fact that inducing morphological disturbances in dwarfs is easier (e.g. via mergers, flybys around massive galaxies or internal stellar feedback) due to their shallower potential wells (e.g. Agertz & Kravtsov, 2016; Martin et al., 2021).

Some interesting differences between dwarfs and massive galaxies arise when considering the interaction fraction in dwarf ETGs. Dust lanes in massive ETGs are signposts of merger activity with lower mass companions (e.g. Kaviraj et al., 2012; Davis et al., 2015), with around 7 per cent of massive ETGs showing these large-scale dust features (Kaviraj, 2010). However, only 1 dwarf ETG (out of 110) shows a dust lane (see Figure 2.2). The incidence of dust lanes therefore appears to be markedly lower in dwarf

ETGs than in their massive counterparts. In a similar vein, the incidence of interactions in massive ETGs tends to be higher (by around a factor of $\sim$1.5) than in massive LTGs, particularly when compared to the interaction fractions in massive Sc and Sd type systems that are morphologically similar to our dwarf LTGs (see Table 1 in Kaviraj (2014a) and Kaviraj (2010)). However, the dwarf regime does not appear to preserve this trend, with dwarf LTGs showing higher interaction fractions, by around a factor of 2, than their ETG counterparts.

Finally, the interaction fractions in dwarf ETGs are significantly lower than that found in the massive ETGs. For example, while the interaction fraction in dwarf ETGs is $\sim$14 per cent (Table 2.1), it is $\sim$70 per cent (i.e. a factor of 5 higher) in massive red ETGs (van Dokkum, 2005) in images that are around 2 magnitudes shallower than the HSC images used in this study. The interaction fraction in massive galaxies could, therefore, be expected to be even higher in images that have similar depth to the ones used in this study. This suggests that the origin of dwarf ETGs may have less to do with interactions than their massive counterparts, in line with the findings of recent observational work that suggests that secular accretion from the cosmic web is the dominant evolutionary channel for these systems (Lazar et al., 2023). We explore this point in more detail in Section 2.5.1.

When large statistical samples are available, the interaction fraction in a population can be combined with the enhancement of the specific SFR (sSFR) in interacting galaxies to estimate the fraction of the star formation activity which is being driven by interactions (Eq. 2 in Kaviraj, 2014a). Large samples likely span the range of properties (e.g. the mass ratio and coalescence timescale of interacting systems) that may influence the interaction-driven sSFR of the system (e.g. Lotz et al., 2010a,b). Following Kaviraj (2014a), we calculate the enhancement of the sSFR in different dwarf morphological classes by calculating the ratio of the median sSFR in interacting systems to the median sSFR in non-interacting systems. Combining these with the interaction fractions in Table 2.1 suggests that around 27, 19 and 14 per cent of the star formation activity in our dwarf LTG, featureless and ETG populations are driven by interactions respectively. Combining these values, weighted by the number fractions in the different dwarf morphological classes, suggests that around 20 per cent of the overall star formation activity in dwarfs is triggered by interactions. It is interesting to note that the fraction of star formation activity in massive nearby LTGs that is likely triggered by interactions is higher (around 40 per cent, see Kaviraj, 2014a). Together with the low frequency of tidal features seen in the dwarf ETGs, this further suggests that, in general, the evolution of dwarf galaxies may have had less to do with interactions than their massive counterparts and that their evolution is likely to be influenced more by secular accretion and internal processes.

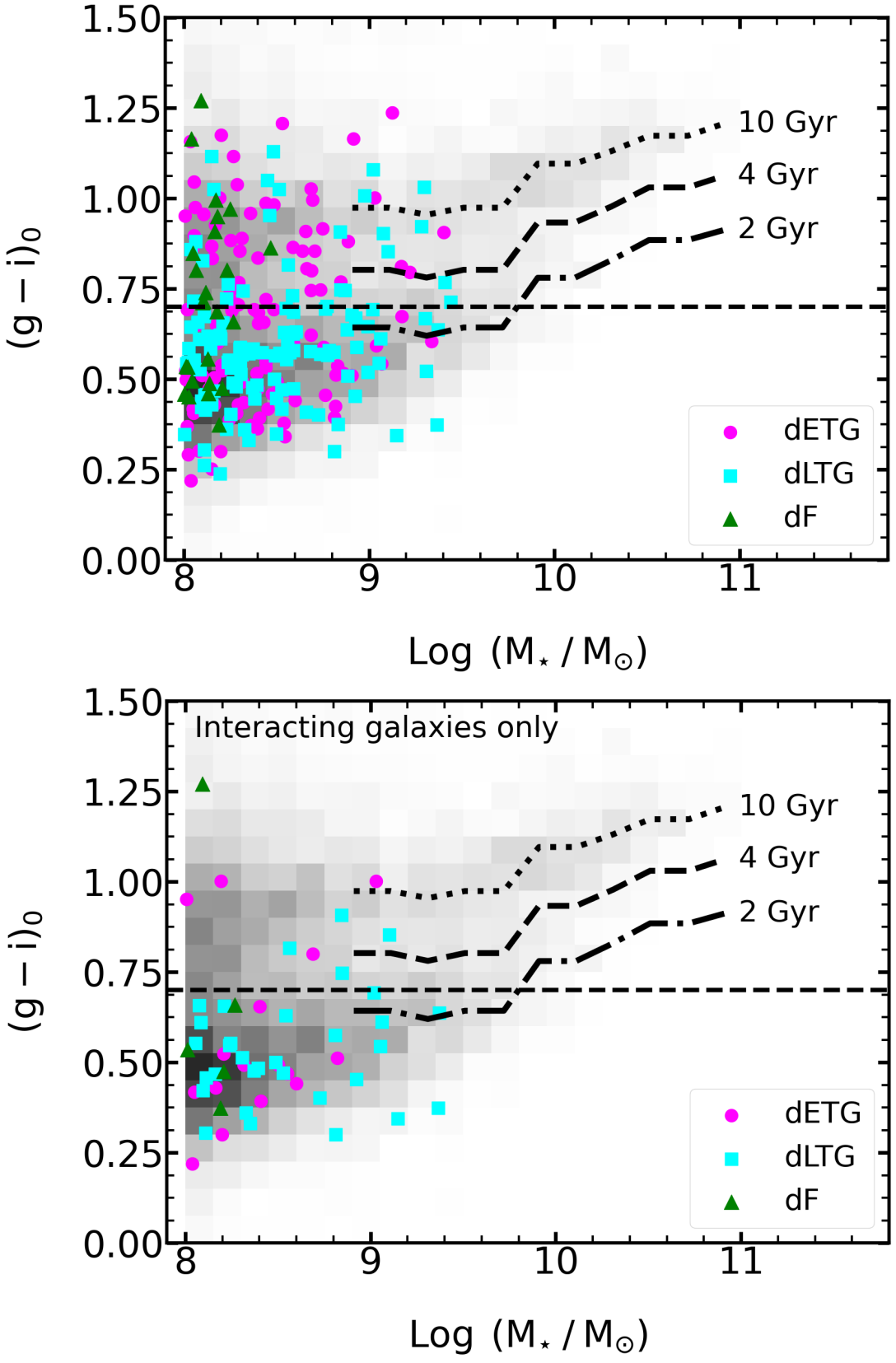

FIGURE 2.8: **Top:** Rest-frame $(g-i)$ colours vs stellar mass of COSMOS2020 galaxies at $z < 0.2$, shown as a heatmap. The dashed lines show simple stellar populations (SSPs) which form at various look-back times (10, 4, and 2 Gyrs). We use the median metallicity as a function of stellar mass Gallazzi et al. (2005) to construct the rest-frame colour for all SSPs. The three principal dwarf morphological classes are shown overplotted in different colours (see legend). **Bottom:** Same as the top panel but for interacting galaxies only. The horizontal dashed line at rest-frame $(g-i) = 0.7$ separates red and blue galaxies.

2.3.4 Rest-frame colours

In Figure 2.8, we explore the rest-frame colours of the different dwarf morphological classes. The top panel shows the rest-frame $(g-i)$ colour vs stellar mass of the COSMOS2020 population at $z < 0.2$ as a heatmap. Galaxies in our dwarf morphological classes are shown overplotted using different colours (see legend). The bottom panel shows the same heatmap as the top panel but this time with only interacting dwarfs overplotted. In the dwarf regime, this heatmap is bimodal around $(g-i) \sim 0.7$ (indicated using the black dashed line), with well-defined red and blue peaks. In both panels we also overlay rest-frame $(g-i)$ colours of simple stellar populations (SSPs) which form at look-back times of 2, 4 and 10 Gyrs (taken from the Bruzual & Charlot (2003) population synthesis models). Since galaxies show a relationship between stellar mass and metallicity (e.g. Gallazzi et al., 2005), these SSPs use the median stellar metallicity of galaxies at the stellar mass in question. While real galaxies do not have star formation histories that correspond exactly to SSPs, these SSP-based colours serve as a useful guide to understanding the broad features of the star-formation histories of galaxies in this colour-mass diagram. For example, they indicate that a very small fraction of dwarfs (less than 8 per cent in any morphological class) are consistent with purely old stellar populations (i.e. a 'monolithic' formation scenario). In a similar vein, dwarfs which reside in the rest-frame $(g-i)$ blue cloud lie blueward of the 4 Gyr SSP, suggesting that they are likely to have had some star formation within the last 4 Gyrs.

Table 2.2 summarises the results in Figure 2.8. The upper sub-table shows the red fractions in each morphological class i.e. the fraction of galaxies that have $(g-i) > 0.7$, while the lower sub-table shows the same for interacting galaxies only. Together with Figure 2.8, we find that the red fractions in interacting dwarfs are lower across all morphological classes, indicating that interactions generally enhance star formation in dwarfs, as is the case in massive galaxies (e.g. Kaviraj, 2014a). The featureless dwarfs are not all red, which indicates that star formation in at least some of these objects is not completely quenched (or has quenched very recently). Around 60 per cent of dwarf ETGs are blue, and comparison to the 4 Gyr-old SSP suggests that these dwarf ETGs are likely to have had some star formation in the last 4 Gyrs.

In Figure 2.9, we compare the colours of our dwarf morphological classes with their counterparts in the massive ($M_\star > 10^{10}$ $M_\odot$) galaxy regime from Kaviraj (2010). While the majority of dwarf ETGs live in the blue cloud, this is qualitatively different from the massive-galaxy regime, where ETGs tend to be preferentially red. This suggests differences between the formation channels of ETGs in the massive and dwarf regimes, a point we study further in Section 2.5.1. Nevertheless, the dwarf ETGs do contain the highest fraction of red objects compared to the other morphological classes.

Red fractions	
dETG	$0.41^{0.04}$
dLTG	$0.23^{0.03}$
dF	$0.52^{0.09}$
Red fractions in interacting galaxies	
dETG	$0.25^{0.09}$
dLTG	$0.13^{0.05}$
dF	$0.20^{0.14}$

TABLE 2.2: **Upper sub-table:** Red fractions of dwarf galaxies in different dwarf morphological classes. A galaxy is defined as red if its rest-frame $(g-i)$ colour is greater than 0.7 (see text for details). dETG = dwarf early-types, dLTG = dwarf late-types, dF = dwarf featureless. Errors, which are shown as superscripts, are calculated following Cameron (2011). **Lower sub-table:** Red fractions of interacting galaxies in different dwarf morphological classes.

In it instructive to compare our red fractions to past work which has probed red/quenched fractions in dwarf galaxies. Some studies, based on the SDSS, have suggested that, at $M_\star < 10^9$ $M_\odot$, the quenched fraction in nearby ($z < 0.055$) dwarfs that reside in low-density environments may be close to zero (e.g. Geha et al., 2012). This is, however, likely to be affected by the selection biases (described in the introduction) whereby red dwarfs preferentially fall out of the selection in shallow surveys outside the local neighbourhood. Indeed, when a similar analysis is performed, using the SDSS, at very low redshift ($z < 0.02$), which mitigates some of this bias, the dwarf population shows a pronounced red sequence, which suggests a much higher red/quenched fraction (e.g. Barazza et al., 2006). This agrees with recent work on nearby low-surface-brightness galaxies (which are dominated by dwarfs) using the Dark Energy Survey (Abbott et al., 2018). The red fraction using the $(g-i)$ colour is around 30 per cent (Tanoglidis et al. (2021), see also Thuruthipilly et al. (2023)), in reasonable agreement with the values we find in this study.

2.3.5 Bars

We complete our visual exploration of dwarf morphologies by considering the incidence of bars in our dwarf galaxy sample. Figure 2.10 shows examples of four barred galaxies in our dwarfs. The bar fraction in face-on late-type dwarfs is $\sim$11 $\pm$ 3 per cent (9 out of 82 face-on late type galaxies). These are likely to be strong bars (e.g. Nair & Abraham, 2010) given that we are using ground-based images of relatively small galaxies. The bar therefore has to extend along a significant fraction of the galaxy to be visible (e.g. Géron et al., 2023).

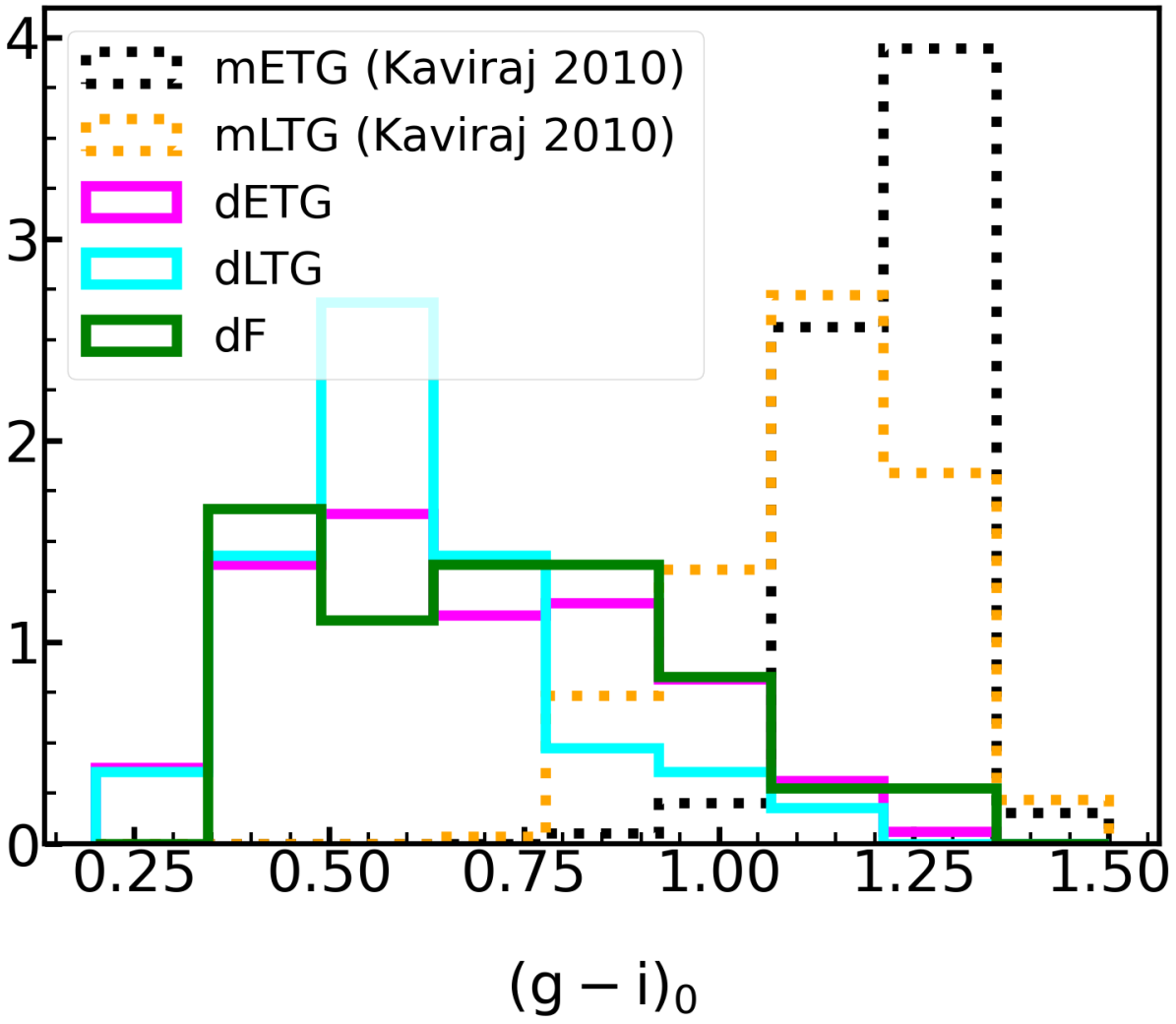

FIGURE 2.9: Comparison between the rest-frame $(g-i)$ colours of our dwarf morphological classes with their counterparts in the massive-galaxy regime from Kaviraj (2010).

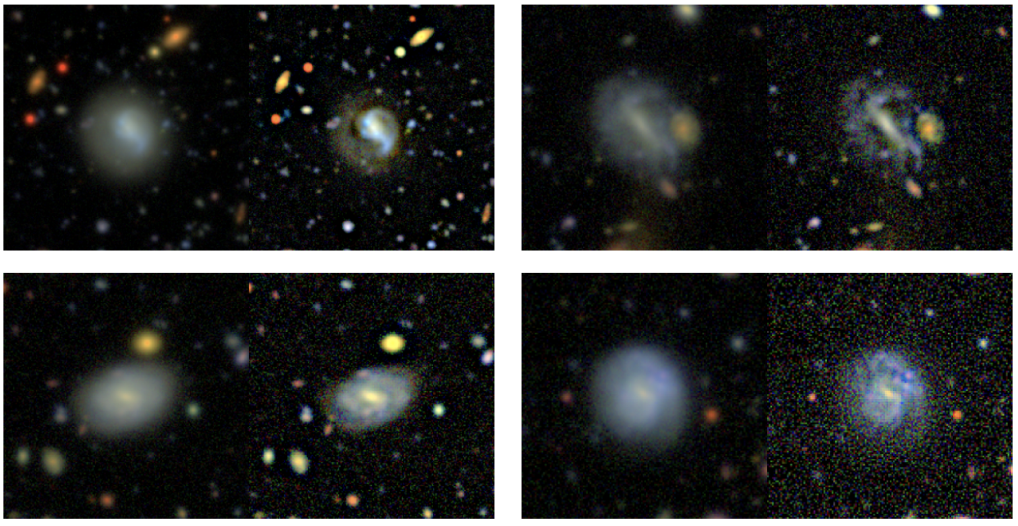

FIGURE 2.10: Examples of galaxies with bars in our dwarf sample. For each galaxy the left-hand panel shows the HSC colour image, while the right-hand panel shows its unsharp-masked counterpart.

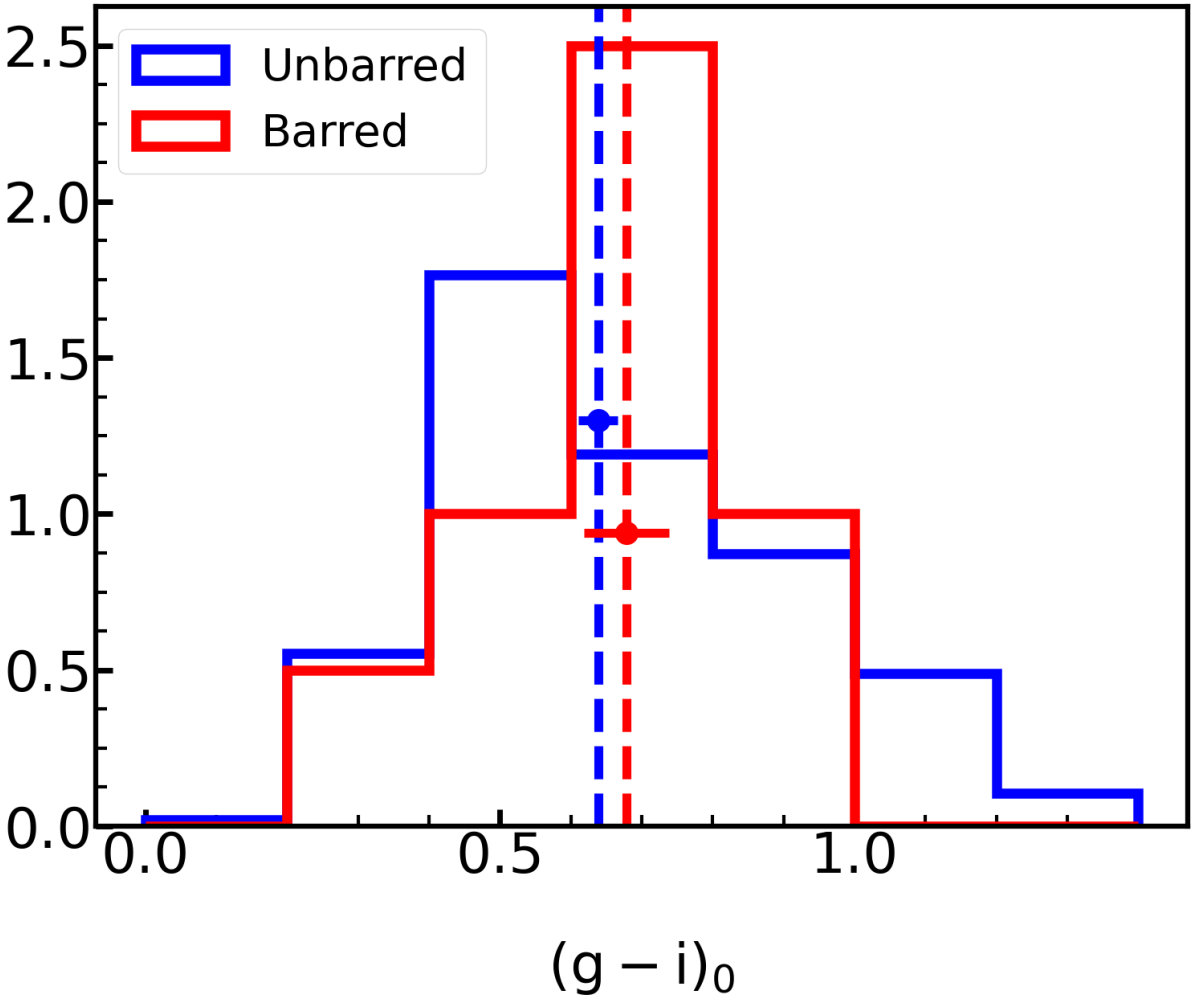

FIGURE 2.11: Rest-frame $(g-i)$ colours of barred and unbarred dwarfs in our sample.

The bar fractions at the upper end of the mass range traced by our dwarf galaxies are similar to values reported by recent work (e.g. Géron et al., 2021). In comparison, the strong bar fraction in massive galaxies (e.g. Masters et al., 2011; Cheung et al., 2013; Géron et al., 2021) is around 20 per cent. The frequency of strong bars therefore decreases as we move from the massive-galaxy to the dwarf regime.

Figure 2.11 shows the rest-frame colours of the barred dwarfs compared to their unbarred counterparts. Barred dwarfs in our sample have a similar colour distribution as unbarred dwarfs with similar median values (which overlap within their uncertainties). A KS test between the two distributions returns a K-S value of 0.22 and p-value of 0.6 indicating that the two distributions are similar. The trends found here appear to be somewhat in contrast to what is found in massive galaxies (e.g. Géron et al., 2021) where strongly-barred galaxies tend to be redder.

2.4 Morphological parameters

We proceed by investigating the differences between dwarf and massive galaxies from the perspective of commonly-used morphological parameters that dominate the recent literature: concentration (C), asymmetry (A), clumpiness (S) (collectively know as the 'CAS' system), M_{20}, the Gini coefficient and the Sérsic index. We compare our results to previous studies that have used optical data in the dwarf and massive-galaxy regimes, both in the very local Universe (within 50 Mpc) and in the general Universe at low redshift ($z < 0.1$).

Prior to the calculation of morphological parameters for our dwarf sample we mask any interloper sources, foreground stars and background and foreground galaxies that may interfere with the measurement of the surface brightness profile or morphology of the object in question. In order to obtain accurate unbiased estimates for the morphological parameters studied in this work, we estimate the masked flux using the interpolation algorithm described in Watkins et al. (2022), which uses the Fourier transform of the azimuthal surface brightness profiles in concentric radial bins to reconstruct the missing flux. The missing flux is reconstructed concentrically until the edge of the image is reached (i.e. where the semi-major axis of the interpolation ellipse reaches the edge of the image) leaving the area outside the interpolation ellipse unchanged. Any masked regions outside the interpolation area are filled with Gaussian noise 1σ away from the RMS of the image. We use the modified interpolated images from the algorithm to calculate effective radii, effective surface brightnesses and central concentrations. The asymmetry, clumpiness, M_{20}, Gini coefficient and Sérsic index are estimated by employing the Python `statmorph` package (Rodriguez-Gomez et al., 2019) on the same interpolated images, using Petrosian radii estimated by `statmorph` as described below. We calculate the morphological parameters and physical properties mentioned above in five HSC bands (*grizy*) for 247 dwarfs from our initial sample of 257. The dwarfs for which parameters cannot be calculated either have relatively low signal to noise, suffer from some light contamination from a nearby bright galaxy or are relatively compact sources.

2.4.1 Concentration

We calculate the concentration parameter (C_{82}), described in Eq. 2.1, as defined by Kent (1985), Bershady et al. (2000) and C03.

$$C_{82} = 5 \times \log_{10}\left(\frac{R_{80}}{R_{20}}\right) \tag{2.1}$$

where R_{80} and R_{20} correspond to the radii enclosing 80% and 20% of the total light of the galaxy.

We calculate R_{80} and R_{20} from the curve of growth, without making any assumptions about the shapes of the light profiles. As opposed to other studies in which set apertures are used, we estimate the total luminosity of galaxies by extrapolating their total luminosities to infinity, where the slope of the curve of growth tends to zero (Muñoz-Mateos et al., 2015).

2.4.2 Asymmetry

The asymmetry parameter as defined by Abraham et al. (1996); Conselice et al. (2000b); Rodriguez-Gomez et al. (2019) is described in Eq. 2.2 below.

$$A = \frac{\sum_{i,j} \mid I_{\rm ij} - I^{180}_{\rm ij} \mid}{\sum_{i,j} \mid I_{\rm ij} \mid} - A_{\rm bgr} \tag{2.2}$$

where $I_{\rm i,j}$ and $I^{180}_{\rm i,j}$ are the pixel values of the original and the rotated images, respectively. We calculate asymmetry within a circular aperture of either 1 or 1.5 $\times$ $R_{\rm petro}$ depending on the dataset that we are comparing our sample with. In Sections 2.4.7 and 2.4.9 all pixels within a circular aperture of radius 1.5 $\times$ $R_{\rm petro}$ are taken into account for the calculation of asymmetry. However, in Section 2.4.8 the aperture changes to 1 $\times$ $R_{\rm petro}$. $A_{\rm bgr}$ is the asymmetry of the background (Lotz et al., 2004). This quantity is calculated as in Eq. 2.3 where $I_{\rm bgr;i,j}$ and $I^{180}_{\rm bgr;i,j}$ are the pixel values of the original and the rotated images respectively, corresponding to the region outside the circular aperture, which we define as the background.

$$A_{\rm bgr} = \frac{\sum_{i,j} \mid I_{\rm bgr;ij} - I^{180}_{\rm bgr;ij} \mid}{\sum_{i,j} \mid I_{\rm bgr;ij} \mid} \tag{2.3}$$

2.4.3 Clumpiness

The clumpiness (sometimes also referred to as the smoothness index), as defined in L04 and Rodriguez-Gomez et al. (2019), is described in Eq. 2.4 below.

$$S = \frac{\sum_{i,j} I_{\rm ij} - I^{\rm S}_{\rm ij}}{\sum_{i,j} I_{\rm ij}} - S_{\rm bgr} \tag{2.4}$$

where $I_{\rm i,j}$ and $I^{\rm S}_{\rm i,j}$ are the pixel values of the original image and its smoothed version, respectively, within circular apertures of 1 or 1.5 $\times$ $R_{\rm petro}$, depending on which dataset we are comparing our sample with. The smoothed image is obtained using a boxcar filter of width σ, which is set to 0.25 $\times$ $R_{\rm petro}$, as in Lotz et al. (2004, L04 hereafter). The calculation is performed only for the pixels corresponding to radii between σ and 1.5 $\times$ $R_{\rm petro}$ (or 1 $\times$ $R_{\rm petro}$ for the comparison with DES data in Section 2.4.8), since the central region is avoided due to most galaxies showing significant central concentration. The pixels with a negative numerator value in Eq. 2.4 are excluded from the summation. $S_{\rm bgr}$ is the background clumpiness, calculated using the background pixels residing outside the segmentation map (Eq. 2.5).

$$S_{\rm bgr} = \frac{\sum_{i,j} I_{\rm bgr,ij} - I^{\rm S}_{\rm bgr,ij}}{\sum_{i,j} I_{\rm bgr,ij}} \tag{2.5}$$

where $I_{\rm bgr;i,j}$ and $I^{\rm S}_{\rm bgr;i,j}$ are the pixel values of the original image and its smoothed version, respectively, corresponding to the region outside the circular aperture, which we define as the background.

2.4.4 Gini coefficient

We calculate the Gini coefficient (Abraham et al., 2003; Lotz et al., 2004; Rodriguez-Gomez et al., 2019) as described in Eq. 2.6.

$$G = \frac{1}{\mid \overline{X} \mid n(n-1)} \sum_{i=1}^{n} (2i - n - 1) \mid X_{\rm i} \mid \tag{2.6}$$

where n is the number of pixels in the image, $\overline{X}$ is the mean of the pixel values and $X_{\rm i}$ corresponds to the value of each pixel. Following Lotz et al. (2004) we use a segmentation map for the calculation of the Gini coefficient. The segmentation map is constructed by convolving the image with a Gaussian kernel of $\sigma = R_{\rm petro}/5$. We then select pixels from the original image that are above the mean surface brightness of the convolved image at $R_{\rm petro}$.

2.4.5 M_{20}

The M_{20} index (L04) is the second order moment of the brightest 20 per cent of the galaxy's flux, divided by the total second order central moment ($M_{\rm tot}$; see Eq. 2.7). The formula for calculating this statistic is given in Eq. 2.8.

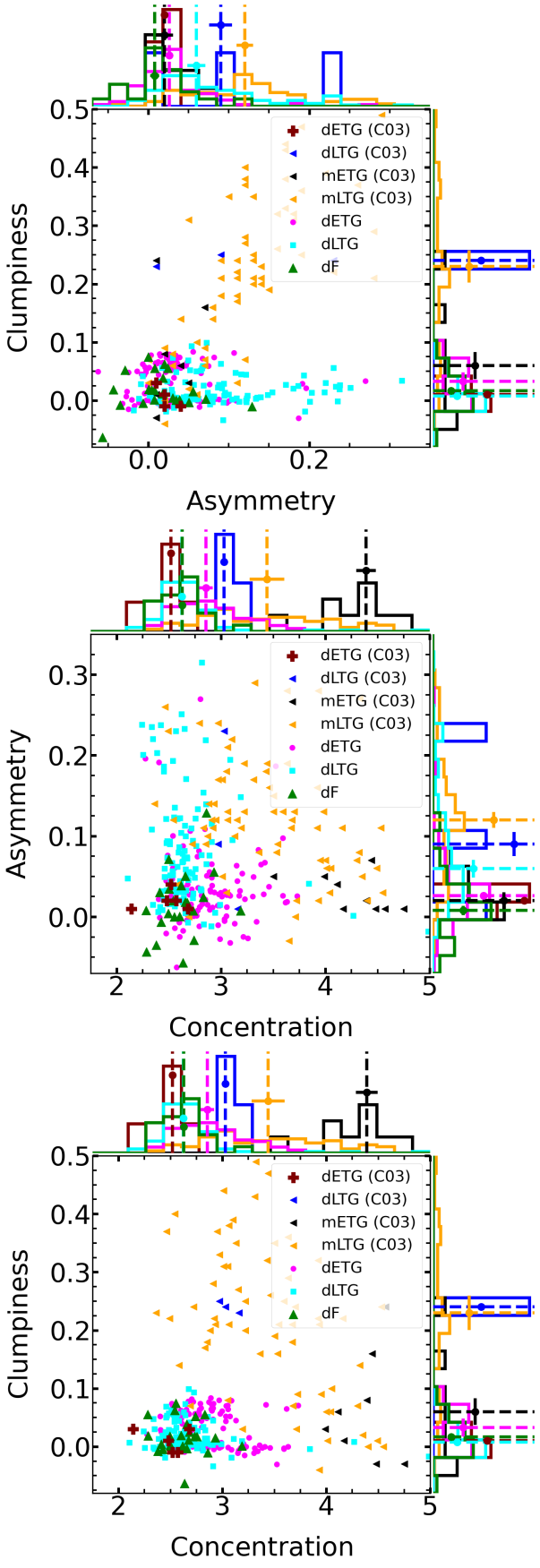

FIGURE 2.12: Combinations of CAS parameters for various morphological classes in our dwarf and the C03 samples. Distributions of parameters and median values (together with their associated uncertainties) are shown on the sides of all panels. 'm' and 'd' correspond to the massive ($M_\star > 10^{10}$ $M_\odot$) and dwarf (10^8 $M_\odot$ $< M_\star < 10^{9.5}$ $M_\odot$) regimes, respectively. ETG = early-type galaxy, LTG = late-type galaxy, F = featureless galaxy.

$$M_{\rm tot} = \sum_{i}^{n} M_{\rm i} = \sum_{i}^{n} f_{\rm i}((x_{\rm i} - x_{\rm c})^2 + (y_{\rm i} - y_{\rm c})^2) \tag{2.7}$$

where $f_{\rm i}$ is the flux of each pixel, $x_{\rm i}$ and $y_{\rm i}$ represent the positions of each pixel, $x_{\rm c}$ and $y_{\rm c}$ correspond to the locations of the galaxy's centre, which are estimated by minimizing $M_{\rm tot}$.

$$M_{20} = \log_{10}\left(\frac{\sum_{\rm i,BRG} M_{\rm i}}{M_{\rm tot}}\right) \tag{2.8}$$

where $\sum_{\rm i,BRG} M_{\rm i}$ represents the summation over the brightest pixels of the image corresponding to 20 per cent of the galaxy's total flux. The pixels which are taken into account for the calculation of this statistic are selected using the Gini segmentation map described in Section 2.4.4. We note that the segmentation maps (i.e. the collections of pixels taken into account) used to calculate concentration, asymmetry and clumpiness are different, because they are either defined based on set apertures using $R_{\rm petro}$ or by extrapolating the total galaxy luminosity to infinity (in the case of concentration).

2.4.6 Sérsic index

The details of the Sérsic fitting procedure in this study can be found in Rodriguez-Gomez et al. (2019) as we use the `statmorph` package to perform the fitting. We fit 2D single component Sérsic profiles using the Python `astropy` modelling package within `statmorph`.

In the following sections we compare the morphological parameters derived for our dwarf galaxies first to dwarfs and massive galaxies in the very local Universe (within 50 Mpc) and then to massive galaxies in the general Universe at low redshift ($z < 0.1$). We then consider whether combinations of parameters exist which can separate interacting dwarfs from their non-interacting counterparts. We perform these comparisons via parameters calculated using HSC *r* band images, which is closest to the filters used in previous studies.

2.4.7 Comparison to dwarf and massive galaxies in the very local Universe (within 50 Mpc)

In Figures 2.12 and 2.13 we compare our results to C03 and L04 who have used the CAS system, M_{20} and the Gini coefficient to study galaxy morphology in the optical wavelengths in the very local Universe (within 50 Mpc), in galaxies that span all Hubble types from Frei et al. (1996). It is worth noting that the

numbers of galaxies in these studies is relatively small, which may affect the reliability of the comparisons. However, the proximity of the galaxies studied by C03 and L04 means that the dwarf samples are relatively unbiased in terms of morphological type, stellar mass and colour distribution, although there is still a bias because the dwarfs from these studies reside in a relatively dense environment. Nevertheless, such a comparison can provide a useful sanity check of the parameter values we have found in our dwarf population. In order to perform this comparison consistently with these previous studies, we consider asymmetry and clumpiness for our dwarfs calculated using pixels that reside within $1.5 \times R_{\rm petro}$, as in C03 and L04. We also restrict dwarfs from C03 and L04 to the stellar mass range spanned by our HSC dwarfs.

It is instructive to note the physical scales that are being traced in the images used to calculate morphological parameters in the different datasets. Given the upper redshift limit of our dwarf sample ($z \sim 0.08$) and a median HSC seeing of 0.6 arcseconds, the physical scales resolved in these galaxies is $\sim$0.9 kpc or better. Given a median seeing of around 2 arcseconds on the Palomar P60 telescope, the Frei et al. images are able to resolve physical scales of $\sim$0.4 kpc or better in these galaxies (within a factor of 2 of our HSC dwarf sample).

The top panel of Figure 2.12 presents the clumpiness vs asymmetry plane for both our dwarfs and the C03 galaxies. The distributions of parameters and median values (together with their associated uncertainties) are shown on the sides of all panels. It is worth noting first that, while the different morphological classes within our dwarf sample show differences in asymmetry, they show significant overlap in clumpiness. The dwarf featureless galaxies show the lowest median value of asymmetry. After that, the galaxies with the lowest asymmetries are dwarf and massive ETGs, while those with the highest asymmetries are late-type systems (consistent with what is seen in other work e.g. Bershady et al., 2000; Taylor-Mager et al., 2007).

Our dwarf ETGs reside in similar regions of this plane as their C03 counterparts. Our dwarf LTGs and the massive LTGs from C03 show reasonable overlap in asymmetry. However our dwarf LTGs show a significant offset in the median values of clumpiness, with both dwarf and massive LTGs from C03 (the latter having clumpiness values which are a factor of $\sim$5 higher than our dwarf LTGs). The difference between the dwarf LTGs in our sample and those in C03 could be driven by the fact that the local Universe galaxies are better resolved, although the very small numbers make it difficult to come to a definitive conclusion. A potential reason for the significant difference in clumpiness between dwarf and massive LTGs is that the SFRs are higher in the massive-galaxy regime, where the potential wells are steeper and gas accretion takes place more efficiently. This is likely to cause the galactic structure to

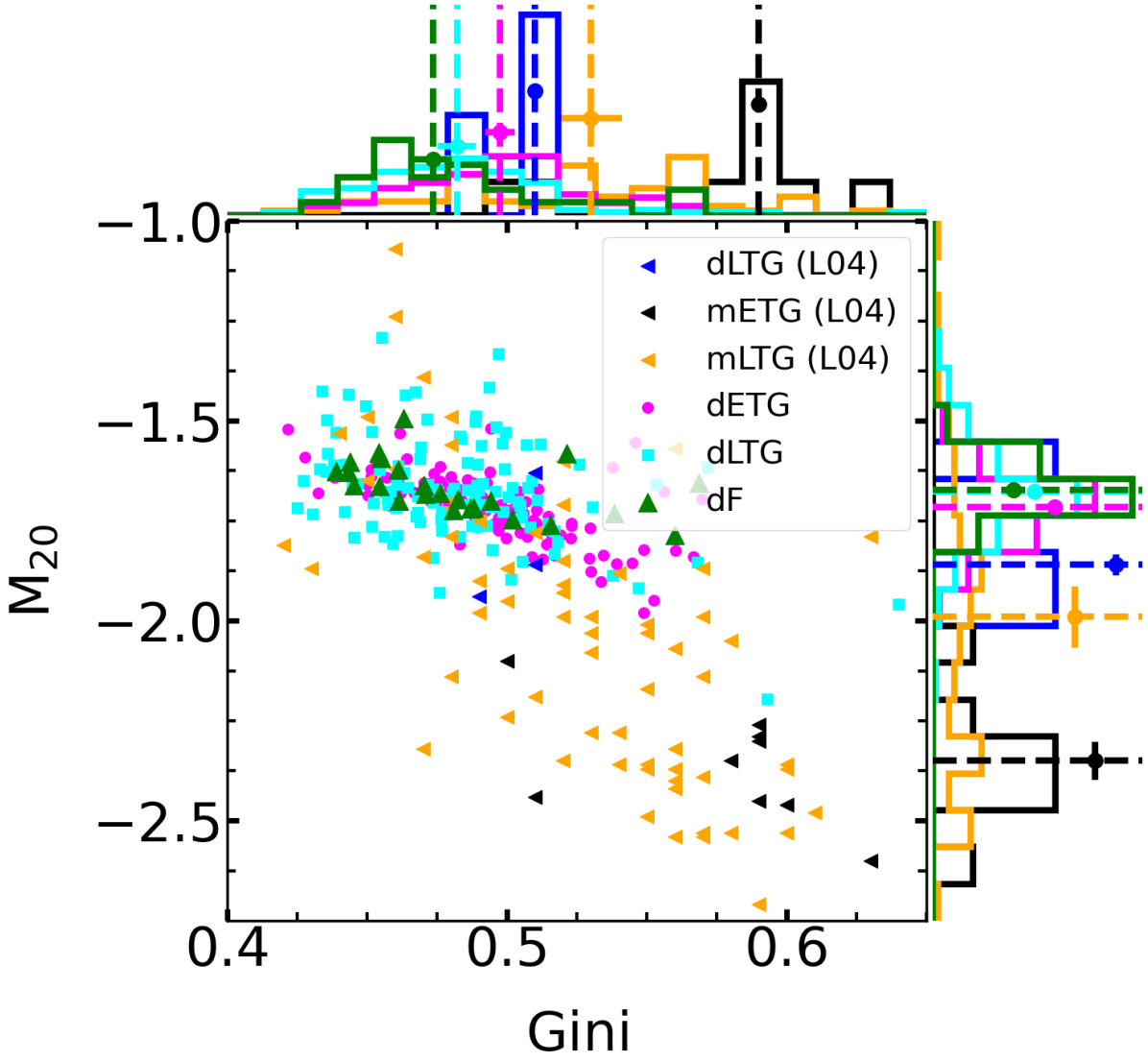

FIGURE 2.13: M_{20} vs Gini coefficient for various morphological classes in our dwarf and the L04 samples. Distributions of parameters and median values (together with their associated uncertainties) are shown on the sides of all panels. 'm' and 'd' correspond to the massive ($M_\star > 10^{10}$ $M_\odot$) and dwarf (10^8 $M_\odot < M_\star < 10^{9.5}$ $M_\odot$) regimes, respectively. ETG = early-type galaxy, LTG = late-type galaxy, F = featureless galaxy.

contain more HII regions which, in turn, would tend to increase the 'patchiness' (or clumpiness) of the system. Interestingly, the same differences are seen when we compare our dwarfs to massive galaxies from the DES survey, which are resolved slightly worse than our HSC dwarfs (see Section 2.4.8 below), suggesting that the trend may indeed be driven by differences in physical properties, such as SFR, than simply the resolution of the images.

The middle and bottom panels of Figure 2.12 present the asymmetry vs concentration and the clumpiness vs concentration planes. The concentrations of our dwarf ETGs are similar to those measured in their counterparts from C03. However, massive ETGs show, on average, higher concentration values than the

dwarf ETGs (both from our sample and C03) by a factor of $\sim$1.7. As noted already in C03, the structural differences between dwarf and massive ETGs suggest that they form via different evolutionary channels. The observed differences in terms of morphological parameters appear aligned with the results of Section 2.3.3, which demonstrates that dwarf ETGs appear to exhibit a lower incidence of interactions than their high mass counterparts. Thus, while massive ETGs are known to experience minor and major merger events, which are likely to increase the concentration of these systems (e.g. Martin et al., 2018a), the lower concentrations and lack of interaction signatures in dwarf ETGs suggest an evolutionary history that is dominated by secular accretion. We explore this point further in Section 2.5.1.

