

Preparing for low surface brightness science with the Vera C. Rubin Observatory: characterisation of tidal features from mock images

G. Martin, A. E. Bazkiaei, M. Spavone E. Iodice, J. C. Mihos, M. Montes, J. A. Benavides, S. Brough, J. L. Carlin, C. A. Collins, P. A. Duc, et al.

▶ **To cite this version:**

G. Martin, A. E. Bazkiaei, M. Spavone E. Iodice, J. C. Mihos, M. Montes, et al.. Preparing for low surface brightness science with the Vera C. Rubin Observatory: characterisation of tidal features from mock images. Monthly Notices of the Royal Astronomical Society, 2022, 10.1093/mnras/stac1003 . insu-03658558

HAL Id: insu-03658558
https://insu.hal.science/insu-03658558

Submitted on 7 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **513,** 1459–1487 (2022) https://doi.org/10.1093/mnras/stac1003
Advance Access publication 2022 April 11

Preparing for low surface brightness science with the Vera C. Rubin Observatory: Characterization of tidal features from mock images

G. Martin [1,2★] A. E. Bazkiaei,[3,4,5★] M. Spavone [6] E. Iodice,[6] J. C. Mihos,[7] M. Montes [8†] J. A. Benavides [9,10] S. Brough [11] J. L. Carlin,[12] C. A. Collins,[13] P. A. Duc,[14] F. A. Gómez [15,16] G. Galaz,[17] H. M. Hernández-Toledo,[18] R. A. Jackson [19] S. Kaviraj [20] J. H. Knapen,[21,22] C. Martínez-Lombilla [11] S. McGee,[23] D. O'Ryan,[24] D. J. Prole,[3,4] R. M. Rich [25] J. Román [21,22] E. A. Shah,[26] T. K. Starkenburg,[27] A. E. Watkins,[20] D. Zaritsky [2] C. Pichon,[28,29,30] L. Armus,[31] M. Bianconi [23] F. Buitrago [32,33] I. Busá [34] F. Davis [20] R. Demarco [35] A. Desmons,[11] P. García,[36] A. W. Graham,[37] B. Holwerda [38] D. S. -H. Hon,[37] A. Khalid,[11] J. Klehammer,[14] D. Y. Klutse,[39] I. Lazar,[20] P. Nair,[40] E. A. Noakes-Kettel,[20] M. Rutkowski,[41] K. Saha,[42] N. Sahu,[37,43] E. Sola,[14] J. A. Vázquez-Mata [18,44] A. Vera-Casanova [16] and I. Yoon [45]

Affiliations are listed at the end of the paper

Accepted 2022 April 7. Received 2022 April 1; in original form 2021 December 8

ABSTRACT

Tidal features in the outskirts of galaxies yield unique information about their past interactions and are a key prediction of the hierarchical structure formation paradigm. The Vera C. Rubin Observatory is poised to deliver deep observations for potentially millions of objects with visible tidal features, but the inference of galaxy interaction histories from such features is not straightforward. Utilizing automated techniques and human visual classification in conjunction with realistic mock images produced using the NEWHORIZON cosmological simulation, we investigate the nature, frequency, and visibility of tidal features and debris across a range of environments and stellar masses. In our simulated sample, around 80 per cent of the flux in the tidal features around Milky Way or greater mass galaxies is detected at the 10-yr depth of the Legacy Survey of Space and Time (30–31 mag arcsec^{-2}), falling to 60 per cent assuming a shallower final depth of 29.5 mag arcsec^{-2}. The fraction of total flux found in tidal features increases towards higher masses, rising to 10 per cent for the most massive objects in our sample ($M_\star \sim 10^{11.5}$ M$_\odot$). When observed at sufficient depth, such objects frequently exhibit many distinct tidal features with complex shapes. The interpretation and characterization of such features varies significantly with image depth and object orientation, introducing significant biases in their classification. Assuming the data reduction pipeline is properly optimized, we expect the Rubin Observatory to be capable of recovering much of the flux found in the outskirts of Milky Way mass galaxies, even at intermediate redshifts ($z < 0.2$).

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

1 INTRODUCTION

Hierarchical structure formation scenarios (e.g. Fall & Efstathiou 1980; van den Bosch et al. 2002; Agertz, Teyssier & Moore 2011) predict that massive galaxies acquire much of their stellar mass through a combination of continuous cold gas accretion and mergers with smaller objects (e.g. Press & Schechter 1974; Moster, Naab & White 2013; Kaviraj et al. 2015; Rodriguez-Gomez et al. 2016; Martin et al. 2018b; Davison et al. 2020; Martin et al. 2021). As a consequence, mergers are also expected to play a significant role in driving the evolution of galaxy properties, for example, by triggering (Schweizer 1982; Mihos & Hernquist 1996; Duc et al. 1997; Elbaz & Cesarsky 2003; Kaviraj et al. 2011; Lofthouse et al. 2017; Martin et al. 2017) or quenching (Schawinski et al. 2014; Barro et al. 2017; Kawinwanichakij et al. 2017; Pontzen et al. 2017) star formation in the host galaxy or by driving its morphological evolution (e.g. Toomre 1977; Conselice, Yang & Bluck 2009; Dekel, Sari & Ceverino 2009; Taranu, Dubinski & Yee 2013; Naab et al. 2014; Fiacconi, Feldmann & Mayer 2015; Graham, Dullo & Savorgnan 2015; Deeley et al. 2017; Gómez et al. 2017; Welker et al. 2017; Martin et al. 2018a; Jackson et al. 2019). Signatures of past mergers take the form of faint extended tidal features such as tails (e.g. Pfleiderer 1963; Toomre & Toomre 1972; Peirani et al. 2010; Kaviraj 2014; Kaviraj, Martin & Silk 2019), or plumes (e.g. Lauer 1988) – which are typically produced by major mergers – and streams (e.g. Johnston, Sigurdsson & Hernquist 1999; Shipp et al. 2018; Martinez-

★ E-mail: garrethmartin@kasi.re.kr, garrethmartin@arizona.edu (GM); amir.ebadati-bazkiaei@mq.edu.au (AEB)
† STScI Prize Fellow
‡ LSSTC DSFP Fellow

© 2022 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

Delgado et al. 2021) or shells (e.g. Malin & Carter 1983; Quinn 1984) – which mainly arise from minor interactions – as well as in the structure of the surrounding diffuse light (e.g. Choi, Guhathakurta & Johnston 2002; Graham 2002; Johnston, Choi & Guhathakurta 2002; Seigar, Graham & Jerjen 2007; Kaviraj et al. 2012; Monachesi et al. 2016, 2019; Iodice et al. 2019; Montes 2019). These features, which arise from many different types of encounter, hold a fossil record of the host galaxy's past interactions and mergers which can be used to reconstruct its assembly history and dynamical history (Johnston et al. 2008; Martínez-Delgado et al. 2009; Belokurov et al. 2017; Montes et al. 2020; Ren et al. 2020; Spavone et al. 2020; Vera-Casanova et al. 2021). However, the majority of tidal features are expected to have surface brightnesses fainter than 30 mag arcsec^{-2} in the *r*-band (Johnston et al. 2008). Although pushing towards these kinds of limiting surface brightnesses remains extremely challenging, it is nevertheless desirable to do so, being necessary to uncover a more detailed history of local Universe. This is not only vital for our understanding of hierarchical galaxy assembly (e.g. Johnston, Sackett & Bullock 2001; Wang et al. 2012), but also serves as a novel galactic scale probe of more fundamental physics such as theories of gravity (e.g. Gentile et al. 2007; Renaud, Famaey & Kroupa 2016) and dark matter (Dubinski, Mihos & Hernquist 1996; Kesden & Kamionkowski 2006; Dumas et al. 2015; van Dokkum et al. 2018; Montes et al. 2020). In particular, tidal structure is a powerful tracer of the underlying galactic halo potential (e.g. Dubinski, Mihos & Hernquist 1999; Varghese, Ibata & Lewis 2011; Bovy et al. 2016; Ibata et al. 2020; Malhan, Valluri & Freese 2021).

Over the last few decades, advances in the sensitivity and field of view of modern instruments (e.g. Kuijken et al. 2002; Miyazaki et al. 2002; Mihos et al. 2005; Diehl & Dark Energy Survey Collaboration 2012; Miyazaki et al. 2012; Abraham & van Dokkum 2014; Torrealba et al. 2018) and increasing sophistication of observational and data-analysis techniques (e.g. Mihos et al. 2005; Akhlaghi & Ichikawa 2015; Pawlik et al. 2016; Morales et al. 2018; Prole et al. 2018; Rich et al. 2019; Tanoglidis, Ćiprijanović & Drlica-Wagner 2021b; Zaritsky et al. 2021) have permitted relatively large studies that concentrate on the low surface brightness (LSB) regime which tidal features inhabit. This has made possible the detailed characterization of the LSB components of galaxies (e.g. Kado-Fong et al. 2018; Bílek et al. 2020) and allowed studies of their prevalence (e.g. Hood et al. 2018).

The 10-yr Legacy Survey of Space and Time (LSST), which will take place at the Vera C. Rubin Observatory (Olivier, Seppala & Gilmore 2008; Ivezić et al. 2019), will lead to a step change in the depth and detail that can be achieved by wide area surveys. Data from the 10-yr survey will vastly increase the number of known objects with tidal features. While deep observations tracing low surface brightness structures and galaxies have been possible previously (e.g. Martínez-Delgado et al. 2009; Kim et al. 2012; Beaton et al. 2014; Duc et al. 2015; Mihos et al. 2015; Kado-Fong et al. 2018; Iodice et al. 2019; Zaritsky et al. 2019; Trujillo et al. 2021), LSST will offer a distinct advantage as these studies have generally been limited to small fields or targeted observations of individual galaxies (typically emphasizing cluster environments) or else do not have the requisite depth to detect a significant fraction of prominent tidal features. LSST uniquely combines very deep imaging (*r*-band depth better than 30.5 mag arcsec^{-2} with 10 arcsec × 10 arcsec binning; Laine et al. 2018; Brough et al. 2020) with a wide area covering the whole Southern sky (18 000 deg^2). This will enable detailed statistical studies of tidal features within a representative volume of the Universe for the first time.

It is expected that the raw data produced by the Rubin Observatory will be of sufficient quality to study low surface brightness features (Robertson et al. 2017; Kaviraj 2020; Trujillo et al. 2021). However, a number of other obstacles still remain if the available data are to be exploited to their full potential. The characterization of tidal features requires not only sufficiently deep imaging but also bespoke data reduction suitable for LSB science and a thorough understanding of biases and uncertainties present in the data.