In Figure 2.13, we complete the comparison of our dwarfs to galaxies in the very local Universe by considering the M_{20} – Gini plane. L04 demonstrate that, in the massive-galaxy regime, ETGs reside at relatively high and low values of Gini and M_{20} respectively, while the opposite is true for the LTG population. In other words, these morphological classes are reasonably well separated in the M_{20} – Gini plane. However, this figure indicates that dwarf ETGs and LTGs show significant overlap in the M_{20} – Gini plane. The lower concentrations of dwarf ETGs, seen in Figure 2.12, are mirrored in lower values of the Gini coefficient, while dwarf LTGs do not share the low and high values of M_{20} and Gini respectively like their massive counterparts. Overall, the M_{20} – Gini plane is unable to cleanly separate the ETG and LTG classes in the dwarf regime like it does in the massive-galaxy regime.

2.4.8 Comparison to massive galaxies in the nearby Universe ($z < 0.1$)

We proceed by comparing our dwarf sample with massive ($M_\star > 10^{10}$ M$_\odot$) galaxies observed in the Dark Energy Survey (DES; Abbott et al. (2018)) at low redshift ($z < 0.1$). Note that, given the relatively small ($\sim$2 deg^2) area of COSMOS (as opposed to $\sim$5000 deg^2 in DES), there are only a handful of massive galaxies in the COSMOS2020 catalogue at $z < 0.08$, making it impossible to do a statistical comparison between dwarfs and massive galaxies within this footprint alone. We restrict our DES comparison to massive galaxies only. This is because, while DES is deeper than the SDSS, the magnitude limits of the training set used for the DES classifications ($16 < i < 18$) mean that DES dwarfs with reliable morphological classifications remain biased towards blue, star-forming objects (in a similar fashion to what is the case for SDSS). As we discuss below, and demonstrate in Appendix A.1, this biases the DES dwarfs in the catalogues we use towards blue, morphologically late-type systems.

Given the upper redshift limit of our DES sample and a median seeing of 0.9 arcseconds for the DES images, the physical scale that is resolved in these images is $\sim$1.6 kpc or better. The DES sample used for our comparisons is constructed by cross-matching three catalogues:

- Zou et al. (2022) provide physical properties of galaxies, including photometric redshifts, stellar masses and SFRs of around 300 million galaxies from DES DR2 data. The photometric redshifts and their uncertainties are calculated using *ugrizyW*1*W*2 photometry, from the SCUSS (Zou et al., 2016), SDSS (Strauss et al., 2002) and WISE (Wright et al., 2010) surveys in a hybrid way via local regression on a spectroscopic training set combined with template fitting. The accuracy of the photometric redshifts for galaxies in our mass range of interest is $\sim$0.024 at $z < 1.2$.

- Tarsitano et al. (2018) provides measurements of morphological parameters including concentration, asymmetry, clumpiness, M_{20}, Gini coefficient and the Sérsic index in the g, r and i bands for 45 million galaxies from DES Year 1 data.

- Cheng et al. (2021) provides one of the largest catalogues of morphological classifications to date, comprising measurements for over 20 million galaxies using the DES Y3 data, based on convolutional neural networks. The training sets used in this work are taken from the Galaxy Zoo 1 (GZ1) catalogue (Lintott et al., 2011) which consist of bright galaxies with $16 < i < 18$ at $z < 0.25$. The accuracy of the 'superior confidence' ($i < 18$) classifications, which we use here, reaches 99 per cent for galaxies when compared to the GZ1 data. This makes this catalogue ideal for comparing the morphological properties of our dwarfs to those of massive galaxies in the nearby Universe. The final catalogue provides probabilities for two galaxy types: ETGs and LTGs.

For the comparison with our HSC dwarf population, we first cross-match and combine the three catalogues mentioned above. We then use this combined catalogue to select galaxies with $\mathrm{M}_\star > 10^{10}$ $\mathrm{M}_\odot$ at $z < 0.1$, which have probabilities of being ETG or LTG higher than 90 per cent, redshift and stellar mass errors less than 0.02 and 0.2 dex respectively and which have the highest ('superior') confidence flag of 4. Importantly, galaxies with the superior confidence flag span the same magnitude range ($16 < i < 18$) as that of the Galaxy Zoo training set used in Cheng et al. (2021) and will, therefore, have the most reliable morphological classifications. Our final sample contains 809 objects out of which 103 are ETGs and 706 are LTGs. For consistency with the DES studies described above, we calculate the morphological parameters for our HSC sample using pixels within R_{petro}, as in Tarsitano et al. (2018).

We begin, in Figure 2.14, by presenting the Sérsic indices of different morphological classes in both the dwarf and massive galaxy regimes. Median values are shown using the dashed vertical lines. The bootstrapped standard errors in the medians are shown using the small horizontal bars. The massive-galaxy regime exhibits the known trends in Sérsic index, with the ETG population (which has a median of

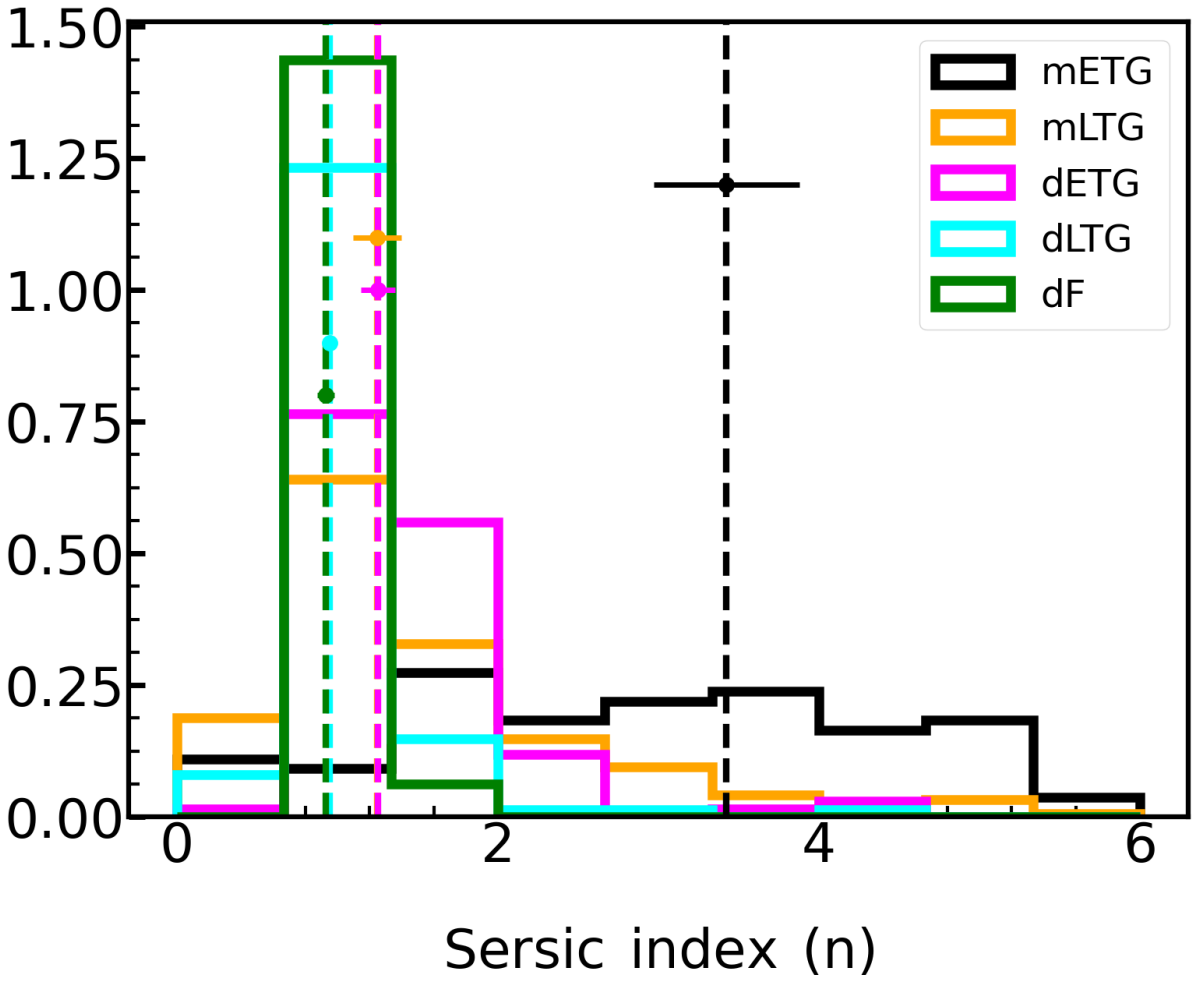

FIGURE 2.14: Sérsic index distribution for our dwarf sample and massive galaxies in DES. Median values (together with their associated uncertainties) are shown using the vertical lines and horoizontal error bars. 'm' and 'd' correspond to the massive ($M_\star > 10^{10}$ $M_\odot$) and dwarf (10^8 $M_\odot < M_\star < 10^{9.5}$ $M_\odot$) regime, respectively. ETG = early-type galaxy, LTG = late-type galaxy, F = featureless galaxy.

$\sim$3.5) exhibiting larger values than their LTG counterparts (which has a median of $\sim$1.3). As originally noted by Faber & Lin (1983), unlike their massive counterparts, dwarf ETGs exhibit relatively low values of Sérsic index (with a median of $\sim$ 1.5 in their study), mirroring the differences in concentration between dwarf and massive ETGs seen earlier in Section 2.4.7). The different dwarf morphological classes span similar values of Sérsic index, indicating that this parameter is not a reliable discriminator between morphological classes in the dwarf regime.

In Figures 2.15 and 2.16 we compare the morphological parameters of our dwarf sample to the massive galaxy population from DES. The trends we find here are broadly similar to those from the local Universe comparison in Section 2.4.7. But the much larger number of DES galaxies enables us to put these trends on a firmer statistical footing. The top panel of Figure 2.15 indicates that dwarf ETGs exhibit similar values of asymmetry as their massive counterparts (similar to our findings for the very local Universe

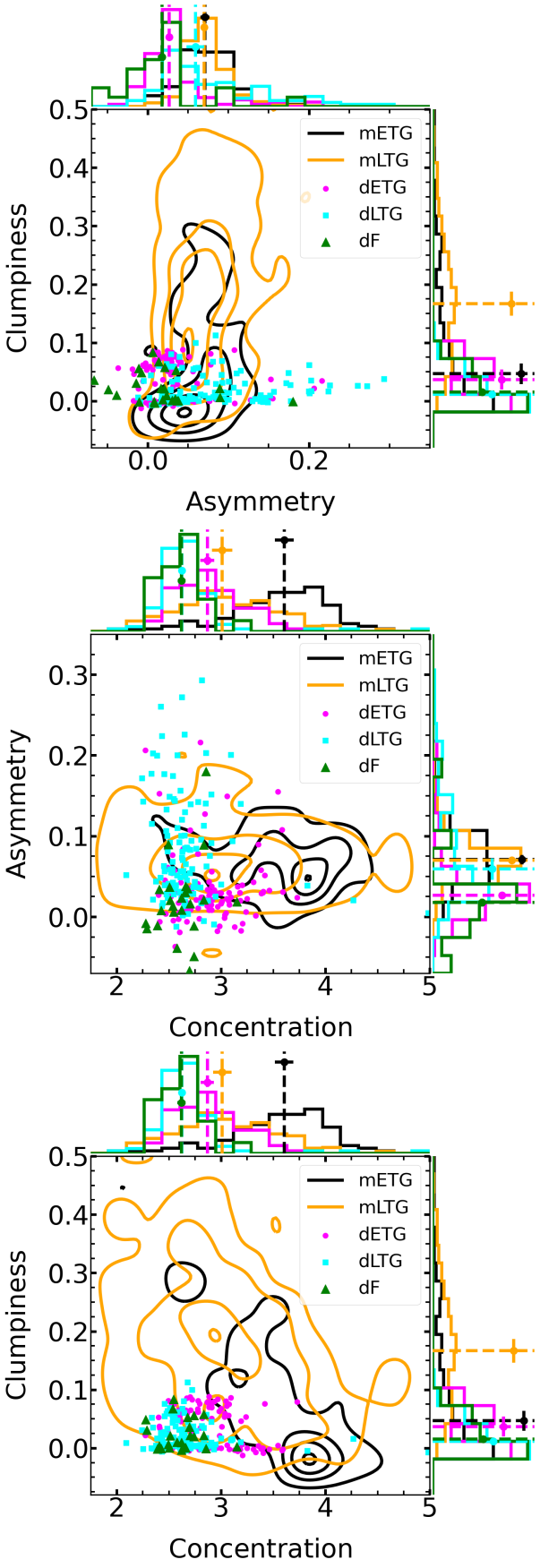

FIGURE 2.15: Combinations of CAS parameters for various morphological classes in our dwarf sample and massive galaxies in the DES sample (contours). Distributions of parameters and median values (together with their associated uncertainties) are shown on the sides of all panels. 'm' and 'd' correspond to the massive ($M_\star > 10^{10}$ $M_\odot$) and dwarf (10^8 $M_\odot < M_\star < 10^{9.5}$ $M_\odot$) regimes, respectively. ETG = early-type galaxy, LTG = late-type galaxy, F = featureless galaxy. Each DES distribution contains 4 equidistant contours starting with 2 counts pixel^{-1} and ending with the maximum pixel value.

in Section 2.4.7). The locations of dwarf LTGs are virtually orthogonal to their massive counterparts in the clumpiness vs asymmetry plane. While dwarf LTGs exhibit relatively high asymmetry and low clumpiness values, massive LTGs show high clumpiness and low asymmetry values. As we noted above, the difference in clumpiness between the dwarf and high mass LTGs, which is seen both in the very local Universe and DES comparisons, could be caused by the higher SFRs in the massive LTGs. Indeed, the median SFR of massive LTGs in DES is a factor of 3 higher than the median SFR of the HSC dwarf LTGs (SFR $_{\rm HSC,dLTG}$ $\sim$0.05 $M_\odot$ yr^{-1}; SFR $_{\rm DES,mLTG}$ $\sim$0.15 $M_\odot$ yr^{-1}). The fact that clumpiness typically correlates strongly with the SFR (as also seen in e.g. C03) suggests that massive LTGs are likely to show higher values of clumpiness.

The middle panel of Figure 2.15 shows that, like in the very local Universe comparison, dwarf ETGs show much lower values of concentration than their massive counterparts, with the median concentration of dwarf ETGs being around a factor of 1.4 smaller than that in the massive ETG population. As noted in the previous section, past studies have shown that the M_{20} – Gini plane can be used to separate ETGs and LTGs in the massive-galaxy regime, using various instruments at different redshifts e.g. HST at intermediate and high redshift (e.g. Lotz et al., 2008; Lee et al., 2013), DECaM via DES at $z < 0.25$ (Cheng et al., 2021) and Spitzer via the S_4G survey (Holwerda et al., 2014)) at distances up to $\sim$40 Mpc. Figure 2.16 shows the ETG – LTG separation criterion (dashed blue line), defined by Lotz et al. (2008), applied to the DES galaxies from Cheng et al. (2021) and confirms that this plane can separate massive ETGs and LTGs well. However, dwarf ETGs and LTGs show a reasonable degree of overlap in the M_{20} – Gini plane, which makes it difficult to cleanly separate these morphological classes in the dwarf regime using these parameters.

2.4.9 Can interacting and non-interacting dwarfs be separated using morphological parameters?

We now explore whether morphological parameters could be used to separate interacting dwarfs from their non-interacting counterparts. In the massive-galaxy regime, an asymmetry threshold of $A = 0.35$ has often been used to identify interacting systems during a significant fraction of the interaction (e.g. Conselice et al., 2000a; Conselice, 2003). While galaxies such as ULIRGs, which are undergoing significant merging, span a wide range of asymmetry values, non-interacting galaxies in the massive regime appear not to extend beyond $A \sim 0.35$, which makes this a reasonable criterion for selecting strongly interacting systems.

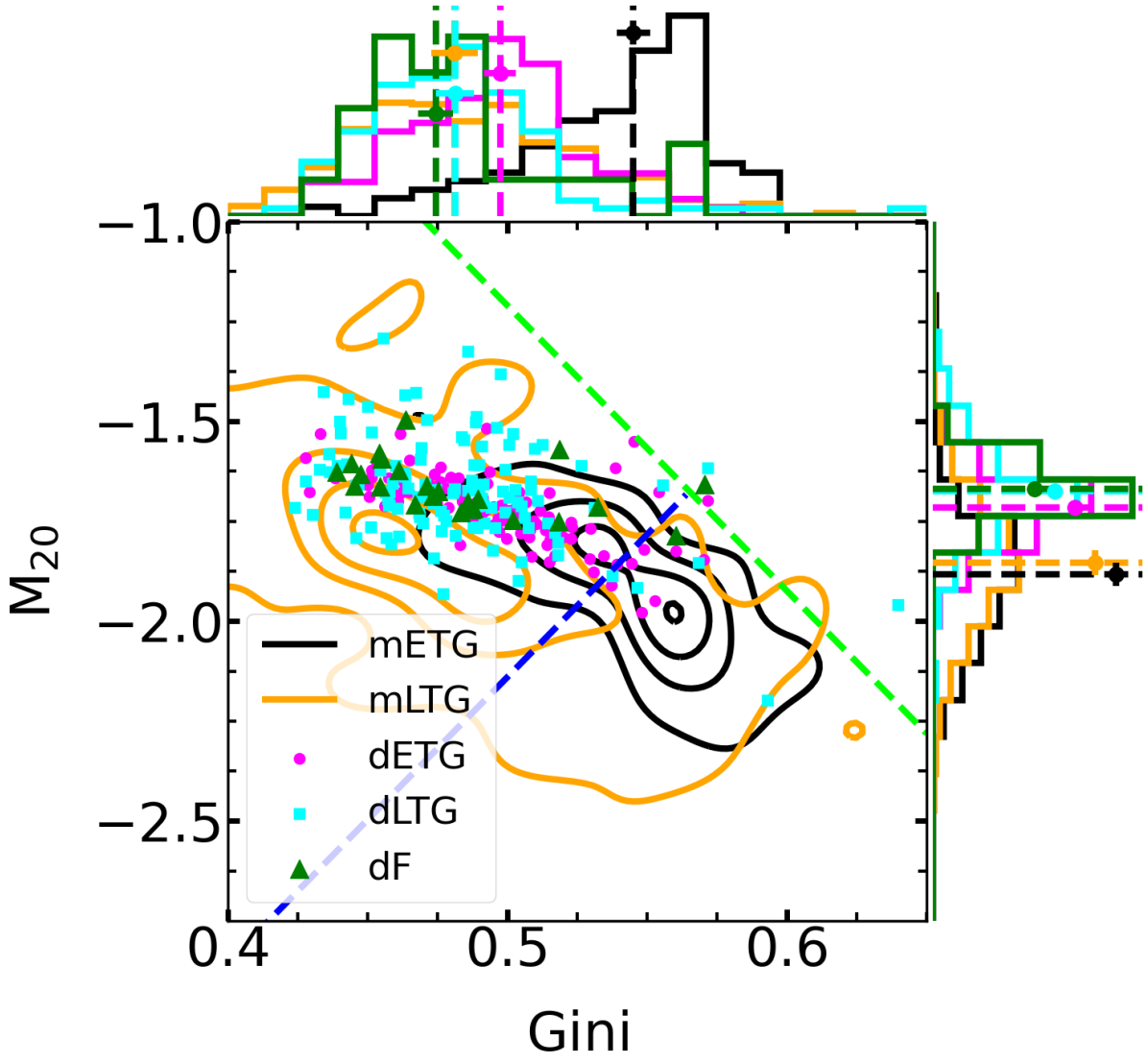

FIGURE 2.16: M_{20} vs Gini for our dwarf sample and massive galaxies in the DES sample (contours). The light green dashed line represents the interaction criterion from Lotz et al. (2008). The dashed blue line represents the separation relation between ETGs and LTGs described in Lotz et al. (2008). Distributions of parameters and median values (together with their associated uncertainties) are shown on the sides of all panels. 'm' and 'd' correspond to the massive ($M_\star > 10^{10}$ $M_\odot$) and dwarf (10^8 $M_\odot < M_\star < 10^{9.5}$ $M_\odot$) regimes, respectively. ETG = early-type galaxy, LTG = late-type galaxy, F = featureless galaxy. Each DES distribution contains 4 equidistant contours starting with 2 counts pixel^{-1} and ending with the maximum pixel value.

In Figure 2.17, we compare the values of asymmetry for dwarf galaxies which have been flagged as being interacting to those that have not. The dwarf ETGs and LTGs which are interacting exhibit median asymmetry values that are larger by around a factor of 4 and 3, respectively, than that of their non-interacting counterparts (see the top panel of this figure). As might be expected, the presence of visually-identifiable morphological disturbances correlates with larger values of asymmetry. There is some overlap between the interacting and non-interacting populations (at least in the relatively small dwarf sample in this study), which means that using asymmetry to cleanly separate disturbed and undisturbed dwarfs is challenging.

Nevertheless, the fraction of interacting galaxies increases for progressively higher thresholds of asymmetry (see the bottom panel of this figure). For example, in dwarf ETGs which have asymmetry values larger than $\sim$0.05, more than 50 per cent of galaxies are interacting. The corresponding value for dwarf LTGs is $\sim$0.08. These thresholds can be used to select samples of dwarfs in which a majority are likely to be interacting (in images that have similar depth and resolution to the ones used in this study). We note that these thresholds may change for systems outside the stellar mass range probed in this work (10^8 M$_\odot$ $< M_\star <$ $10^{9.5}$ M$_\odot$) or at higher redshift.

In a similar vein, L04 and Lotz et al. (2008) have formulated an interaction criterion in the M_{20} – Gini plane for massive galaxies, which is represented by the green dashed line in Figure 2.18. Using ULIRGs as proxies for strongly interacting galaxies, they have shown that most ULIRGs reside on the right-hand side of this line. However, this discriminator does not separate interacting and non-interacting dwarfs and performs worse than asymmetry.

2.5 Implications for the evolutionary histories of dwarf morphological classes

In this section, we complete our study by bringing together our results and discussing their implications for the evolutionary histories of the various dwarf morphological classes.

2.5.1 Dwarf ETGs have different evolutionary histories to massive ETGs

Section 2.3.3 indicates several interesting differences between dwarf and massive ETGs. Contrary to what is found in the massive-galaxy regime, the incidence of interactions in dwarf ETGs is lower than that in dwarf LTGs. The interaction fraction in dwarf ETGs appears to be around a factor of 5 lower than

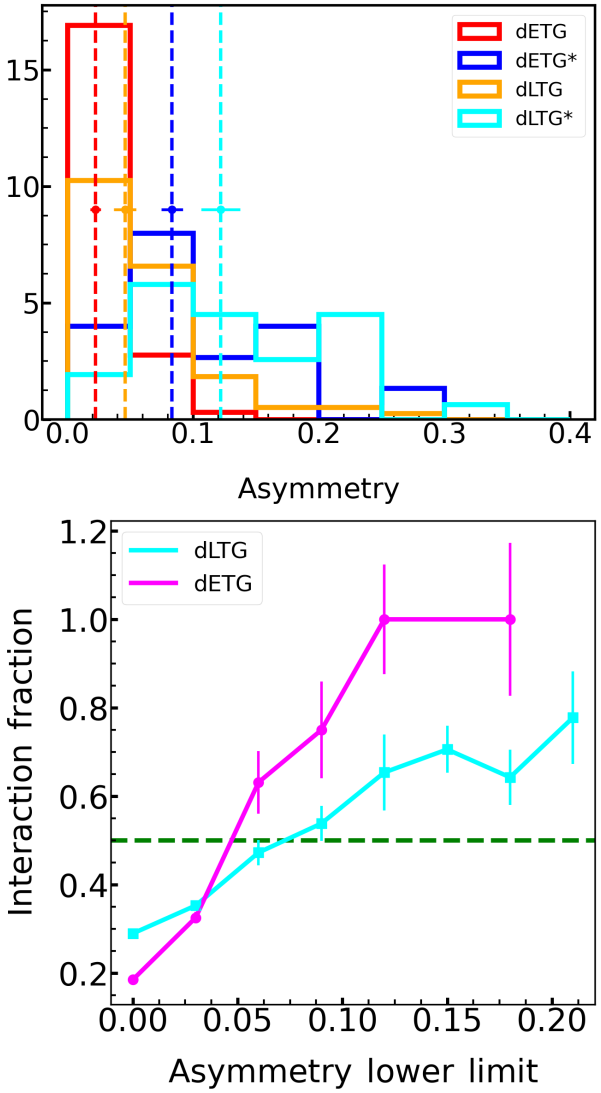

FIGURE 2.17: **Top:** Asymmetry distributions for interacting (starred in legend) and non-interacting (not starred) early (dETG) and late (dLTG) type dwarf galaxies. Median values (together with their associated uncertainties) are shown using vertical lines and horizontal error bars. **Bottom:** Fractions of interacting galaxies for different lower asymmetry thresholds, shown on the x-axis. The green horizontal dashed line indicates an interaction fraction of 50 per cent.

that in their massive counterparts. In a similar vein, the frequency of dust lanes, which are signposts of interactions with lower mass companions, are several factors lower in dwarf ETGs than in massive ETGs. While the majority of massive ETGs have redder colours than their LTG counterparts, the majority of dwarf ETGs are blue like their LTG counterparts (Section 2.3.4), due to a high incidence of features such as blue cores. Finally, Section 2.4 demonstrates that dwarf ETGs have lower concentrations, Sérsic indices and Gini coefficients (all of which are measures of how concentrated the light profile is within a galaxy) than massive ETGs. The strong differences in both the structural properties and rest-frame colours (which trace recent star formation histories) suggest that the evolutionary histories of dwarf ETGs is likely to be different from that of their massive counterparts.

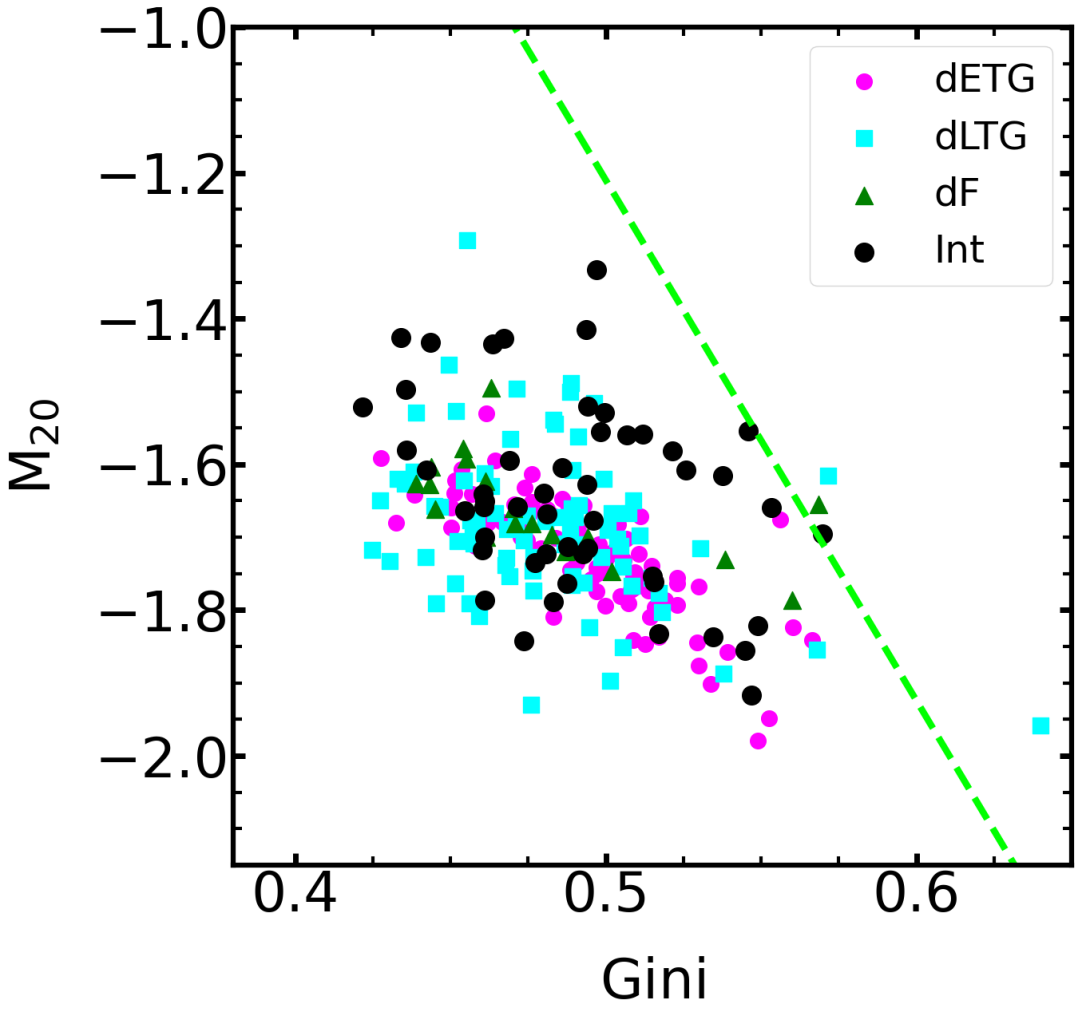

FIGURE 2.18: M_{20} vs Gini coefficient for our dwarf sample. The light green dashed line represents the criterion from Lotz et al. (2008) which is often used to separate interacting galaxies in the massive-galaxy regime. The black circles represent dwarf galaxies which are interacting.

The lower incidence of interactions and the less concentrated light profiles both suggest that the origin of dwarf ETGs might have less to do with interactions and merging than their massive counterparts. This appears to be in line with the findings of recent observational work that suggests that secular accretion from the cosmic web is the dominant evolutionary channel for these systems (Lazar et al., 2023). This observational picture also appears consistent with theoretical studies which have explored the formation mechanisms of dwarf ETGs.

For example, some studies postulate the formation of dwarf ETGs through monolithic collapse, with the subsequent star formation history regulated by internal processes like stellar and supernova feedback (e..g Yoshii & Arimoto, 1987; Chiosi & Carraro, 2002). Other studies postulate a 'harassment' scenario, in which dwarf ETGs are created from dwarf LTGs as they orbit massive galaxies. This takes place both via tidal perturbations, which remove angular momentum and make the system more dispersion

dominated, and the tidal stripping of gas (e.g. Mayer et al., 2001; Marcolini et al., 2003). While such a scenario is more likely in dense environments like groups and clusters, it is plausible that it also operates is relatively less dense environments like the ones explored in this study. A further potential channel of dwarf ETG formation is in the tidal tails created by the gas-rich interactions of massive galaxies (e.g. Duc et al., 2004), although the overall fraction of dwarfs that is likely to be created via this channel is expected to be very small (e.g. Kaviraj et al., 2012) and a tidal origin does not easily explain the formation of large numbers of dwarf ETGs in low-density environments (such as the ones studied here).

2.5.2 Featureless dwarfs in low-density environments are formed via baryonic feedback

Section 2.3 indicates that a minority ($\sim$10 per cent) of our dwarfs fall in the featureless category. Interestingly, all of these systems are in the lower half of the dwarf stellar mass range considered in this study (10^8 M$_\odot$ $< M_\star <$ $10^{8.75}$ M$_\odot$), suggesting that the formation of these featureless galaxies becomes progressively easier as stellar mass decreases. This, in turn, suggests that the formation of featureless galaxies is driven by processes whose impact becomes stronger as the depth of the gravitational potential well becomes shallower e.g. environmental processes like tidal perturbations and ram pressure or baryonic feedback from supernovae or AGN.

However, the featureless systems are found at similar local densities as the other morphological classes (Section 2.3.2) and we have shown that the COSMOS2020 footprint does not contain dense environments in the redshift range of our study (Section 2.2.1). Finally, Section 2.3.4 shows that the featureless galaxies exhibit a range of rest-frame colours, indicating that they are not quenched as a population and that many show evidence for recent star formation. These points suggest that featureless galaxies in low-density environments, at least in our dwarf sample, are likely to be created, not by environmental processes but by internal mechanisms such as baryonic feedback, either from supernovae and/or AGN (e.g. Kaviraj et al., 2019; Koudmani et al., 2022).

2.5.3 Shallower potential wells cause late-type dwarfs to be structurally more asymmetric than massive late-types

Figure 2.15 indicates that dwarf LTGs show higher median asymmetry than their massive counterparts. They also extend to much larger asymmetry values than massive LTGs. This is likely caused by the fact that the shallower potential wells in the dwarf LTGs make them more susceptible to internal displacement of material. For example, baryonic feedback may be able to move gas around more easily, making the gas

reservoir more asymmetric. Stars that form from this gas will then inherit this asymmetry. Alternatively, tidal perturbations due to nearby large-scale structure are likely to be able to alter the stellar distribution in dwarfs more easily, also inducing larger asymmetries.

2.6 Summary

Dwarf galaxies dominate the galaxy number density at all epochs, making them fundamental to understanding the evolution of the Universe. While dwarfs have been studied in detail in the very local Universe (e.g. within around 50 Mpc), typical dwarfs are difficult to study at cosmological distances in past large surveys (e.g. the SDSS), because they are too shallow. Outside the local neighbourhood, shallow surveys typically detect dwarfs that have high SFRs, which temporarily boost the luminosity of these galaxies, lifting them above the detection threshold of shallow surveys. However, this also makes the small subset of dwarfs that exist in such surveys biased which, in terms of morphology, skews these samples towards predominantly blue late-type galaxies (as we have shown in Appendix A.1).

Quantifying the morphological mix of dwarfs, outside our immediate neighbourhood, in low-density environments requires surveys that are both deep and wide and which can therefore provide unbiased statistical samples of dwarf galaxies. Here, we have constructed such a sample of 257 dwarfs, which lie in the stellar mass range 10^8 M$_\odot$ $< M_\star <$ $10^{9.5}$ M$_\odot$ and in the redshift range $z < 0.08$. We have first performed visual inspection, using ultra-deep HSC images (and their unsharp-masked counterparts) that are around 5 magnitudes deeper than standard-depth SDSS imaging, to establish the principal morphological classes in the dwarf regime. We have then explored the local densities, the role of interactions, rest-frame colours and the incidence of bars in these morphological classes. This is followed by an exploration of commonly-used morphological parameters in the dwarf regime and a comparison to both dwarf and massive galaxies in the very local Universe (within 50 Mpc) and the massive-galaxy population in the nearby Universe, using DES at $z < 0.1$. The overall aim has been to provide a pilot study that offers a useful benchmark for the study of dwarf morphology using new and forthcoming deep-wide surveys such as Euclid and LSST. Our main results are as follows:

- Visual inspection reveals three broad morphological classes in the dwarf population in our stellar mass range of interest (10^8 M$_\odot$ $< M_\star <$ $10^{9.5}$ M$_\odot$). Around 43 and 45 per cent of our dwarfs exhibit early-type (elliptical/S0) and late-type (spiral) morphology respectively. Around 10 per cent of dwarfs populate a ‘featureless’ class, that lacks both the central light concentration seen in early-types and any spiral structure. While the dwarf ETGs and LTGs are visually similar to the

ETG and LTG classes found in massive galaxies, the featureless class does not have a counterpart in the massive-galaxy regime.

- Dwarf ETGs diverge strongly from their massive counterparts in both their structural and photometric properties. They show an incidence of interactions and dust lanes which are several factors lower than that seen in massive ETGs. They are also significantly less concentrated and, unlike massive ETGs, share similar rest-frame colours as dwarf LTGs. This suggests that, unlike their massive counterparts, the formation of dwarf ETGs may be driven less by interactions and more by secular processes over cosmic time.

- The COSMOS footprint does not contain large groups or clusters at $z < 0.08$ and the local density of the various dwarf morphological classes, traced via projected distances to the first to the tenth nearest massive neighbours, do not show significant differences. This implies that the creation of the featureless dwarfs in low-density environments is likely to be driven by internal baryonic (stellar or AGN) feedback rather than by environmental processes.

- Around 20 per cent of the star formation activity in dwarfs, in the stellar mass and redshift range probed in this study, is likely to be driven by interactions.

- The (strong) bar fraction in dwarf galaxies is around 11 per cent, consistent with the recent literature and lower than that found in the massive-galaxy regime ($\sim$20 per cent). Unlike the massive galaxy regime, the median rest-frame colour of (strongly) barred dwarfs is not significantly redder than that of unbarred dwarfs.

- It is challenging to separate different dwarf morphological classes using commonly-used parameters like the Sérsic index, CAS, M_{20} and the Gini coefficient. This is largely driven by the fact that dwarf ETGs are less concentrated than their massive counterparts. Thus, the light concentration, which is a key separator between morphological classes in the massive-galaxy regime, does not offer the same level of discrimination in the dwarf regime.

- The asymmetry of dwarf ETGs and LTGs that are interacting is larger, by approximately factors of 4 and 3 respectively, than their non-interacting counterparts. Asymmetry thresholds of 0.05 and 0.08 respectively are able to select samples of dwarf ETGs and LTGs in which more than 50 per cent of the galaxies are interacting (in images that have similar resolution and depth to the ones used in this study).

Broadly speaking, there is evidence that a transition in morphological behaviour occurs between the dwarf and massive regimes (i.e. around $M_\star \sim 10^{9.5}$ $\mathrm{M}_\odot$). While similar morphological classes (ETGs

and LTGs) exist in both regimes, they have different formation mechanisms and star formation histories. In addition, new morphological classes, such as the featureless objects emerge in the dwarf regime, which do not share the characteristics of the classical ETG and LTG populations. It is likely that galaxies may show even more morphological diversity at stellar masses lower than the limit of our study (i.e. at $M_\star < 10^8$ $M_\odot$).

Our study poses several questions that need to be addressed in future work. What is the relationship between the different dwarf morphological classes? Can dwarf LTGs transform into dwarf ETGs as is seen in the massive-galaxy regime and, if so, via what processes? Could the featureless dwarfs be short-lived systems, as might be suggested by their low number fractions? Or do they transform, over short timescales, into the other morphological types? How do these relationships vary as a function of stellar mass and environment? Answering these questions requires statistical explorations of how the morphological mix of dwarfs evolves over cosmic time. While this is beyond the scope of this study, it will form the basis of future investigations using both ground and space-based data from surveys like LSST and Euclid.

Given the time-consuming nature of visual inspection, the challenges in separating dwarf morphological classes using parameters and the fact that millions of dwarfs will be imaged by forthcoming deep-wide surveys (e.g. LSST and Euclid), future statistical studies of dwarf morphology will likely need to leverage machine-learning techniques. Given the differences between the dwarf and massive regimes, there is a need to construct training sets (e.g. using systems like Galaxy Zoo) that can be used for dwarf classifications based on supervised machine-learning (e.g. Walmsley et al., 2022). Unsupervised machine learning techniques (e.g. Martin et al., 2020; Lazar et al., 2023), which can perform morphological clustering of arbitrarily large galaxy populations from wide-area surveys, are also likely to be important for future morphological studies in the dwarf regime.

To conclude, studying the evolution of dwarfs over the lifetime of the Universe, using statistically significant samples of dwarf galaxies from deep-wide surveys like LSST and Euclid, is likely to become a significant endeavour in the coming decades. Such studies will be capable of bringing revolutionary advances in our understanding of the physical processes that drive galaxy evolution.

Chapter 3

The structure of nearby dwarf galaxies as a function of stellar mass and morphology – size, effective surface brightness and colour gradients

3.1 Introduction

Dwarf galaxies ($M_\star < 10^{9.5}$ $M_\odot$) dominate the galaxy census in all environments and at all redshifts (e.g. Ferguson & Binggeli, 1994; Baldry et al., 2012; Alavi et al., 2016), making them important for a complete comprehension of galaxy evolution. The physical characteristics of dwarfs make them particularly useful for probing the physics of galaxy evolution. For example, their shallow potential wells make dwarfs more sensitive probes of key physical processes, such as baryonic feedback, tidal perturbations, and ram pressure than their massive counterparts (e.g. Martin et al., 2019; Jackson et al., 2021a). Similarly, the high dark-matter fractions in these systems (e.g. Kirby et al., 2015; Jackson et al., 2021b; Bruce et al., 2023; Jackson et al., 2024), make them good laboratories for probing the nature of dark matter (e.g. Ackermann et al., 2015).

In spite of their utility as cosmological probes, much of our current understanding of dwarf galaxies comes from studies in the very nearby Universe. This is because typical dwarfs are not bright enough to be detectable in the shallow surveys, e.g. the SDSS (Alam et al., 2015), that have dominated the astrophysical landscape over the past few decades (e.g. Lazar et al., 2024). The dwarfs that do appear

in these surveys tend to have high star formation rates (SFRs) which make them brighter and detectable in shallow surveys (e.g. Jackson et al., 2021c). However, this also biases them towards objects that tend to be bluer and may skew the morphological mix towards late-type systems (e.g. Lazar et al., 2024). Unbiased studies of the dwarf population outside our local neighbourhood requires surveys that are both deep and wide (as is the case here).

A rich literature exists on key structural properties in massive galaxies, such as galaxy size (typically parametrised by the half-light or 'effective' radius), the effective surface-brightness (defined as the mean surface brightness within the effective radius) and colour profiles and gradients. In the massive-galaxy regime, these structural parameters tend to change both as a function of stellar mass and morphology. For example, around $M_\star \sim 10^{10}$ $M_\odot$, massive early-type galaxies (ETGs) have effective radii that are around a factor of 2 smaller and effective surface brightnesses that are around a factor of 2.5 brighter than late-type galaxies (LTGs). However, at higher stellar masses ($M_\star \sim 10^{11}$ $M_\odot$) massive ETGs and LTGs exhibit similar radii and effective surface brightnesses (e.g. Shen et al., 2003; Graham & Worley, 2008), while beyond this value ETGs become larger than their LTG counterparts. The colour gradients in massive ETGs are generally flat or negative i.e. ETGs are typically either red throughout or redder in their centres (e.g. De Propris et al., 2005; La Barbera et al., 2012; Ferreras et al., 2017), with the bluer outer regions attributed to late-epoch satellite accretion. Massive LTGs, on the other hand, generally have negative colour gradients in the inner regions and flat or positive gradients in the outskirts (e.g. D'Souza et al., 2014; Tortora et al., 2010; Liao & Cooper, 2023). The progressive reddening in the outskirts is attributed to non-axisymmetric features, such as spirals or bars and/or a star formation cut off due to a surface density threshold (e.g. Roškar et al., 2008; Martin & Kennicutt, 2001).

While dwarf galaxies in shallow surveys which reside outside the local neighbourhood are likely to show the biases outlined above, relatively unbiased explorations of dwarfs are possible in the Local Group (e.g. McConnachie, 2012) or in the very nearby Universe via surveys such as MATLAS (e.g. Duc et al., 2015), NGVS (Ferrarese et al., 2012, 2020) and FDS (e.g. Venhola et al., 2017, 2019) which identify dwarfs either around nearby massive galaxies or in nearby groups and clusters. By construction, such studies offer insights into dwarf galaxy evolution in relatively high-density environments. In such environments, past work has found four main morphological classes: dwarf ellipticals, which are systems with central light concentrations and smooth light distributions like those found in the massive-galaxy regime (e.g. Graham & Guzmán, 2003; Janz et al., 2017), dwarf 'spheroidals', which are diffuse low-surface-brightness systems which lack a central light concentration (e.g. Irwin et al., 1990; Wilkinson et al., 2006; Simon & Geha, 2007; Koda et al., 2015), dwarf spirals, which are rotationally-supported systems with spiral structure (e.g. Schombert et al., 1995; Graham & Guzmán, 2003) and dwarf irregulars

which appear chaotic in their structural appearance and are thought to host significant amounts of gas (e.g. Gallagher & Hunter, 1984; Hunter et al., 2024).

Several studies (e.g. McConnachie, 2012; van Dokkum et al., 2015; Eigenthaler et al., 2018; Lim et al., 2020; Ferrarese et al., 2020; Poulain et al., 2021) have compared the structural properties of these types. For example, at a stellar mass of $M_\star \sim 10^8$ M$_\odot$ the effective radii of the dwarf ellipticals and irregulars are similar ($\sim$ 1 kpc in the g-band) with the dwarf spheroidals typically larger by around a factor of 2. At such stellar masses the typical effective surface brightnesses of dwarf ellipticals and irregulars (in the g-band) are $\sim$ 24 mag arcsec^{-2}, while dwarf spheroidals are a factor of $\sim$4 fainter. At lower stellar masses, e.g. $M_\star \sim 10^7$ M$_\odot$, all dwarf morphological types show a similar large scatter around effective radii and surface brightnesses values of R$_{e,g} \sim 0.7$ and $\mu_{e,g} \sim 25$ mag arcsec^{-2} respectively. The colour gradients in dwarfs that reside in high-density environments are found to mostly vary from flat or negative at the upper end of the stellar masses spanned by dwarfs ($M_\star \sim 10^{9.5}$ M$_\odot$), to flat or positive at lower stellar masses ($M_\star \sim 10^8$ M$_\odot$, e.g. Vader et al., 1988; Bremnes et al., 2000; Jansen et al., 2000; Parodi et al., 2002; Tortora et al., 2010; den Brok et al., 2011). Dwarf irregulars, in particular, generally have positive gradients (e.g. Parodi et al., 2002).

While the studies described above have shaped our understanding of dwarfs in nearby high-density regions, much less is known about the bulk of the dwarf population that lives in low-density environments. To address this, we have constructed, in a recent paper (Lazar et al., 2024), a mass-complete, unbiased sample of 257 dwarf (10^8 M$_\odot$ $< M_\star <$ $10^{9.5}$ M$_\odot$) galaxies at $z < 0.08$ in the COSMOS field which, at these redshifts hosts galaxies in groups and the field i.e. low-density environments. Visual inspection of ultra-deep optical images of these dwarfs from the Hyper Suprime-Cam reveals three principal morphological classes in dwarfs that inhabit low-density environments: ETGs, i.e. elliptical/S0 systems, LTGs, which show evidence of disks and 'featureless' dwarfs which show neither the central light concentration seen in ETGs nor any spiral structure that typifies LTGs. 43, 45 and 10 per cent of dwarfs correspond to the ETG, LTG and featureless class, while a small fraction (2 per cent) are morphologically irregular. The featureless class is akin to the dwarf spheroidals in high-density regions. However, Lazar et al. (2024) label them as featureless rather than spheroidal because, at least in the field, the featureless systems are not produced by cluster-specific processes that are thought to give rise to the dwarf spheroidals in high-density environments. Indeed, as we show later in this study, the featureless dwarfs deviate strongly from the ETGs in their structural properties, which suggests that the featureless galaxies are not a subset of the ETG (i.e. spheroidal) population. It is worth noting that, while the dwarf ETGs and LTGs are akin to the well-established morphologies seen in the massive-galaxy regime (e.g.

Hubble, 1936; Lintott et al., 2011; Kaviraj, 2014a), the featureless class is essentially missing in the massive-galaxy regime.

Lazar et al. (2024) (or Chapter 2 in this thesis) have used this sample to study the frequency of dwarfs in each morphological class and their key properties e.g. colours, local environments, incidence of interactions and the extent to which the visually classified morphologies can be separated using standard morphological parameters. This chapter is a companion study, which focuses on exploring the structural properties of the Lazar et al. dwarfs, as a function of their stellar mass and morphology. In particular, we probe typical structural parameters that underpin similar work in the massive-galaxy regime: size (parameterised by the effective radius), effective surface brightness and colour profiles and gradients. The overall aim of this study, when combined with Lazar et al. (2024), is to establish a low-redshift benchmark for the morphological and structural properties of dwarf galaxies in low-density environments.

This chapter is organised as follows. In Section 3.2, we describe the sample of nearby dwarf galaxies, from Lazar et al. (2024), that underpins this study. In Sections 3.3 and 3.4, we explore the effective radii (R_e) and the effective surface brightness (μ_e) of our dwarfs as a function of stellar mass and morphology[1]. In Section 3.5, we study colour gradients within our dwarf population as a function of stellar mass and morphology. We summarise our findings in Section 3.6.

3.2 A sample of nearby dwarf galaxies

Our study is based on the dwarf galaxy sample constructed by Lazar et al. (2024). In this section we describe aspects of the creation of this catalogs that are relevant to our study and direct readers to Lazar et al. (2024) for further details on its construction. The sample is assembled using the Classic version of the COSMOS2020 catalogue (Weaver et al., 2022). This catalogue provides accurate physical parameters (e.g. photometric redshifts, stellar masses and SFRs) for $\sim$1.7 million sources in the $\sim$2 deg^2 COSMOS field (Scoville et al., 2007). The parameter estimation employs deep photometry in 40 broad-band filters spanning the UV through to the the mid-infrared, from the following instruments: GALEX (Zamojski et al., 2007), MegaCam/CFHT (Sawicki et al., 2019), ACS/HST (Leauthaud et al., 2007), Hyper Suprime-Cam (Aihara et al., 2019), Subaru/Suprime-Cam (Taniguchi et al., 2007, 2015), VIRCAM/VISTA (McCracken et al., 2012) and IRAC/Spitzer (Ashby et al., 2013; Steinhardt et al., 2014; Ashby et al., 2015, 2018). Aperture photometry in the optical and infrared filters is extracted using the SExtractor and IRACLEAN codes respectively. The physical parameters are then calculated using

[1]Recall that the effective radius is defined as the radius that contains half the light of the galaxy.

	Median R_e (arcsec)				
Morphology	All galaxies	Red	Blue	Interacting	Non-interacting
dETG	$1.18\,^{0.08}_{0.03}$	$1.27\,^{0.12}_{0.14}$	$1.17\,^{0.03}_{0.03}$	$1.65\,^{0.17}_{0.17}$	$1.14\,^{0.03}_{0.04}$
dLTG	$2.06\,^{0.07}_{0.07}$	$2.09\,^{0.16}_{0.28}$	$2.05\,^{0.08}_{0.08}$	$2.22\,^{0.17}_{0.17}$	$2.00\,^{0.08}_{0.05}$
dF	$1.37\,^{0.06}_{0.11}$	$1.37\,^{0.18}_{0.22}$	$1.37\,^{0.12}_{0.11}$	$1.50\,^{0.17}_{0.17}$	$1.33\,^{0.09}_{0.20}$
All morphologies	$1.61\,^{0.04}_{0.07}$	$1.40\,^{0.08}_{0.05}$	$1.66\,^{0.08}_{0.05}$	$1.94\,^{0.11}_{0.11}$	$1.49\,^{0.08}_{0.05}$

TABLE 3.1: Median effective radius in the *i*-band for different galaxy populations. The errors on the medians (calculated via bootstrapping) are indicated using superscripts and subscripts. The threshold between the red and blue populations is at rest-frame $(u-i)$ = 1.7 (see justification in Appendix B.1).

	Median μ_e (mag arcsec^{-2})				
Morphology	All galaxies	Red	Blue	Interacting	Non-interacting
dETG	$22.9\,^{0.1}_{0.3}$	$23.1\,^{0.2}_{0.1}$	$22.5\,^{0.3}_{0.1}$	$22.8\,^{0.6}_{0.5}$	$22.9\,^{0.1}_{0.3}$
dLTG	$23.2\,^{0.1}_{0.1}$	$22.9\,^{0.3}_{0.2}$	$23.2\,^{0.1}_{0.1}$	$23.2\,^{0.1}_{0.2}$	$23.2\,^{0.2}_{0.1}$
dF	$24.0\,^{0.1}_{0.2}$	$24.4\,^{0.2}_{0.2}$	$23.6\,^{0.1}_{0.1}$	$23.8\,^{0.4}_{0.3}$	$24.0\,^{0.1}_{0.2}$
All morphologies	$23.2\,^{0.1}_{0.1}$	$23.4\,^{0.3}_{0.3}$	$23.1\,^{0}_{0.1}$	$23.3\,^{0.1}_{0.2}$	$23.2\,^{0.1}_{0.1}$

TABLE 3.2: Median effective surface brightness in the *i*-band for different galaxy populations. The errors on the medians (calculated via bootstrapping) are indicated using superscripts and subscripts. The threshold between the red and blue populations is at rest-frame $(u-i)$ = 1.7 (see justification in Appendix B.1).

the LEPHARE SED-fitting algorithm (Arnouts et al., 2002; Ilbert et al., 2006). The wide wavelength baseline results in photometric redshift accuracies better than ∼1 and ∼4 per cent for bright ($i < 22.5$ mag) and faint ($25 < i < 27$ mag) galaxies respectively.

The dwarf galaxy sample is constructed by selecting objects that are classified as galaxies by LEPHARE ('type' = 0 in the COSMOS2020 catalogue), exhibit an extendedness of 1 in the HSC *griz* filters, have stellar masses in the range 10^8 $M_\odot < M_\star < 10^{9.5}$ $M_\odot$, redshifts in the range $z < 0.08$ and lie outside masked regions. The final sample created by Lazar et al. (2024) contains 257 dwarf galaxies, with median redshift and stellar mass errors of 0.02 and 0.08 dex respectively. As noted in Lazar et al. (2024), given the depth of the data, this dwarf galaxy sample is mass complete and therefore offers an unbiased statistical sample of galaxies which can be used to study the morphological and structural properties of the nearby dwarf population.

3.2.1 Morphological classification via visual inspection

The galaxies in this final sample are then visually classified, using optical *gri* colour-composite images, and their unsharp-masked counterparts, from the HSC-SSP Ultra-deep layer. This layer has a 5σ point

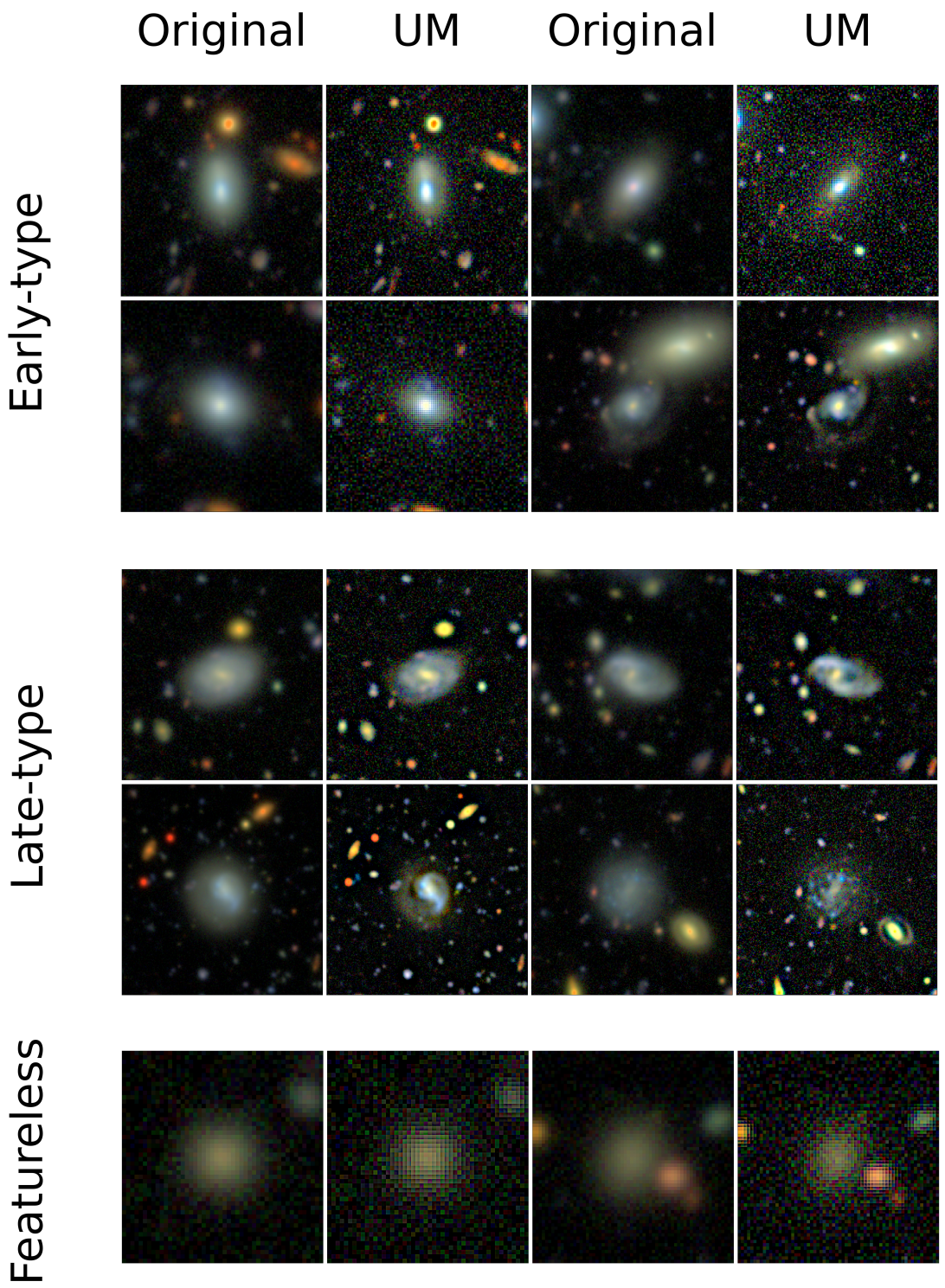

FIGURE 3.1: Examples of the three principal morphological classes – early-type, late-type and featureless galaxies – found by Lazar et al. (2024) in nearby ($z < 0.08$) dwarf galaxies. For each galaxy, we show the original *gri* colour-composite image constructed using ultra-deep Hyper Suprime-Cam images and its unsharp-masked counterpart (in the columns labelled 'UM'). The Hyper Suprime-Cam images have a point-source depth of $\sim$28 mag, around 6 magnitudes deeper than standard-depth SDSS imaging. Examples of interactions include internal asymmetries (e.g. row 1, cols 1/2 and row 4, cols 1/2), faint tidal features (e.g. row 1, cols 3/4) and tidal bridges in ongoing interactions (e.g. row 2, cols 3/4).

.

source depth of ∼28 magnitudes, around 5 magnitudes deeper than standard depth SDSS imaging and almost 10 magnitudes deeper than the detection limit of the SDSS spectroscopic main galaxy sample. The median HSC seeing is ∼0.6 arcseconds, around a factor of 2 better than the SDSS. The visual inspection is used to classify the dwarfs into three principal morphological classes: early-type galaxies (ETGs) i.e. elliptical and S0 systems, late-type galaxies (LTGs) which show evidence of disks and featureless galaxies which show neither the central light concentrations seen in ETGs or any disk structure that typifies LTGs. The inspection is also used to flag dwarfs that show evidence of interactions e.g. internal asymmetries, tidal features and tidal bridges that connect galaxies in ongoing mergers. We direct readers to Section 3 in Lazar et al. (2024) for more details of the morphological classification. Figure 3.1 shows an abridged version of the images presented in Lazar et al. (2024), with examples of dwarfs in the different morphological classes. Finally, as noted in Lazar et al. (2024), the COSMOS2020 galaxy population, in our redshift range of interest, resides preferentially in low-density environments (i.e. galaxies in groups and the field).

3.2.2 Masking, PSF correction and construction of surface-brightness profiles

In each band, we manually mask flux from interloper sources, such as background galaxies and foreground stars. We use the azimuthal flux interpolation method from Watkins et al. (2022) to reconstruct the missing flux in the masked regions (see Lazar et al. (2024) for further details). Figure 3.2 presents examples of the original *gri* colour-composite image (column 1) and the final image produced using this method (column 2), which is then used to calculate structural parameters. The morphology of each galaxy is indicated in the lower left corner of each *gri* galaxy image (dETG = dwarf ETG, dLTG = dwarf LTG, and dF = dwarf featureless).

We construct surface-brightness profiles for our dwarfs using the PYTHON-based galaxy ellipse fitter `photutils.isophote.Ellipse` (Jedrzejewski, 1987; Busko, 1996). The profiles are corrected for light smearing due to the PSF in the following way:

- We first fit 2D single-component Sérsic profiles (Sérsic, 1963), using the `astropy` module within the PYTHON package `statmorph`.
- We use the resultant Sérsic index and the half-light radius obtained from the fitting to produce a 2D model for each galaxy using `GALFIT` (Peng et al., 2002).
- We convolve this model with the Hyper Suprime-Cam PSF calculated by Montes et al. (2021). We use *g* and *i* band PSFs for the *g* and *i* band galaxy images and the *g*-band PSF for images in

other filters (because only g and i-band PSFs are available from Montes et al. (2021)). Note that, while the results presented in this study only involve the g and i bands, structural parameters are calculated in other bands (r and z) and are available to the reader upon request.

- We calculate surface-brightness profiles from the convolved and unconvolved models using the method described above and calculate the ratio between the two in each isophotal bin. This ratio is a measure of the effect of the PSF on the surface-brightness profile.
- Finally, we correct for light smearing due to the PSF by multiplying the initial surface-brightness profile in each isophotal bin by the ratio calculated in the previous step.

Examples of the final surface-brightness profiles for different dwarf morphologies are presented in column 3 of Figure 3.2. The corresponding $(g-i)$ colour profiles are presented in column 4. We also calculate curves of growth using `photutils.isophote.Ellipse` and apply the same PSF correction procedure to these profiles. Using the corrected curves of growth, we then calculate the effective radius (studied in Section 3.3) and effective surface brightness (studied in Section 3.4) for each dwarf galaxy.

Finally, we note that before the PSF correction procedure is applied, we eliminate, from our final sample, galaxies with a Sersic index higher than 4 and effective radii smaller than the PSF (i.e. 3 pixels). This is because such objects are either unresolved or exhibit sharply peaked surface brightness profiles which hinder the performance of the `GALFIT` 2D modelling. Our final sample of objects comprises 211 dwarf galaxies, out of the original sample of 257 in Lazar et al. (2024).

3.3 Size

We begin in Figure 3.3 by presenting galaxy sizes, traced using effective radii ($R_{\rm e}$), as a function of stellar mass for our dwarf population. Different morphological classes are shown using different symbols, while galaxies are colour-coded using their rest-frame $(u-i)$ colour. Interacting systems are indicated using crosses. In our redshift and stellar mass ranges of interest, the galaxy population is bimodal around $(u-i) \sim 1.7$ (see Figure B.1 in Appendix B.1). Red and blue galaxies are defined as those with rest-frame $(u-i)$ greater and less than 1.7 respectively. Table 3.1 summarises the median values of the effective radius for different galaxy populations. Note that, since we use photometric redshifts in this study, the physical sizes (in kpc) incur large errors when the uncertainties in the photometric redshifts are propagated through the conversion from angular and physical size. Therefore, we largely base our analysis in this section on the angular sizes (left-hand panel of this figure) which, in turn restricts us to

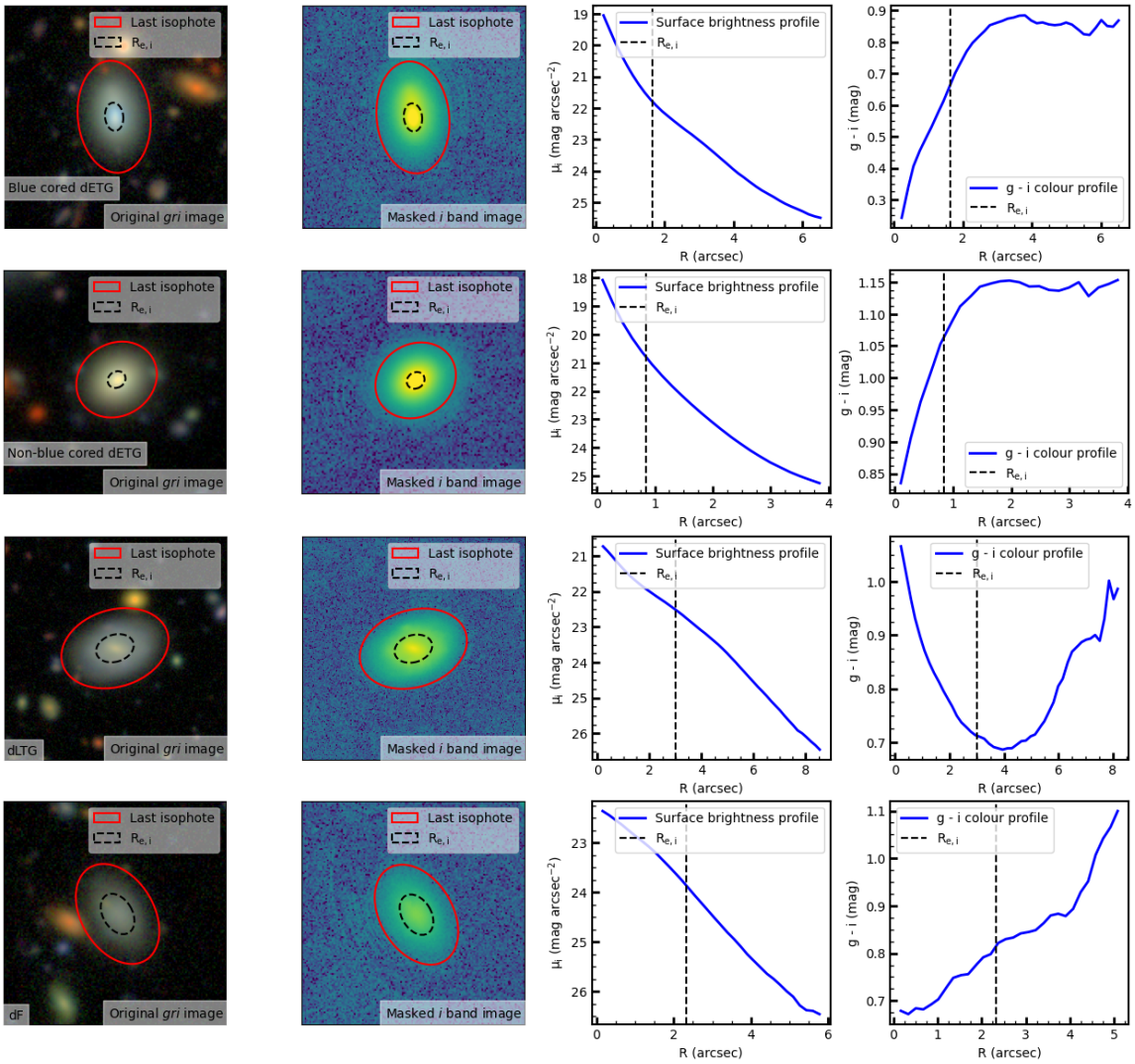

FIGURE 3.2: Examples of surface brightness and $(g-i)$ colour profiles of dwarfs in different morphological classes (shown in the different rows). The first column (from the left) shows the original *gri* colour-composite image of the galaxy. The second column shows the masked *i*-band image. The third column presents the *i*-band surface brightness profile ending at the last isophote (shown by the red ellipse in the first 2 columns). The last column presents the $(g-i)$ colour profile ending at the last isophote. The black dashed vertical line (corresponding to the black dashed ellipse in the second column) represents the *i*-band half-light radius (i.e. the effective radius). The morphology of each galaxy is indicated in the lower left corner of each *gri* galaxy image (dETG = dwarf ETG, dLTG = dwarf LTG, and dF = dwarf featureless).

studying the *relative* sizes between different morphological classes (rather than physical sizes). However, we also show the physical sizes for completeness in the right-hand panel of this figure.

Table 3.1 shows the median effective radii for different dwarf populations. When taking into account the whole dwarf sample Figure 3.3 and Table 3.1 together indicate that ETGs and LTGs are well separated in effective radius, with the median effective radius of ETGs being around a factor of 2 smaller than that of LTGs. The featureless dwarfs lie in between the ETGs and LTGs, with a median effective radius that is around a factor of 1.2 larger than their ETG counterparts. The effective radii of red (i.e. quenched) and blue (i.e. star-forming) galaxies do not show significant differences within each morphological class. When all morphological classes are considered together, red galaxies show a weak trend of being more compact than blue galaxies.

We note that since the redshift range in this study is z=0-0.08 the median effective radii calculated above may be biased by the significant change in physical scales probed. As a result, we recalculated the median effective radii for each morphological class for a narrow redshift range of z=0.06-0.07 which contains the highest statistical sample set (i.e. 44 LTGs, 38 ETGs and 12 featureless) and the R_e values obtained are similar as for the whole sample, within statistical error: $1.88^{0.13}_{0.16}$ arcsec, $1.12^{0.05}_{0.04}$ arcsec and $1.31^{0.09}_{0.2}$ arcsec for LTGs, ETGs and featureless, respectively. In order to investigate further the impact of the change in physical scales on our results we also split our sample in redshift bins in Figure 3.3. Here we show that for the highest redshift bin (z=0.06-0.08) we confidently retrieve the trends seen for the whole sample due to the fact that $\sim$70 per cent of our sample resides in that redshift range. For the redshift bin z=0.04-0.06 we notice that the difference between the median effective radii for the ETG and LTG populations loses significance. This may be caused by a statistical imbalance in the two populations (difference of a factor of 4) with 6 ETGs and 22 LTGs available for that redshift range. For the lowest redshift bin the distributions appear to be shifted to higher R_e values, as expected, since physical scales probed become smaller for lower redshifts. Furthermore, the difference in the medians of the effective radii for this redshfit range is similar as in the case using the whole sample, where the imbalance in statistical samples is now lower (difference of a factor of 2) between the two morphological sets with 10 ETGs and 20 LTGs. Therefore, the trends seen in effective radii for different morphologies do not change significantly when we probe the same physical scales unless there are statistical imbalances present. In the lower redshift bins we do not take into account the featureless galaxies since we find only 4 of them corresponding to these redshifts.

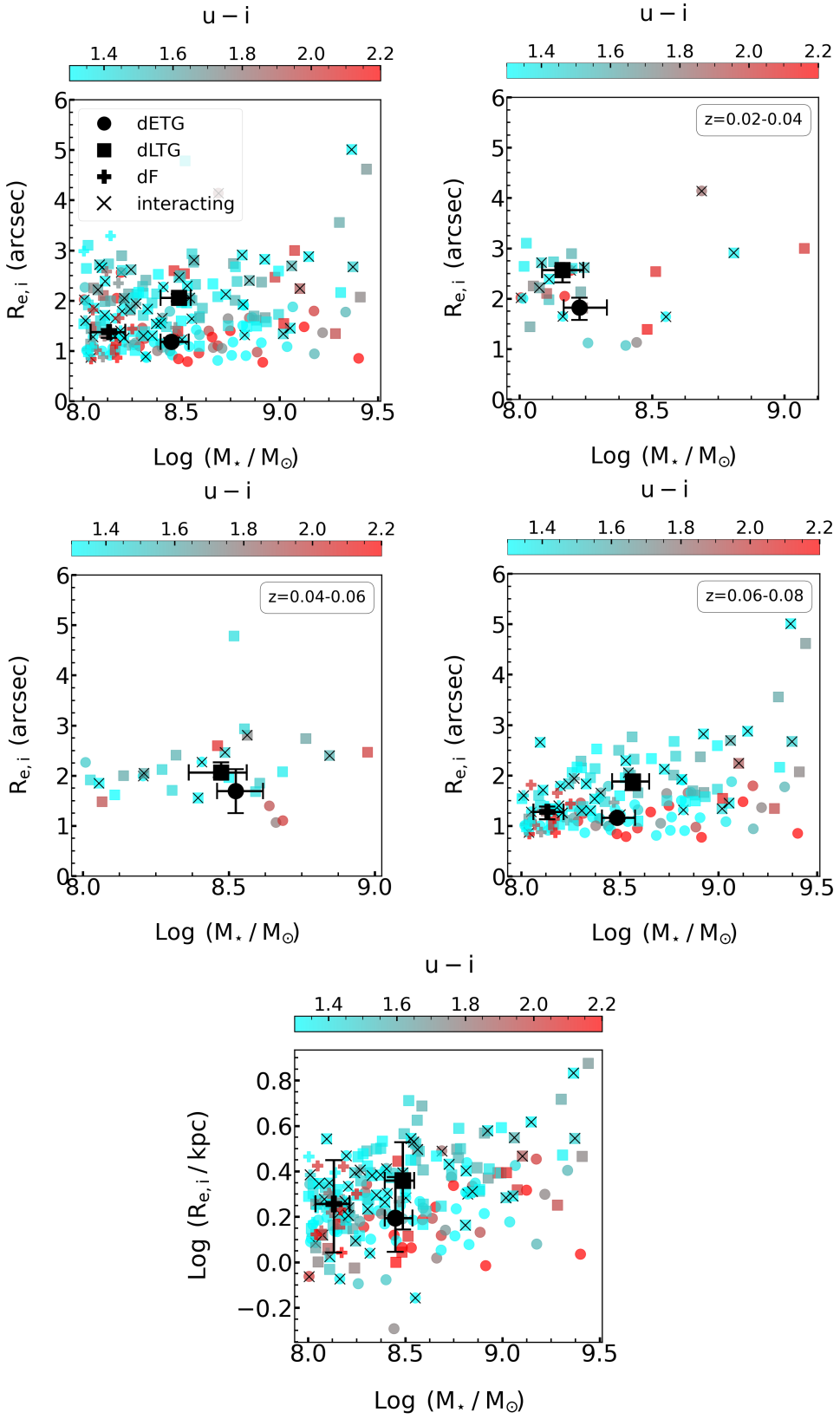

FIGURE 3.3: Effective radius in arcseconds (top and middle panels) and in kpc (bottom panel) vs stellar mass for our dwarf galaxies. The top-right and the middle panels show the sample split in redshift bins (redshift bin is shown in the top right corner of the panels). If there is no redshift bin specified in the panel then the whole sample is taken into account. Our sample does not contain objects with z<0.02. Different morphological classes are shown using different symbols, while galaxies are colour-coded using their rest-frame $(u-i)$ colour. Galaxies that are interacting are shown using crosses. The median values (and their uncertainties) for each morphological class are shown using the black symbols. Since we use photometric redshifts in this study, the physical sizes (in kpc) incur large errors when the uncertainties in the photometric redshifts are propagated through the conversion from angular to physical effective radii.

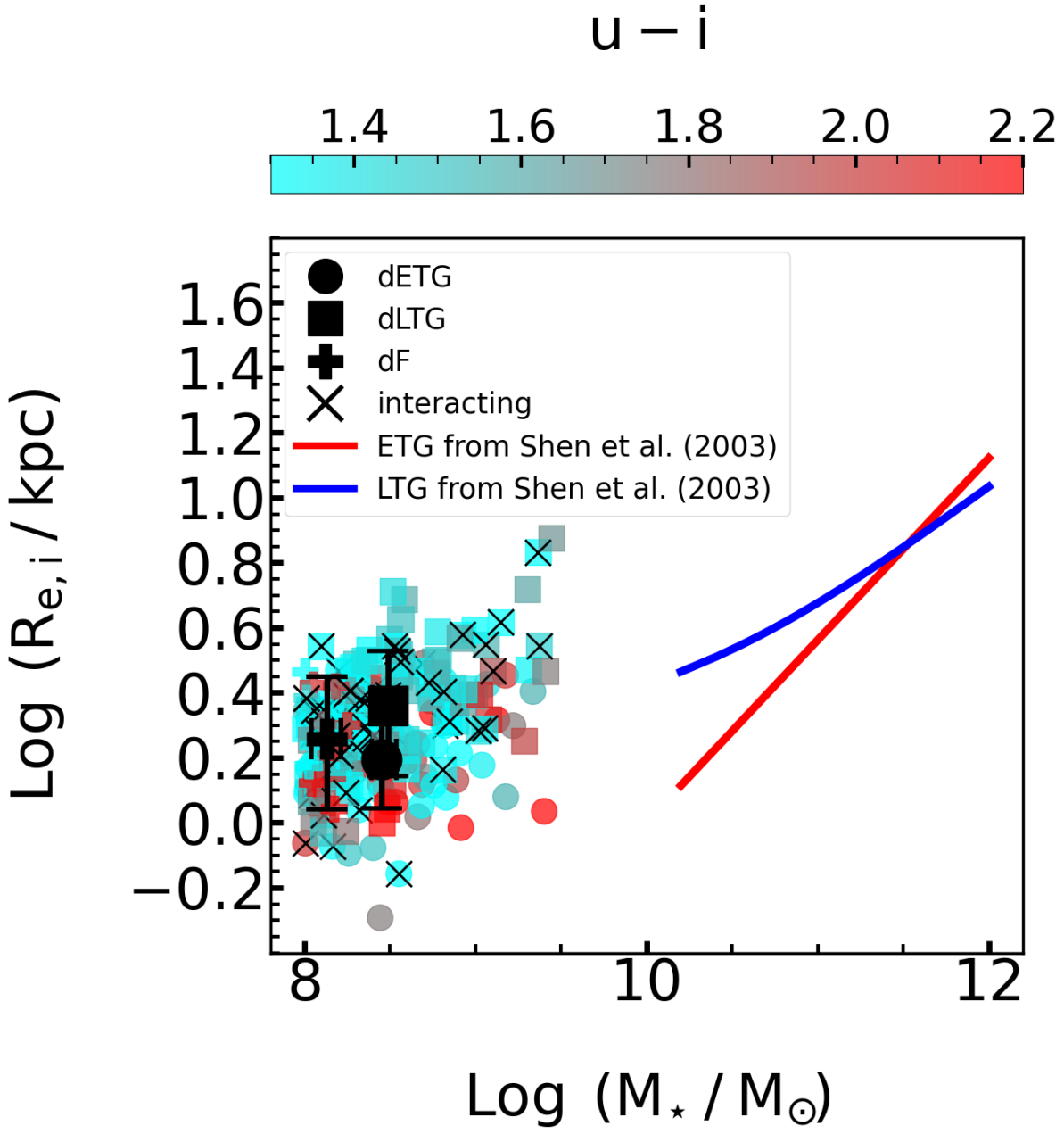

FIGURE 3.4: Same as top-left panel from Figure 3.3 but now including the least square fits of the ETG (in red) and LTG (in blue) effective radii distributions as a function of stellar mass from Shen et al. (2003) which use z-band data from SDSS.

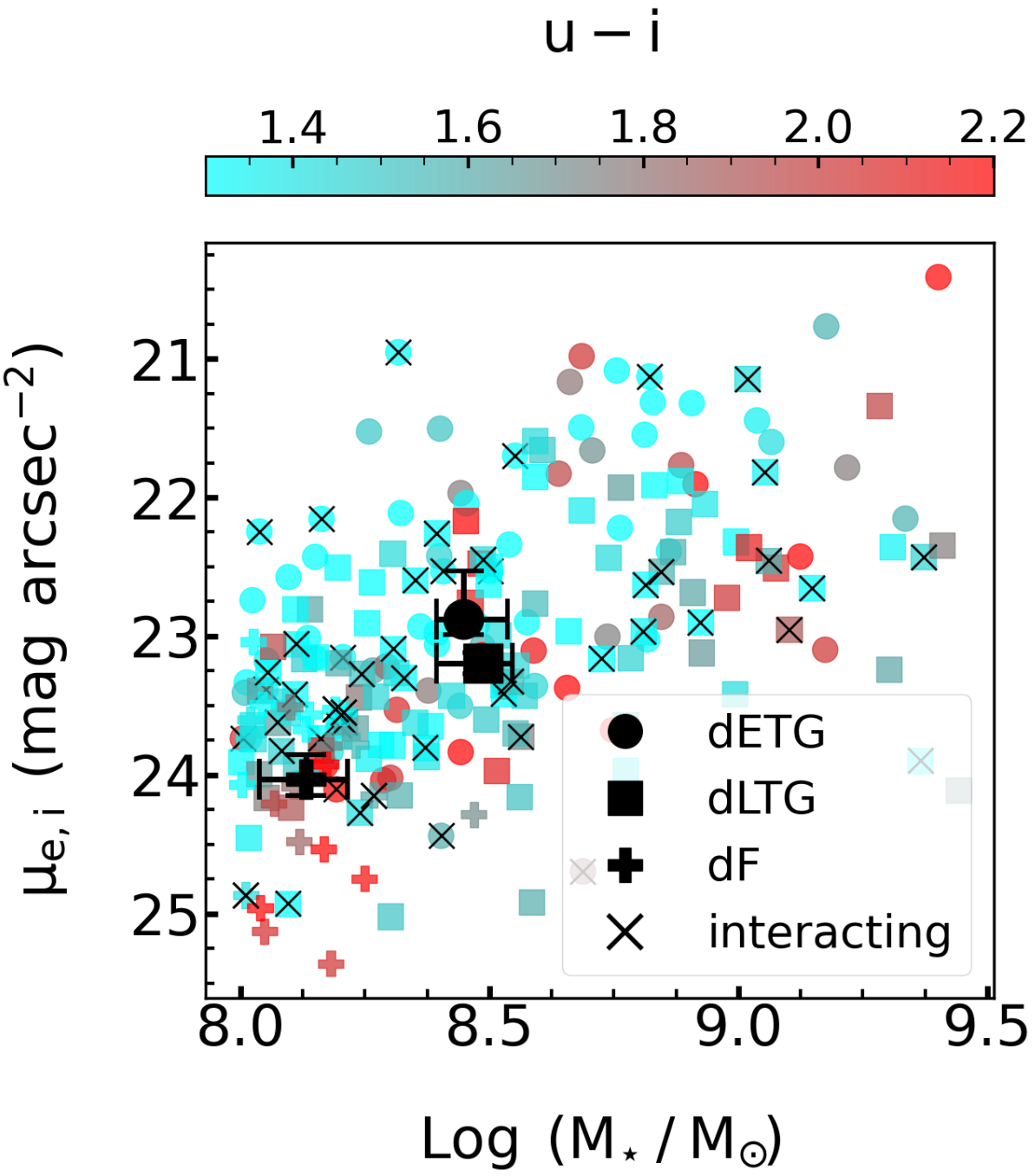

FIGURE 3.5: Effective surface brightness vs stellar mass for our dwarf galaxies. Different morphological classes are shown using different symbols, while galaxies are colour-coded using their rest-frame $(u-i)$ colour. Galaxies that are interacting are shown using crosses. The median values (and their uncertainties) for each morphological class are shown using the black symbols.

Table 3.1 indicates that interacting galaxies show modest differences in median effective radius compared to their non-interacting counterparts. The median difference when all morphologies are considered together is ∼33 per cent, suggesting that interactions act to puff up dwarfs and increase their sizes. The dwarf ETGs show the largest discrepancy between interacting and non-interacting galaxies, with the interacting ETGs being ∼42 per cent larger than their non-interacting counterparts.

We proceed by comparing the trends we find in our dwarf population to those that are known in the massive galaxy regime. Recent studies (e.g. Shen et al., 2003; Bernardi et al., 2014) have used SDSS data of massive ($M_\star > 10^{9.5}$ $M_\odot$) galaxies in the nearby Universe to compare galaxy sizes as a function of morphology. They find that the median effective radius of galaxies with $M_\star \sim 10^{10}$ $M_\odot$ is a factor of 2 higher in massive LTGs than in their ETG counterparts, similar to the ratio seen in our dwarf sample. However, the sizes of massive LTGs and ETGs become similar at $M_\star \sim 10^{11}$ $M_\odot$, beyond which massive ETGs actually exhibit larger effective radii than the LTGs.

It is worth considering our findings in conjunction with the results from these studies. We note first that past studies already indicate that the slope of the effective radius – stellar mass relation in ETGs becomes flatter as we move from the massive to the dwarf regime (see e.g. Figure 4 in Shen et al., 2003). In Figure 3.4 we combine our top-left panel from Figure 3.3 with the least square fits of the effective radii distributions from Shen et al. (2003). Figure 3.4 suggests that the slope of this relation is indeed much shallower in dwarf ETGs than in their massive counterparts. This apparent discontinuity in the slope of the effective radius – stellar mass relation in ETGs supports the notion (see e.g. Lazar et al., 2024) that the principal processes that dominate the evolution of ETGs are different in the dwarf and massive regimes. As noted already by Lazar et al. (2024), dwarf ETGs are likely to evolve primarily via secular process like gas accretion, while massive ETGs are influenced more by interactions. The general trend of LTG effective radii from the massive to the dwarf regime seems to change towards a shallower slope at around $M_\star \sim 10^{10.5}$ $M_\odot$ and does not present such discontinuity.

We conclude this section by considering the differences between red and blue galaxies in the massive and dwarf regimes. Recent work in the literature (e.g. Lange et al., 2015) suggests that in nearby ($z < 0.1$) galaxies with $M_\star \sim 10^{10}$ $M_\odot$ the median effective radius for star forming (blue) galaxies is a factor of 2 higher than the one for the quiescent (red) galaxies. The median effective radii for these sub-populations become similar at $M_\star \sim 10^{11}$ $M_\odot$. In the massive-galaxy regime, the relative trends between star-forming/blue and quiescent/red galaxies closely mirror those between the ETGs and LTGs described above. This is largely driven by the fact that, in this regime, ETGs and LTGs dominate the red and blue populations respectively. On the other hand, our dwarf sample does not show significant differences in

the median effective radius between red and blue galaxies because more than 50 per cent of our dwarf ETGs are optically blue, in contrast with the massive regime where optically blue ETGs are rare (e.g. Fukugita et al., 2004; Schawinski et al., 2009a; Kaviraj et al., 2007).

3.4 Effective surface brightness

Figure 3.5 presents the effective surface brightness[2] (μ_e) vs stellar mass for our dwarf galaxies. As in Figure 3.3 above, different morphological classes are shown using different symbols, while galaxies are colour-coded using their rest-frame $(u-i)$ colour. Interacting systems are indicated using crosses. A scaling relation is present, with μ_e becoming brighter with stellar mass. Table 3.2 summarises the median values of effective surface brightness for different galaxy populations. The dwarf ETGs are marginally brighter in this quantity that their LTG counterparts, with the featureless class being around 1 mag arcsec^{-2} fainter.

Not expectedly, red galaxies within each morphological class have fainter median μ_e values than their blue counterparts, with the difference being largest in the featureless galaxies. While interacting systems are larger, as described above, they have a similar median effective surface brightness as their non-interacting counterparts. This suggests that, as already noted in Lazar et al. (2024) (see also the discussion in Section 3.5 below), that interactions trigger star formation which boosts the surface brightness and offsets the increase in size. It is worth noting that the featureless class is similar to ETGs in terms of R_e but differs from the ETG population in terms of μ_e. We notice that we obtain similar results for a narrow mass range of 10^8 M$_\odot$ $<$ M$_\star$ $<$ $10^{8.5}$ M$_\odot$ as well, where our featureless galaxies mostly reside. The spatial distributions of baryons within these classes therefore show strong differences which suggests that their formation histories are distinct. In other words, the featureless class in low-density environments should not be considered a subset of the ETG population.