Follow-up observations for the full population of galaxies with LSB features that will be revealed by LSST will be intractable, especially as tidal features and disturbed morphologies are expected to be ubiquitous in massive galaxies (e.g. Tal et al. 2009; Cibinel et al. 2019) and likely remain at least somewhat common in lower mass galaxies (e.g. Martínez-Delgado et al. 2012; Martin et al. 2021). Analysis of the majority of galaxies will therefore be based primarily on available 2D photometric information. This means that additional information such as spectroscopy and multi-wavelength data, which can reveal important information about the distances, 3D distribution, kinematics, environments, baryonic content, and stellar populations of galaxies (e.g. Bournaud et al. 2004; Kadowaki, Zaritsky & Donnerstein 2017; Junais et al. 2020; Karunakaran et al. 2020), will be unavailable for a majority of objects. Analysis of the majority of galaxies will therefore be limited to Rubin Observatory data.

With regards to the characterization of the tidal features themselves, automated methods (e.g. Grillmair et al. 1995; Conselice, Bershady & Jangren 2000; Rockosi et al. 2002; Lotz, Primack & Madau 2004; Hendel et al. 2019; Pearson et al. 2021) can help us to define the structure of galaxies and identify merging systems and their tidal features, but a full characterization of these galaxies and their tidal features befitting the quality of the available photometric data will require detailed visual inspection by human classifiers (e.g. Darg et al. 2010; Bílek et al. 2020). Visual inspection relies on a significant level of domain knowledge and physical intuition for interpretation. This inevitably introduces some level of subjectivity, especially in the absence of precise redshifts, kinematics, or other 3D information. While machine learning and machine vision techniques can help alleviate reliance on human classifiers (e.g. Beck et al. 2018; Hendel et al. 2019; Walmsley et al. 2019), continuous human intervention will likely still be required. Training sets, will still need to be constructed and labelled by human classifiers, and as the coadded LSST images become deeper they will need to be routinely updated (Martin et al. 2020). Some level of bias is therefore unavoidable and its nature may evolve with a number of factors including limiting surface brightness, galaxy mass, and orientation (e.g. Mantha et al. 2019; Müller, Vudragović & Bílek 2019; Blumenthal et al. 2020; Lambrides et al. 2021).

Some sources of bias, such as the effect of projection, are intrinsic to observations, while others like image depth, can be improved with longer exposure times. For example, depending on the angle at which a given LSB structure is observed, the efficiency with which they are detected can be impacted (Mancillas et al. 2019; Vera-Casanova et al. 2021) or their nature can change so that the same structures appear either stream-like or shell-like from different angles (Hendel & Johnston 2015; Greco et al. 2018a). Additionally different classes of tidal features may become more or less detectable over time, can persist over differing time-scales (Johnston et al. 1999; Bullock & Johnston 2005; Mancillas et al. 2019; Vera-Casanova et al. 2021) or else may transform into different classes of tidal structures (Foster et al. 2014; Hendel & Johnston 2015). Other unrelated structures like galactic cirrus (Miville-Deschênes et al. 2016; Román, Trujillo & Montes 2020) or instrumental artefacts (Chang et al. 2021; Tanoglidis et al. 2021b) can be misclassified or otherwise inhibit the detection of tidal features. At higher redshifts, it can also become increasingly

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

difficult to interpret images as the angular scale of objects decreases and they become more poorly resolved.

The purpose of this paper is to evaluate the expected performance of the LSST at recovering all forms of tidal features and diffuse light occurring in the outskirts of galaxies based on realistic mock images produced using the NEWHORIZON simulation (Dubois et al. 2021). We make predictions as a function of a galaxy's physical properties, redshift, and imaging depth. We then explore how stellar mass, ex-situ mass fraction, redshift, limiting surface brightness and orientation may affect or bias the visual characterization of galaxies by expert human classifiers across different types of tidal features.

(i) In Section 2 we present an overview of the NEWHORIZON simulation along with the relevant physics and the method for producing mock images and merger trees as well as outline our visual classification scheme.

(ii) In Section 3 we explore the properties of the extended light around galaxies using automated techniques to separate the LSB components. We study the spatial and surface brightness distributions as well as the fraction of tidal flux that we expect to detect at various limiting surface brightnesses and redshifts. We additionally consider how tidal flux fraction evolves with galaxy mass and accretion history.

(iii) In Section 4 we present the results of visual classifications of our mock images by human classifiers. We discuss the frequency of different classes of tidal feature as a function of galaxy mass and limiting surface brightness and look at how limiting surface brightness, redshift, and projection can introduce biases.

(iv) In Section 5 we summarize our results.

Throughout this paper we adopt a ΛCDM cosmology consistent with Komatsu et al. (2011) ($\Omega_m = 0.272$, $\Omega_\Lambda = 0.728$, $\Omega_b = 0.045$, $H_0 = 70.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$) and we primarily assume a Salpeter (1955) initial mass function (IMF).

2 METHOD

We employ the state-of-the-art cosmological hydrodynamical simulation, NEWHORIZON, in order to produce realistic mock observations of galaxies and their outskirts, companions and satellites within a self-consistent cosmological context. These objects have known properties and interaction histories which can be used to test the efficacy of observational assumptions and techniques.