In the massive-galaxy regime, around a stellar mass of M$_\star$ $\sim$ 10^{10} M$_\odot$, ETGs are typically brighter in effective surface-brightness by a factor of 2.5 than LTGs (see e.g. Capaccioli et al., 1992; Bershady et al., 2000; Shen et al., 2003; Graham & Worley, 2008). As the stellar mass approaches M$_\star$ $\sim$ $10^{11.5}$ M$_\odot$, the median effective surface brightness for ETGs and LTGs becomes progressively more similar. The segregation of ETGs and LTGs, in terms of their effective surface brightness, is therefore more significant in the massive regime (M$_\star$ $>$ 10^{10} M$_\odot$) than for dwarfs, suggesting that the processes that dominate the evolutionary histories of these morphological types may be different in the two mass regimes.

[2]Recall that the effective surface brightness is defined as the mean surface brightness within the effective radius (R_e).

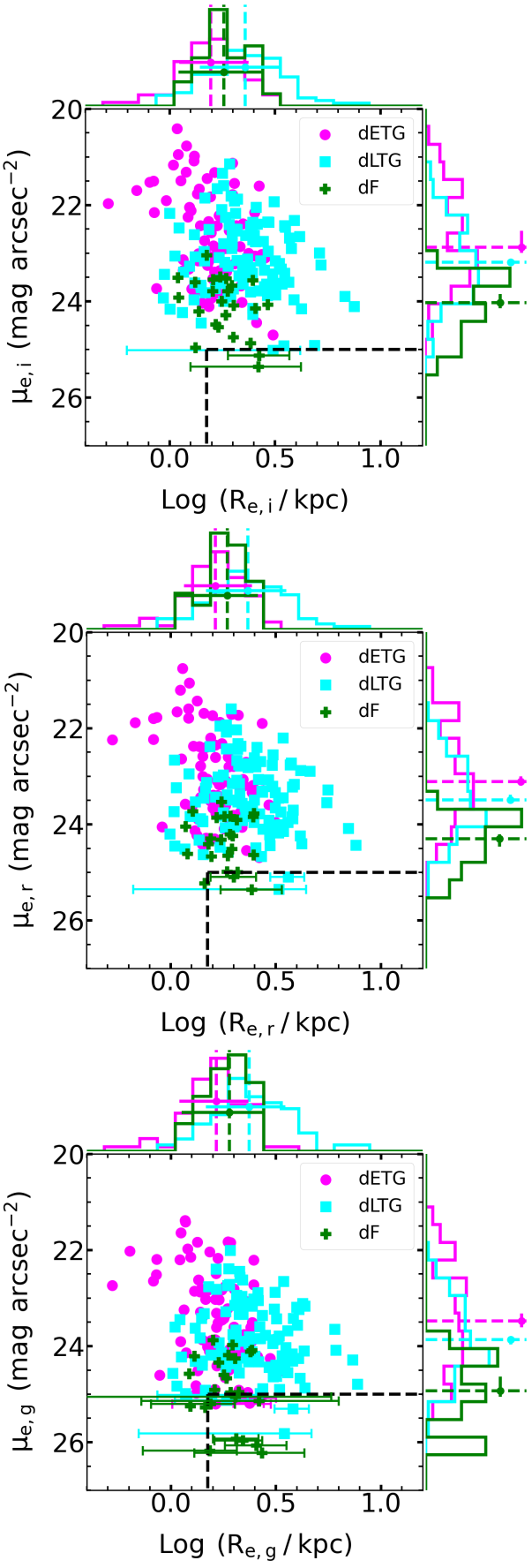

FIGURE 3.6: Effective surface brightness as a function of effective radius for our dwarf galaxies. The top, middle and bottom panels correspond to measurements in the *i*, *r* and *g* bands respectively. Different morphological classes are shown using different colours. The dashed line corresponds to the UDG selection criterion (e.g. Conselice, 2018). Error bars of galaxies that fall in the UDG region are shown (and omitted for other galaxies for clarity).

It is interesting to consider how our galaxies relate to 'ultra-diffuse galaxies' (UDGs), a population of faint, diffuse galaxies that exist across a variety of environments (e.g. van Dokkum et al., 2015; Koda et al., 2015; Benavides et al., 2021; Marleau et al., 2021) and have sometimes been considered to be a new class of object. In Figure 3.6, we plot μ_e vs R_e for our dwarfs in both the g and i bands, with typical UDG criteria indicated using a dashed rectangle ($\mu_e > 25$ mag arcsec^{-2} and $R_e > 1.5$ kpc, see e.g. Conselice, 2018). Our lower mass threshold ($M_\star > 10^8$ M$_\odot$) results in very few of our dwarfs satisfying the UDG criterion (the parameter space defined by this criterion is likely to be populated more by galaxies that have $M_\star < 10^8$ M$_\odot$ which will be fainter). However, the galaxies in our sample that reside in this region include members of all three morphological classes i.e. these galaxies are not morphologically distinct from the rest of the galaxy population. Furthermore, galaxies in this region appear to be a continuous extension of the galaxy population towards lower values of μ_e. This suggests that, rather than being a novel class of object, galaxies in low-density environments that satisfy the UDG criterion are simply part of the fainter and more diffuse end of the overall galaxy population. We note that a similar conclusion has been drawn about UDG-type systems in high-density environments by Conselice (2018). Finally, we note that, while the results presented in Figures 3.5 and 3.3 are based on i-band images, the relative trends do not change if other filters (e.g. g, r or z) are used.

3.5 Colour profiles and gradients

In Figure 3.7, we present $(g-i)$ colour profiles and colour gradients in our dwarf galaxies. Note that, throughout this section, to ensure that our results are not noisy, we only calculate the colour in a given radial bin if it has at least 10 data points. The top row presents these quantities for the full mass range spanned by our sample (10^8 M$_\odot > M_\star > 10^{9.5}$ M$_\odot$), while the middle and bottom rows correspond to the lower and upper halves of our mass range respectively. The left and right-hand columns show the $(g-i)$ colour and its gradient as a function of the radius normalised by R_e. As shown in Lazar et al. (2024), the featureless galaxies only appear in the lower half of our stellar mass range, as result of which this morphological class is missing in the bottom row. In each panel the solid line represents the running median value, while the shaded region indicates the error in the running median.

We first consider dwarfs across our full stellar mass range. Dwarf LTGs exhibit a 'U' shaped profile with a negative gradient in their central regions out to $\sim$1.5 R_e. At this point the $(g-i)$ colour reaches a plateau, which represents the bluest region of the galaxy. Beyond $\sim$2 R_e the colour progressively reddens and the gradient becomes positive. On the other hand, the dwarf ETGs exhibit positive colour gradients until $\sim$2.5 R_e, beyond which the colour gradient becomes negative and the colours in the outskirts of the

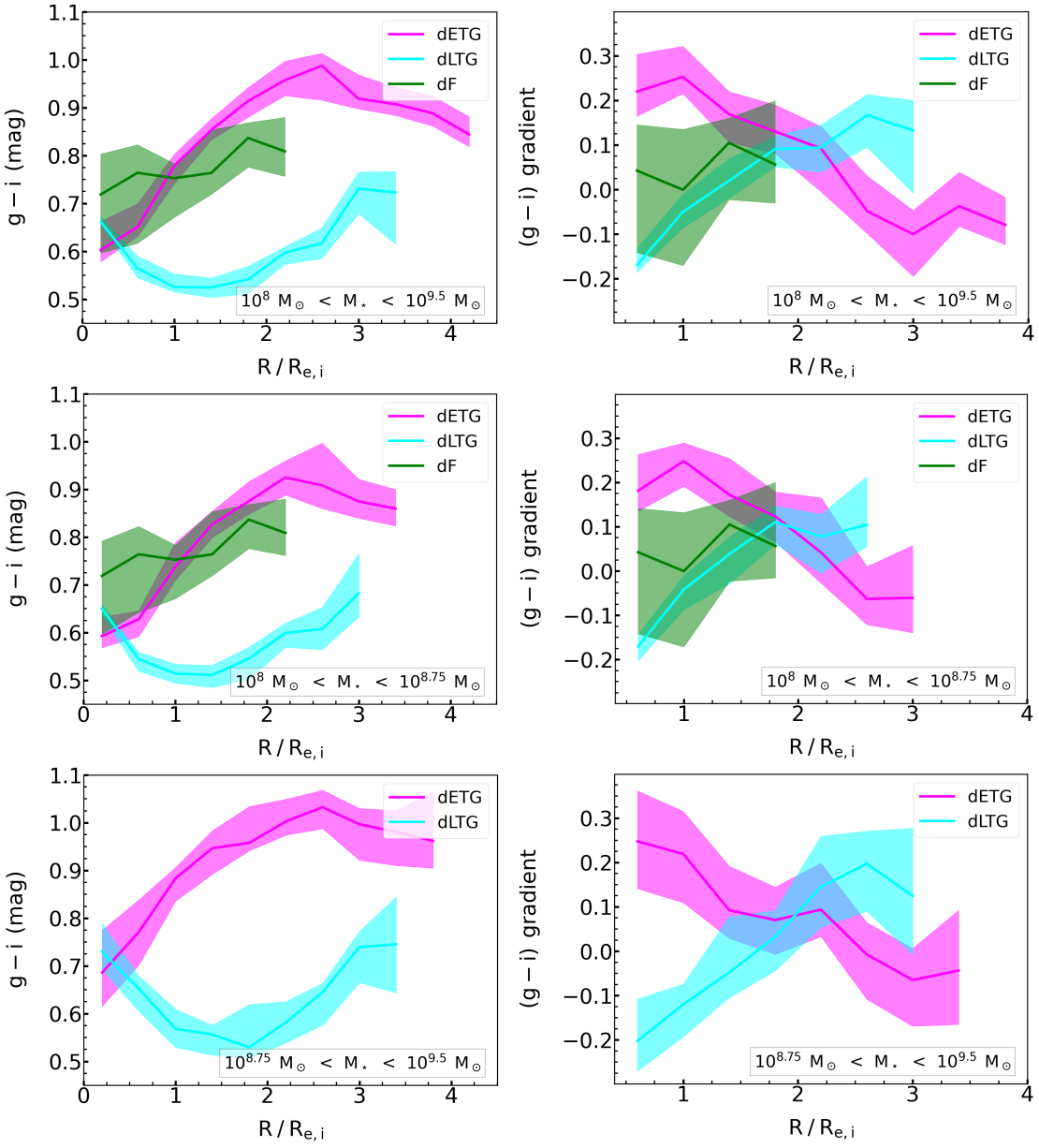

FIGURE 3.7: Median $(g-i)$ colour profiles (left-hand column) and colour gradients (right-hand column) in our dwarf galaxies. The top row presents these quantities for the full mass range spanned by our sample (10^8 M$_\odot$ $< M_\star <$ $10^{9.5}$ M$_\odot$), while the middle and bottom rows correspond to the lower and upper halves of our mass range respectively. The left and right-hand columns show the $(g-i)$ colour and its gradient as a function of the radius normalised by R_e, respectively. As shown in Lazar et al. (2024), all featureless galaxies occupy the lower half of our stellar mass range, as result of which this morphological class is missing in the bottom row. In each panel the solid line represents the running median, while the shaded region indicates its uncertainty. The gradient has been calculated by dividing the difference of the colour values of adjacent radial bins.

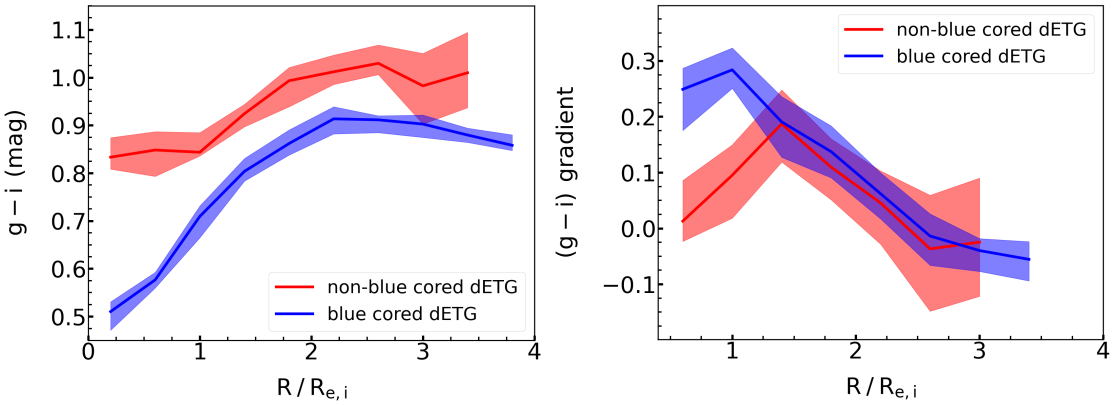

FIGURE 3.8: Median $(g-i)$ colour profiles (left-hand panel) and colour gradients (right-hand panel) in blue-cored (blue) and non blue-cored (red) dwarf ETGs. In each panel the solid line represents the running median, while the shaded region indicates its uncertainty.

dwarf ETGs become progressively bluer. It is worth noting that even though their colour gradients are the opposite of each other, dwarf ETGs and LTGs show similar colours in their central regions (i.e. at $<$ 0.5 R_e). Dwarf featureless galaxies show a shallow positive (i.e. close to flat) gradient throughout their colour profiles. They generally exhibit redder central regions (out to around $R < R_e$) than both ETGs and LTGs, suggesting the existence of comparatively older stellar populations. At larger radii ($R > R_e$) the featureless galaxies are slightly bluer than ETGs but significantly redder than their LTG counterparts.

The middle and bottom rows of Figure 3.7 show that there are no qualitative differences between the two mass bins in their colour profiles and gradients. However, we note that both the ETGs and LTGs have redder central regions (by $\sim$0.1 mag at $R < R_e$) in the higher mass bin compared to their counterparts in the low mass bin. Dwarf ETGs in the high mass bin also exhibit redder colours at all radii than ETGs in the lower mass bin.

We proceed by comparing our findings to what is known in the massive-galaxy regime. The majority of massive LTGs have negative colour gradients at $R < R_e$ which flatten or turn positive at outer radii (e.g. Moth & Elston, 2002; Taylor et al., 2005; Bakos et al., 2008; Liu et al., 2009; Tortora et al., 2010; D'Souza et al., 2014; Watkins et al., 2019; Liao & Cooper, 2023). This behaviour resembles that seen in our dwarf LTGs. The physical processes that cause the reddening of the colour in the outskirts (i.e. a "U" shaped color profile) are still a matter of debate but some models postulate that this behaviour could be driven by the outward radial migration of evolved stars via intrinsic secular processes (e.g. through spirals or bars Debattista et al., 2006; Roškar et al., 2008; Muñoz-Mateos et al., 2013) combined with a star formation cut-off at a certain surface density threshold in the outer regions of the galaxy (e.g. Martin & Kennicutt, 2001).

Massive ETGs ($M_\star > 10^{9.5}$ $M_\odot$) typically have flat or negative colour gradients (e.g. Kormendy & Djorgovski, 1989; Peletier et al., 1990; De Propris et al., 2005; Marian et al., 2018). These gradients and the resultant bluer colours in the outskirts of these systems are thought to be caused by satellite accretion events (e.g. Kaviraj et al., 2009; La Barbera et al., 2012; Ferreras et al., 2017; Jackson et al., 2023), with the stars from the satellites, which are typically bluer, settling in the outskirts of the massive ETGs. It is worth noting, however, that at the lower mass end of the massive regime ($10^{9.5}$ $M_\odot < M_\star < 10^{10.5}$ $M_\odot$), a minority (between 10 and 30 per cent) of ETGs in relatively low-density environments do show positive gradients (e.g. Ferreras et al., 2005; Jiang et al., 2011; Suh et al., 2010; Tortora et al., 2010), driven by recent star formation activity in their cores within the last $\sim$1 Gyr. When considered together with our results, this suggests that the incidence of blue ETGs (driven by blue central regions) does indeed increase with decreasing stellar mass.

The differences in the colour gradients between massive and dwarf ETGs suggests that the principal mode of stellar mass growth is likely to be different. The evolution of massive ETGs is likely to be driven by 'inside-out' growth, where the core of the galaxy forms earlier in cosmic time and material is then accreted through minor mergers in the outskirts. However, the opposite behaviour of dwarf ETGs as compared to massive ETGs in their colour gradients may suggest that ETGs in the dwarf regime could experience 'outside-in' growth, which is likely to be driven less by local environment and interactions and more by stellar feedback and secular processes (e.g. Pipino & Matteucci, 2004; Zhang et al., 2012; Pérez et al., 2013; Cheng et al., 2020; Ge et al., 2024).

This 'outside-in' scenario is thought to proceed via star formation starting in the outskirts of the galaxy (e.g. via gas accretion) and propagating inwards causing gas heating and expansion due to supernova events. Since low mass galaxies have shallow potential wells, some of the gas is blown away from the outer regions of the galaxy. The gas that remains cools down and sinks deeper towards the galactic centre giving rise to further star formation events. As a result, the galaxy is left with older stellar populations in the outer regions (since the gas reservoir is depleted due to the supernova-driven galactic wind) and younger stellar populations in the galactic centre. Note that our results are consistent with the conclusions of past studies (e.g. Pan et al., 2015) that have suggested that the transition between inside-out and outside-in growth takes place somewhere between $M_\star \sim 10^{10}$ $M_\odot$ and $M_\star \sim 10^{10.5}$ $M_\odot$.

It is worth exploring the colour profiles of two interesting sub-populations of dwarfs separately - ETGs which have visually identified blue cores and dwarfs that have been flagged as interacting in Lazar et al. (2024). As noted in that study, significant blue cores appear to be common in dwarf ETGs, with around 46 per cent of these systems showing cores that are clearly visible by eye (two such examples can be

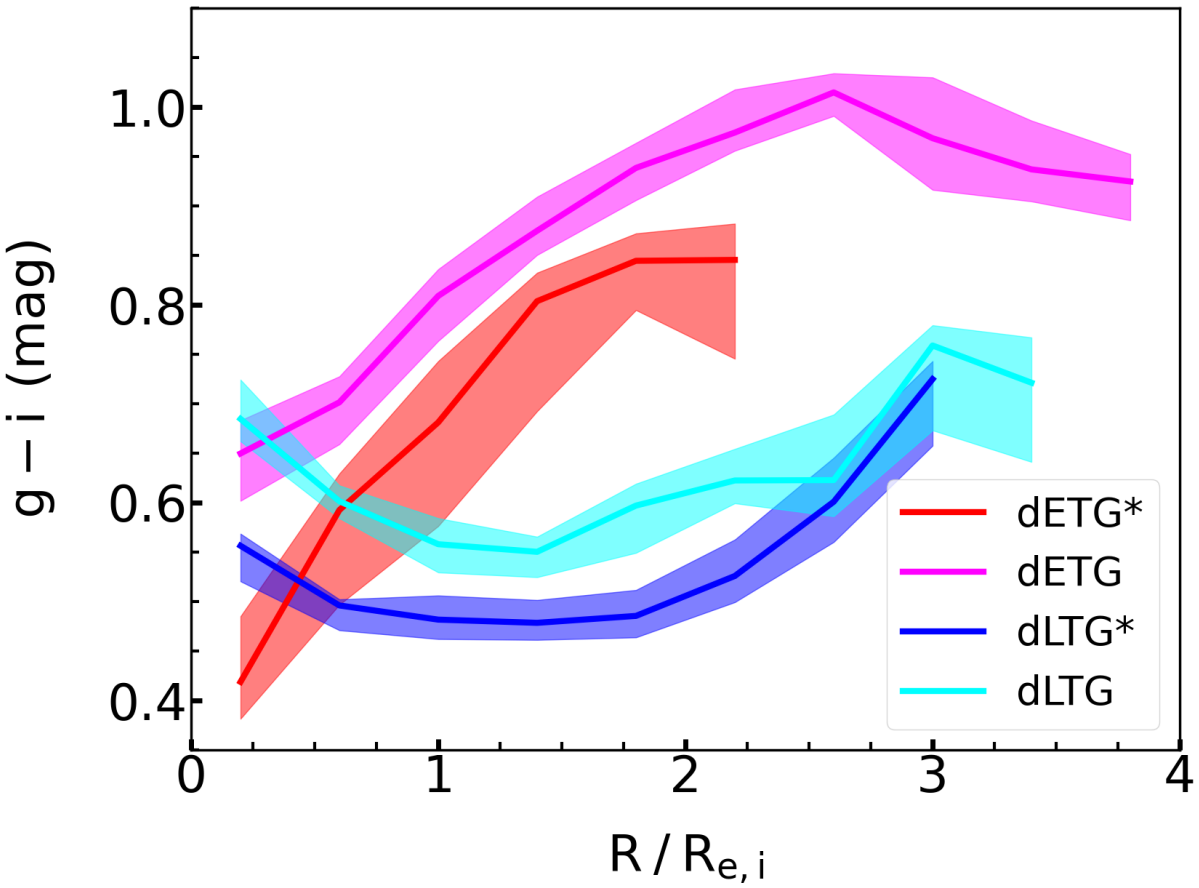

FIGURE 3.9: Median $(g-i)$ colour profiles in interacting (shown using a '*') and non-interacting dwarf ETGs and LTGs. In each panel the solid line represents the running median, while the shaded region indicates its uncertainty.

seen in the top row of Figure 3.1). Figure 3.8 shows that the blue cored dwarf ETGs are bluer than their non-cored counterparts at all radii. However, beyond $R \sim 1.5\ R_e$, the colour gradients become indistinguishable in the cored and non-cored populations (i.e. the colour profiles are roughly parallel to each other). Within this region the colour of the blue-cored ETGs diverges significantly, reaching an offset of -0.35 mag compared to the non-cored systems in the very centre. Most of the light from the blue core therefore appears to originate from the region enclosed within $\sim 1.5\ R_e$ in the cored systems.

We conclude our study by exploring dwarfs that have been flagged as interacting in the visual inspection performed in Lazar et al. (2024). Figure 3.9 shows that, regardless of morphology, interacting dwarfs tend to be bluer than their non-interacting counterparts at all radii. This is consistent with the finding in Lazar et al. (2024) that the median integrated colour of interacting dwarfs is typically bluer than their non-interacting counterparts. The bluer colours suggest that the interactions are enhancing the star formation activity in these systems, as has been suggested in the recent literature (e.g. Stierwalt et al., 2015; Martin et al., 2021). In particular, the consistent offset between the colour profiles of the

interacting and non-interacting dwarfs suggests that star formation may be enhanced across the entire extent of the galaxy (rather than being localised in, say, the central regions). This appears consistent with recent studies (e.g. Privon et al., 2017) which show that star formation enhancement via interactions in dwarfs could indeed be spatially extended and triggered by large-scale tidal compressions in the inter-stellar medium which act across the inner and outer regions of the galaxy. Note that the apparent truncation in the $(g-i)$ colour for interacting dwarf ETGs is artificial and caused by the fact that the number of data points for $R > 2\ R_{\rm e,i}$ drops below 10 (the threshold we apply for calculating the median colour in a given radial bin).

3.6 Summary

We have used an unbiased sample of 211 nearby ($z < 0.08$) dwarf ($10^8\ {\rm M}_\odot < {\rm M}_\star < 10^{9.5}\ {\rm M}_\odot$) galaxies, to study the structure of the dwarf population in the nearby Universe. In particular, we have studied galaxy size (parameterised via the effective radius), the effective surface brightness and the $(g-i)$ colour profile and colour gradient, as a function of stellar mass and morphology. Our main conclusions are as follows.

- Dwarf ETGs and LTGs are well separated in effective radius, with the median effective radius of LTGs being around a factor of 2 larger than that of the ETGs. The featureless dwarfs lie in between the ETGs and LTGs, with a median effective radius that is around a factor of 1.2 larger than their ETG counterparts.

- The effective radii of red (i.e. quenched) and blue (i.e. star-forming) galaxies do not show significant differences within each dwarf morphological class. When all morphological classes are considered together, there is a weak trend of red galaxies being more compact than blue galaxies.

- Dwarf ETGs are marginally brighter in median effective surface brightness that their LTG counterparts, with the featureless class being around 1 mag arcsec^{-2} fainter.

- The colour profiles and gradients of dwarf ETGs differ significantly from their massive counterparts. Dwarf ETGs typically show positive gradients (i.e. bluer central regions), while massive ETGs either have red colours throughout or a red core with blue outer regions (i.e. negative or flat gradient values). The divergence in colour profiles suggest that ETGs have different formation channels in the dwarf and massive regimes. Massive ETGs likely evolve 'inside-out', as a result of minor mergers adding stars to their outskirts. Dwarf ETGs, on the other hand, are likely

that they evolve through 'outside-in' growth, as a result of stellar-feedback driven galactic winds being more effective at quenching their outskirts while star formation continues in their central regions. Since dwarf ETGs have shallower potential wells than massive ETGs the effects of stellar feedback are also enhanced and so gas is dispersed easily from the outskirts of the galaxy and redistributed in the central region.

- Interacting systems are larger but have a similar median effective surface brightness as their non-interacting counterparts. This suggests that interactions trigger star formation, which boosts the surface brightness and offsets the increase in size. Regardless of morphology, the colour profiles of interacting dwarfs are bluer than their non-interacting counterparts at all radii. This indicates that the enhancement of star formation due to interactions takes places across the entire body of the galaxy, rather in a specific region (e.g. the centre).

- The featureless and ETG classes do not differ significantly in terms of their median effective radius but do differ in their median effective surface brightness. The distributions of baryons within these classes therefore show strong differences, suggesting that their formation histories are different. Thus, the featureless class in low-density environments should not be considered a subset of the ETG population.

- Dwarfs in our sample that reside in the 'UDG' region in the μ_e vs R_e parameter space include members of all morphological classes and are a continuous extension of the galaxy population towards lower values of μ_e. This suggests that, rather than being a novel class of object, galaxies in low-density environments that satisfy the UDG criterion are simply part of the fainter and more diffuse end of the overall galaxy population.

- Dwarf late type galaxies resemble color profiles similar as massive LTGs ($M_\star > 10^{10}\,M_\odot$) showing negative color gradient values in their central regions ($R < 1.5\,R_e$) and positive gradients in the outer regions ($R > 2\,R_e$).

- The prominent blue cores visually identified in around 46 per cent of the dwarf ETG population extend out to $\sim 1.5\,R_e$.

Chapter 4

Autonomous classification of galaxy morphologies in the HSC-SSP using unsupervised machine learning

4.1 Introduction

The study of galaxy morphology has been a cornerstone of our understanding of galaxy evolution since the mid 1800s. By this time the scientific community had realized that galaxies ('nebulae') exist in a variety of shapes, colours and morphologies. Massive galaxies are typically grouped into two major morphological types: 'early-type' galaxies (ETGs), which show smooth, spheroidal shapes with high mass concentrations in the central region and 'late type' galaxies (LTGs), which exhibit disky structures (typically evidenced by the presence of spiral arms). The processes that create these different morphologies is still not completely understood, almost a century after Hubble's revolutionary classification work (Hubble et al., 1926).

ETGs are thought to form through a combination of galaxy interactions such as major and minor mergers (e.g. Cooper et al., 2006; Kaviraj, 2010, 2014b; Peng et al., 2010) which can create or increase the stochasticity of the stellar orbits in these systems, environmental processes such as ram pressure stripping, starvation and strangulation (e.g. Balogh et al., 2004; Boselli et al., 2006; Font et al., 2008; Tasca et al., 2009; Berrier et al., 2009; Wetzel et al., 2013; Boselli & Gavazzi, 2014; Martin et al., 2018b) which remove gas and are particularly efficient in dense environments like galaxy clusters, and intrinsic processes like stellar and AGN feedback that act to maintain the ETGs in a quenched state (e.g. Hopkins

et al., 2006; Croton et al., 2006; Guo et al., 2011; Maiolino et al., 2012; Kaviraj et al., 2017; Beckmann et al., 2017; Martin et al., 2018b). LTGs, on the other hand, are thought to form primarily through secular mechanisms such as steady gas accretion and a more quiet interaction history dominated by minor mergers (e.g. Sparre & Springel, 2017; Aumer et al., 2013; Kaviraj, 2014b; Jackson et al., 2020; Jackson et al., 2022; Zeng et al., 2021; Martig et al., 2021).

In the massive-galaxy regime ETGs typically have higher masses, older stellar populations and redder colours, dominating the 'red sequence' in the colour-mass plane, while LTGs tend to be less massive and host higher star formation rates (SFRs) which makes them bluer (e.g. Baldry et al., 2004; Bremer et al., 2018). The region in colour-mass space which links the red sequence and blue cloud (sometimes termed the 'green valley') contains galaxies with properties somewhat in between these two groups such as lenticulars (S0s) and blue ETGs (e.g. Salim et al., 2007; Wyder et al., 2007; Angthopo et al., 2019). These galaxies are thought to be transitioning from being star-forming to passive (Schawinski et al., 2009a, 2014; Martin et al., 2007; Nogueira-Cavalcante et al., 2017).

While measurements of morphology are fundamental to studies of galaxy evolution, morphological information also plays an important role in facilitating and improving many other areas of astrophysical science. For example, knowledge of the morphology of galaxies can be used to improve the accuracy of their photometric redshift measurements (e.g. Soo et al., 2018; Menou, 2018), provide important contextual data for transient light-curve classification (e.g. Djorgovski et al., 2012; Wollaeger et al., 2018) and can offer insights into the mechanisms that trigger AGN activity (e.g. Kaviraj, 2014b; Schawinski et al., 2014).

An extensive literature exists on techniques for the morphological classification of galaxies. Many traditional techniques include fitting light distributions using Sérsic, de Vaucouleurs or exponential profiles (e.g. Sérsic, 1963; Simard et al., 2002; Odewahn et al., 2002; Lackner & Gunn, 2012), the 'CAS' system (e.g. Abraham et al., 1994; Conselice, 2003; Menanteau et al., 2006), the Gini-M_{20} coefficients (e.g. Lotz et al., 2004; Scarlata et al., 2007; Peth et al., 2016) and the MID system (e.g. Freeman et al., 2013). Over the last couple of decades, the advent of large space and ground-based surveys, e.g. the SDSS (Alam et al., 2015) and HST-CANDELS (Koekemoer et al., 2011) has driven a strong interest within the scientific community towards machine learning (ML) as a key tool for morphological classification (e.g. Lahav et al., 1995). For instance, supervised ML, which relies on accurate training sets, has been popular for classifying galaxy morphologies in large survey datasets (e.g. Huertas-Company et al., 2015; Ostrovski et al., 2017; Schawinski et al., 2017; D'Isanto & Polsterer, 2018; An et al., 2018; Cheng et al., 2020; Walmsley et al., 2022, 2023).

While the methods described above provide algorithmic classification of galaxies in large datasets, they are often benchmarked against visual inspection, which offers an accurate method of morphological classification (e.g. Lintott et al., 2011; Kaviraj, 2014b; Simmons et al., 2017). Massively distributed classification platforms, such as Galaxy Zoo (e.g. Lintott et al., 2011), can provide comprehensive databases of 'crowd-sourced' visual classifications from thousands of citizen scientists, for large surveys like the SDSS (Willett et al., 2017) and the legacy surveys performed by the HST (Simmons et al., 2017).

Forthcoming big data surveys, such as LSST (Ivezić et al., 2019), Euclid (Percival et al., 2019) and JWST (Gardner et al., 2006), which will operate in the peta and exabyte scale regimes, present a unique challenge to morphological classification work. The sheer size of such datasets may make them intractable for visual inspection, even via massively-distributed platforms like Galaxy Zoo (e.g. Lintott et al., 2011). In a similar vein, creating training sets that fully capture the morphological diversity of the galaxy population that is needed for supervised ML may be difficult. The short cadence of surveys like LSST, where the images will change over timescales of days, may pose an additional challenge, as it may become impractical to repeatedly produce training sets on short timescales. New techniques, which either combine visual and automated classification (e.g. Beck et al., 2018; Dickinson et al., 2019; Uzeirbegovic et al., 2020) or perhaps remove the need for visual classification altogether (e.g. Siudek et al., 2018; Hocking et al., 2018; Hendel et al., 2018; D'Isanto et al., 2018), will be crucial in dealing with the unprecedented data volumes expected from these new and forthcoming surveys.

While its implementation for morphological classification is still in its infancy (e.g. Hocking et al., 2018; Siudek et al., 2018; Ay et al., 2020; Martin et al., 2020), unsupervised ML (UML), coupled with post-processed visual inspection, offers an ideal route for classifying galaxy morphologies in big data surveys. UML does not need training sets and is capable of autonomously grouping together morphologically-similar objects in such suveys. For example, a prototype algorithm presented in Martin et al. (2020) utilizes hierarchical clustering and growing neural gas networks to group together image patches from a survey that have similar visual properties (characterised by their spatial power spectrum), followed by a clustering of objects (e.g. galaxies) that are constructed from these patches. In principle, such an algorithm can reduce an arbitrarily large number of galaxies into a small number ($\sim$150) of 'morphological clusters' which can then be visually labelled. In this study we further develop the prototype methodology from Martin et al. (2020) and implement it on the Deep layer of the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP; Aihara et al., 2018) Data Release 3 (DR3) to create a catalogue of morphological classifications which will be of broad interest for the astronomical community.

This chapter is structured as follows. In Section 4.2, we outline the properties of the HSC-SSP and the

ancillary data used in this study. In Section 5.2, we describe the UML algorithm that underpins this work. In Section 4.4, we explore the robustness of the classifications, by comparing the properties of galaxies in different morphological groups to known trends in these properties as a function of morphology. We summarise our results in Section 4.5.

4.2 Data: the HSC-SSP

The HSC-SSP is a deep-wide optical survey in five broad-band (*grizy*) and four narrow-band filters (Aihara et al., 2018). The Hyper Suprime-Cam offers a 1.5 degree field of view, with a median *i*-band seeing of ~ 0.6 arcsec. The survey provides Wide, Deep and Ultra-deep layers, with areas and *r*-band point source depths of 1200, 27 and 3.5 deg^2 and $\sim$ 26, 27, 28 mag respectively. The surface-brightness limit of the Deep layer (which we use in this study) is at least ~ 30.5 mag arcsec^{-2} (Huang et al., 2018). The catalogs containing photometric redshifts and other physical properties are obtained from Aihara et al. (2022).

The HSC-SSP provides an unprecedented combination of depth and area compared to the benchmark wide-area surveys of the past. For example, the shallowest layer of the HSC-SSP is four magnitudes deeper than standard-depth imaging from the SDSS. The HSC-SSP offers the deepest footprint of any optical survey with a comparable area to date. The data reduction pipeline for this survey is identical to the one that will be used to process the LSST, making it a useful precursor dataset. The deep-wide nature and relatively high spatial resolution of this survey makes it a useful pathfinder for forthcoming deep-wide surveys and enables us to perform morphological classification, at least for massive galaxies, out to $z \sim 1$.

4.3 Methodology: unsupervised machine learning algorithms for morphological analysis

The UML method used in this study is based on, but incorporates several improvements to, the technique described in Martin et al. (2020). We first outline the basic details of the algorithm and then describe the improvements that have been made to it for this work. Martin et al. (2020) represents one of the first attempts at using a high-dimensional feature space extracted from survey images to autonomously segregate galaxy morphologies, with high precision, at low and intermediate redshift. The main goal of

this algorithm is to minimise the amount of human intervention required in the morphological classification process in extremely large datasets, such as those expected from upcoming surveys such as LSST, Euclid and JWST.

This optimisation is achieved by grouping together morphologically similar galaxies into a small number ($\sim$100-280) of 'morphological clusters' with the clusters (rather than individual galaxies) being visually inspected to label them into standard morphological types (e.g. ellipticals, spirals etc.). The overall benefit is that, if the morphological clusters are relatively pure, then an arbitrarily large sample of galaxies can be reduced to a small number of morphological clusters. The gain is higher for progressively larger datasets and becomes potentially indispensable for morphological classification in surveys like LSST.

A variant of this method has been employed on space-based images from the HST-CANDELS survey in Hocking et al. (2018) and a prototype of what is used here has been implemented on ground-based images from the HSC in a pilot study in Martin et al. (2020). These studies show that the automated classifications exhibit high levels of purity, with individual morphological clusters correlating strongly with known Hubble types, as labelled by either expert classifiers or platforms like Galaxy Zoo.

4.3.1 Feature space

The feature space is an N dimensional matrix on which ML models are trained. This can be based on parametric or non-parametric methods used to measure morphology or just imaging data. In this case the feature data matrix is extracted from HSC pixel data and is built in such a way that the feature vectors that describe individual galaxies are invariant with respect to galaxy orientation, rotation or scale.

The first step is to obtain the detection (i.e. segmentation) maps for each galaxy using the HSC image processing pipeline (Bosch et al., 2018). Using these detection maps we extract square sub-image patches, each of which has a size that is approximately twice the PSF which, in the case of our HSC images, is 8×8 pixels for each object in the *griz* bands. We do not use the *y*-band because it has relatively low signal-to-noise. We note that the patch size has to be large enough to sample the spatial scales of the smallest flux variations.

After the patch extraction, we calculate the radially-averaged pixel-intensity power spectrum of each patch, using a 2D fast fourier transform (FFT; Ballard & Brown, 1982). We multiply this by its conjugate in order to remove the phase information and retrieve only the amplitude of the spectrum. The radial profile of the power spectrum has five bins for each of the four bands and provides a rotationally-invariant depiction of each patch. The radial profiles corresponding to each band are then concatenated together

into a 1D array, the 'feature vector' **p**, which is composed of 20 values (number of bands $\times$ number of bins). This feature vector holds information about colour, pixel intensity and spatial frequency for each patch. The **p** vectors corresponding to all patches are then added to form a final data matrix **P**. For the next step, in order to increase the variety of final galaxy morphological clusters obtained, we sort **P** in such a way so that we pick the patch vectors with the highest dispersion from the mean.

4.3.2 Growing Neural Gas

Each feature vector in matrix **P** resembles a log normal function. In order to increase the efficiency of the clustering we take the log of the matrix and convert the feature arrays to a Gaussian function. The next step is to normalize each column **P** using Equation 4.1.

$$y = \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{4.1}$$

where x_{min} and x_{max} are the minimum and maximum column values, respectively. The data is then reduced in size by employing a Growing Neural Gas (GNG) algorithm (first implemented by Fritzke, 1995) using the Python function `algorithms.competitive.growing_neural_gas.GrowingNeuralGas` from the **NeuPy** library, to produce a topological map of nodes (or neurons), where each node is an N dimensional vector representing a group of similar patches. The GNG method is an iterative process that uses a distance minimization technique, where an N dimensional map is constructed and then expanded from large scale to small scale features by learning from the positions of the nodes created in the previous iterations. In order to form the final network (i.e. the topological map) we iterate through the first 2 million patch vectors in matrix **P** until the network reaches a maximum number of nodes of 100000 (`max_nodes`). The 2 million patch vectors are split into 10 batches. For each batch (hereafter referred to as **B**) the GNG algorithm operates as follows:

(a) The network is populated with two randomly selected patch feature vectors (also called nodes) from matrix batch **B**. A new random feature vector is extracted from the same batch.

(b) The closest nodes (from the batch sample **B**) to the sample vector are identified by calculating their euclidean distances. If an edge connecting the two nodes does not already exist, it is created. Whenever two nodes are connected by an edge, the edge is assigned an age which is initially set to zero, and the ages corresponding to the edges which share the same nodes are incremented by 1.

TABLE 4.1: Parameters used for the Growing Neural Gas (GNG) Python function. See Section 4.3.2 for parameter explanations.

Parameter	Value
`max_nodes`	100000
`n_iter_before_neuron_added`	100
`max_edge_age`	50
`step`	0.2
`neighbour_step`	0.006
`after_split_error_decay_rate`	0.5
`error_decay_rate`	0.995

(c) The closest node to the sample vector is assigned an error based on the square of their separation and, together with its node neighbours, are moved towards the position of the random sample vector by a fraction (`step` and `neighbour_step`, respectively). Edges with ages older than a certain number (`max_edge_age`) and nodes without any neighbours in their close proximity are removed.

(d) Steps (a)-(c) are repeated and more nodes are added to the network until the algorithm has gone through a certain number of random sample vectors (`n_iter_before_neuron_added`).

(e) Additional nodes are inserted at half the distance between neighbouring nodes with the highest errors in order to 'draw' even further the small scale details of the topological map. The initial edges connecting neighbouring nodes with the highest errors are removed and the errors of the initial nodes reduced by a certain fraction (`after_split_error_decay_rate` and `error_decay_rate` for the error reduction to node with the largest error and for the error reduction to all nodes, respectively).

The values of the parameters discussed above are shown in Table 4.1. The positions of the nodes are then outputted to a multi-dimensional data matrix **N**.

4.3.3 Hierarchical clustering

The next step is to employ agglomerative hierarchical clustering (HC; Johnson, 1967) on the data matrix **N** to gather the nodes within the GNG network into representative groups and use the resulting model on the original dataset (**P**) to assign a 'type' label to each feature vector. The HC method is a 'bottom-up' grouping approach, where each node initially starts in its own cluster. The following steps join the closest clusters to each other iteratively, thus creating new and bigger multi-node clusters after each iteration.

The process ends with one cluster that contains all the other sub-level clusters. The affinity metric used to compute the linkage between clusters is euclidean. Other metrics (such as pearson or cosine methods) can also be used but Hocking et al. (2018) and Martin et al. (2020) show that they yield almost the same results.

In order to carry out hierarchical clustering, we use the Python function `cluster.AgglomerativeClustering` from the **SkLearn** library. The grouping process is run until 1500 clusters is reached (shown to provide reliable results in terms of morphological variety in Martin et al., 2020). Other cluster numbers can also be used depending on the complexity of the input data. Since this clustering method is not inductive (i.e. cannot be used to predict new data samples directly), we use the Python Random Forest (RF) classifier (`ensemble.RandomForestClassifier`) from the same library (**SkLearn**) to train a new model on the same input data but this time using the predicted labels from the HC step. We then predict 'type' labels for each of the patch feature vectors from matrix **P** using the resulting models from the HC and RF classifiers.

The RF algorithm outputs an averaged model among an ensemble of decision trees (Breiman, 2001), where each decision tree is trained using a different group of features associated with corresponding labels. Depending on the feature space, the trees contain a number of nodes and at each one of these points a threshold condition is enforced. If this condition is upheld (or not), a portion of the data is passed through to the next condition (i.e. another tree 'branch' is created along with a corresponding new node). This process repeats itself until there are no more training data left to model. The resulting tree predictions are then compared with each other and the most predicted model between all trees is chosen.

The RF algorithm is a highly accurate and robust method for our purposes, since the data can be partitioned using a certain number of trees and modelled in detail accordingly. This method also avoids overfitting, mainly because it takes the average between the tree predictions and thus cancels out any biases. In order to balance the accuracy achieved and the CPU power used we choose a number of 1000 trees for our study.

4.3.4 Final object clustering

Once each feature vector from the original data matrix **P** has been assigned a patch 'type' by means of the HC, the algorithm generates a histogram for each object containing all its constituent patches (the 'object feature vector'), where each bin represents a different patch type. The number of clusters

produced through HC and the number of bins for each histogram must match. As a result, each histogram corresponding to an object has 1500 bins (or patch 'types'). In order to account for the increased number of patches for objects with large angular sizes, we normalize all histograms by dividing each bin by the sum of all bin values. The object feature vectors are appended into an array of dimensions $N_{obj} \times N_{bins}$ (number of objects $\times$ number of bins) which represents the final feature vector of all galaxies. This creates a feature space where each object is defined by a vector encoding basic visual properties.