2.1 The NEWHORIZON simulation

The NEWHORIZON simulation[1] (Dubois et al. 2021) is a zoom-in of the $(142\ \mathrm{Mpc})^3$ parent Horizon-AGN simulation (Dubois et al. 2014; Kaviraj et al. 2017). Initial conditions are generated using cosmological parameters that are compatible with *WMAP7* ΛCDM cosmology (Komatsu et al. 2011) ($\Omega_m = 0.272$, $F\Omega_\Lambda = 0.728$, $\sigma_8 = 0.81$, $\Omega_b = 0.045$, $H_0 = 70.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and $n_s = 0.967$). Within the original Horizon-AGN volume, a spherical volume with a diameter of 20 Mpc and an effective resolution of 4096^3 is defined, corresponding to a dark matter (DM) mass resolution and initial gas mass resolution of $m_{DM} = 1.2 \times 10^6\ \mathrm{M_\odot}$ and $m_{gas} = 2 \times 10^5\ \mathrm{M_\odot}$. NEWHORIZON uses the adaptive mesh refinement (AMR) code RAMSES (Teyssier 2002) and gas is evolved with a second-order Godunov scheme and the approximate Harten-Lax-Van Leer-Contact (Toro 1999) Riemann solver with linear interpolation of the cell-centred quantities at cell interfaces.

[1] http://new.horizon-simulation.org

NEWHORIZON combines high stellar mass ($1.3 \times 10^4\ \mathrm{M_\odot}$) and spatial resolution (∼34 pc), with a contiguous volume of $(16\ \mathrm{Mpc})^3$. The volume probes field and group environments, but does not extend to dense clusters (the maximum halo mass is $M_h \sim 10^{13}\ \mathrm{M_\odot}$). Given the diffuse nature of galaxy stellar haloes, the trade off between resolution and volume is an important consideration. We find NEWHORIZON to be a better compromise than similar simulations like Illustris TNG50 (Nelson et al. 2019) or ROMULUS25 (Tremmel et al. 2017), both of which trade larger volumes for lower mass resolution. The closest similar simulation in terms of mass resolution is TNG50, with a volume of $(50\ \mathrm{Mpc})^3$ and a stellar mass resolution ($8.5 \times 10^4\mathrm{M_\odot}$). For comparison, most observed tidal features individually account for ∼0.1 per cent to a few per cent of the total stellar mass of a system (e.g. van Dokkum et al. 2019; Fensch et al. 2020), meaning these tidal features would resolved with only ∼100–1000 particles and ∼800–8000 particles for a galaxy of $M_\star = 10^{10}\ \mathrm{M_\odot}$ by TNG50 and NEWHORIZON, respectively.

The additional resolution of NEWHORIZON is therefore important in order to sample as much as possible the LSB outskirts of galaxies at surface brightness limits that contemporary or forthcoming instrument will be capable of targeting (see Section 2.1.1). NEWHORIZON also has sufficient volume to yield a reasonable sample of massive galaxies ($M_\star > 10^{10}\ \mathrm{M_\odot}$) and provides a realistic distribution of galaxies, as well as fully simulating the cosmological context required to produce galaxies with ab initio realistic interaction and formation histories (as opposed to zoom-in simulations of individual haloes, where the zoom region must be carefully selected to avoid bias).

NEWHORIZON reproduces key galaxy properties with good agreement to observed quantities. The galaxy stellar mass function, galaxy size–mass relation, halo mass–stellar mass relation as well as the evolution of galaxy morphology and cosmic star formation rate densities show fair agreement with observationally derived relations. There is however significant uncertainty from cosmic variance owing to size of the simulated region. Relevant to this study, NEWHORIZON appears to deviate from observations at the high mass end or the galaxy size–mass relation and at the low mass end of the halo mass–stellar mass relation. Galaxies appear somewhat more compact than expected at $M_\star \gtrsim 10^{11}\ \mathrm{M_\odot}$ and have stellar masses that are relatively too massive for halo masses of $M_h \lesssim 10^{11}\ \mathrm{M_\odot}$. We refer readers to Sections 3.2., 3.6., 3.7, and 3.9 of Dubois et al. (2021) for a more detailed description of the galaxy stellar mass function, halo mass–stellar mass relation, size–mass relation and kinematics, respectively.

2.1.1 *Numerical resolution limit for detecting tidal features*

Because the stellar particle mass resolution of a simulation places limits on its ability to resolve structures, we first attempt to estimate the numerical limits that the resolution of NEWHORIZON places on our ability to resolve tidal features. We restrict our analysis to shells, which we are typically fainter than tails (see Fig. 14) and therefore more susceptible to resolution effects. In Appendix A we describe an analytical method using analytical shell profiles (Sanderson & Helmi 2013) and Sérsic (1968) profiles. We find that this is dependent on a number of factors including the galactocentric radius of the tidal feature and the shape and brightness of the galaxy profile.

Even with its relatively high stellar mass resolution, for the most massive galaxies in our sample, we do not expect NEWHORIZON to resolve shells with surface brightnesses comparable to LSST 10-yr depth ($\mu_r^{\mathrm{lim}}(3\sigma, 10\,\mathrm{arcsec} \times 10\,\mathrm{arcsec}) \approx 30.5\ \mathrm{mag\ arcsec^{-2}}$) close to the central regions of galaxies ($r < 4.5\ R_{\mathrm{eff}}$). For less massive galaxy models, it is possible to detect faint shells at

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

significantly smaller radii but, for the full sample, we would likely require significantly better mass resolution to enable us to detect all tidal features. As the underlying radial and surface brightness distribution of shells and other tidal features is not known, it is difficult to estimate how significantly this affects our results, but shells are typically resolved down to sufficiently small radii so as to have negligible impact on the number of detected tidal features for depths realistically achievable by LSST. At significantly higher limiting surface brightnesses more care is needed in interpreting results. Shells which would be observationally detectable may not be sufficiently resolved at radii significantly larger than 10 $R_{\rm eff}$, meaning the frequency of tidal features at very faint limiting surface brightens is likely underestimated, particularly around more massive galaxies. We refer to Appendix A for a more detailed discussion.