In order to keep the classifications as accurate as possible, we exclude objects with angular sizes smaller than 20×20 pixels (approximately five times the median PSF FWHM) where the object size is defined by the width and height of the object segmentation map. Objects smaller than these sizes are generally poorly resolved and may introduce more uncertainty in the classifications. We then group the objects by their similarity using k-means clustering (MacQueen et al., 1967), separating them into k morphological clusters. This step is an iterative process where we run the k-means algorithm using different numbers of clusters, from 100 to 280 (based on previous studies using the same method, e.g. Hocking et al., 2018; Martin et al., 2020) and choose the model with the highest overall silhouette score. The silhouette score of an object is a metric that measures the similarity between the object feature vector and the vector corresponding to the centre of the morphological cluster containing the object. The value of the silhouette score varies from -1 to 1, where a value of 1 corresponds to an object feature vector being at the very centre of its cluster. The overall silhouette score is the mean of the scores corresponding to all objects. In this way, by choosing the model with the highest silhouette score we are able to filter out any outliers which are at the edges of the clusters and obtain cleaner classifications.

The algorithm described above takes under 30 hours to process the HSC DR3 deep fields (30 deg^2) for $z < 1$ using 20 threads of execution. It is written entirely in Python and relies mainly on the `NUMPY`, `ASTROPY`, `NEUPY` and `SKLEARN` libraries. The last phase of this classification method is to take a representative object sample from each morphological cluster and visually inspect it in order to assign a morphological type (e.g. elliptical, spiral) to each cluster. The representative sample contains 10 objects with the highest silhouette score, 10 random objects and 10 objects with the lowest score. This is done by producing *gri* object images containing only the detected pixels (from the available HSC segmentation maps).

4.3.5 Improvements made to the previous versions of the technique

In this section we describe several significant changes that have been introduced in this work compared to the prototype version presented in Martin et al. (2020):

- Parts of algorithm written in C# have been re-written in PYTHON. This makes it easier and faster to install and run the code (the overall compute time is lower by around a factor of 2). Currently, the algorithm can process 1 deg^2 of sky, for objects at $z < 1$, in $\sim$1 hour using 30 CPUs.

- The algorithm now only takes into account pixels from the galaxy segmentation maps, which reduces the overall uncertainties of the final ML models since interloper objects are masked. It also results in faster processing speeds compared to the prototype version which processed all the pixels in the survey in question.

- The selection of the patch vectors to be inputted into the GNG network has been modified so that it maximizes the variety of galaxy morphological types that are found in the final k-means clusters. The code now also contains a feature where the number of patch vectors inputted into the GNG network is kept the same, which enables it to deal with large volumes of data from upcoming peta-exabyte level surveys.

- By making use of an inductive clustering technique, in combination with the RF algorithm and the HC model, we have introduced a way of directly predicting the patch labels.

- The process of picking the best k-means model in the last clustering step has been optimised. The code creates models with a number of clusters varying between 100 and 280 and automatically selects the one with the highest overall silhouette score.

- Each morphological cluster now has a purity value (**Pur**; described in Equation 4.2 below) which can be used as a metric to estimate the accuracy of the ML model.

4.4 Galaxy morphologies in the HSC-SSP DR3 DEEP

The main advantage that this methodology is that the visual inspection step that is required for labelling is performed on a small number (100 to 280) of morphological clusters rather than the (at least) several hundred thousands of individual galaxies that are detected in modern large-area surveys. The visual inspection thus becomes tractable for individual astronomers and requires only a few hours of effort. As long as the purity of the clusters is high, this compression reduces an arbitrarily large dataset into a manageable number of morphological clusters, which will be essential when dealing with peta and exabyte scale surveys of the future. It is worth noting that, in comparison, supervised ML methods may be more time-consuming because of the need to construct training sets. They could also potentially be less accurate if the training set does not span all the existing types of objects in the survey.

We proceed by performing a study of galaxies restricted to the redshift and stellar mass ranges of $z < 1$ and $M_\star > 10^8$ $M_\odot$, respectively, in the Deep fields of the HSC-SSP DR3. We restrict ourselves to these stellar mass and redshift ranges because past experience from Martin et al. (2020) suggests that morphological classification becomes difficult outside these ranges given the depth and resolution of the HSC images. Note that Martin et al. (2020) only focused on relatively massive galaxies. While we probe lower mass galaxies that their study, it is possible that the morphological classifications in our lower stellar mass bin (see list below) have to be treated with some caution.

We filter out stars at the outset by using the extendedness parameter in the *griz* bands from the HSC DR3 catalogue i.e. we only consider objects which have an extendedness of 1 in all filters. Nevertheless, as we show below, extendedness does not appear to completely remove stars. However, the algorithm is able to group stars into morphological clusters and we are able to provide an extra layer of star-galaxy separation using it. Finally, in order to increase the accuracy of the algorithm, and since the individual galaxy morphology fractions can change as a function of stellar mass and redshift, we run the code for nine separate models which correspond to the stellar mass and redshift ranges listed below:

- 10^8 $M_\odot < M_\star < 10^{9.5}$ $M_\odot$ and $0.0 < z < 0.3$
- 10^8 $M_\odot < M_\star < 10^{9.5}$ $M_\odot$ and $0.3 < z < 0.6$
- 10^8 $M_\odot < M_\star < 10^{9.5}$ $M_\odot$ and $0.6 < z < 1.0$

- $10^{9.5}$ $M_\odot < M_\star < 10^{10.5}$ $M_\odot$ and $0.0 < z < 0.3$
- $10^{9.5}$ $M_\odot < M_\star < 10^{10.5}$ $M_\odot$ and $0.3 < z < 0.6$
- $10^{9.5}$ $M_\odot < M_\star < 10^{10.5}$ $M_\odot$ and $0.6 < z < 1.0$

- $M_\star > 10^{10.5}$ $M_\odot$ and $0.0 < z < 0.3$
- $M_\star > 10^{10.5}$ $M_\odot$ and $0.3 < z < 0.6$
- $M_\star > 10^{10.5}$ $M_\odot$ and $0.6 < z < 1.0$

Since for each redshift and mass bin we are selecting the number of morphological clusters by choosing the k-means model with the highest overall silhouette score and since the feature space for each bin is different, the number of morphological clusters varies for each redshift and mass bin. As a result,

we obtain between 120 and 280 morphological clusters for each redshift and mass bin. As described in Section 4.3.4, a representative sample of 30 objects is taken from each morphological cluster and visually inspected. If any morphological cluster contains less than 30 objects, then all objects in that cluster are visually inspected. For this study, the visual classifications have been performed by one expert classifier (Sugata Kaviraj). We assign general morphological labels for our clusters as follows:

- Early-type galaxy (ETG): clusters comprising galaxies that are elliptical or S0.
- Late-type galaxy (LTG): clusters comprising galaxies that show spiral structure (and excludes S0s).
- Featureless galaxies (F): this morphological type is essentially found only in the low mass i.e. dwarf regime and exhibits a smooth morphology with a lack of both a central light concentration and any signs of spiral structure, as described in Chapter 2 and Lazar et al. (2024).
- Mixed galaxy types (MIX): clusters which do not have a dominant galaxy type (ETG, LTG or F). We indicate a dominant galaxy type if its abundance in that specific cluster is over 60 per cent.
- Bad data (N): clusters comprising shredded HII regions, noisy, unresolved galaxies or stars.

In order to quantify the uncertainty of the galaxy classification, each morphological cluster is assigned a purity value (**Pur**), described in Equation 4.2 below. This is defined as the percentage of the morphological type that dominates the representative sample that is used to perform the cluster classification.

$$\mathbf{Pur} = \frac{\mathrm{N_d}}{\mathrm{N}} \times 100, \tag{4.2}$$

where N is the total number of objects in the representative sample and $\mathrm{N_d}$ is the number of objects from the dominant morphology in the sample. We consider a morphological cluster to have a dominant morphological type if its purity value is higher than 60 per cent. If **Pur** is less than 60 per cent then the morphological cluster is labelled as mixed.

Figure 4.1 shows the purity distribution of our clusters. When all clusters are considered together, the median purity across all morphological types is around 90 per cent. In other words, on average, around 10 per cent of objects from a certain morphological cluster are misclassified by the algorithm. The algorithm is also able to isolate bad data, such as shredded HII regions, noisy or unresolved galaxies (i.e. morphological clusters labelled as 'N' in Figure 4.1) with an accuracy of around 90 per cent.

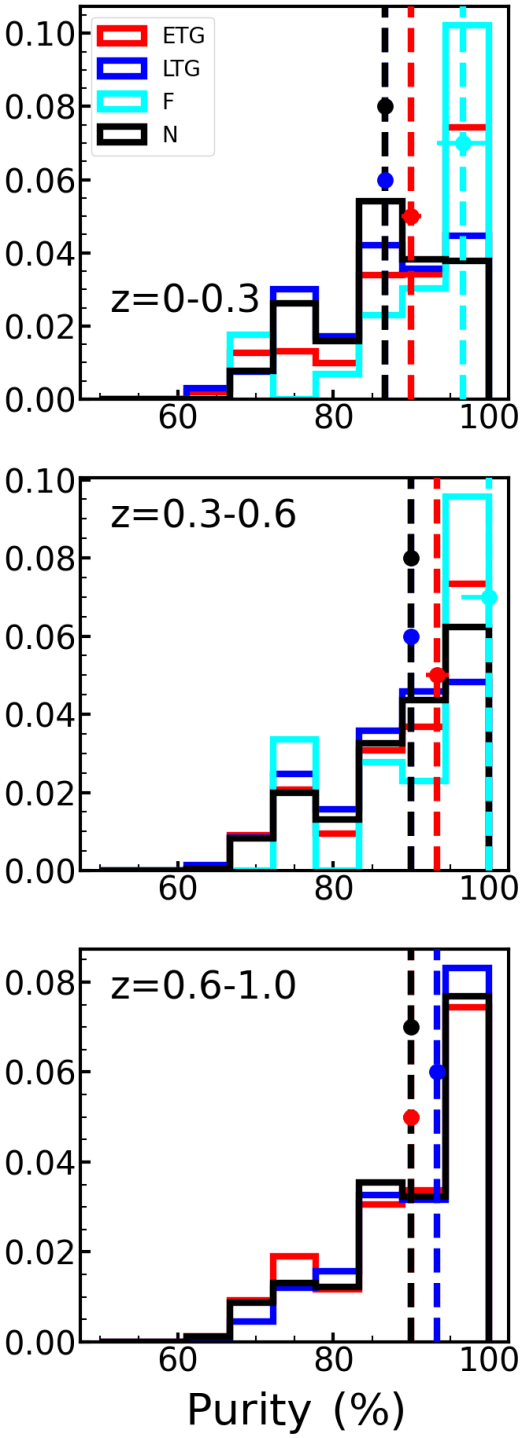

FIGURE 4.1: Purity of morphological clusters as a function of redshift for the ETG, LTG, F and N morphological labels described in section 4.4.

The total number of objects in our sample is ∼177,000. Around 37 per cent of these falls in clusters labelled as 'N' and around 13 per cent in mixed clusters. In the analysis presented below, we remove all morphological clusters labelled as MIX or N, which results in a dataset that can be used for galaxy morphological studies. Figure 4.2 shows examples of galaxies in each general morphological class.

The percentages of morphological types as a function of redshift and mass are shown in Figure 4.3. We note that, as discussed in Jackson et al. (2021c) and Lazar et al. (2024), the low-mass population (10^8 $M_\odot < M_\star < 10^{9.5}$ $M_\odot$) in this study is unlikely to be complete for $z > 0.3$ and will be skewed towards bluer objects. Therefore, the morphological classifications in our lowest mass bin may contain biases for $z > 0.3$. In addition, morphological classification tends to become more difficult for low mass galaxies because they are less extended and not as well resolved. The morphological percentages in the low mass regime outside the nearby Universe ($z > 0.3$) should therefore be treated with caution. Indeed, while these percentages are presented in this thesis for completeness, we may decide not to release them in the eventual catalogue paper.

The galaxy fractions obtained at different redshifts for the high mass bin ($M_\star > 10^{10.5}$ $M_\odot$) are in agreement with the results from recent work. For example, Buitrago et al. (2013) obtain ETG fractions of ∼75 per cent for $M_\star > 10^{11}$ $M_\odot$ out to $z \sim 1$ using a combined sample of ∼1000 objects from the SDSS and HST surveys. In the very nearby Universe ($z < 0.06$), we obtain an LTG fraction of ∼76 per cent which is in good agreement with studies such as Conselice (2006a) who obtain LTG fractions of ∼75 per cent probing a large mass range ($10^8 < M_\star < 10^{12}$ $M_\odot$) out to $z \sim 0.05$ using more then 22,000 galaxies from the Third Reference Catalogue of Bright Galaxies (RC3; de Vaucouleurs et al., 1991). The ETG fractions increase with increasing stellar mass at all cosmological epochs until $z \sim 1$, generally consistent with the findings of the recent literature (e.g. Kauffmann et al., 2003; Brinchmann et al., 2004; Bamford et al., 2009; Pozzetti et al., 2010; Vulcani et al., 2011).

In Figure 4.4 we present the star formation main sequence, which shows the known LTG – ETG bimodality in the SFR – mass relation, where LTGs reside primarily in the 'blue cloud' and ETGs populate the 'red sequence' (e.g. Visvanathan, 1981; Baldry et al., 2004, 2006; Faber et al., 2007). Spirals present significant star formation activity defining the star formation main sequence (Schiminovich et al., 2007; Bell et al., 2007; Schawinski et al., 2014; Wilkinson et al., 2021) which increases in SFR with redshift compared to ETGs which typically lie off this locus.

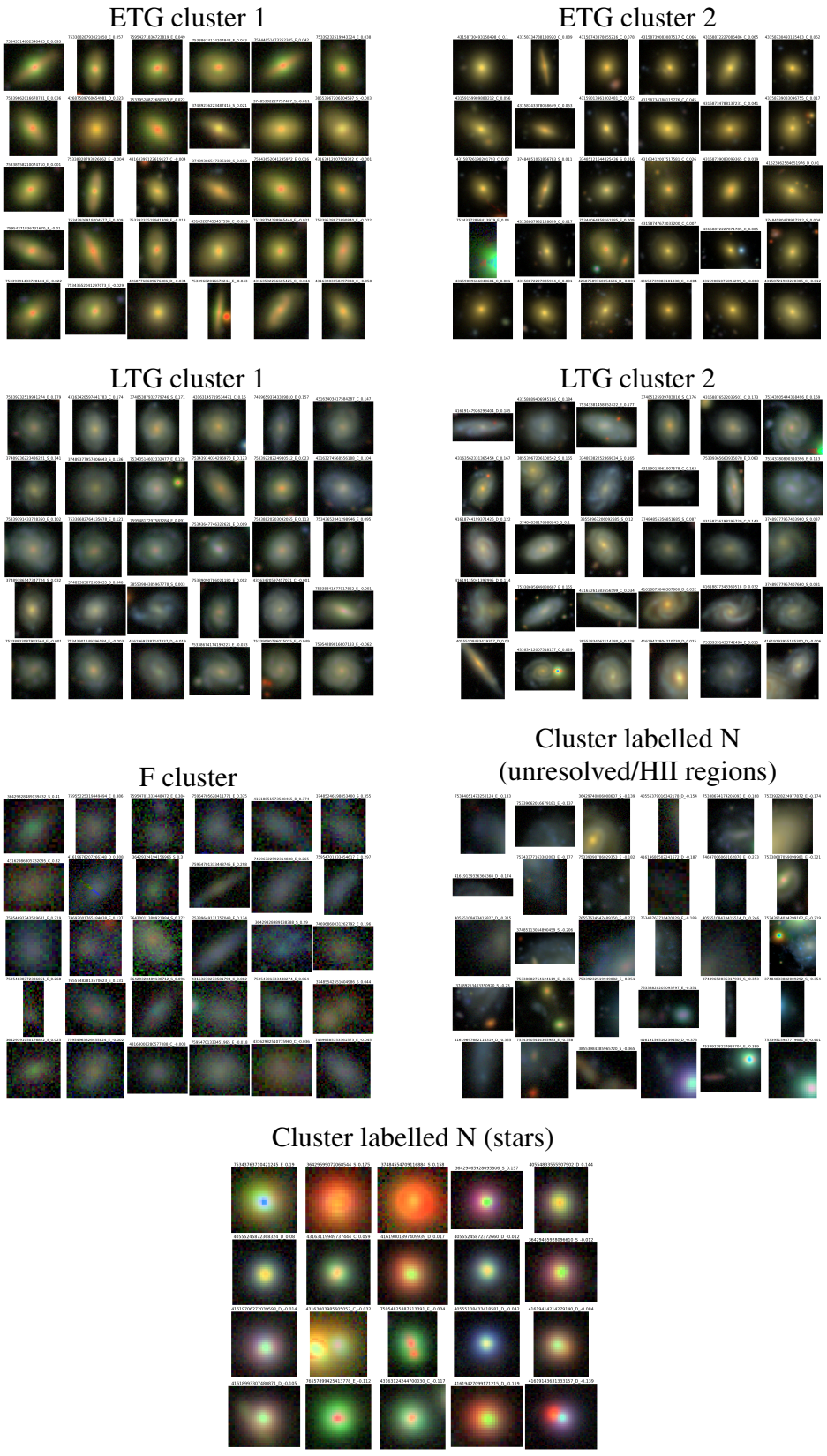

FIGURE 4.2: Examples of morphological clusters produced by the algorithm.

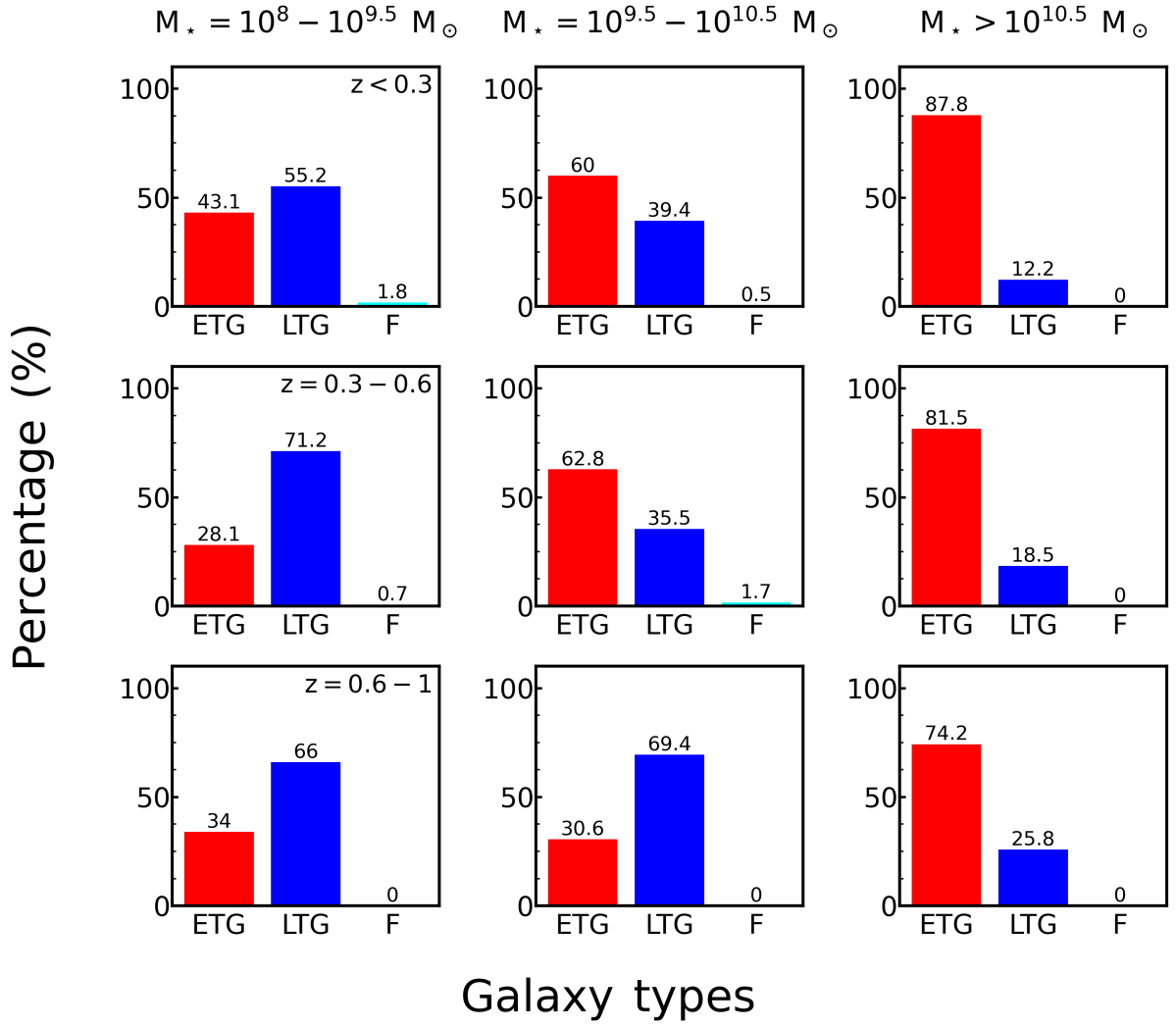

FIGURE 4.3: Morphology percentages as a function of redshift and stellar mass.

4.4.1 An example of the autonomous identification of specific morphological types - blue elliptical galaxy

Blue elliptical galaxies are a subclass of ellipticals with SFRs of up to several tens of $M_{\odot}$ yr^{-1}, which are significantly higher than in their more typical redder counterparts. At low redshift, these systems have generally been found in the low and intermediate stellar-mass regime ($10^8 < M_{\star} < 10^{10.5}$ $M_{\odot}$), typically in low-density environments (e.g. Schawinski et al., 2007, 2009a; Lazar et al., 2023). The processes that drive the SFRs in these systems are unclear. Some studies suggest that the star formation episode is driven by sustained gas accretion before the SFR presumably declines and these objects move to the red sequence (e.g. Fukugita et al., 2004). However, other studies argue that the dominant cause for the enhanced SFR are tidal interactions and mergers (e.g. Schawinski et al., 2009a). Therefore, studying these objects can provide novel insights into the formation of elliptical galaxies.

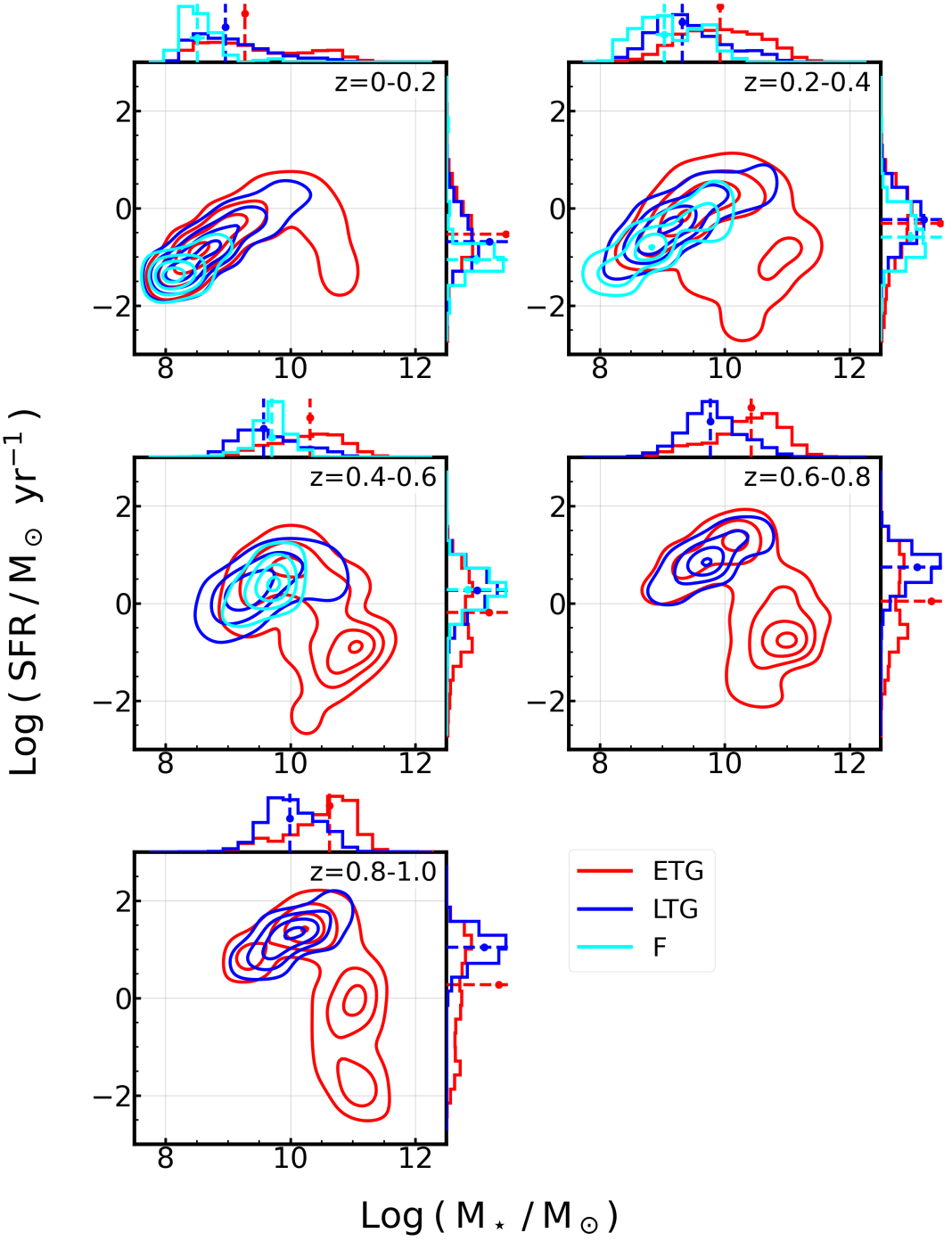

FIGURE 4.4: 2D distributions of SFR vs stellar mass as a function of redshift for ETGs, LTGs and featureless galaxies. The density bins are weighted by the purity value.

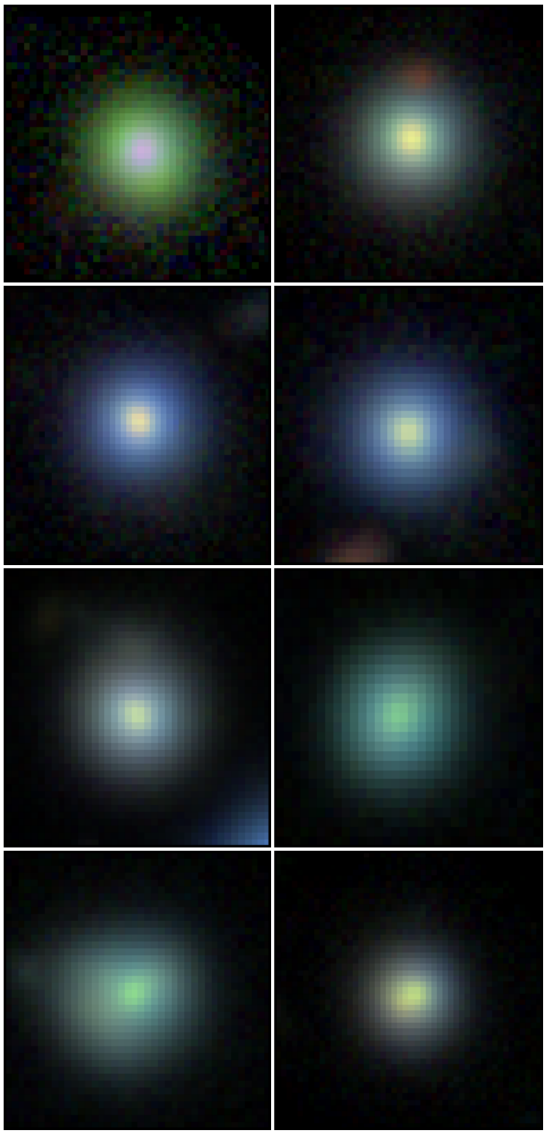

FIGURE 4.5: *gri* HSC example images of the blue elliptical type galaxies found using the unsupervised ML algorithm.

The UML algorithm autonomously finds morphological clusters corresponding to this galaxy type (see Figure 4.5 for examples). It is interesting to note that the blue colour of these galaxies is not confined just to a circumnuclear starburst as is often the case in elliptical galaxies which show typical levels of star formation (e.g. Crockett et al., 2011) but extends spatially across the entire galaxy, similar to the blue ellipticals found by Schawinski et al. (2009a). As we show in the next chapter, our blue ellipticals reside in the low to intermediate mass range and have star formation rates and colours comparable to LTG and starburst galaxies, even reaching 5 $M_{\odot}$ yr^{-1} (also seen in Schawinski et al., 2007, 2009a). Interestingly, as we discuss in the next chapter, most of these systems do not exhibit any tidal features, even in the ultra-deep HSC images which are $\sim$4 mag deeper than those from the SDSS. This suggests that the star formation in these galaxies largely does not involve either major or minor mergers, contrary

to our classical understanding of how ellipticals are supposed to form. We present a detailed study of the blue ellipticals identified by our algorithm in the next chapter.

4.5 Summary

The measurement of galaxy morphology is a fundamental tool for uncovering the evolutionary history of galaxies. Morphology is a key quantity, not only for galaxy-evolution studies, but also for a plethora of other astrophysical science, such as photometric redshift pipelines and transient light-curve classifications. Over the past few decades, with the arrival of large surveys, there has been a significant shift, away from traditional galaxy morphological classification methods such as direct visual classification to more automated techniques (e.g. supervised ML). However, future large surveys will provide peta and exascale data volumes, often coupled with short cadences (as is the case for LSST). This may make even supervised ML difficult to deploy on these datasets, as it may become impractical to construct training sets on short timescales for such surveys. UML, which does not need labelled training data can, in principle, compress hundreds of thousands of galaxies into a small ($\sim$100-280) number of morphological clusters, which can then be collectively labelled via visual inspection. This offers an ideal solution for the morphological analysis of galaxies in this new era of big data surveys.

In this study, we have used such a UML algorithm on the HSC-SSP Deep fields to morphologically classify the galaxy population. Our algorithm is able to separates the main Hubble types (ETGs and LTGs) relatively cleanly, identified bad data (such as shredded regions and stars) and automatically isolates some specific morphological classes such as blue elliptical galaxies. The main conclusions from our study are as follows:

- We are able to classify the HSC-SSP galaxy population into ETGs, LTGs and featureless galaxies at $z < 1$ with an accuracy of $\sim$90 per cent.
- The algorithm is also able to isolate 'bad data' such as shredded objects, HII regions, noisy or unresolved galaxies with an accuracy of around 90 per cent. The percentage of such objects in the HSC DEEP layer is $\sim$37 per cent.
- Galaxies in different morphological clusters reproduce the known trends in SFR vs stellar mass.
- Morphological fractions and trends in SFR as a function of stellar mass agree with those in the literature, whereby we see the star formation main sequence defined by spirals, with the elliptical galaxies residing outside this region.

- The algorithm is able to autonomously group together galaxies belonging to specific morphological classes such as blue elliptical galaxies.

Chapter 5

Relaxed blue ellipticals: accretion-driven stellar growth is a key evolutionary channel for low mass elliptical galaxies

5.1 Introduction

Our currently-accepted paradigm of structure formation postulates a hierarchical build up of stellar mass, driven by interactions and merging (e.g Press & Schechter, 1974; White & Rees, 1978; Blumenthal et al., 1984). Within this picture, elliptical galaxies, at least those with high stellar masses, are considered to be the natural dynamical end-points of galaxy evolution, intimately linked to galaxy merging (e.g. Dressler, 1980; Moore et al., 1996; Dekel & Birnboim, 2006; Faber et al., 2007). The strong gravitational torques created during galaxy interactions can destroy discs, by removing angular momentum and randomising stellar orbits, thus producing the smooth, relatively featureless morphologies typically found in elliptical systems (e.g. Martin et al., 2018c; Park et al., 2019).

Various pieces of evidence appear to support this picture in the high-mass regime. Both theoretical and observational studies suggest an increase in the fraction of massive elliptical systems in regions of higher density, where the frequency of merging is expected to be higher (e.g. the central regions of clusters), with the morphological mix being increasingly dominated by late-type (i.e. disc-dominated) galaxies as local density decreases (e.g. Dressler, 1980; Brough et al., 2006; Hopkins et al., 2010; Li et al., 2012; Deeley et al., 2017; Sengupta et al., 2022; Kolokythas et al., 2022). The tendency for merging to create, or at least reinforce, the spheroidal elements of a massive galaxy appears consistent

with a large variety of morphological and kinematical phenomena that are present in many nearby ($z < 1$) massive ellipticals (Quilley & de Lapparent, 2022) and which are signposts of galaxy interactions, e.g. tidal features (Toomre & Toomre, 1972; Peirani et al., 2010; Kaviraj, 2014b; Kaviraj et al., 2019), shells (Quinn, 1984; Cooper et al., 2011; Bílek et al., 2022), stellar streams (Johnston et al., 1999) or kinematically decoupled cores (Jedrzejewski & Schechter, 1988; Miller et al., 1990; Forbes et al., 1995; Cappellari et al., 2011; Bois et al., 2011; Guérou et al., 2015).

While the dynamical changes in the structure of today's massive ellipticals appears to have taken place via merging over cosmic time (e.g. Bender, 1988; Boylan-Kolchin et al., 2006; Martin et al., 2018c), the bulk of the stellar mass in these galaxies appears to have formed relatively rapidly at high redshift, as evidenced by their red optical colours (e.g Larson, 1974; Chiosi & Carraro, 2002; Ferreras et al., 2005; Baldry, 2007; Kaviraj et al., 2007; Bamford et al., 2009), which exhibit little scatter (e.g. Stanford et al., 1998; van Dokkum et al., 2000) and high alpha-to-iron ratios (e.g. Thomas et al., 2005).

Both theory (e.g. Martin et al., 2019) and observations (e.g. Lofthouse et al., 2017) suggest, however, that the majority ($\sim$80 per cent) of the old stellar populations that form in the early Universe and dominate massive ellipticals at the present day form through gas accretion, rather than mergers. This appears to be driven both by the relative paucity of merger events over cosmic time (particularly major mergers, e.g. Conselice, 2003) and by the fact that the early Universe is extremely gas rich and mergers do not enhance star formation rates beyond what is possible by gas accretion alone (e.g. Kaviraj et al., 2013; Lofthouse et al., 2017; Fensch et al., 2017). Stellar assembly continues in massive ellipticals over cosmic time (e.g. Fukugita et al., 2004; Kannappan et al., 2009; Yi et al., 2005; Kaviraj et al., 2007; Kelvin et al., 2014), with a significant minority (20 – 30 per cent) of stellar mass forming at $z < 1$ (Kaviraj et al., 2008). At late epochs, the star formation activity in massive ellipticals appears to be more influenced by merging (e.g. George, 2017; Omar & Paswan, 2018), with most of the interactions being unequal mass ratio (i.e. minor) mergers (e.g. Kaviraj et al., 2011).

While massive ellipticals have been extensively studied, both via observations and theory, much less is known about the evolutionary pathways of their low mass counterparts. While important studies on this subject have been done in the Local Group or for z<0.02 (e.g. Binggeli et al., 1984; Faber et al., 1992; Burstein et al., 1993), this is especially true for these populations outside the very nearby Universe, because past large surveys (like the SDSS) offer large footprints but are relatively shallow, making it more difficult to detect lower luminosity objects outside the very local Universe (e.g. Jackson et al., 2021c; Davis et al., 2022). Nevertheless, a growing literature has recently probed the evolution of local ellipticals in the low mass ($M_\star < 10^{9.5}$ $M_\odot$) regime. Many of these studies have focused on blue

ellipticals which are brighter and easier to detect in surveys like the SDSS. Apart from their increased detectability, blue ellipticals are also a particularly efficient route to understanding the processes that dominate the stellar assembly of these systems, since they are in a phase where they are actively building significant fractions of their stellar mass.

These studies have postulated several mechanisms that may contribute to the formation of blue low-mass ellipticals, including secular gas accretion (e.g. Mahajan et al., 2018), major and minor mergers (e.g. Schawinski et al., 2009b; Chung et al., 2019) or a combination of both (e.g. Moffett et al., 2019). It is worth noting that the unavailability, in past studies, of deep images, that are required to reveal merger-driven tidal features (e.g. Kaviraj, 2014b), has been a significant impediment to conclusively probing the role of mergers in the formation of these objects. Since blue ellipticals will become members of the general (red) elliptical population when they quench (e.g. Meyer et al., 2014), these galaxies are valuable as progenitor systems which can provide insights into the processes that form the low mass elliptical population as a whole.

Here, we study a large sample of low-redshift ($z < 0.3$) blue elliptical galaxies, which have low stellar masses, with a median stellar mass of $10^{8.7}$ $M_{\odot}$. We explore the physical processes that drive their stellar assembly, using multi-wavelength data in the COSMOS field, which includes extremely deep imaging from the Hyper Suprime-Cam Subaru Strategic Program (Aihara et al., 2022, ; HSC-SSP). Our study offers several novelties which allow us to select blue ellipticals, probe their properties and constrain their link to normal ellipticals with higher precision and in greater detail than has been possible in previous work.

Our blue ellipticals are selected from HSC-SSP images using a novel unsupervised machine-learning algorithm, which is designed to autonomously separate objects with similar morphological properties (e.g. colour, texture and luminosity) in large survey datasets. As described below, this automated selection is then augmented by Sersic fitting to remove stars and ensure that the objects being selected have the strongly peaked profiles expected in ellipticals. The HSC-SSP images used in this study have point-source detection limits of 28.1, 27.7 and 27.4 magnitudes in the *g*, *r* and *i* bands respectively, around 6 magnitudes deeper than standard depth imaging from the SDSS. These represent the deepest survey data used for the analysis of such blue ellipticals to date. This allows us to probe the role of mergers in driving the formation of our blue ellipticals, by looking for merger-induced tidal features more effectively than has been possible in previous work.

The exquisite deep, multi-wavelength (UV to infrared) photometric data available in the central regions of the HSC-SSP COSMOS field results in extremely precise photometric redshifts (with accuracies

better than 1 per cent for massive galaxies) and physical parameters such as stellar masses and star formation rates (SFRs), available via the COSMOS2020 catalog (Weaver et al., 2022). Finally, the accurate redshifts in this field enable us to use the DisPerSE (Sousbie, 2011) filament-finding algorithm to construct density maps and the filamentary structure of the cosmic web. This is then used to study the local density of our blue ellipticals and their topological locations (i.e. proximity to nodes, large-scale filaments and voids) with higher precision than has been possible in past studies.

This chapter is structured as follows. In Section 5.2 we describe the machine-learning algorithm used to extract our blue ellipticals and the properties of the HSC-SSP on which the algorithm has been implemented. In Section 5.3 we describe the COSMOS2020 catalog from which we extract photometric redshifts and physical parameters for our galaxies. In Section 5.4 we describe the algorithm used to construct maps of local density and topological structure that we use to study the locations of our galaxies in the cosmic web. In Section 5.5 we explore the role of different processes (e.g. mergers and gas accretion) in the stellar assembly of our blue ellipticals and explore their connection to normal ellipticals (i.e. those that are not blue) to make a broad comment on how the bulk of the stellar mass in low mass elliptical galaxies is likely to form. We summarise our findings in Section 5.6.

5.2 Galaxy morphologies via unsupervised machine learning

Our blue ellipticals are identified using an unsupervised machine learning (UML) algorithm (Lazar et al. in prep), which has previously been implemented on images from both the *Hubble Space Telescope* (Hocking et al., 2018) and the HSC-SSP (Martin et al., 2020). The algorithm samples a large number of random patches (8 by 8 pixels in size) in the footprints of the objects of interest (in our case, galaxies), in multi-band survey imaging. Here, we use HSC-SSP images of galaxies in the *griz* bands. The algorithm then uses a radial Fourier transform to construct a 'feature vector' for each patch, that holds information about its physical properties (principally colour, texture and intensity). Patches are clustered via a growing neural gas algorithm (Fritzke, 1995) and hierarchical clustering (Johnson, 1967), and galaxies with similar patch properties are then grouped together using k-means clustering. Arbitrarily large galaxy samples are rapidly and autonomously compressed into $\sim$150 'morphological clusters', with typical purities greater than 90 per cent. $\sim$150 clusters are needed because identical morphologies at different redshifts can occupy separate clusters (since galaxies change in size with redshift) and because this allows for a better morphological diversity between different clusters.

Each of these clusters is then associated, via visual inspection, with standard Hubble types. To perform this labelling, we visually inspect a sample of 30 objects from each cluster, whose feature vectors have varying Euclidean distances from the cluster centre in feature space - we select 10 which are closest to the centre, 10 which are furthest and 10 random objects in between. We label clusters into broad Hubble types: ellipticals, blue ellipticals and early- and late-type spirals (rows 1 - 4 in Figure 5.1). As described in Martin et al. (2020), these clusters obey known morphological trends in the literature e.g. stellar mass functions and the relationships between stellar mass, SFR and colour as a function of morphology. Peculiar objects such as blue ellipticals (row 4 in Figure 5.1) are naturally segregated into separate morphological clusters. We refer readers to Martin et al. (2020) for more details of the algorithm and note that morphological catalogs for the entire HSC-SSP Public Data Release 3 (PDR3) will be released shortly in a forthcoming paper (Lazar in prep.).

Note that, prior to inputting galaxy footprints into our UML classification algorithm, objects that are likely to be stars are initially filtered out by only selecting objects which have an HSC `extendedness` flag set to 1 (indicating that they are extended objects and not stars) and which have the `type` parameter in the COSMOS2020 catalog set to 0 (indicating, again, that they are galaxies). We perform additional filtering for stars when we select our final sample of blue ellipticals, as described in Section 5.5.1. Note also that we focus, in this particular study, on the nearby Universe ($z < 0.3$), where galaxies are extended, redshift errors are small and surface-brightness dimming is minimised, which aids the visual identification of tidal features that are signposts for recent mergers and interactions.

5.3 COSMOS2020: accurate physical parameters for galaxies

In this work we use physical parameters such as photometric redshifts, stellar masses, rest-frame colours and SFRs for our galaxies from the COSMOS2020 catalog (Weaver et al., 2022). COSMOS2020 is a high precision value-added catalog of over 1.7 million sources in the central region of the COSMOS field (Scoville et al., 2007). The physical parameters are estimated using photometry in 40 broad and medium band filters, from deep multi-band images taken using the following instruments: GALEX, MegaCam, ACS/HST, Hyper Suprime-Cam, Suprime-Cam, VIRCAM and IRAC. Here, we use parameters derived using the 'Classic' method which uses the `SExtractor` and `IRACLEAN` codes, with fluxes extracted within circular apertures, after the survey images are homogenized to a common point-spread function. The physical parameters are calculated using the `LePhare` SED-fitting algorithm (Ilbert et al., 2013). The photometric redshifts have accuracies better than 1 and 4 per cent for bright ($i < 22.5$ mag) and faint ($25 < i < 27$ mag) galaxies respectively, making this catalog ideally suited for our purposes.

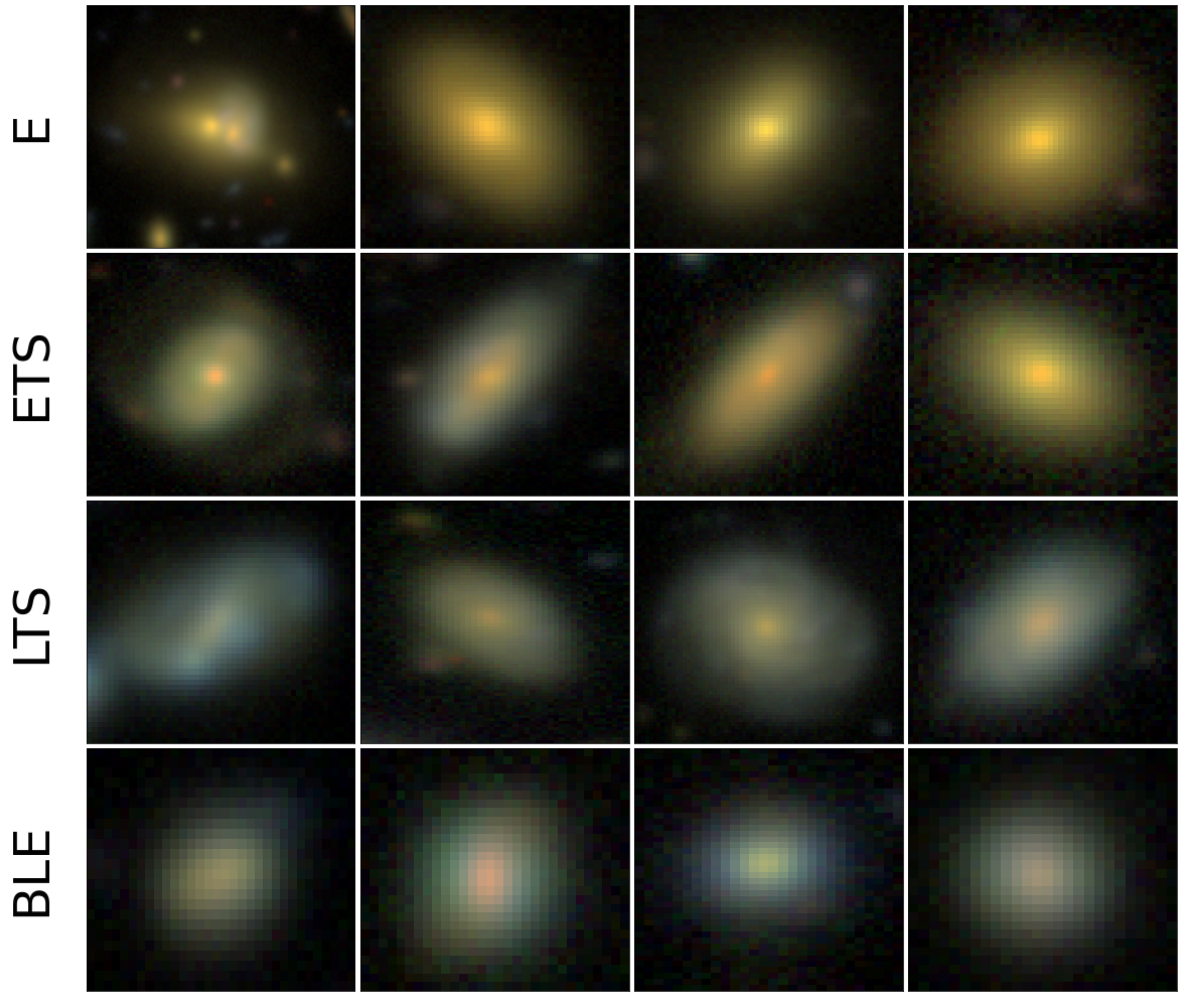

FIGURE 5.1: Examples of morphological clusters (show in different rows), outputted by the unsupervised machine learning algorithm, which have been visually associated with Hubble types: ellipticals (E), early type spirals (ETS), late type spirals (LTS) and blue ellipticals (BLE).

While accurate photometric redshifts are available from COSMOS2020 for all objects, where possible we use spectroscopic redshifts from the literature (which have been compiled and released as part of the HSC-SSP DR3) only as a way to validate our objects as galaxies.

5.4 Measurement of local density and topological structure

We use the `DisPerSE` algorithm (Sousbie, 2011) to measure local density and the locations of galaxies in the cosmic web e.g. their proximity to nodes and large-scale filaments. `DisPerSE` measures the density field, derived using Delaunay tessellations computed from galaxy positions (Schaap & van de Weygaert, 2000), to connect topological saddle points with peaks (i.e. local maxima or nodes) in the

density map via connecting segments, forming a set of ridges which constitute the 'skeleton'. The 'skeleton' describes the network of large-scale filaments that define the cosmic web, with the stationary points in density maps calculated by `DisPerSE` (minima, maxima and saddles) denoting the locations of voids, nodes and the centres of large-scale filaments. We refer readers to Sousbie (2011) for further details about the algorithm.

The properties of the skeleton are determined by a 'persistence' parameter, which is used to set a threshold value for defining critical pairs within the density map. A persistence of 'N' creates a skeleton where all critical pairs with Poisson probabilities below $N\sigma$ from the mean are removed. For our density analysis, we follow the methodology of Laigle et al. (2018), who have implemented `DisPerSE` on redshift slices of similar widths, constructed from the COSMOS2015 (Laigle et al., 2016) catalog, where the thicknesses of the slices are driven by the redshift uncertainties of the galaxy sample. They have also used the Horizon-AGN cosmological simulation (Dubois et al., 2014; Kaviraj et al., 2017) to show that datasets like COSMOS2015, which offer high photometric redshift precision (and those like COSMOS2020 which provide higher precision), can retrieve the broad 3D properties of the cosmic web from 2D projected density maps. We use a persistence of 2, which removes ridges close to the noise level, where structures are likely to be spurious.

The highly accurate redshifts in the COSMOS2020 catalog enable us to use well-defined and relatively narrow redshift slices to build density maps. We only use massive ($M_{\star} > 10^{10}$ M$_{\odot}$) galaxies to build these maps, as they dominate the local gravitational potential well and have the smallest redshift errors. Figure 5.2 indicates the physical distance associated with the median redshift error of massive galaxies, as a function of redshift. The redshift error reaches a broad minimum in the redshift range $0.18 < z < 0.3$, corresponding to a physical distance of $\sim$55 – 65 Mpc. In Section 5.5 we restrict the analysis of the environmental properties of our galaxies to this redshift range.

Following Laigle et al. (2018), we define the slice width to be twice the median redshift error of the massive galaxies, which results in slice widths of $\sim$ 100 Mpc in our redshift range of interest ($0.18 < z < 0.3$). We produce 9 density slices, each overlapping by 50 Mpc. When constructing each density map, every galaxy is weighted by the area under its redshift probability density function that is contained within the redshift limits of the slice in question. This takes into account the fact that the photometric redshifts of massive galaxies, although very accurate in COSMOS2020, do have associated errors. Figure 5.3 shows an example redshift slice for the reshift range $0.218 < z < 0.243$, with the associated density map and skeleton calculated using `DisPerSE`, and galaxies of various morphologies identified using our UML algorithm overplotted. Other redshift slices used in this study are shown in Appendix C.1.

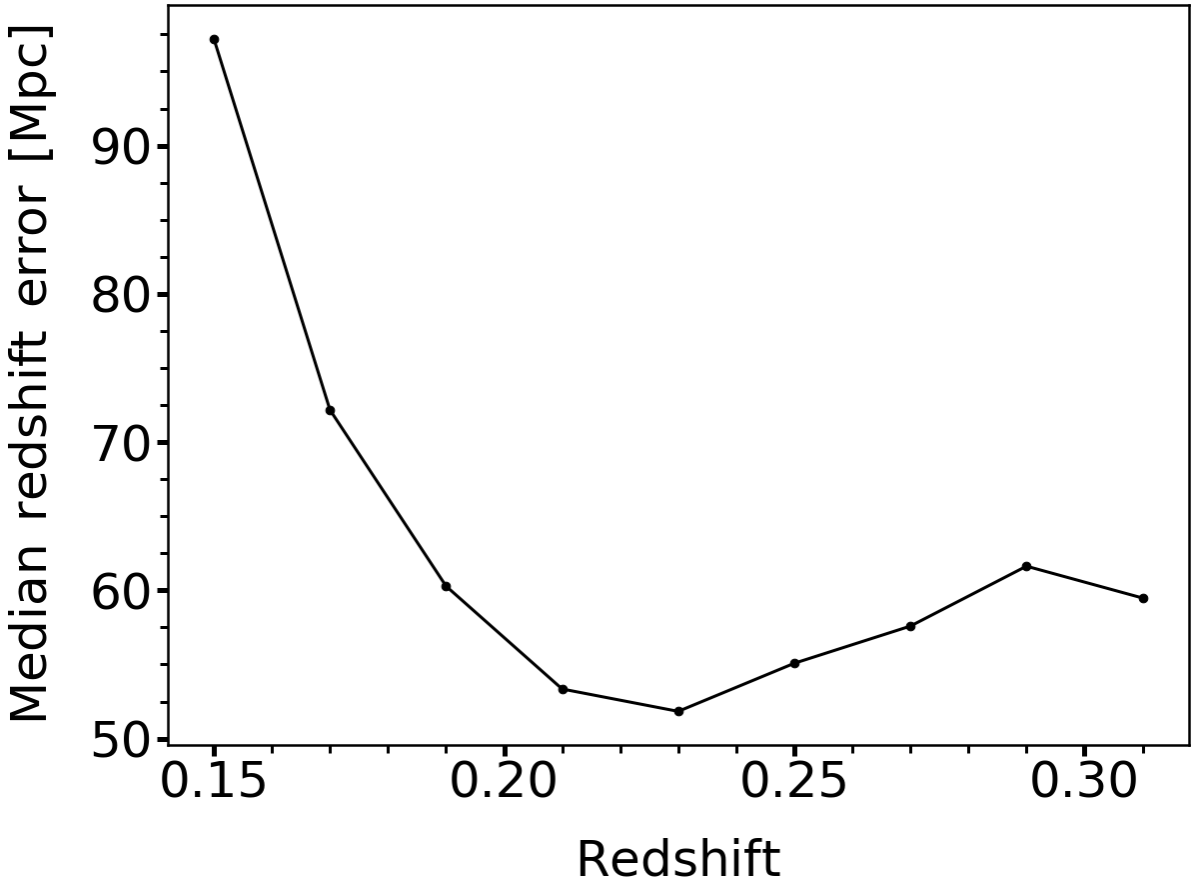

FIGURE 5.2: Median redshift error of massive ($M_\star > 10^{10}$ M$_\odot$) galaxies converted to a physical distance in Mpc, using the standard cosmology, as a function of redshift. As noted in the text, only galaxies with stellar masses greater than 10^{10} M$_\odot$ are used to construct density maps, as they dominate the local gravitational potential well and have the smallest redshift errors.

It is instructive to consider the large-scale structure that is enclosed within the COSMOS2020 region. The COSMOS2020 field has approximate dimensions of 1.6 $\times$ 2 degrees (Figure 5.3). At the lower ($z \sim 0.18$) and upper ($z \sim 0.3$) ends of the redshift range we probe here, 1.8 degrees corresponds to a linear size of $\sim$20 and $\sim$30 Mpc respectively. To get insights into the types of structures that might be present in a field of this size, we consider the NewHorizon cosmological simulation (Dubois et al., 2021), which has a similar size to the COSMOS2020 field at these redshifts. The largest dark matter halo in NewHorizon has a mass close to that of the Fornax cluster. It is, likely, therefore that the large scale structure within the COSMOS2020 field contains small clusters like Fornax, along with smaller groups of galaxies.

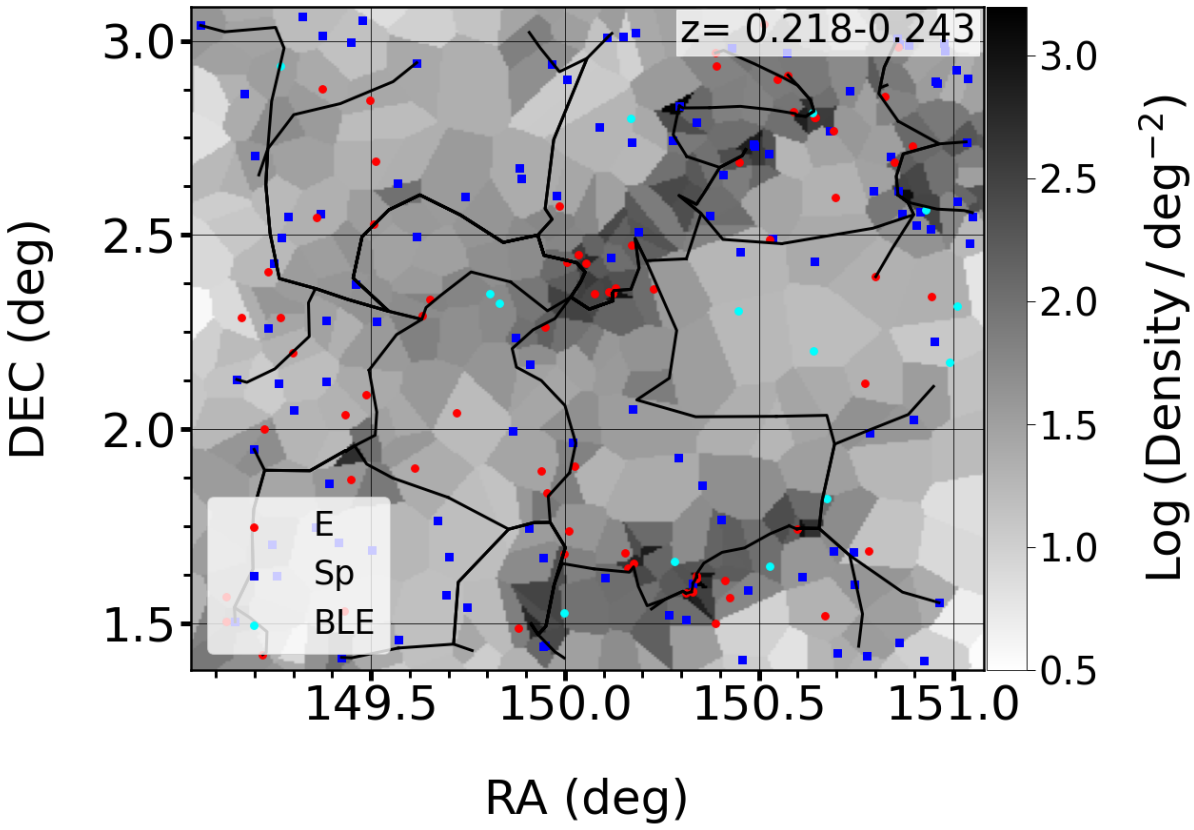

FIGURE 5.3: Example density map, with the associated skeleton calculated using `DisPerSE` shown using the black solid lines. Different morphological classes (red: ellipticals, blue: spirals, cyan: blue ellipticals) are shown overlaid. This particular density map corresponds to the redshift range $0.218 < z < 0.243$. All other density maps used in this study are shown in Appendix C.1.

5.5 Blue ellipticals: the importance of secular gas accretion and the insignificance of mergers

In this section, we first describe how blue ellipticals are identified using the morphological clusters produced by the UML algorithm described in Section 5.2. We then combine their physical and morphological properties with their locations in the cosmic web to explore the processes that are likely to give rise to these systems. Finally, we consider their connection to the general elliptical population and study the implications of our findings on the evolution of low mass elliptical galaxies as a whole.

5.5.1 Identification

As noted in Section 5.2, morphological clusters outputted by our classification algorithm are associated via visual inspection with standard Hubble types - broadly ellipticals and spirals. Elliptical galaxies which are blue naturally end up in their own morphological clusters, as the classification algorithm

uses colour as one of its primary discriminants for separating objects of different morphological types. Galaxies in these clusters represent a sample of blue elliptical *candidates* which form the starting point for selecting our final sample of galaxies (amounting to $\sim$1100 initial candidates), as described below. For these candidate systems we (1) perform further filtering for stars and (2) select a subset of objects that have strongly peaked profiles that are typical of elliptical morphologies (e.g. de Vaucouleurs, 1948; Kormendy, 1977). This subset forms the final sample of blue ellipticals that underpins our study.

We first fit the g band HSC image of each blue elliptical candidate with a single-component Sersic profile using `PyImfit` (a Python wrapper for `Imfit` from Erwin, 2015), convolved with the point spread function. We then remove objects which have Sersic indices less than 4 (since these may contain a disc component, e.g. de Vaucouleurs, 1959, 1977; Blanton et al., 2003; Pannella et al., 2006; De Propris et al., 2016). Although we perform an initial filtering for stars using two catalog parameters at the outset, before any galaxies are inputted into the classification algorithm (as described above), we take additional steps to ensure that our blue elliptical sample is not contaminated by stars. Note that 40 objects have spectroscopic redshifts, confirming that they are galaxies, and are therefore excluded from the filtering process described below.

To perform additional filtering for stars, we use the catalog of Bailer-Jones et al. (2019), using the maximum value from the residual map of the Sersic fits in the g-band. The Bailer-Jones et al. (2019) catalog contains star, galaxy and quasar classifications for objects in the GAIA Data Release 2 (DR2). The authors make use of a supervised classifier, based on Gaussian Mixture Models, which use the GAIA DR2, cross-matched with spectroscopic classifications from the SDSS as a training set.

The maximum residual values of objects that are identified as stars and quasars in Bailer-Jones et al. (2019) are significantly higher than those for objects that are identified as galaxies. We define a conservative threshold of 1.5 for this maximum residual value, in order to differentiate between stars and galaxies. This threshold value corresponds to the lowest maximum residual value among all objects labelled as stars in the Bailer-Jones et al. (2019) catalog. Appendix C.2 shows examples of blue elliptical candidates which have Sersic indices greater than 4 and objects classified as stars by Bailer-Jones et al. (2019). This leaves a final sample of 108 blue ellipticals which forms the basis of our study (of which 40 i.e. $\sim$37 per cent have spectroscopic redshifts).

It is worth noting here that the Sersic index (n) threshold which defines elliptical galaxies may vary as a function of stellar mass, with studies in the very nearby Universe (e.g. S^4G sample Watkins et al., 2022) suggesting that a lower value than $n \sim 4$ may be appropriate in the low mass regime. Nevertheless, we use $n \sim 4$ as our threshold here, as the S^4G galaxies probe significantly lower redshifts than our

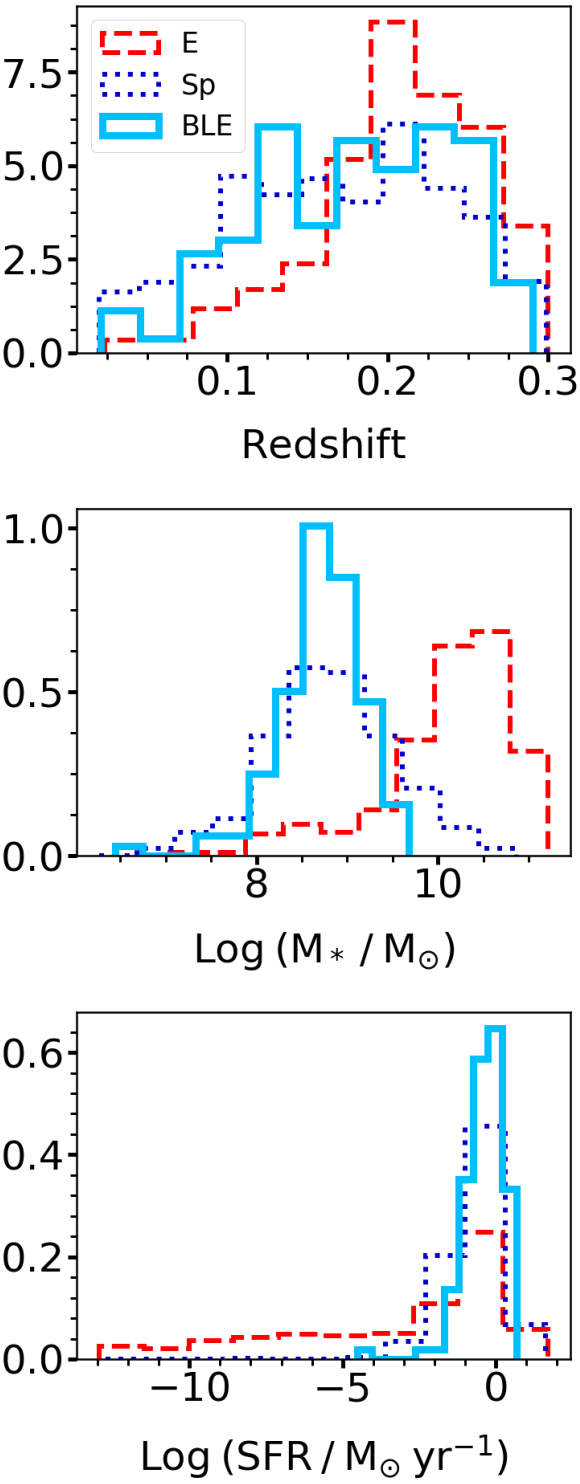

FIGURE 5.4: Redshift (top), stellar mass (middle) and star fomation rate (bottom) distribution for our blue ellipticals (BLE) and other morphological types: ellipticals (E) and spirals (Sp).

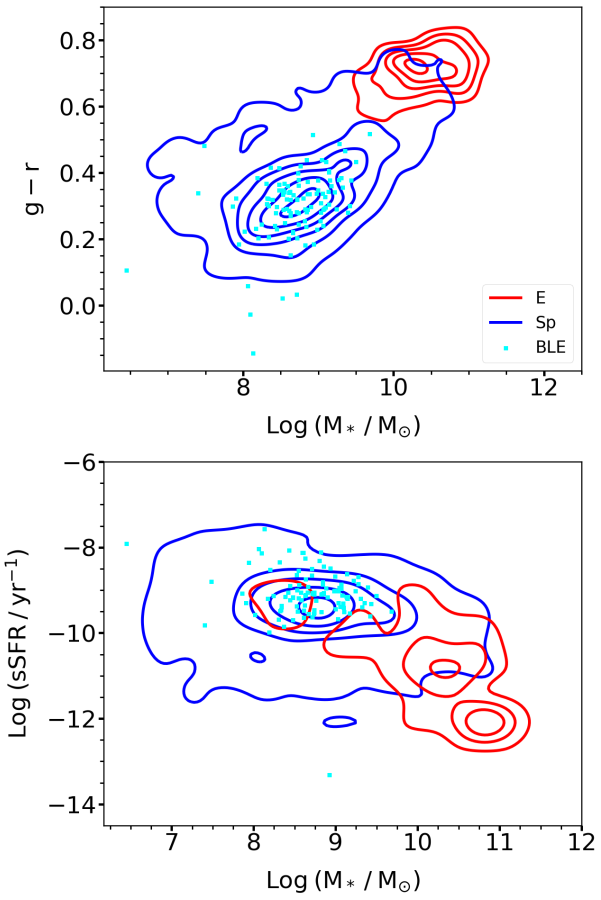

FIGURE 5.5: Rest-frame $g-r$ colour (top) and specific star formation rates (bottom) of our blue ellipticals (BLE) and other morphological types: ellipticals (E) and spirals (Sp).

objects and because statistical morphological studies of large samples of low mass galaxies outside the local Universe, which would require deep-wide surveys, are still largely missing. The existence of low mass ellipticals with high concentrations is therefore not ruled out. It is also worth noting that the high central concentrations in our blue ellipticals are, at least in part, driven by central ongoing star formation in these galaxies. The profiles are likely to become shallower as the star formation subsides and the galaxies quench.

5.5.2 Physical properties

The top and bottom panels of Figure 5.4 show the redshift and stellar mass distributions respectively of our blue ellipticals and other morphological classes in our galaxy sample. Our blue ellipticals have a median stellar mass of $10^{8.7}$ $M_{\odot}$ (similar to blue ellipticals studied in some past work e.g. Kannappan et al., 2009; Mahajan et al., 2018; Moffett et al., 2019). The median redshift of our sample is 0.19. Our blue ellipticals exhibit specific star formation rates and rest-frame colours that are comparable to objects

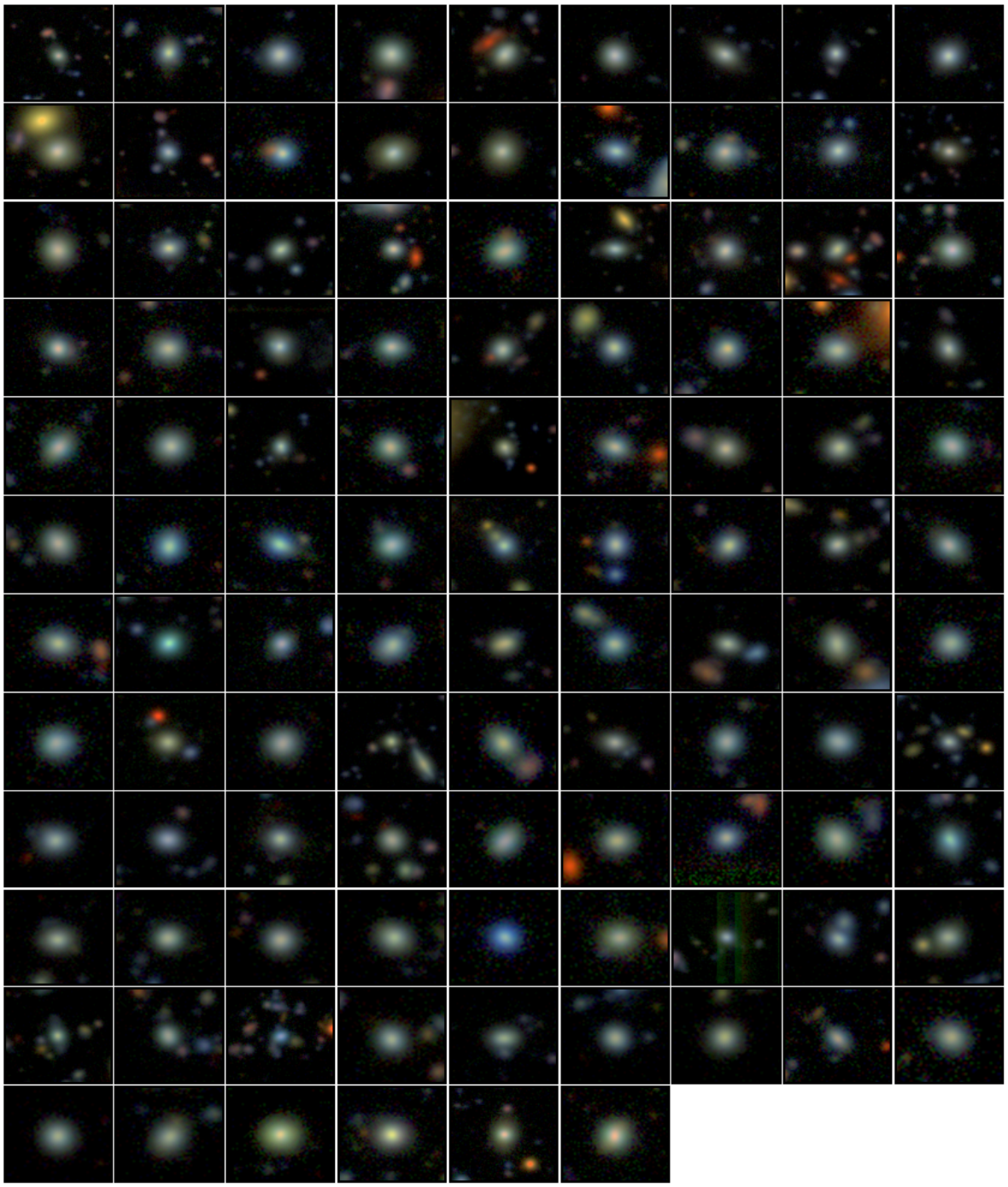

FIGURE 5.6: Blue ellipticals which are relaxed i.e. show no evidence for tidal features in the deep HSC images.

on the main sequence of star-forming galaxies, which is dominated by late-type (i.e. spiral) systems (Figure 5.5). As expected, the population of normal elliptical galaxies, in contrast, dominates the red sequence. The behaviour seen here is similar to that in Schawinski et al. (2007) and Schawinski et al. (2014) who have also studied low-mass blue ellipticals in the nearby Universe.

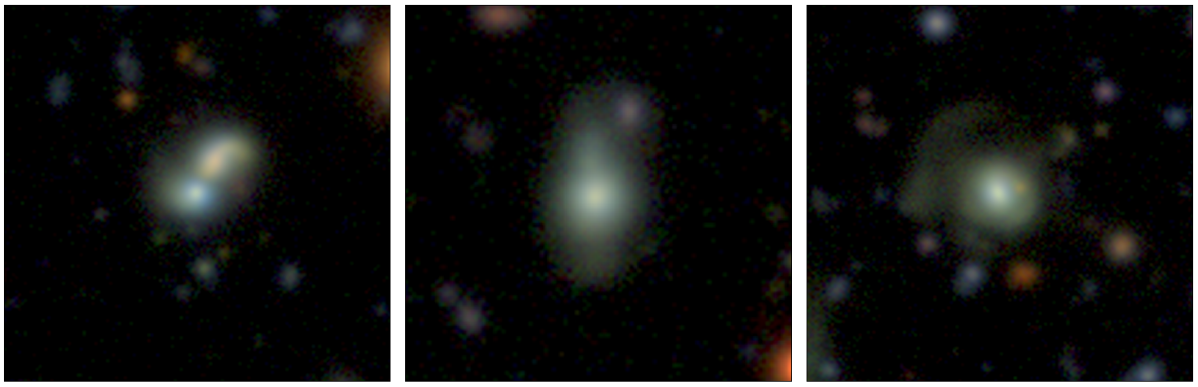

FIGURE 5.7: Blue ellipticals which exhibit either evidence of an ongoing merger (left-hand panel) or tidal features indicative of a recent interaction (other panels).

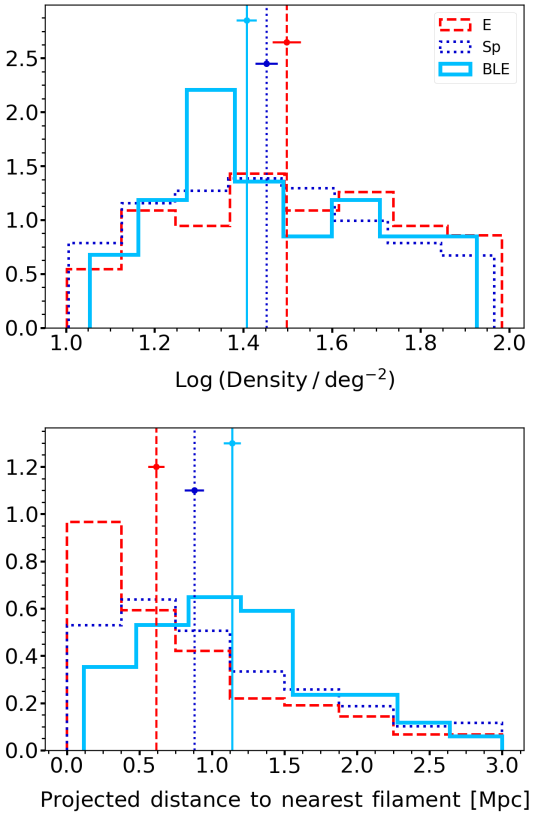

FIGURE 5.8: The distributions of local density (top) and projected distance from the nearest filament (i.e. saddle point within the skeleton) (bottom) for different morphological types (red: ellipticals, dark blue: spirals, light blue: blue ellipticals). The medians of the distributions are represented by dashed lines. The standard error in these median values are calculated using 100 bootstrapped iterations and are shown using the solid horizontal lines.

5.5.3 Absence of tidal features - blue ellipticals typically do not evolve via galaxy interactions

We first consider whether the star formation seen in our blue ellipticals is likely to be driven by a recent merger or interaction. The HSC images of our blue ellipticals offer an unprecedented combination of depth and resolution, compared to past surveys like the SDSS (on which many previous studies of blue ellipticals have been based). Recent mergers or interactions produce tidal features (e.g. Toomre & Toomre, 1972; Schweizer, 1982; Mihos & Hernquist, 1996; Springel & Hernquist, 2005; Yoon et al., 2022; Martin et al., 2022), which are best identified by direct visual inspection of galaxy images (e.g. Darg et al., 2010). Indeed, tidal features are readily detected in images from surveys like the SDSS Stripe 82 (e.g. Kaviraj, 2014b), which are around 3 magnitudes shallower than the images we use here.

Visual inspection of the HSC images indicates that only 3 out of the 108 blue ellipticals ($2.8^{+2.6}_{-0.9}$ per cent) show either a tidal feature indicative of a recent interaction or an ongoing merger. To study how this interaction fraction compares to that in the general galaxy population, we select a random sample of 108 galaxies within the COSMOS field which have the same stellar mass and redshift distribution as our blue ellipticals. We then visually inspect the HSC images of this general sample of galaxies. The tidal feature fraction in this general sample is $5.6^{+3.1}_{-1.5}$ per cent. The errors in the tidal feature fractions have been calculated following Cameron (2011)[1]. Our blue ellipticals therefore show a lower tidal feature fraction than the general galaxy population, indicating that mergers are not a preferential channel for their recent evolution.[2] Figures 5.6 and 5.7 show the blue ellipticals which are relaxed and show no tidal perturbations and the three objects that do, respectively.

Note that, if the star formation is indeed merger-driven, then the chances of detecting tidal features is expected to be *higher* when the galaxy is blue, since the star formation episode has started relatively recently and not much time has elapsed in which the tidal features can fade (e.g. Lofthouse et al., 2017). The negligible incidence of tidal features in our blue ellipticals indicates that the star formation observed in these systems is not driven by mergers and interactions but rather by secular gas accretion. This result is consistent with the findings of Mahajan et al. (2018) who combine optical and HI data from the GAMA (Driver et al., 2016) and ALFALFA (Giovanelli et al., 2005) surveys to reach a similar conclusion about

[1]Cameron (2011) calculate accurate Bayesian binomial confidence intervals using the quantiles of the beta distribution. These are considered to be more accurate than simpler methods, like using the normal approximation, which misrepresents the statistical uncertainty under the sampling conditions (e.g. small number counts) that are often encountered in astronomical surveys.

[2]It is worth noting that tidal features are also readily visible, in these deep HSC images, around dwarf galaxies which have significantly smaller stellar masses than the blue ellipticals in this study. Since tidal debris tends to become fainter when lower mass galaxies are involved, this strengthens the conclusion that, if they were present, tidal features would indeed be visible around our blue ellipticals.

the star formation activity in such systems. Note, however, that while Mahajan et al. (2018) arrive at this conclusion due to blue ellipticals being rich in HI and residing in low-density environments, we are able to probe this directly through the lack of tidal features, by virtue of the depth of our optical images. An important study that could be done in the future is to investigate how visible tidal features are in dwarfs at a given redshift by visually inspecting simulated dwarfs with added noise matched for a specific instrument depth.

We conclude this section by considering whether our blue ellipticals could have a tidal origin i.e. whether they could have formed in the tidal debris of mergers between two gas-rich massive galaxies (e.g. Bournaud & Duc, 2006; Ploeckinger et al., 2018). The images of our blue ellipticals (Figure 5.6) demonstrate that none of these galaxies exhibit tidal bridges with nearby galaxies, consistent with the fact that these objects reside further away from the massive galaxies that define the nodes and filaments of the cosmic web (see Section 5.5.4 below). It is also worth noting that the fraction of low mass galaxies that have a tidal origin is likely to be quite low ($\sim$6 per cent, see e.g. Kaviraj et al., 2012). Thus, it is unlikely that our blue ellipticals formed in the tidal debris of mergers between massive galaxies.

5.5.4 Blue ellipticals reside further away from nodes and large-scale filaments

Recall that Figure 5.3 shows the density map and skeleton calculated using `DisPerSE` in the redshift range $0.22 < z < 0.24$, with galaxies of various morphologies identified using our UML algorithm overplotted (other redshift slices are presented in Appendix C.1). Visual inspection of these density maps suggests that, unlike their normal (red) counterparts which tend to populate dense regions (e.g. those around nodes), the blue elliptical population appears to reside preferentially in low-density regions, further away from nodes and large-scale filaments than e.g. normal elliptical galaxies.

These results indicate that our blue ellipticals populate low-density regions of the Universe, at larger projected distances from the cosmic web than other morphological types. In Figure 5.8, we quantify the locations of the various morphological types within the cosmic web, in terms of local density and the 2D projected physical distance from the nearest nodes and large-scale filaments. The distances of galaxies from the nearest nodes and filaments are obtained by calculating the projected euclidean distance between the object in question and local maxima in the case of nodes and saddle points in the case of large scale filaments (which form the skeleton). Note that this analysis uses all density maps (which span the redshift range $0.18 < z < 0.31$). While the different morphological types exhibit similar distributions in local density, with their medians coinciding within their statistical uncertainties (top panel of Figure 5.8), strong differences are found between the different morphological classes in terms

of the projected distances from the nearest nodes and large-scale filaments. We find that blue ellipticals completely avoid the (projected) inner ∼0.5 Mpc around nodes. And while the distribution for normal ellipticals peaks at 0 – 0.3 Mpc from their nearest large-scale filaments, the corresponding range of values for the blue ellipticals is 0.8 – 1.3 Mpc (bottom panel of Figure 5.8).

Even though such low density regions are relatively poor in gas (Krumm & Brosch, 1984; Burchett et al., 2020; Li et al., 2022), they are likely to offer the best conditions for star formation in such low-mass objects to be driven via gas accretion. This is because these objects will dominate their immediate environment and will not have to compete with the larger gravitational potential wells of more massive galaxies (e.g. Kreckel et al., 2012; Beygu et al., 2013; Das et al., 2014, 2015; Florez et al., 2021; Jian et al., 2022), which are found closer to large-scale filaments and are more likely to preferentially accrete the gas available in these regions (e.g. Song et al., 2021). While we do not have information about the past assembly history of our blue ellipticals, it is interesting to note that direct gas accretion could also impart their elliptical morphology. This can happen, for example, if the object in question is fed by multiple streams from the cosmic web. The net angular momentum of the gas sourced from multiple streams can be low because the angular momenta imparted by individual streams can largely cancel out (e.g. Crone Odekon et al., 2018; Song et al., 2021), resulting in new stars with a mass profile resembling that found in ellipticals (rather than a disc).

5.5.5 Blue ellipticals as progenitors of the general low mass elliptical population

We complete our study by exploring the connection between our blue ellipticals and the general low-mass elliptical galaxy population. We consider the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentile values of the distributions of projected distances from nearest filaments and nodes of the blue ellipticals, which translates to distances between 0.7 and 1.6 Mpc. The ratio of blue to normal ellipticals that lie both within these distances and within the stellar mass range of the blue ellipticals ($10^8 < M_\star < 10^{9.5}$ M$_\odot$), is around a factor of 2.

Kaviraj et al. (2007) show that galaxies in the blue cloud, in which star formation has been completely shut off, migrate out of the blue cloud over timescales of ∼1.5 Gyrs. This is consistent with the findings of Schawinski et al. (2007), who suggest that blue ellipticals likely remain in the blue cloud for at least around a Gyr, after which they migrate towards the red sequence, either as a result of natural or possibly feedback-driven gas exhaustion. In Figure 5.9 we show the stellar mass growth that can be expected for our blue ellipticals if their current measured SFR remains constant over 1 Gyr. In this estimation we take into account a stellar mass loss rate due to stellar evolutionary processes (such as red super-giant winds, asymptotic giant branch winds and SNe events) of 0.3 M$_\odot$ yr^{-1} from Leitner & Kravtsov (2011)

which has been obtained from SSPs evolved until the present time assuming a Chabrier (2003) IMF. The distribution of mass growth values indicates that some blue ellipticals can increase their stellar mass by 100 per cent or more i.e. they are capable of more than doubling their stellar mass over a timescale of around a Gyr. However, as shown in Figure 5.9, the number of galaxies with a stellar mass growth of more than 100 per cent over 1 Gyr is very small. We note that this is only meant to provide a rough mass growth estimation over a 1 Gyr timescale since constant SFRs are usually unlikely to happen for large timescales. For example, the SFRs of low mass galaxies ($M_{\star}\sim10^{9}$ $M_{\odot}$) in cosmological simulations can have variations of more than 0.5 $M_{\odot}$ yr^{-1} over timescales of 3 Gyrs (e.g. Wang et al., 2024) in the local Universe ($z<1$) depending on whether they had significant starburst histories or not.

The number of blue ellipticals is sufficient for these galaxies to feed the population of normal ellipticals which reside in similar environments and have similar (or somewhat higher) stellar masses. In addition, in the context of hierarchical structure formation, galaxies will tend to migrate towards higher-density environments, so that ellipticals which currently reside in low-density environments will progressively populate higher-density environments over cosmic time. In that sense, our blue ellipticals may also feed the elliptical population in denser environments within the cosmic web.

It is worth noting here that theoretical studies of the formation of elliptical galaxies, using cosmological simulations (e.g. Kobayashi, 2004, 2005; Feldmann et al., 2010; Merlin et al., 2012; Rodriguez-Gomez et al., 2016b), appear aligned with our observational results. These studies show that, while most massive ($M_{\star} > 10^{10.5}$ $M_{\odot}$) ellipticals undergo (major) mergers in their formation histories, their lower mass counterparts have relatively relaxed, merger-free assembly histories. Taken together with our results, this suggests that accretion-driven stellar assembly, without recourse to mergers, could be a significant channel for stellar mass assembly in the low mass elliptical galaxy population.

5.6 Summary

The evolution of elliptical galaxies is a key topic in observational cosmology. Within the standard hierarchical paradigm, the formation of these systems, in the high-mass regime, is strongly linked to galaxy mergers and interactions. However, the low mass domain remains less well explored and empirical studies that probe the physical processes that drive the evolution of low-mass elliptical galaxies are highly desirable. In particular, studying populations of ellipticals when they are blue (and therefore actively building significant fractions of their stellar mass) can offer important insights into their formation.

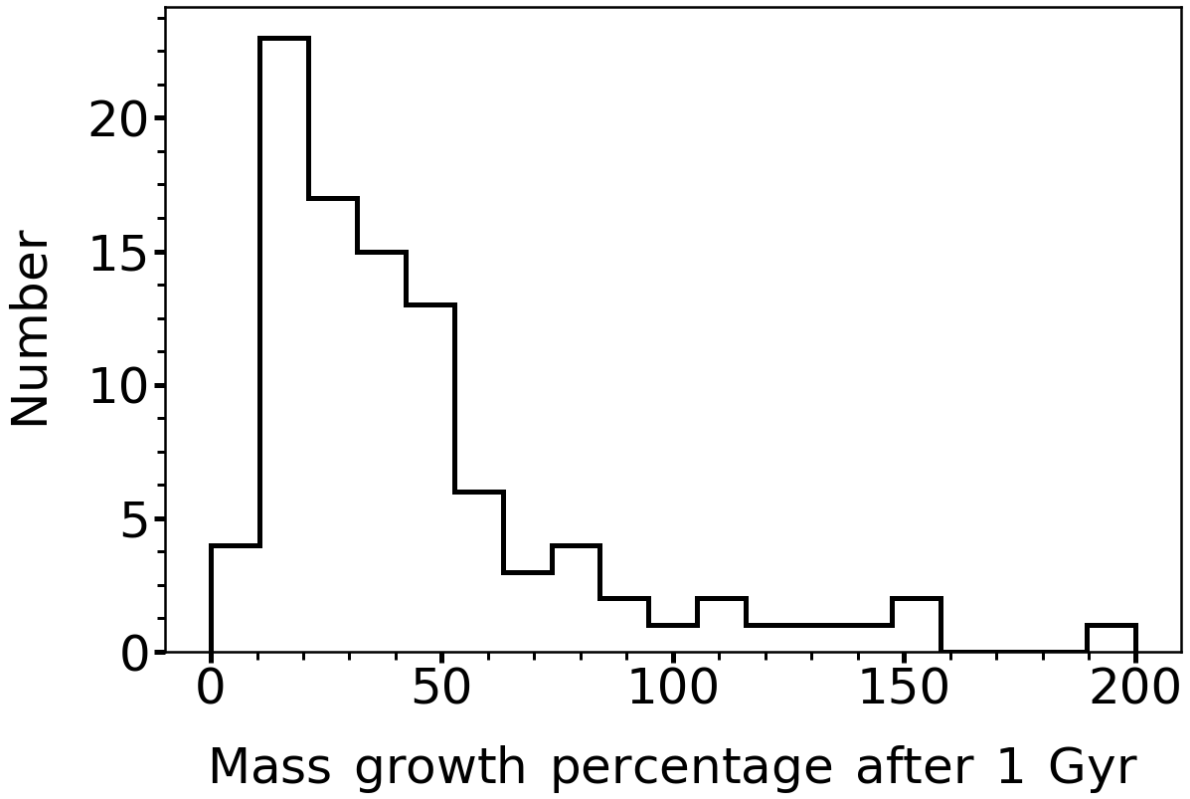

FIGURE 5.9: Mass growth percentage in our blue ellipticals, relative to their current stellar mass, if their current measured SFRs remain constant for 1 Gyr. This estimate takes into account the stellar mass loss due to stellar evolutionary processes calculated by (Leitner & Kravtsov, 2011).