2.2 Galaxy sample

We use the structure finder ADAPTAHOP (Aubert, Pichon & Colombi 2004) to separately detect both galaxies, haloes along with their respective substructures based on the distribution of dark matter and star particles in the simulation box, respectively. The centre of each galaxy or halo is recursively determined by seeking the centre of mass in a shrinking sphere, while decreasing its radius by 10 per cent recurrently down to a minimum radius of 0.5 kpc (Power et al. 2003). We impose a minimum structure size of 100 dark matter particles and 50 star particles as well as requiring and an average overdensity of 80 times the critical density for dark matter haloes and 160 times the critical density for galaxies (see Aubert et al. 2004, for details). Halo virial masses and radii are obtained by computing the kinetic and gravitational energy within ellipsoids, stopping once virial equilibrium is sufficiently well verified (Dubois et al. 2021).

We select 30 host galaxies with stellar masses greater than 10^{10} $M_\odot$ with a supplementary sample of seven host galaxies with stellar masses of $10^{9.5}$ $M_\odot$ < $M_\star$ < 10^{10} $M_\odot$, which were selected to better probe trends for lower mass galaxies. We do not include any galaxies whose haloes are contaminated by low-resolution dark matter particles from outside of the high-resolution zoom region. In total, this sample consists of 37 objects at $z = 0.2$[2] and their progenitors at $z = 0.4$, 0.6, and 0.8 giving a total of 148 objects across four different redshifts. Fig. 1 shows the stellar mass ($M_\star$) and halo mass (M_h) distribution of host galaxies in our sample presented as a scatter plot and stacked histograms for each redshift. All galaxies in our sample are resolved with a minimum of ∼250 000 star particles and an average of $\sim 10^6$ star particles. We select galaxies only based on the criteria above, making no attempt to preferentially select galaxies with prominent tidal features. Apart from environmental bias due to the size of the simulated volume (see Section 2.1), the sample presented in this paper is therefore unbiased with respect to accretion history and representative of the intermediate and high mass populations found in the simulation at low-to-intermediate redshift as a whole.

The thick black line and filled region indicate the median halo mass–stellar mass relation and its 1σ scatter at $z = 0.2$. While there is good qualitative agreement for more massive central haloes compared to best-fitting semi-empirical relations from Behroozi, Wechsler & Conroy (2013) and Moster et al. (2013) and compared to the empirical model of Behroozi et al. (2019), below the knee of the relation there is significant overestimation in baryon conversion

[2]The lowest redshift to which the simulation had been run at the time of analysis.

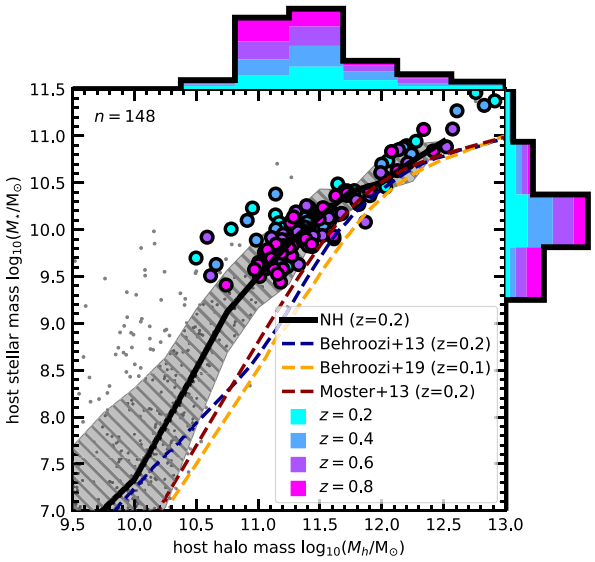


Figure 1. Scatter plot and stacked histograms showing the distribution of host galaxy halo masses and stellar masses selected at $z = 0.2$ along with their progenitors at $z = 0.4$, 0.6, and 0.8. There are a total of 148 objects (37 unique objects at four different snapshots). Objects in our sample are indicated by coloured points while all other objects are indicated by smaller grey points. The thick black line and shaded region indicate the median halo mass–stellar mass relation and its 1σ scatter. Also indicated by coloured dashed lines show relations from the literature (Behroozi et al. 2013, 2019; Moster et al. 2013).

efficiency. This is at least partially a consequence of the limited volume of NEWHORIZON, which lacks clusters, rich group haloes, and very rarefied environments. Note that in denser environments the median halo mass–stellar mass relation tends towards lower star formation efficiencies – possibly a result of earlier formation times driving more efficient feedback and more self-regulation or by environmental quenching (e.g. Garrison-Kimmel et al. 2019; Samuel et al. 2022). Better agreement with observations (an improvement of 0.3 to 0.6 dex at halo masses smaller than 10^{11} $M_\odot$) is achieved when we weight our sample to account for the fact that underdense environments are overrepresented in NEWHORIZON compared with the parent Horizon-AGN simulation. As a result of the discrepancy, the total accreted stellar mass in central haloes is likely overestimated compared with a more representative sample. For a given halo mass, this may result in elevated tidal feature strength or greater quantities of diffuse light around NEWHORIZON galaxies compared with their observed counterparts.

Another important consideration, which we do not investigate here, is how resolution effects and implementation of subgrid physics impact the orbital substructures that are produced in our synthetic galaxies and their haloes. One example is the over- or under-production of bars, explored in Reddish et al. (2021), which could potentially inhibit the detection of tidal features or otherwise result in misclassification. Perhaps more important to this study are the orbits and phase-space correlations of the satellite galaxies that are responsible for producing tidal features (e.g. Pawlowski 2021). We defer a full discussion of agreement with observed quantities and phase-space analysis to an upcoming paper (Uzeirbegovic, in preparation).

In Fig. 2, we present g, r, i false colour images of each object in our sample for the snapshot corresponding to a redshift of $z = 0.4$ in the context of the larger cosmic structure and with the same scale. The distribution of LSB structure is shown out to 1 $R_{\rm vir}$ for each galaxy,

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

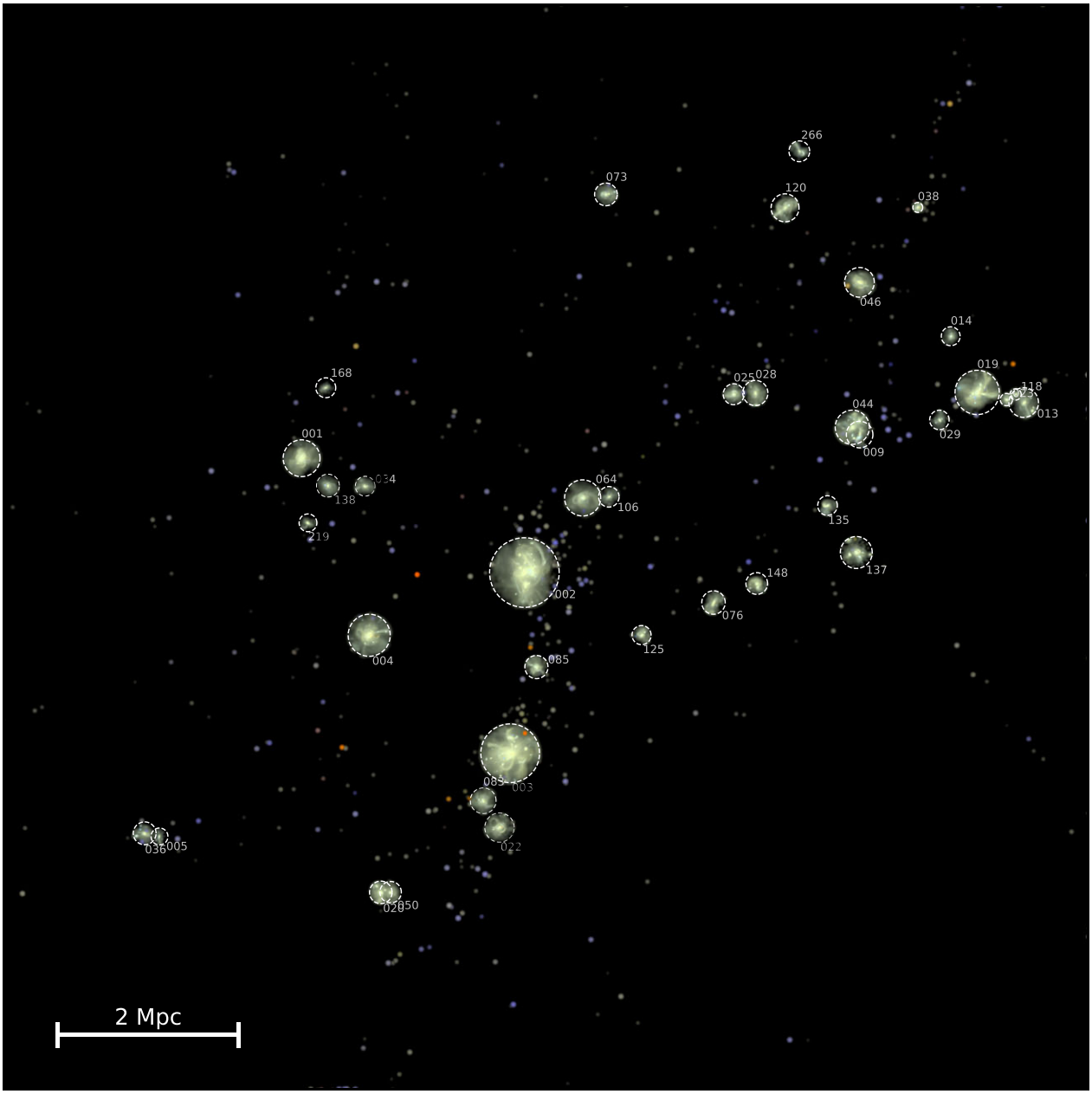


Figure 2. *g*, *r*, *i* false colour mock image showing the distribution of light within 1 $R_{\rm vir}$ of each of the massive galaxies in our sample in the context of the $\sim (16\,{\rm Mpc})^3$ simulation volume at a single simulation snapshot corresponding to $z = 0.4$. The dashed circles enclose the virial radius of each object and the numbers indicate their object ID. Less massive objects which we do not use in our sample ($10^{7.5} < M_\star/{\rm M}_\odot < 10^{9.5}$) are indicated by coloured points with the brightness and colour corresponding to mass (brighter objects are more massive) and specific star formation rate (bluer objects are more star forming), respectively.

with every other galaxy in the simulation with $10^{7.5} < M_\star/{\rm M}_\odot < 10^{9.5}$ shown as a point source whose brightness and colour correspond to their mass and specific star formation rate, respectively. The images are stretched so that black corresponds to 35 mag arcsec^{-2}. The process of producing these images is described in the next section.