Here, we have studied a sample of 108 blue ellipticals in the low mass regime (with a median mass of $\sim 10^{8.7}$ $M_{\odot}$) at $z < 0.3$. The novelties of this study include (1) making use of extremely deep images from the HSC-SSP (which are around 6 magnitudes deeper than standard-depth imaging from the SDSS), (2) exploiting 40 filter photometry that yields accurate physical parameters for galaxies (e.g. redshifts, stellar masses, SFRs and rest-frame colours) via the COSMOS2020 catalog and (3) quantitatively investigating the local density and locations in the cosmic web of our blue elliptical sample, using the DisPerSE structure finding algorithm. Our main results are as follows:

- Less than 3 per cent of our blue ellipticals show signs of tidal features (indicative of recent mergers and interactions), which are expected to be readily visible at the depth of the images used in this study. The incidence of tidal features in the blue ellipticals is also a factor of 2 lower than in the general galaxy population, strongly implying that the star formation activity is likely to be driven, not by mergers or interactions, but by secular gas accretion.

- The blue ellipticals typically inhabit low-density environments and reside further away from nodes and filaments of the cosmic web compared to other morphological types. Such environments are

likely to facilitate sustained and efficient gas accretion on to these systems because they are able to dominate their local environment, as opposed to higher density regions where nearby massive galaxies will preferentially accrete the available gas by virtue of their deeper potential wells.

- The number ratios of blue and normal (non-blue) ellipticals that have similar stellar masses and reside in similar environments is around a factor of 2. In addition, given their high SFRs, our blue ellipticals have the potential to significantly increase their stellar mass over timescales of ~ 1 Gyr. Thus, these blue ellipticals are likely progenitors of typical ellipticals at low (and slightly higher) stellar masses.

- Taken together, our results suggest that stellar mass growth driven by secular gas accretion plays a significant (and possibly dominant) role in the stellar mass build up of elliptical galaxies in the low mass regime.

Chapter 6

Conclusions

In this thesis, we have investigated the morphological diversity and structural properties of dwarf galaxies in the nearby Universe and we have explored the differences in the processes that drive stellar mass assembly in the dwarf and massive-galaxy regimes. We have also specifically analysed the role of mergers and environmental processes in shaping the evolution of dwarf elliptical galaxies. Furthermore, we have explored if traditional morphological parameters can be used to separate visually-classified morphological classes in the dwarf regime in the same way as they have been employed in massive galaxies. Finally, we have developed and presented an unsupervised machine learning algorithm for galaxy morphological classification aimed at deep-wide surveys that will dominate the astrophysical landscape in the next decades.

6.1 The morphological mix and structural properties of dwarf galaxies in the nearby Universe

The evolution of dwarf galaxies outside the very local Universe ($z > 0.02$) remains poorly understood. The principal reason behind this is that past large surveys like the SDSS are too shallow to detect typical dwarfs outside the local Universe. The dwarfs in these datasets at cosmological distances have high star formation rates and blue colours, as the fainter, red dwarfs preferentially fall out of the selection. As a result dwarf populations outside the local Universe in past surveys show strong biases, which prevents a comprehensive exploration of dwarf evolution.

In Chapters 2 and 3 we have studied a complete, unbiased sample of over 200 dwarf galaxies (10^8 M$_\odot$ $< M_\star < 10^{9.5}$ M$_\odot$) in the COSMOS field at $z < 0.08$, in combination with ultradeep optical imaging from

the Hyper Suprime-Cam and accurate physical properties from the COSMOS2020 catalog. We have used these datasets to quantify the morphological diversity, structural properties, interaction fractions, colours and environmental properties of our dwarfs as a function of stellar mass.

Following visual inspection of deep images, we have found three main morphological classes in the dwarf regime: early type galaxies (ETGs), late type galaxies (LTGs) and featureless systems with number fractions of 43, 45 and 10 per cent respectively. The ratio of dwarf early to late type galaxies is around 1:1 which is similar to what is found in the massive regime.

The median effective radius of dwarf LTGs is around a factor of 2 larger than that of the ETGs and both morphologies are found to have similar median effective surface brightness values. The dwarf featureless type presents a smooth diffuse structure and has a lack of central light concentration. This galaxy type is not found in the massive regime and, while the median effective surface brightness of the dwarf ETGs and LTGs is similar, the featureless class is much fainter. Given that we find dwarf featureless galaxies in an environment relatively devoid of overdensities, this work suggests that featureless dwarfs in low-density environments are more likely to have been created via internal baryonic feedback, rather than by environmental processes. Furthermore, compared to dwarf ETGs, the dwarf featureless class seems to have a similar median effective radius but fainter median effective surface brightness by $\sim$1 mag arcsec^{-2}. This suggests that their stellar mass assembly is significantly different. It is worth noting that dwarfs residing within the 'UDG' region in $\mu_{\rm e}$ vs $R_{\rm e}$ space are a continuous extension of the galaxy population towards lower values of effective brightness, indicating that UDGs are not a distinct class of objects. We have shown that interacting dwarfs are larger and bluer at all radii while their non-interacting counterparts show the opposite behaviour which suggests that interactions enhance star formation across the entire galaxy.

Traditional morphological parameters such as 'CAS', M_{20} and the Gini coefficient are not as efficient in separating ETGs and LTGs in the dwarf regime as they are in the massive-galaxy regime. However, the asymmetry parameter proves to be useful in distinguishing interacting from their non-interacting counterparts.

6.1.1 Differences in stellar mass assembly in dwarf and high mass early type galaxies

The underlying physical processes that shape early type galaxies are a key element in understanding their formation and evolution. Although, there is a rich literature on this subject in the massive-galaxy regime, the dwarf regime ($M_{\star} < 10^{9.5}$ M$_{\odot}$) remains mostly unexplored.

We have found significant differences between the massive and dwarf regimes in terms of the structural and photometric properties of early type galaxies. Compared to their massive counterparts, dwarf early type galaxies show a significantly lower incidence of tidal interactions, they are less centrally concentrated and have similar rest frame colours as dwarf late type galaxies.

By analysing and comparing the effective radius - stellar mass plane from our study with studies in the massive regime (e.g. Shen et al., 2003) we have found a higher dependence of effective radius on stellar mass (i.e. steeper slope) for $M_{\star} > 10^{10}$ $M_{\odot}$ than for the galaxies in our study with $M_{\star} < 10^{9.5}$ $M_{\odot}$, where the slope of this relation is close to zero. Furthermore, the colour profiles of massive ETGs have either flat or negative gradients in contrast with dwarf galaxies which show positive gradients and often exhibit extended blue cores suggesting recent star formation in their centres.

These findings suggest that the stellar mass assembly of massive and dwarf early types may transition from a scenario where 'inside-out' stellar mass growth is more dominant in the massive regime (generally via minor mergers) to a framework where 'outside-in' mass assembly regulated by baryonic feedback dominates in the dwarf regime. Dwarf ETGs are, therefore, shaped mainly by secular processes while massive ETGs have a more violent history dominated by mergers.

We have presented some direct evidence of this hypothesis in Chapter 5, where we have used a sample of $\sim$100 dwarf blue elliptical galaxies (found using the algorithm described in Chapter 4) which have a median stellar mass of $10^{8.7}$ $M_{\odot}$. We have used ultradeep images from the HSC survey to investigate their interaction fractions, environmental properties and their connection to their red counterparts. We find that only 3 per cent of these objects have visible tidal features and when compared to a control sample which has the same mass and redshift distribution the incidence of tidal features in the blue elliptical galaxies is lower by a factor of 2.

These objects reside in relatively low density environments and when compared with normal (red) low mass ellipticals in the same environments (low density regions outside filaments) and the same stellar mass range, we have found that blue ellipticals are a factor of 2 more numerous. This suggests that the blue dwarf ellipticals may be the precursors of red dwarf ellipticals, as the blue systems migrate to overdense regions of the cosmic web and begin quenching. Moreover, these blue objects may also act as precursors for intermediate mass red ellipticals due to their significant star formation rates.

These findings support the view that mergers are not as important in the evolution of dwarf ellipticals as in their massive counterparts, with secular gas accretion likely to dominate the stellar mass assembly of ellipticals in the low-mass regime (as pointed out in Chapters 2 and 3).

6.2 Galaxy morphological classification in the HSC-SSP survey using unsupervised machine learning

Upcoming big data surveys such as LSST and Euclid will offer large sky areas at unprecedented depths. The unprecedented data volumes from these surveys mean that using traditional techniques in an efficient manner will be a significant challenge especially for galaxy morphological classification. Even techniques involving supervised machine learning may be challenging to use, since repeatedly building training sets, especially for datasets that have a short cadence will be time consuming. However, unsupervised machine learning (UML), does not require training on labelled data, making it well-suited to galaxy morphological classification for large surveys.

In Chapter 4 we have presented and implemented a UML technique on survey data from the HSC-SSP. The algorithm starts from an initial feature space characterized by the power spectrum of patches (of around twice the size of the PSF), which is used in a series of clustering and size reduction techniques in order to group together thousands of galaxies into $\sim$150 morphological clusters, each containing galaxies of a specific morphology.

We have used the algorithm on the HSC-SSP deep layer ($\sim$27 deg^2) for galaxies with $M_\star > 10^8$ M$_\odot$ at $z < 1$. Using the classifications found by the method we reproduce expected morphological fractions and star formation rate - stellar mass trends from the literature. The average accuracy of the method in separating different morphological classes is around $\sim$90 per cent. Furthermore, the algorithm can also identify bad data, such as shredded objects (i.e. HII regions), unresolved/noisy images or stars with high accuracy. We note that due to the observational biases described above, and the fact that smaller galaxies are not well resolved in ground based images, the classifications of galaxies with low stellar mass and high redshift need to be treated with some caution. Finally, in addition to a general morphological classification, the algorithm can isolate objects with specific morphologies and colours. We have presented an example of such a morphological class by studying a sample of blue ellipticals found by the method in Chapter 5.

Chapter 7

Future work

LSST and Euclid will reach a point source depth of $\sim$26.4 mag after 1 year of operation with angular resolution of $\sim$0.6 and $\sim$0.1 arcsec in the optical respectively. The depth of the images will allow us to detect complete samples of hundreds of thousands dwarfs with stellar masses greater than 10^8 $M_\odot$ at $z < 0.3$. Given the faint source depth and large statistical samples available from LSST and Euclid, and taking into account the fact that Euclid offers an angular resolution of around 0.1 arcsec (which translates to physical scales of $\sim$0.5 kpc at $z = 0.3$) the synergy between LSST and Euclid will be ideal for dwarf galaxy evolution studies in the nearby Universe. Two example avenues for future work in the near term are as follows:

- Implement the UML algorithm presented in Chapter 3 on the first 50 deg^2 of Euclid data which is expected as a early quick data release by December 2024 and/or on LSST commissioning data (expected early 2025) and then on future data releases from these two surveys. The post-process labelling of morphological clusters could be performed either using individual experts (like in this thesis) or by using Galaxy Zoo. This can be combined with structural parameters such as CAS, M_{20} and the Gini coefficient and environmental parameters to perform a similar study to Chapters 2, 3 and 5 but with unprecedented statistics. Some of the questions that can be answered are as follows: what are the dominant morphological classes in dwarf galaxies? what are the physical processes that dominate the evolution of each dwarf morphological class? how does the morphological mix change with redshift in the nearby Universe?

- Compare the results obtained using LSST and Euclid data to cosmological hydro-dynamical simulations such as NewHorizon (Dubois et al. 2021) and TNG50 (Nelson et al. 2019), which have

the required sub kpc resolution to properly resolve dwarfs. I will be able to trace the evolutionary histories of different dwarf morphologies in relation to their environment in these simulations and compare how this correlates with the real data. This project could provide the first statistical test of the evolution of dwarfs from a morphological perspective in relation to the predictions of cosmological hydrodynamical simulations. Some of the questions that can be answered are as follows: how do our empirical results compare to simulations such as New Horizon and TNG50? More importantly, can we build better simulations using these results?

Appendix A

A.1 Dwarfs in the DES comparison sample are biased towards blue, late-type galaxies

As noted in Section 2.4.8, DES dwarfs that conform to the selection criteria we use to construct the DES comparison sample are strongly biased towards, blue galaxies which are dominated by late-type morphologies. This prevents us from comparing our dwarfs to those in DES (as a result of which our DES comparison is restricted to massive galaxies only). Here, we demonstrate these biases. Figure A.1 presents the rest-frame $(g-i)$ colour vs stellar mass for HSC galaxies in the COSMOS2020 catalogue (shown using the heatmap) compared to the DES galaxies which satisfy the selection criteria described in Section 2.4.8 i.e. which have probabilities of being ETG or LTG higher than 90 per cent, redshift and stellar mass errors less than 0.02 and 0.2 dex respectively and the highest confidence flag for morphological classifications of 4. The DES population is shown using the orange contours. A random 1 per cent of the DES galaxies are shown overplotted using red points. While massive galaxies are visible in DES regardless of whether they are red or blue, this is not the case for dwarf galaxies. The DES galaxies become progressively bluer at lower stellar masses. Indeed, towards the lower end of the stellar mass range considered in this study, the small fraction of DES objects that do appear in the DES comparison sample are some of the most extreme in terms of their blue colour.

Figure A.2 shows a random sample of 35 DES dwarfs. Visual inspection of a random sample of 300 galaxies in this DES dwarf population (using *gri* colour composite images) indicates that the ETG fraction is $\sim$11 per cent, a factor of 4 lower than the ETG fraction observed in the significantly deeper COSMOS2020 catalogue ($\sim$44 per cent, see Section 2.3). Not unexpectedly, a consequence of the DES dwarfs being dominated by blue galaxies is that the morphological mix is also skewed towards late-type systems. Note that reducing the `confidence_flag` parameter to 3 does not change these results.

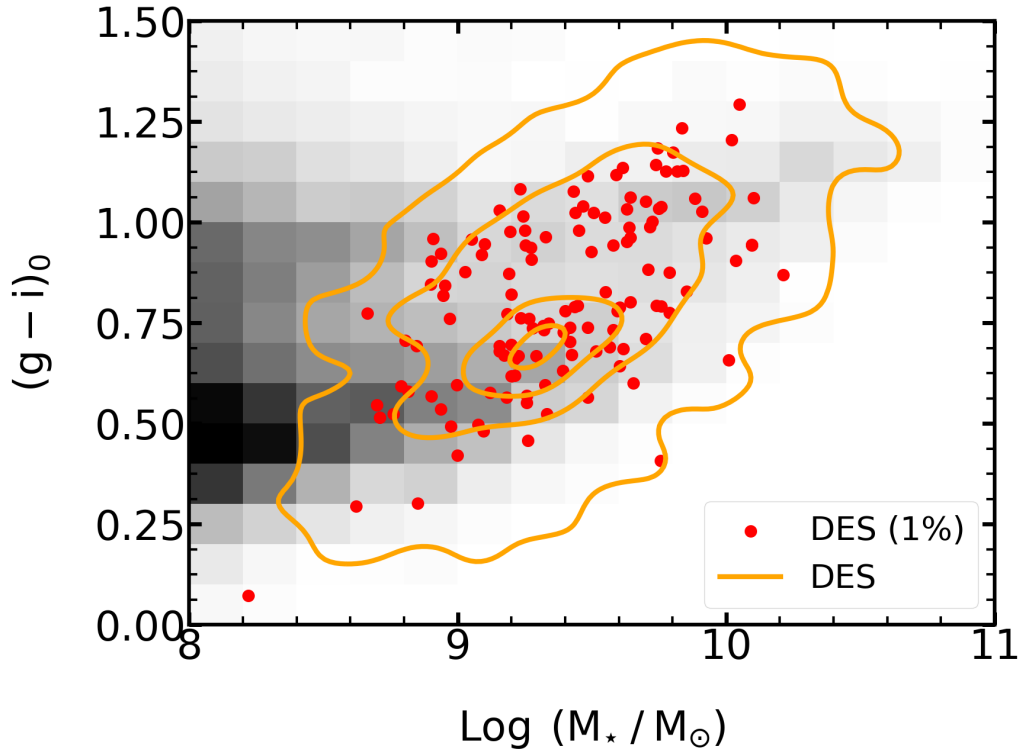

FIGURE A.1: $(g-i)$ colour vs stellar mass for galaxies in the COSMOS2020 catalogue (shown using the heatmap) compared to the DES galaxies (shown using the orange contours) described in Section 2.4.8 (i.e. those which have probabilities of being ETG or LTG higher than 90 per cent, redshift and stellar mass errors less than 0.02 and 0.2 dex respectively and which have the highest morphological confidence flag of 4). A random 1 per cent of the DES galaxies are shown overplotted using red points.

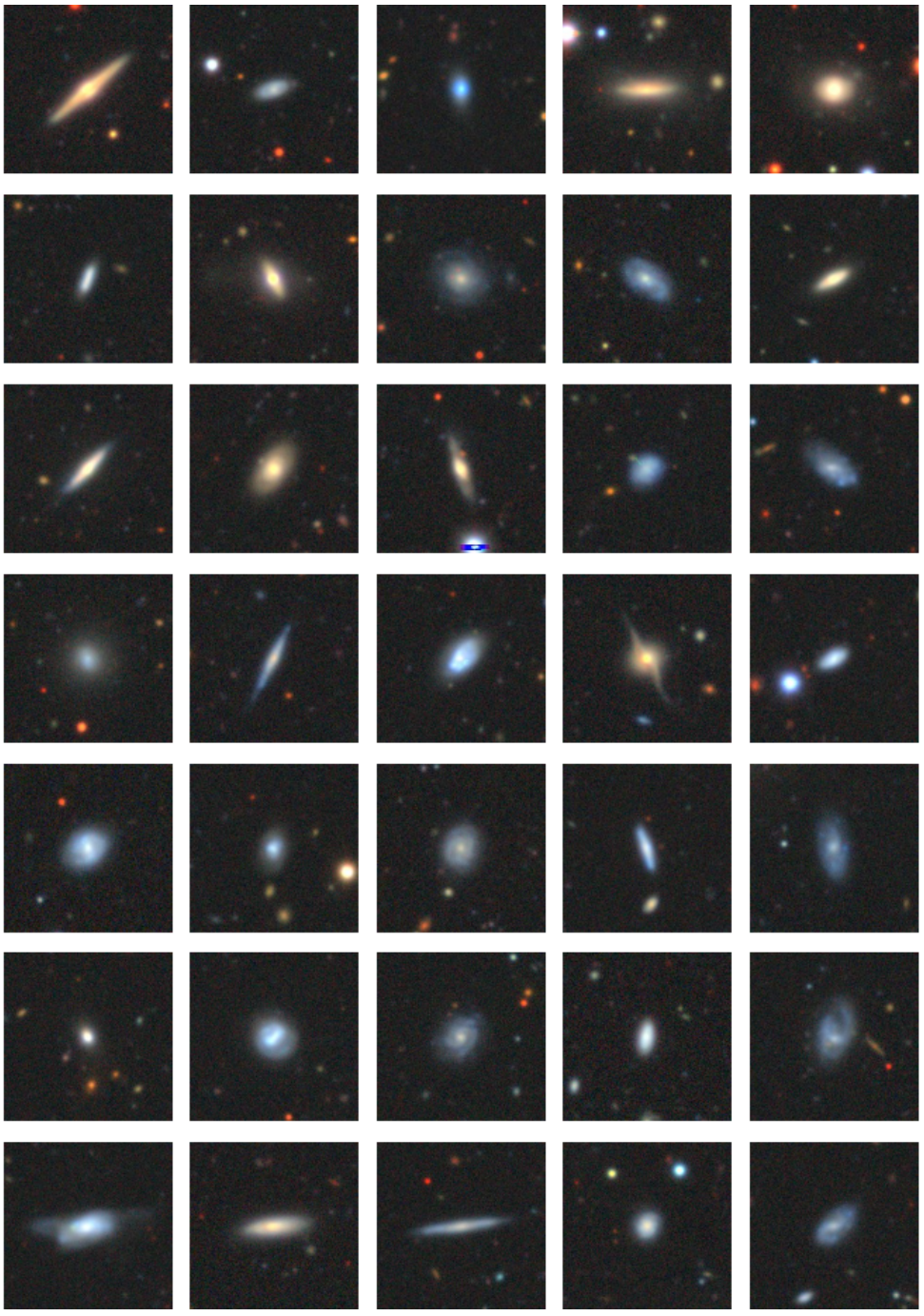

FIGURE A.2: *gri* colour composite images of a random sample of DES dwarfs which satisfy the selection criteria for the DES comparison sample i.e. galaxies which have probabilities of being ETG or LTG higher than 90 per cent, redshift and stellar mass errors less than 0.02 and 0.2 dex respectively and the highest confidence flag for morphological classifications of 4.

Appendix B

B.1 Colour bimodality in the nearby dwarf population

Figure B.1 presents the rest-frame $(u-i)$ colours of our dwarf galaxies. The distribution shows a bimodality around rest-frame $(u-i) = 1.7$. This value is therefore used to separate red and blue galaxies in this study.

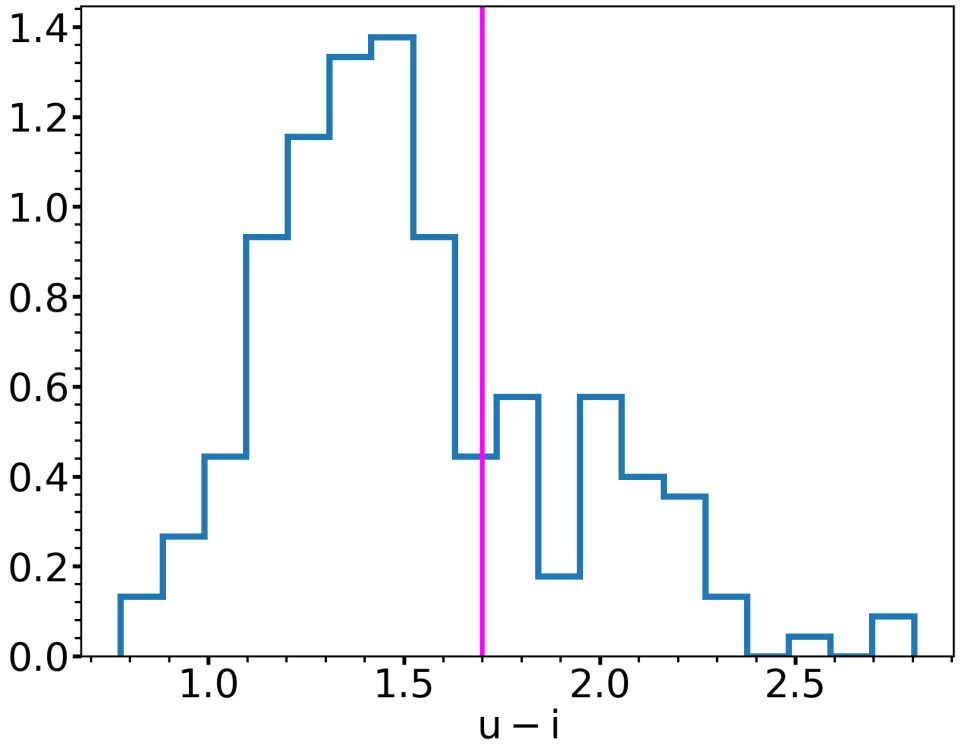

FIGURE B.1: The restframe $(u-i)$ colour distribution for our dwarf sample. The magenta vertical line represents the threshold between red and blue galaxies at restframe $(u-i)$=1.7.

Appendix C

C.1 Density maps

In this section we present all density maps used in our study (see text in Section 5.4 for details about their construction). Note that the density map for the redshift range $0.218 < z < 0.243$ is shown in Figure 5.3.

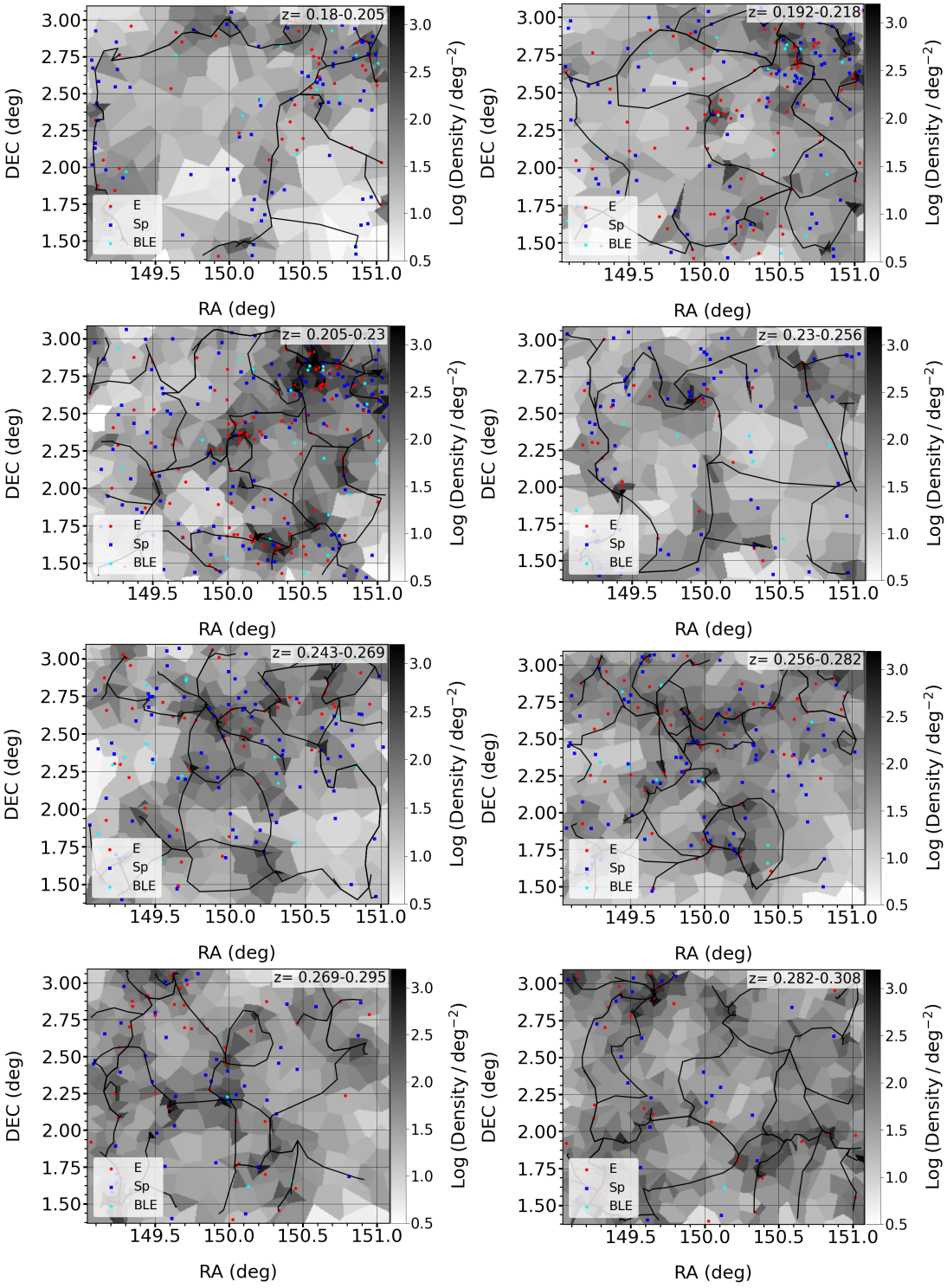

FIGURE C.1: Density maps and skeletons (shown using black solid lines) used in this study. Different morphological classes (red: ellipticals, blue: spirals, cyan: blue ellipticals) are shown overlaid.

C.2 Examples of sersic fits

In this section we show Sersic fits for typical examples of blue ellipticals which have both photometric and spectroscopic redshifts (Figure C.2), blue ellipticals which have only photometric redshifts available (Figure C.3) and objects classified as stars (Figure C.4) by our method.

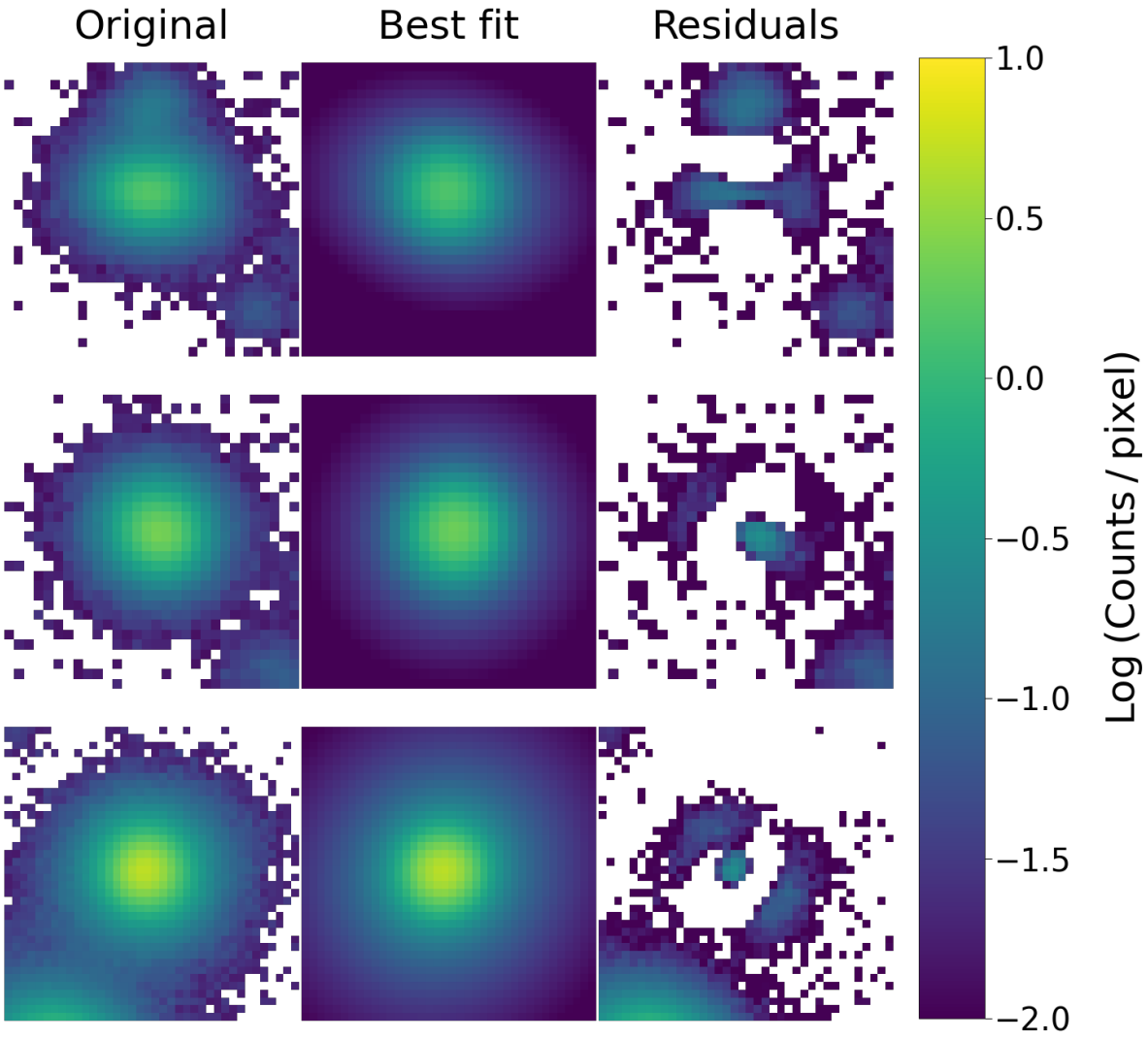

FIGURE C.2: Sersic fits for three random blue ellipticals which have spectroscopic redshifts available. From top to bottom the fitted Sersic indices and maximum residual values are 4.56, 7, 5.13 and 0.1, 0.3 and 0.7 counts pixel^{-1} respectively.

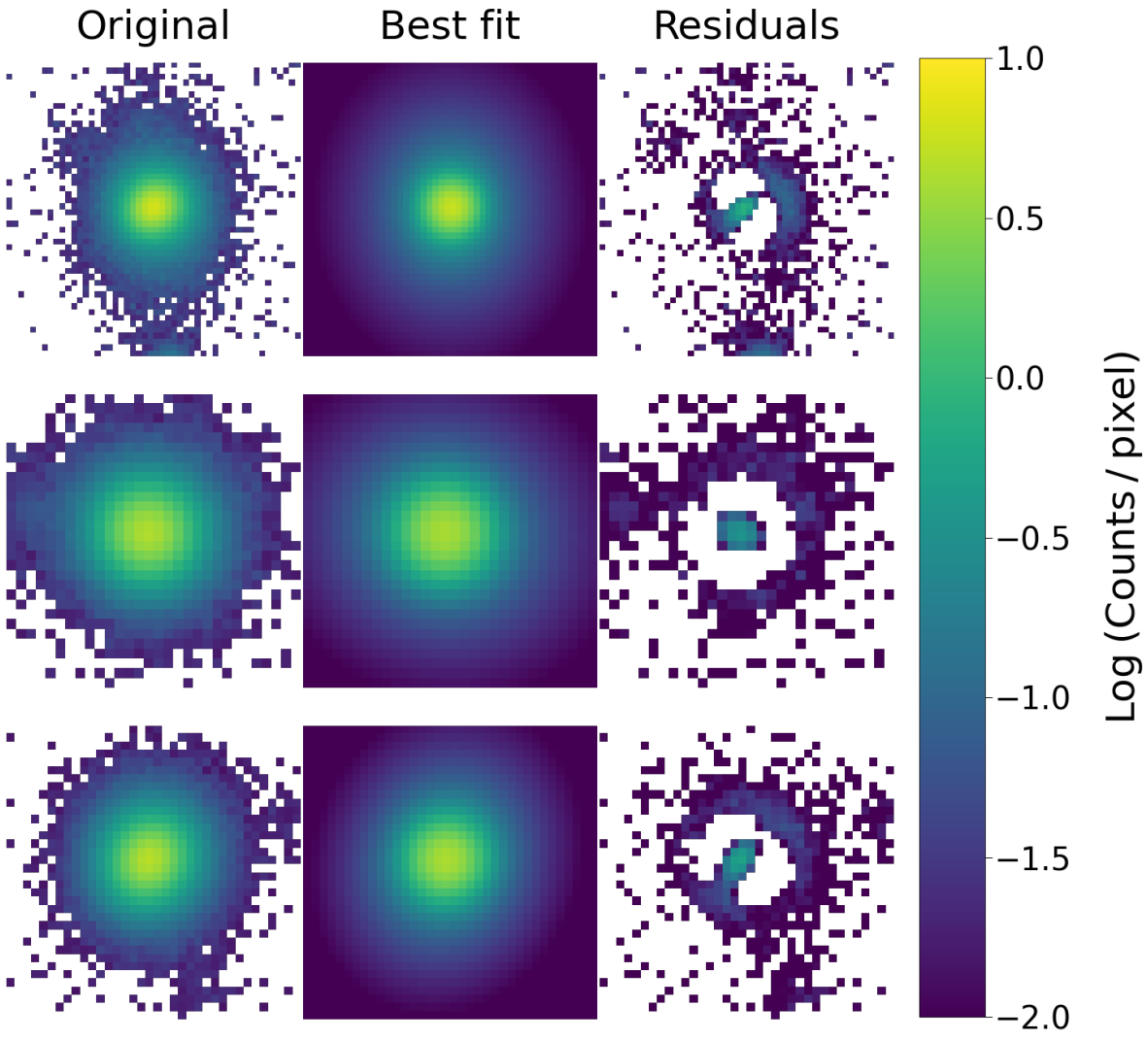

FIGURE C.3: Sersic fits for three blue ellipticals which have only photometric redshifts available. From top to bottom the fitted Sersic indices and maximum residual values are 7, 6.5, 4.4 and 0.8, 0.3 and 0.5 counts pixel^{-1} respectively.

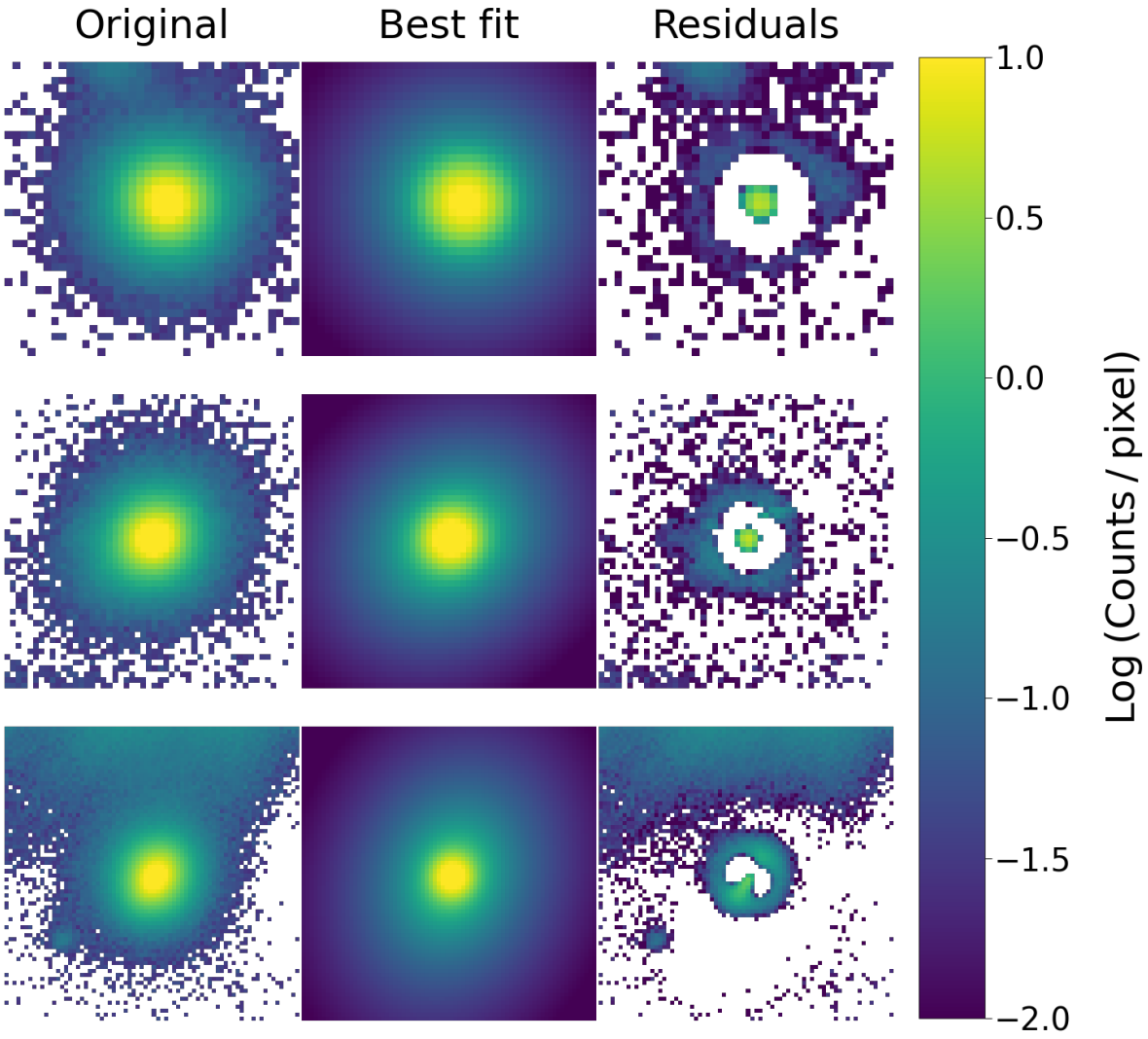

FIGURE C.4: Sersic fits for three objects classified as stars by our method. From top to bottom the fitted Sersic indices and maximum residual values are 7, 5, 6.1 and 4.7, 6.1 and 1.9 counts pixel^{-1} respectively.