2.3 Mock images

The analysis of mock observations (e.g. Jonsson 2006; Naab et al. 2014; Choi et al. 2018; Camps & Baes 2020; Olsen et al. 2021) is the most direct method of comparing models and making predictions based on theoretical or synthetic data. In the following section, we describe how we produce Rubin-like mock images for each of the galaxies in our sample.

2.3.1 *Star particle fluxes*

We produce mock images by first extracting star particles in a $(1\ {\rm Mpc})^3$ cube centred around each host galaxy. Spectral energy distributions (SEDs) for each star particle are calculated from a grid of Bruzual & Charlot (2003, BC03 hereafter) simple stellar population (SSP) models interpolated to the age and metallicity of each star particle. We assume a single Salpeter (1955) IMF for all

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

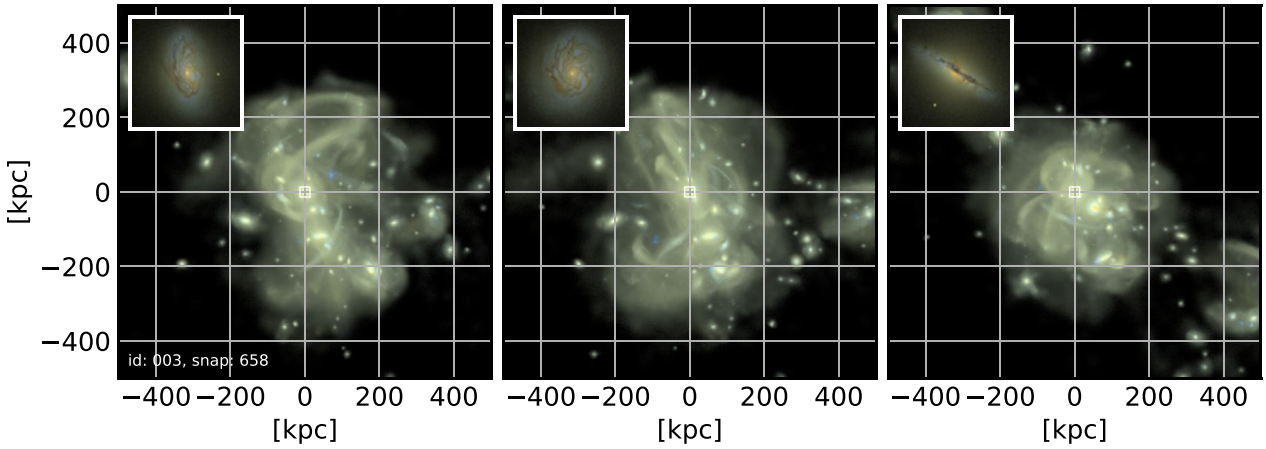


Figure 3. *g*, *r*, *i* false colour mock images of the same 1 Mpc field. From left to right, the panels show projections in *xy*, *xz*, and *yz*. The total stellar mass enclosed in the image is 3.2×10^{11} $M_\odot$ with the host accounting for 2.9×10^{11} $M_\odot$ (90 per cent). Images are produced using the arcsin stretch scheme from Lupton et al. (2004) with the *g*, *r*, and *i* bands rescaled by a factor of 1, 0.7, and 0.5, respectively, black corresponds to surface brightnesses fainter than ∼35 mag arcsec^{-2}. Inset panels show an enlarged image of the host galaxy produced using only particles detected by ADAPTAHOP as part of the primary structure. The corresponding location and scale of the inset plot (∼50 kpc) is indicated by the small white box at the centre of each panel. An interactive version of this plot showing multiple examples can be found at garrethmartin.github.io/files/example_images.html. A video showing this object rotated through multiple projections and with different limiting surface brightnesses can be found at youtube.com/watch?v=hZg_5FbnnyE. At such high limiting surface brightnesses, the morphology of the extended light can appear radically different depending on projection.

objects.[3] If we instead consider a Chabrier (2003) IMF, this changes the brightness of the central galaxy and its tidal features roughly equally so that they are both become slightly brighter overall. There is not, therefore, any qualitative impact on our results other than to increase surface brightnesses by roughly 0.6 mag arcsec^{-2} (or equivalently reducing the limiting surface brightness by the same amount) with negligible scatter introduced. Changing the IMF from Salpeter (1955) to Chabrier (2003) confers a less than a 2 per cent change in the quantities presented in Section 3.1.2.

We account for the effects of dust via a dust screen model in front of each star particle, so that the dust column density in each AMR gas cell is given by:

$$N_{\rm cell} = \rho\, Z\, \Delta r \times {\rm GDR}, \quad (1)$$

where ρ is the gas density of the cell, Z is the metallicity, Δr is the length of the cell along a given line of sight, and GDR is the gas-to-dust ratio, for which we assume a value of 0.4 (e.g. Draine et al. 2007). The total column density in front of each star particle, N, is calculated by summing along the line of sight. By calculating dust attenuation separately for each particle, we ensure that the geometry of the spatial distribution of dust within and around the galaxy is taken into account. Note that, since we focus on the outskirts of galaxies where there is very little gas and dust, modelling dust attenuation is only relevant for observational predictions for the flux of the host galaxy.