Bibliography

Abbott T. M. C., et al., 2018, ApJ, 239, 18

Abraham R. G., Valdes F., Yee H. K. C., van den Bergh S., 1994, ApJ, 432, 75

Abraham R. G., Tanvir N. R., Santiago B. X., Ellis R. S., Glazebrook K., van den Bergh S., 1996, MNRAS, 279, L47

Abraham R. G., van den Bergh S., Nair P., 2003, ApJ, 588, 218

Ackermann M., et al., 2015, , 115, 231301

Agertz O., Kravtsov A. V., 2016, ApJ, 824, 79

Ahn C. P., et al., 2014, ApJ, 211, 17

Aihara H., et al., 2018, Publications of the ASJ, 70, S8

Aihara H., et al., 2019, Publications of the ASJ, 71, 114

Aihara H., et al., 2022, Publications of the ASJ, 74, 247

Alam S., et al., 2015, ApJ, 219, 12

Alavi A., et al., 2016, ApJ, 832, 56

Alberts S., et al., 2016, ApJ, 825, 72

Almeida C., Baugh C. M., Lacey C. G., 2007, MNRAS, 376, 1711

An F. X., et al., 2018, ApJ, 862, 101

Angthopo J., Ferreras I., Silk J., 2019, Monthly Notices of the Royal Astronomical Society: Letters, 488, L99

Arnouts S., et al., 2002, MNRAS, 329, 355

Ashby M. L. N., et al., 2013, ApJ, 769, 80

Ashby M. L. N., et al., 2015, ApJ, 218, 33

Ashby M. L. N., et al., 2018, ApJ, 237, 39

Athanassoula E., Rodionov S. A., Peschken N., Lambert J. C., 2016, ApJ, 821, 90

Aumer M., White S. D., Naab T., Scannapieco C., 2013, Monthly Notices of the Royal Astronomical Society, 434, 3142

Ay F., Ince G., Kamaşak M. E., Ekşi K. Y., 2020, MNRAS, 493, 713

Bailer-Jones C. A. L., Fouesneau M., Andrae R., 2019, MNRAS, 490, 5615

Bakos J., Trujillo I., Pohlen M., 2008, ApJ, 683, L103

Baldry I. K., 2007, in Metcalfe N., Shanks T., eds, Astronomical Society of the Pacific Conference Series Vol. 379, Cosmic Frontiers. p. 276

Baldry I. K., Balogh M. L., Bower R., Glazebrook K., Nichol R. C., 2004, THE NEW COSMOLOGY: Conference on Strings and Cosmology; The Mitchell Symposium on Observational Cosmology. AIP Conference Proceedings, Volume 743, pp. 106-119 (2004)., 743, 106

Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, Monthly Notices of the Royal Astronomical Society, Volume 373, Issue 2, pp. 469-483., 373, 469

Baldry I. K., et al., 2012, MNRAS, 421, 621

Ball N. M., Loveday J., Brunner R. J., 2008, MNRAS, 383, 907

Ballard D. H., Brown C. M., 1982, J: Prentice Hall

Balogh M. L., et al., 2004, ApJL, 615, L101

Bamford S. P., et al., 2009, Monthly Notices of the Royal Astronomical Society, Volume 393, Issue 4, pp. 1324-1352., 393, 1324

Barazza F. D., et al., 2006, ApJ, 643, 162

Barnes J. E., 2004, MNRAS, 350, 798

Barsanti S., et al., 2022, MNRAS, 516, 3569

Beck M. R., et al., 2018, MNRAS, 476, 5516

Beckmann R. S., et al., 2017, MNRAS, 472, 949

Bell E. F., de Jong R. S., 2000, MNRAS, 312, 497

Bell E. F., Zheng X. Z., Papovich C., Borch A., Wolf C., Meisenheimer K., 2007, The Astrophysical Journal, Volume 663, Issue 2, pp. 834-843., 663, 834

Benavides J. A., et al., 2021, Nature Astronomy, 5, 1255

Bender R., 1988, A&A, 202, L5

Bernardi M., et al., 2003, AJ, 125, 1817

Bernardi M., Meert A., Vikram V., Huertas-Company M., Mei S., Shankar F., Sheth R. K., 2014, MNRAS, 443, 874

Berrier J. C., et al., 2009, ApJ, 690, 1292

Bershady M. A., Jangren A., Conselice C. J., 2000, AJ, 119, 2645

Besla G., Martínez-Delgado D., van der Marel R. P., Beletsky Y., Seibert M., Schlafly E. F., Grebel E. K., Neyer F., 2016, ApJ, 825, 20

Beygu B., Kreckel K., van de Weygaert R., van der Hulst J. M., van Gorkom J. H., 2013, AJ, 145, 120

Bezanson R., van Dokkum P. G., Tal T., Marchesini D., Kriek M., Franx M., Coppi P., 2009, ApJ, 697, 1290

Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846

Bílek M., et al., 2020, MNRAS, 498, 2138

Bílek M., Fensch J., Ebrová I., Nagesh S. T., Famaey B., Duc P.-A., Kroupa P., 2022, A&A, 660, A28

Binggeli B., Sandage A., Tarenghi M., 1984, AJ, 89, 64

Blanton M. R., et al., 2003, ApJ, 594, 186

Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517

Bois M., et al., 2011, MNRAS, 416, 1654

Bosch J., et al., 2018, PASJ, 70, S5

Boselli A., Gavazzi G., 2014, Astronomy and Astrophysics Review, 22, 1

Boselli A., Gavazzi G., Boselli A., Gavazzi G., 2006, PASP, 118, 517

Boselli A., Boissier S., Cortese L., Gavazzi G., 2008, ApJ, 674

Bottrell C., Simard L., Mendel J. T., Ellison S. L., 2019, MNRAS, 486, 390

Bouché N., et al., 2016, ApJ, 820, 121

Bournaud F., 2010, in Smith B., Higdon J., Higdon S., Bastian N., eds, Astronomical Society of the Pacific Conference Series Vol. 423, Galaxy Wars: Stellar Populations and Star Formation in Interacting Galaxies. p. 177 (arXiv:0909.1812), doi:10.48550/arXiv.0909.1812

Bournaud F., Duc P. A., 2006, A&A, 456, 481

Bournaud F., Jog C. J., Combes F., 2007, A&A, 476, 1179

Boylan-Kolchin M., Ma C.-P., Quataert E., 2006, MNRAS, 369, 1081

Breiman L., 2001, Machine Learning 2001 45:1, 45, 5

Bremer M. N., et al., 2018, Monthly Notices of the Royal Astronomical Society, 476, 12

Bremnes T., Binggeli B., Prugniel P., 2000, A&A Supp., 141, 211

Brennan R., et al., 2015, MNRAS, 451, 2933

Brinchmann J., et al., 2004, MNRAS, 351, 1151

Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, ApJ, 694, 396

Brough S., Forbes D. A., Kilborn V. A., Couch W., 2006, MNRAS, 370, 1223

Bruce J., Li T. S., Pace A. B., Heiger M., Song Y.-Y., Simon J. D., 2023, ApJ, 950, 167

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Buck T., Obreja A., Ratcliffe B., Lu Y., Minchev I., Macciò A. V., 2023, MNRAS, 523, 1565

Buitrago F., Trujillo I., Conselice C. J., Häußler B., 2013, MNRAS, 428, 1460

Burchett J. N., Elek O., Tejos N., Prochaska J. X., Tripp T. M., Bordoloi R., Forbes A. G., 2020, ApJ, 891, L35

Burstein D., Bender R., Faber S. M., 1993, in Danziger I. J., Zeilinger W. W., Kjär K., eds, European Southern Observatory Conference and Workshop Proceedings Vol. 45, European Southern Observatory Conference and Workshop Proceedings. p. 31

Busko I. C., 1996, in Jacoby G. H., Barnes J., eds, Astronomical Society of the Pacific Conference Series Vol. 101, Astronomical Data Analysis Software and Systems V. p. 139

Buta R., Combes F., 1996, , 17, 95

Buta R. J., et al., 2010, ApJ, 190, 147

Buta R. J., et al., 2015, ApJ, 217, 32

Cameron E., 2011, , 28, 128

Capaccioli M., Caon N., D'Onofrio M., 1992, MNRAS, 259, 323

Cappellari M., et al., 2011, MNRAS, 416, 1680

Cappellari M., et al., 2013, MNRAS, 432, 1862

Carlsten S. G., Greene J. E., Greco J. P., Beaton R. L., Kado-Fong E., 2021, ApJ, 922, 267

Carlsten S. G., Greene J. E., Beaton R. L., Danieli S., Greco J. P., 2022, ApJ, 933, 47

Chabrier G., 2003, Publications of the Astronomical Society of the Pacific, 115, 763

Chen Y.-C., Ho S., Blazek J., He S., Mandelbaum R., Melchior P., Singh S., 2019, MNRAS, 485, 2492

Cheng C., et al., 2020, A&A, 633, A105

Cheng T.-Y., et al., 2021, MNRAS, 507, 4425

Cheung E., et al., 2013, ApJ, 779, 162

Chiosi C., Carraro G., 2002, MNRAS, 335, 335

Chung J., Rey S.-C., Sung E.-C., Kim S., Lee Y., Lee W., 2019, ApJ, 879, 97

Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781

Collins M. L. M., Read J. I., 2022, Nature Astronomy, 6, 647

Combes F., 2017, Frontiers in Astronomy and Space Sciences, 4, 10

Conselice C. J., 2003, ApJ, 147, 1

Conselice C. J., 2006a, MNRAS, 373, 1389

Conselice C. J., 2006b, ApJ, 638, 686

Conselice C. J., 2014, Annual Review of Astronomy and Astrophysics, 52, 291

Conselice C. J., 2018, Research Notes of the American Astronomical Society, 2, 43

Conselice C. J., Bershady M. A., Gallagher J. S. I., 2000a, A&A, 354, L21

Conselice C. J., Bershady M. A., Jangren A., 2000b, ApJ, 529, 886

Conselice C. J., et al., 2007, MNRAS, 381, 962

Conselice C. J., Yang C., Bluck A. F. L., 2009, MNRAS, 394, 1956

Conselice C. J., Ownsworth J., Mortlock A., Bluck A. F. L., 2013, in Thomas D., Pasquali A., Ferreras I., eds, Vol. 295, The Intriguing Life of Massive Galaxies. pp 45–48, doi:10.1017/S1743921313004225

Conselice C. J., Bluck A. F. L., Mortlock A., Palamara D., Benson A. J., 2014, MNRAS, 444, 1125

Conselice C. J., et al., 2024, arXiv e-prints, p. arXiv:2405.00376

Cooper M. C., et al., 2006, MNRAS, 370, 198

Cooper A. P., et al., 2011, ApJ, 743, L21

Coupon J., Czakon N., Bosch J., Komiyama Y., Medezinski E., Miyazaki S., Oguri M., 2018, Publications of the ASJ, 70, S7

Crockett R. M., et al., 2011, ApJ, 727, 115

Crone Odekon M., Hallenbeck G., Haynes M. P., Koopmann R. A., Phi A., Wolfe P.-F., 2018, ApJ, 852, 142

Croton D. J., et al., 2006, MNRAS, 365, 11

Croton D. J., et al., 2016, ApJ, 222, 22

D'Isanto A., Polsterer K. L., 2018, A&A, 609, A111

D'Isanto A., Cavuoti S., Gieseke F., Polsterer K. L., 2018, A&A, 616, A97

D'Onofrio M., Marziani P., Buson L., 2015, Frontiers in Astronomy and Space Sciences, 2, 4

D'Souza R., Kauffman G., Wang J., Vegetti S., 2014, MNRAS, 443, 1433

Daddi E., et al., 2005, ApJ, 631, L13

Dalcanton J. J., Spergel D. N., Gunn J. E., Schmidt M., Schneider D. P., 1997, AJ, 114, 635

Danovich M., Dekel A., Hahn O., Teyssier R., 2012, MNRAS, 422, 1732

Danovich M., Dekel A., Hahn O., Ceverino D., Primack J., 2015, MNRAS, 449, 2087

Darg D. W., et al., 2010, MNRAS, 401, 1043

Das M., Iono D., Saito T., Subramanian S., 2014, in Astronomical Society of India Conference Series. pp 299–301

Das M., Saito T., Iono D., Honey M., Ramya S., 2015, ApJ, 815, 40

Davis T. A., et al., 2015, MNRAS, 449, 3503

Davis F., et al., 2022, MNRAS, 511, 4109

De Propris R., Colless M., Driver S. P., Pracy M. B., Couch W. J., 2005, MNRAS, 357, 590

De Propris R., Bremer M. N., Phillipps S., 2016, MNRAS, 461, 4517

Debattista V. P., Mayer L., Carollo C. M., Moore B., Wadsley J., Quinn T., 2006, ApJ, 645, 209

Deeley S., et al., 2017, MNRAS, 467, 3934

Dekel A., Birnboim Y., 2006, MNRAS, 368, 2

Dekel A., Sari R., Ceverino D., 2009, ApJ, 703, 785

Diamond-Stanic A. M., Coil A. L., Moustakas J., Tremonti C. A., Sell P. H., Mendez A. J., Hickox R. C., Rudnick G. H., 2016, ApJ, 824, 24

Dickinson H., Fortson L., Scarlata C., Beck M., Walmsley M., 2019, arXiv e-prints, p. arXiv:1903.07776

Dimauro P., et al., 2022, MNRAS, 513, 256

Djorgovski S. G., Mahabal A. A., Donalek C., Graham M. J., Drake A. J., Moghaddam B., Turmon M., 2012, arXiv e-prints,

Domínguez Sánchez H., Huertas-Company M., Bernardi M., Tuccillo D., Fischer J. L., 2018, MNRAS, 476, 3661

Domínguez Sánchez H., et al., 2023, MNRAS, 521, 3861

Doroshkevich A. G., 1970, Astrophysics, 6, 320

Dressler A., 1980, ApJ, 236, 351

Dressler A., et al., 1997, ApJ, 490, 577

Dressler A., Smail I., Poggianti B. M., Butcher H., Couch W. J., Ellis R. S., Oemler Augustus J., 1999, ApJ, 122, 51

Drinkwater M. J., Gregg M. D., Colless M., 2001, ApJ, 548, L139

Driver S. P., et al., 2016, MNRAS, 455, 3911

Dubois Y., et al., 2014, MNRAS, 444, 1453

Dubois Y., et al., 2021, A&A, 651, A109

Duc P. A., Bournaud F., Masset F., 2004, A&A, 427, 803

Duc P.-A., et al., 2015, MNRAS, 446, 120

Eigenthaler P., et al., 2018, ApJ, 855, 142

Erwin P., 2015, ApJ, 799, 226

Faber S. M., Lin D. N. C., 1983, ApJ, 266, L17

Faber S. M., Worthey G., Gonzales J. J., 1992, in Barbuy B., Renzini A., eds, IAU Symposium Vol. 149, The Stellar Populations of Galaxies. p. 255

Faber S. M., et al., 2007, The Astrophysical Journal, Volume 665, Issue 1, pp. 265-294., 665, 265

Fabian A. C., 2012, Annual Review of Astronomy and Astrophysics, 50, 455

Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189

Fan X., et al., 2006, AJ, 132, 117

Farouki R. T., Shapiro S. L., 1982, ApJ, 259, 103

Feldmann R., Carollo C. M., Mayer L., Renzini A., Lake G., Quinn T., Stinson G. S., Yepes G., 2010, ApJ, 709, 218

Fensch J., et al., 2017, MNRAS, 465, 1934

Ferguson H. C., Binggeli B., 1994, , 6, 67

Ferrarese L., et al., 2012, ApJ, 200, 4

Ferrarese L., et al., 2020, ApJ, 890, 128

Ferreras I., Lisker T., Carollo C. M., Lilly S. J., Mobasher B., 2005, ApJ, 635, 243

Ferreras I., et al., 2017, MNRAS, 468, 607

Finoguenov A., et al., 2007, ApJ, 172, 182

Florez J., et al., 2021, ApJ, 906, 97

Font A. S., et al., 2008, MNRAS, 389, 1619

Forbes D. A., Franx M., Illingworth G. D., 1995, AJ, 109, 1988

Fouqué P., Solanes J. M., Sanchis T., Balkowski C., 2001, A&A, 375, 770

Fraser-McKelvie A., Cortese L., 2022, ApJ, 937, 117

Freeman P. E., Izbicki R., Lee A. B., Newman J. A., Conselice C. J., Koekemoer A. M., Lotz J. M., Mozena M., 2013, MNRAS, 434, 282

Frei Z., Guhathakurta P., Gunn J. E., Tyson J. A., 1996, AJ, 111, 174

Fritzke B., 1995, in Advances in neural information processing systems. pp 625–632

Fukugita M., Nakamura O., Turner E. L., Helmboldt J., Nichol R. C., 2004, ApJ, 601, L127

Fukugita M., et al., 2007, AJ, 134, 579

Gallagher John S. I., Hunter D. A., 1984, Annual Review of Astronomy and Astrophysics, 22, 37

Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41

Gardner J. P., et al., 2006, SSRv, 123, 485

Gavazzi R., Adami C., Durret F., Cuillandre J. C., Ilbert O., Mazure A., Pelló R., Ulmer M. P., 2009, A&A, 498, L33

Ge X., Wang H.-T., Lei C.-L., Guo Y.-J., Jiang Y.-L., Cao X.-X., 2024, Research in Astronomy and Astrophysics, 24, 035006

Geha M., Blanton M. R., Yan R., Tinker J. L., 2012, ApJ, 757, 85

Geha M., et al., 2017, ApJ, 847, 4

Genzel R., et al., 2010, MNRAS, 407, 2091

George K., 2017, A&A, 598, A45

George M. R., et al., 2011, ApJ, 742, 125

Géron T., Smethurst R. J., Lintott C., Kruk S., Masters K. L., Simmons B., Stark D. V., 2021, MNRAS, 507, 4389

Géron T., et al., 2023, MNRAS, 521, 1775

Ghosh A., et al., 2023, ApJ, 953, 134

Giovanelli R., et al., 2005, AJ, 130, 2598

Gozaliasl G., et al., 2019, MNRAS, 483, 3545

Graham A. W., Guzmán R., 2003, AJ, 125, 2936

Graham A. W., Worley C. C., 2008, MNRAS, 388, 1708

Greco J. P., et al., 2018, ApJ, 857, 104

Guérou A., et al., 2015, ApJ, 804, 70

Guo Q., et al., 2011, MNRAS, 413, 101

Hardcastle M. J., et al., 2019, A&A, 622, A12

Hashemizadeh A., et al., 2021, MNRAS, 505, 136

Hayward C. C., Hopkins P. F., 2017, MNRAS, 465, 1682

Hendel D., Johnston K. V., Patra R. K., Sen B., 2018, arXiv e-prints, p. arXiv:1811.10613

Hernquist L., Mihos J. C., 1995, ApJ, 448, 41

Hester J. A., 2006, ApJ, 647, 910

Ho S. H., Martin C. L., Kacprzak G. G., Churchill C. W., 2017, ApJ, 835, 267

Hocking A., Geach J. E., Sun Y., Davey N., 2018, MNRAS, 473, 1108

Hogg D. W., et al., 2002, AJ, 124, 646

Holmberg E., 1937, Annals of the Observatory of Lund, 6, 1

Holwerda B. W., et al., 2014, ApJ, 781, 12

Holwerda B. W., Phillipps S., Weerasooriya S., Bovill M. S., Brough S., Brown M. J. I., Robertson C., Cook K., 2023, MNRAS, 526, 4575

Hopkins P. F., et al., 2006, ApJS, 163, 1

Hopkins P. F., Cox T. J., Kereš D., Hernquist L., 2008, ApJ, 175, 390

Hopkins P. F., Bundy K., Hernquist L., Wuyts S., Cox T. J., 2010, MNRAS, 401, 1099

Huang S., Leauthaud A., Greene J. E., Bundy K., Lin Y.-T., Tanaka M., Miyazaki S., Komiyama Y., 2018, MNRAS, 475, 3348

Hubble E. P., 1936, Realm of the Nebulae

Hubble E. P., Hubble P. E., 1926, ApJ, 64, 321

Huertas-Company M., et al., 2015, ApJ, 221, 8

Huertas-Company M., et al., 2024, A&A, 685, A48

Hunter D. A., Elmegreen B. G., Madden S. C., 2024, arXiv e-prints, p. arXiv:2402.17004

Ilbert O., et al., 2006, A&A, 457, 841

Ilbert O., et al., 2009, ApJ, 690, 1236

Ilbert O., et al., 2013, A&A, 556, A55

Irwin M. J., Bunclark P. S., Bridgeland M. T., McMahon R. G., 1990, MNRAS, 244, 16P

Ivezić Ž., et al., 2019, ApJ, 873, 111

Jackson R. A., Martin G., Kaviraj S., Laigle C., Devriendt J. E., Dubois Y., Pichon C., 2020, Monthly Notices of the Royal Astronomical Society, 494, 5568

Jackson R. A., et al., 2021a, MNRAS, 502, 4262

Jackson R. A., et al., 2021b, MNRAS, 502, 4262

Jackson R. A., et al., 2021c, MNRAS, 502, 4262

Jackson R. A., Kaviraj S., Martin G., Devriendt J. E. G., Noakes-Kettel E. A., Silk J., Ogle P., Dubois Y., 2022, arXiv e-prints, p. arXiv:2201.08855

Jackson T. M., Pasquali A., La Barbera F., More S., Grebel E. K., 2023, MNRAS, 520, 1155

Jackson R. A., et al., 2024, MNRAS, 528, 1655

Jansen R. A., Franx M., Fabricant D., Caldwell N., 2000, ApJ, 126, 271

Janz J., Penny S. J., Graham A. W., Forbes D. A., Davies R. L., 2017, MNRAS, 468, 2850

Jedrzejewski R. I., 1987, MNRAS, 226, 747

Jedrzejewski R., Schechter P. L., 1988, ApJ, 330, L87

Jian H.-Y., et al., 2022, ApJ, 926, 115

Jiang F.-Z., Huang S., Gu Q.-S., 2011, Research in Astronomy and Astrophysics, 11, 309

Johnson S. C., 1967, Psychometrika, 32, 241

Johnston K. V., Zhao H., Spergel D. N., Hernquist L., 1999, ApJ, 512, L109

Kannappan S. J., Guie J. M., Baker A. J., 2009, AJ, 138, 579

Kauffmann G., et al., 2003, MNRAS, 346, 1055

Kaviraj S., 2010, MNRAS, 406, 382

Kaviraj S., 2014a, MNRAS, 440, 2944

Kaviraj S., 2014b, MNRAS, 440, 2944

Kaviraj S., et al., 2007, ApJ, 173, 619

Kaviraj S., et al., 2008, MNRAS, 388, 67

Kaviraj S., Peirani S., Khochfar S., Silk J., Kay S., 2009, MNRAS, 394, 1713

Kaviraj S., Tan K.-M., Ellis R. S., Silk J., 2011, MNRAS, 411, 2148

Kaviraj S., Darg D., Lintott C., Schawinski K., Silk J., 2012, MNRAS, 419, 70

Kaviraj S., et al., 2013, MNRAS, 428, 925

Kaviraj S., et al., 2017, MNRAS, 467, 4739

Kaviraj S., Martin G., Silk J., 2019, MNRAS, 489, L12

Kawinwanichakij L., et al., 2021, ApJ, 921, 38

Kelvin L. S., et al., 2014, MNRAS, 444, 1647

Kennicutt Jr. R. C., 1998, ApJ, 498, 541

Kent S. M., 1985, ApJ, 59, 115

Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2

Khoperskov S. A., Vasiliev E. O., 2017, MNRAS, 468, 920

Kimm T., Devriendt J., Slyz A., Pichon C., Kassin S. A., Dubois Y., 2011, arXiv e-prints, p. arXiv:1106.0538

Kirby E. N., Cohen J. G., Simon J. D., Guhathakurta P., 2015, ApJ, 814, L7

Kobayashi C., 2004, MNRAS, 347, 740

Kobayashi C., 2005, MNRAS, 361, 1216

Koda J., Yagi M., Yamanoi H., Komiyama Y., 2015, ApJ, 807

Koekemoer A. M., et al., 2011, ApJ, 197, 36

Kolokythas K., Vaddi S., O'Sullivan E., Loubser I., Babul A., Raychaudhury S., Lagos P., Jarrett T. H., 2022, MNRAS, 510, 4191

Kormendy J., 1977, ApJ, 218, 333

Kormendy J., Djorgovski S., 1989, Annual Review of Astronomy and Astrophysics, 27, 235

Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, ApJ, 182, 216

Koudmani S., Sijacki D., Smith M. C., 2022, MNRAS, 516, 2112

Kreckel K., Platen E., Aragón-Calvo M. A., van Gorkom J. H., van de Weygaert R., van der Hulst J. M., Beygu B., 2012, AJ, 144, 16

Krumm N., Brosch N., 1984, AJ, 89, 1461

Kumar A., Das M., Kataria S. K., 2021, MNRAS, 506, 98

La Barbera F., Ferreras I., de Carvalho R. R., Bruzual G., Charlot S., Pasquali A., Merlin E., 2012, MNRAS, 426, 2300

Lackner C. N., Gunn J. E., 2012, MNRAS, 421, 2277

Lahav O., et al., 1995, Science, 267, 859

Laigle C., et al., 2016, ApJ, 224, 24

Laigle C., et al., 2018, MNRAS, 474, 5437

Laishram R., Kodama T., Morishita T., Faisst A., Koyama Y., Yamamoto N., 2024, ApJ, 964, L33

Lange R., et al., 2015, MNRAS, 447, 2603

Larson R. B., 1974, MNRAS, 166, 585

Laureijs R. J., Duvet L., Escudero Sanz I., Gondoin P., Lumb D. H., Oosterbroek T., Saavedra Criado G., 2010, in Oschmann Jacobus M. J., Clampin M. C., MacEwen H. A., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 7731, Space Telescopes and Instrumentation 2010: Optical, Infrared, and Millimeter Wave. p. 77311H, doi:10.1117/12.857123

Laurikainen E., Salo H., Buta R., Knapen J. H., Comerón S., 2010, MNRAS, 405, 1089

Lazar I., Kaviraj S., Martin G., Laigle C., Watkins A., Jackson R. A., 2023, MNRAS, 520, 2109

Lazar I., Kaviraj S., Watkins A. E., Martin G., Bichang'a B., Jackson R. A., 2024, MNRAS, 529, 499

Leauthaud A., et al., 2007, ApJ, 172, 219

Lee B., et al., 2013, ApJ, 774, 47

Leitner S. N., Kravtsov A. V., 2011, ApJ, 734, 48

Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782

Li I. H., Yee H. K. C., Hsieh B. C., Gladders M., 2012, ApJ, 749, 150

Li R., et al., 2022, ApJ, 936, 11

Liao L.-W., Cooper A. P., 2023, MNRAS, 518, 3999

Lim S., et al., 2020, ApJ, 899, 69

Lin L., et al., 2008, ApJ, 681, 232

Lintott C., et al., 2011, MNRAS, 410, 166

Liu C.-Z., Shen S.-Y., Shao Z.-Y., Chang R.-X., Hou J.-L., Yin J., Yang D.-W., 2009, Research in Astronomy and Astrophysics, 9, 1119

Lofthouse E. K., Kaviraj S., Conselice C. J., Mortlock A., Hartley W., 2017, MNRAS, 465, 2895

Lotz J. M., Primack J., Madau P., 2004, AJ, 128, 163

Lotz J. M., et al., 2008, ApJ, 672, 177

Lotz J. M., Jonsson P., Cox T. J., Primack J. R., 2010a, MNRAS, 404, 575

Lotz J. M., Jonsson P., Cox T. J., Primack J. R., 2010b, MNRAS, 404, 590

Lundmark K., 1926, Arkiv for Matematik, Astronomi och Fysik, 19B, 1

Lundmark K., 1927, Nova Acta Regiae Soc. Sci. Upsaliensis Ser. V, pp 1–127

Lupton R., Blanton M. R., Fekete G., Hogg D. W., O'Mullane W., Szalay A., Wherry N., 2004, PASP, 116, 133

MacQueen J., et al., 1967, in Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. pp 281–297

Madau P., Dickinson M., 2014, Annual Review of Astronomy and Astrophysics, 52, 415

Mahajan S., et al., 2018, MNRAS, 475, 788

Maiolino R., et al., 2012, Monthly Notices of the Royal Astronomical Society: Letters, 425, L66

Malin D. F., 1977, AAS Photo Bulletin, 16, 10

Malin D. F., Carter D., 1983, ApJ, 274, 534

Man A. W. S., Toft S., Zirm A. W., Wuyts S., van der Wel A., 2012, ApJ, 744, 85

Mao Y.-Y., Geha M., Wechsler R. H., Weiner B., Tollerud E. J., Nadler E. O., Kallivayalil N., 2021, ApJ, 907, 85

Marcolini A., Brighenti F., D'Ercole A., 2003, MNRAS, 345, 1329

Marian V., Ziegler B., Kuchner U., Verdugo M., 2018, A&A, 617, A34

Marleau F. R., et al., 2021, A&A, 654, A105

Martig M., Bournaud F., Teyssier R., Dekel A., 2009, ApJ, 707, 250

Martig M., et al., 2021, Monthly Notices of the Royal Astronomical Society, 508, 2458

Martin C. L., Kennicutt Robert C. J., 2001, ApJ, 555, 301

Martin D. C., et al., 2007, ApJS, 173, 342

Martin G., Kaviraj S., Devriendt J. E. G., Dubois Y., Pichon C., 2018a, MNRAS, p. 1855

Martin G., et al., 2018b, MNRAS, 476, 2801

Martin G., Kaviraj S., Devriendt J. E. G., Dubois Y., Pichon C., 2018c, MNRAS, 480, 2266

Martin G., et al., 2019, MNRAS, 485, 796

Martin G., Kaviraj S., Hocking A., Read S. C., Geach J. E., 2020, MNRAS, 491, 1408

Martin G., et al., 2021, MNRAS, 500, 4937

Martin G., et al., 2022, MNRAS, 513, 1459

Masters K. L., et al., 2011, MNRAS, 411, 2026

Mateo M. L., 1998, Annual Review of Astronomy and Astrophysics, 36, 435

Mayer L., Governato F., Colpi M., Moore B., Quinn T., Wadsley J., Stadel J., Lake G., 2001, ApJ, 559, 754

McConnachie A. W., 2012, AJ, 144, 4

McCracken H. J., et al., 2012, A&A, 544, A156

Menanteau F., Ford H. C., Motta V., Benítez N., Martel A. R., Blakeslee J. P., Infante L., 2006, AJ, 131, 208

Menou K., 2018, arXiv e-prints, p. arXiv:1811.06374

Merlin E., Chiosi C., Piovan L., Grassi T., Buonomo U., La Barbera F., 2012, MNRAS, 427, 1530

Meyer H. T., Lisker T., Janz J., Papaderos P., 2014, A&A, 562, A49

Mihos J. C., Hernquist L., 1996, ApJ, 464, 641

Miller R. H., Roelofs G. R., Smith B. F., 1990, in Sulentic J. W., Keel W. C., Telesco C. M., eds, NASA Conference Publication Vol. 3098, NASA Conference Publication. pp 549–554

Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319

Moffett A. J., et al., 2019, MNRAS, 489, 2830

Montes M., Brough S., Owers M. S., Santucci G., 2021, ApJ, 910, 45

Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, Nature, 379, 613

Moore B., Governato F., Quinn T., Stadel J., Lake G., 1998, ApJ, 499, L5

Moth P., Elston R. J., 2002, AJ, 124, 1886

Muñoz-Mateos J. C., et al., 2013, ApJ, 771, 59

Muñoz-Mateos J. C., et al., 2015, ApJ, 219, 3

Muñoz R. P., et al., 2015, ApJ, 813

Nair P. B., Abraham R. G., 2010, ApJ, 714, L260

Negroponte J., White S. D. M., 1983, MNRAS, 205, 1009

Nersesian A., Zibetti S., D'Eugenio F., Baes M., 2023, A&A, 673, A63

Noeske K. G., et al., 2007, ApJ, 660, L43

Nogueira-Cavalcante J. P., Gonçalves T. S., Menéndez-Delmestre K., Sheth K., 2017, MNRAS, 473, 1346

Odewahn S. C., Cohen S. H., Windhorst R. A., Philip N. S., 2002, ApJ, 568, 539

Oemler Augustus J., 1974, ApJ, 194, 1

Omar A., Paswan A., 2018, MNRAS, 477, 3552

Onodera S., et al., 2010, ApJ, 722, L127

Ostrovski F., et al., 2017, MNRAS, 465, 4325

Pan Z., Li J., Lin W., Wang J., Fan L., Kong X., 2015, ApJ, 804, L42

Pannella M., Hopp U., Saglia R. P., Bender R., Drory N., Salvato M., Gabasch A., Feulner G., 2006, ApJ, 639, L1

Papovich C., et al., 2012, ApJ, 750, 93

Park M.-J., et al., 2019, ApJ, 883, 25

Parodi B. R., Barazza F. D., Binggeli B., 2002, A&A, 388, 29

Pawlik A. H., Milosavljević M., Bromm V., 2011, ApJ, 731, 54

Pearson W. J., Wang L., Brough S., Holwerda B. W., Hopkins A. M., Loveday J., 2021, A&A, 646, A151

Pearson W. J., et al., 2022, A&A, 661, A52

Peebles P. J. E., 1969, ApJ, 155, 393

Peirani S., Crockett R. M., Geen S., Khochfar S., Kaviraj S., Silk J., 2010, MNRAS, 405, 2327

Peletier R. F., Davies R. L., Illingworth G. D., Davis L. E., Cawson M., 1990, AJ, 100, 1091

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266

Peng Y.-j., et al., 2010, The Astrophysical Journal, Volume 721, Issue 1, pp. 193-221 (2010)., 721, 193

Percival W., et al., 2019, clrp, 2020, 20

Pérez E., et al., 2013, ApJ, 764, L1

Peth M. A., et al., 2016, MNRAS, 458, 963

Pipino A., Matteucci F., 2004, MNRAS, 347, 968

Ploeckinger S., Sharma K., Schaye J., Crain R. A., Schaller M., Barber C., 2018, MNRAS, 474, 580

Pohlen M., Trujillo I., 2006, A&A, 454, 759

Poulain M., et al., 2021, MNRAS, 506, 5494

Powell L. C., Bournaud F., Chapon D., Devriendt J., Slyz A., Teyssier R., 2011, in Carignan C., Combes F., Freeman K. C., eds, Vol. 277, Tracing the Ancestry of Galaxies. pp 234–237 (arXiv:1102.4195), doi:10.1017/S1743921311022848

Pozzetti L., et al., 2010, A&A, 523, A13

Prescott M. K. M., Martin C. L., Dey A., 2015, ApJ, 799, 62

Press W. H., Schechter P., 1974, ApJ, 187, 425

Privon G. C., et al., 2017, ApJ, 846, 74

Qu Y., Di Matteo P., Lehnert M. D., van Driel W., 2011, A&A, 530, A10

Qu Y., et al., 2017, MNRAS, 464, 1659

Quilley L., de Lapparent V., 2022, arXiv e-prints, p. arXiv:2206.04707

Quinn P. J., 1984, ApJ, 279, 596

Rahman N., et al., 2012, ApJ, 745, 183

Reaves G., 1956, AJ, 61, 69

Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541

Renaud F., Bournaud F., Kraljic K., Duc P. A., 2014, MNRAS, 442, L33

Robotham A. S. G., et al., 2014, MNRAS, 444, 3986

Rodighiero G., et al., 2010, A&A, 518, L25

Rodriguez-Gomez V., et al., 2016a, MNRAS, 458, 2371

Rodriguez-Gomez V., et al., 2016b, MNRAS, 458, 2371

Rodriguez-Gomez V., et al., 2019, MNRAS, 483, 4140

Roškar R., Debattista V. P., Stinson G. S., Quinn T. R., Kaufmann T., Wadsley J., 2008, ApJ, 675, L65

Salim S., et al., 2007, The Astrophysical Journal Supplement Series, 173, 267

Salmon B., et al., 2015, ApJ, 799, 183

Sampaio V. M., de Carvalho R. R., Ferreras I., Aragón-Salamanca A., Parker L. C., 2022, MNRAS, 509, 567

Sandage A., 1961, The Hubble Atlas of Galaxies

Sawicki M., et al., 2019, MNRAS, 489, 5202

Sazonova E., et al., 2020, ApJ, 899, 85

Scarlata C., et al., 2007, ApJ, 172, 406

Schaap W. E., van de Weygaert R., 2000, A&A, 363, L29

Schawinski K., Thomas D., Sarzi M., Maraston C., Kaviraj S., Joo S.-J., Yi S. K., Silk J., 2007, MNRAS, 382, 1415

Schawinski K., et al., 2009a, MNRAS, 396, 818

Schawinski K., et al., 2009b, Monthly Notices of the Royal Astronomical Society, 396, 818

Schawinski K., et al., 2014, MNRAS, 440, 889

Schawinski K., Zhang C., Zhang H., Fowler L., Santhanam G. K., 2017, MNRAS, 467, L110

Schiminovich D., et al., 2007, The Astrophysical Journal Supplement Series, 173, 315

Schombert J. M., Pildis R. A., Eder J. A., Oemler Augustus J., 1995, AJ, 110, 2067

Schweizer F., 1982, ApJ, 252, 455

Schweizer F., 2000, in Astronomy, physics and chemistry of H^+_3. p. 2063 (arXiv:astro-ph/0002263), doi:10.1098/rsta.2000.0630

Scoville N., et al., 2007, ApJ, 172, 1

Sengupta A., Keel W. C., Morrison G., Windhorst R. A., Miller N., Smith B., 2022, ApJ, 258, 32

Sérsic J. L., 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Shapley H., 1938, Harvard College Observatory Bulletin, 908, 1

Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978

Sheth K., et al., 2010, PASP, 122, 1397

Silk J., 2011, in Carignan C., Combes F., Freeman K. C., eds, Vol. 277, Tracing the Ancestry of Galaxies. pp 273–281 (arXiv:1102.0283), doi:10.1017/S1743921311022939

Simard L., et al., 2002, ApJ, 142, 1

Simmons B. D., et al., 2017, MNRAS, 464, 4420

Simon J. D., Geha M., 2007, ApJ, 670, 313

Siudek M., et al., 2018, A&A, 617, A70

Somerville R. S., Davé R., 2015, Annual Review of Astronomy and Astrophysics, 53, 51

Sommer-Larsen J., Gelato S., Vedel H., 1999, ApJ, 519, 501

Song H., et al., 2021, MNRAS, 501, 4635

Soo J., et al., 2018, Monthly Notices of the Royal Astronomical Society, 475, 3613

Sousbie T., 2011, MNRAS, 414, 350

Sparre M., Springel V., 2017, MNRAS, 470, 3946

Springel V., Hernquist L., 2005, ApJ, 622, L9

Stanford S. A., Eisenhardt P. R., Dickinson M., 1998, ApJ, 492, 461

Steinhardt C. L., et al., 2014, ApJ, 791, L25

Stern J., Fielding D., Hafen Z., Su K.-Y., Naor N., Faucher-Giguère C.-A., Quataert E., Bullock J., 2024, MNRAS, 530, 1711

Stevens A. R. H., Lagos C. d. P., Contreras S., Croton D. J., Padilla N. D., Schaller M., Schaye J., Theuns T., 2017, MNRAS, 467, 2066

Stewart K. R., Brooks A. M., Bullock J. S., Maller A. H., Diemand J., Wadsley J., Moustakas L. A., 2013, ApJ, 769, 74

Stewart K. R., et al., 2017, ApJ, 843, 47

Stierwalt S., Besla G., Patton D., Johnson K., Kallivayalil N., Putman M., Privon G., Ross G., 2015, ApJ, 805, 2

Strauss M. A., et al., 2002, AJ, 124, 1810

Suh H., Jeong H., Oh K., Yi S. K., Ferreras I., Schawinski K., 2010, ApJ, 187, 374

Taniguchi Y., et al., 2007, ApJ, 172, 9

Taniguchi Y., et al., 2015, Publications of the ASJ, 67, 104

Tanoglidis D., et al., 2021, ApJ, 252, 18

Tarsitano F., et al., 2018, MNRAS, 481, 2018

Tasca L. A. M., et al., 2009, A&A, 503, 379

Taylor-Mager V. A., Conselice C. J., Windhorst R. A., Jansen R. A., 2007, ApJ, 659, 162

Taylor V. A., Jansen R. A., Windhorst R. A., Odewahn S. C., Hibbard J. E., 2005, ApJ, 630, 784

Tenorio-Tagle G., 1996, AJ, 111, 1641

Thacker R. J., Couchman H. M. P., 2001, ApJ, 555, L17

Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, ApJ, 621, 673

Thuruthipilly H., Junais Pollo A., Sureshkumar U., Grespan M., Sawant P., Malek K., Zadrozny A., 2023, arXiv e-prints, p. arXiv:2310.13543

Tolstoy E., Hill V., Tosi M., 2009, Annual Review of Astronomy and Astrophysics, 47, 371

Tonini C., Mutch S. J., Croton D. J., Wyithe J. S. B., 2016, MNRAS, 459, 4109

Toomre A., 1977, in Tinsley B. M., Larson D. Campbell R. B. G., eds, Evolution of Galaxies and Stellar Populations. p. 401

Toomre A., Toomre J., 1972, ApJ, 178, 623

Tortora C., Napolitano N. R., Cardone V. F., Capaccioli M., Jetzer P., Molinaro R., 2010, MNRAS, 407, 144

Trujillo I., et al., 2021, A&A, 654, A40

Uzeirbegovic E., Geach J. E., Kaviraj S., 2020, MNRAS, 498, 4021

Vader J. P., Vigroux L., Lachieze-Rey M., Souviron J., 1988, A&A, 203, 217

Veilleux S., Bolatto A., Tombesi F., Meléndez M., Sturm E., González-Alfonso E., Fischer J., Rupke D. S. N., 2017, ApJ, 843, 18

Venhola A., et al., 2017, A&A, 608

Venhola A., et al., 2019, A&A, 625, A143

Visvanathan N., 1981, A&A, 100, L20

Vulcani B., et al., 2011, MNRAS, 412, 246

Walker I. R., Mihos J. C., Hernquist L., 1996, ApJ, 460, 121

Walmsley M., et al., 2022, MNRAS, 509, 3966

Walmsley M., et al., 2023, MNRAS, 526, 4768

Wang Y., Dong C., Ruan H., Lin Q., Zhang Y., Chen S., 2024, ApJ, 968, 115

Watkins A. E., Laine J., Comerón S., Janz J., Salo H., 2019, A&A, 625, A36

Watkins A. E., et al., 2022, A&A, 660, A69

Weaver J. R., et al., 2022, ApJ, 258, 11

Wellons S., et al., 2016, MNRAS, 456, 1030

Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013, MNRAS, 432, 336

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wilkinson M. I., Kleyna J. T., Evans N. W., Gilmore G. F., Read J. I., Koch A., Grebel E. K., Irwin M. J., 2006, in Mamon G. A., Combes F., Deffayet C., Fort B., eds, EAS Publications Series Vol. 20, EAS Publications Series. pp 105–112 (arXiv:astro-ph/0602186), doi:10.1051/eas:2006055

Wilkinson A., Almaini O., Wild V., Maltby D., Hartley W. G., Simpson C., Rowlands K., 2021, Monthly Notices of the Royal Astronomical Society, Volume 504, Issue 3, pp.4533-4550, 504, 4533

Willett K. W., et al., 2013, MNRAS, 435, 2835

Willett K. W., et al., 2017, MNRAS, 464, 4176

Wolf M., 1908, Publikationen des Astrophysikalischen Instituts Koenigstuhl-Heidelberg, 3, 109

Wollaeger R. T., et al., 2018, MNRAS, 478, 3298

Wright E. L., et al., 2010, AJ, 140, 1868

Wyder T. K., et al., 2007, The Astrophysical Journal Supplement Series, 173, 293

Yamauchi C., et al., 2005, AJ, 130, 1545

Yi S. K., et al., 2005, ApJL, 619, L111

Yoon Y., Park C., Chung H., Lane R. R., 2022, ApJ, 925, 168

Yoshii Y., Arimoto N., 1987, A&A, 188, 13

Zamojski M. A., et al., 2007, ApJ, 172, 468

Zavala J., Okamoto T., Frenk C. S., 2008, MNRAS, 387, 364

Zeng G., Wang L., Gao L., 2021, Monthly Notices of the Royal Astronomical Society, 507, 3301

Zhang H.-X., Hunter D. A., Elmegreen B. G., Gao Y., Schruba A., 2012, AJ, 143, 47

Zou H., et al., 2016, AJ, 151, 37

Zou H., et al., 2022, Research in Astronomy and Astrophysics, 22, 065001

de Vaucouleurs G., 1948, Journal des Observateurs, 31, 113

de Vaucouleurs G., 1959, Handbuch der Physik, 53, 275

de Vaucouleurs G., 1977, in Tinsley B. M., Larson Richard B. Gehret D. C., eds, Evolution of Galaxies and Stellar Populations. p. 43

de Vaucouleurs G., de Vaucouleurs A., Corwin Herold G. J., Buta R. J., Paturel G., Fouque P., 1991, Third Reference Catalogue of Bright Galaxies

den Brok M., et al., 2011, MNRAS, 414, 3052

van Dokkum P. G., 2005, AJ, 130, 2647

van Dokkum P. G., Franx M., Fabricant D., Illingworth G. D., Kelson D. D., 2000, ApJ, 541, 95

van Dokkum P. G., Abraham R., Merritt A., Zhang J., Geha M., Conroy C., 2015, ApJ, 798

van de Voort F., Schaye J., Booth C. M., Haas M. R., Dalla Vecchia C., 2011, MNRAS, 414, 2458

van den Bergh S., 1959, Publications of the David Dunlap Observatory, 2, 147