Using the $R = 3.1$ Milky Way dust grain model of Weingartner & Draine (2001), we then produce the dust attenuated SED:

$$I(\lambda)_{\rm attenuated} = I(\lambda) e^{-\kappa(\lambda) N}, \quad (2)$$

where $I(\lambda)$ is the SED's luminosity density as a function of wavelength and $\kappa(\lambda)$ is the dust opacity as a function of wavelength from Weingartner & Draine (2001). The luminosity of each star particle is calculated by first summing the resultant luminosity of the attenuated SEDs once they have been redshifted and convolved with the LSST *u*, *g*,*r*,*i*, *z*, and *y* bandpass transmission functions (Olivier et al. 2008). The apparent magnitude of each star particle is calculated taking into account mass-loss from stellar winds and the distance modulus.

2.3.2 *Smoothing*

Where the density of star particles falls below a few particles per 0.2 arcsec pixel of the Rubin Observatory LSSTCam, it is necessary to apply smoothing in order to better represent the distribution of stellar mass in phase space and remove unrealistic variation between adjacent pixels (usually only an issue in the extreme outskirts of galaxies). To achieve this, we use an adaptive smoothing scheme[4] following a similar procedure to the ADAPTIVEBOX method employed by Merritt et al. (2020).

We first create a supersample from the original star particles by splitting them into a large number of smaller particles and then distribute them according to the local density as follows:

(i) Calculate the distance to the 5th nearest neighbour for each star particle, $d_{k=5}$.

(ii) Split each star particle into 500 equal flux particles whose positions are drawn from a Gaussian distribution about the centre of the original particle and with a standard deviation equal to $d_{k=5}$ such that $P(x, y, z) \sim \mathcal{N}([x_0, y_0, z_0], \sigma = d_{k=5})$.

(iii) Create a 2D image by collapsing the particles along one of the axes and summing the flux across a 2D grid with elements of 0.2 arcsec × 0.2 arcsec.

Fig. 3 shows an example of a false colour smoothed mock image for one of our simulated galaxies in three different projections. In these images, black corresponds to a surface brightness fainter than

[3]Note that, for the purposes of calculating stellar feedback and mass-loss, the NEWHORIZON simulation assumes a Chabrier (2003) IMF (see Section 2.4 of Dubois et al. 2021).

[4]The adaptive smoothing code used in this paper is available from github.com/garrethmartin/smooth3d

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

$\sim$35 mag arcsec^{-2}. At very low surface brightnesses (significantly in excess of those currently accessible), almost all objects in our sample display multiple distinct tidal features, often with complex morphologies. Viewed at different angles, the shape and number of visible tidal features can change radically. Examples of additional objects in different projections can be seen in a supplementary interactive version of Fig. 3, found at garrethmartin.github.io/files/example_images.html. We return to the issue of how robustly tidal features are classified in multiple projections later in Section 4.1.

2.3.3 *Mock observations*

For every object we produce smoothed mock images in three projections (*xy*, *xz*, and *yz*) and at distances corresponding to a range of redshifts ($z = 0.05$ to $z = 0.8$) as described above. Each image is then convolved with a point spread function (PSF).[5] We use the *g*-band Hyper Suprime-Cam (HSC; Miyazaki et al. 2012) 1D PSF measured by Montes et al. (2021)[6] as we find that a Gaussian or Moffat distribution do not adequately describe the shape of the PSF at large radii (see Appendix B for discussion of the suitability of various PSF models and the possible effect of the PSF on our ability to detect tidal features). We note that the full-width at half-maximum (FWHM) of the PSF measured by Montes et al. (2021) is slightly broader than the expected median FWHM of the Rubin Observatory PSF (0.7 arcsec in the *r*-band (Ivezić et al. 2019) compared to a FWHM 1.07 arcsec in the *g*-band obtained by Montes et al. (2021) for HSC) and therefore slightly overestimates the likely effect of the PSF.

Finally, we add random Gaussian noise, $\mathcal{N}(0, \sigma_{\rm noise})$, where $\sigma_{\rm noise}$ is the standard deviation of the noise per pixel. We assume that the background is perfectly subtracted so that there is no variation in the noise level across the image. In reality, this is not a realistic assumption as the detection of LSB sources is sensitive to a host of factors. These include sky estimation (e.g. see Section 4.1 and fig. 5 of Aihara et al. 2019) and observing techniques, how CCD artefacts and biases (e.g. Baumer, Davis & Roodman 2017) are accounted for, as well as source extraction and masking methods. We also choose to neglect other instrumental and astrophysical contaminants (e.g. foreground and background objects, Galactic cirrus, scattered light, ghosts, and diffraction spikes) which may be present in real imaging. Although it is possible to mitigate some of this contamination (e.g. Slater, Harding & Mihos 2009; Fliri & Trujillo 2016; Román, Trujillo & Montes 2020; Tanoglidis et al. 2021a), visibility will always be reduced under realistic conditions.[7] Our results therefore represent a best case estimate. Predictions for the LSST final limiting surface brightness vary fairly significantly between $\sim$29.5 mag arcsec^{-2} (P. Yoachim, private communication) and $\sim$31 mag arcsec^{-2} (e.g. Laine et al. 2018; Brough et al. 2020) and up to 32 mag arcsec^{-2} (Brough et al. 2020) in the deep drilling fields and these limits may differ between objects dependent on the severity of sky subtraction bias.

Of course, the final depth achieved by LSST will depend strongly on how well the data are reduced and optimized for LSB science. Given a typical sky brightness in the *r*-band (21.2 mag arcsec^{-2}; Ivezić et al. 2019) and a limiting surface brightness 31 mag arcsec^{-2} requires that the sky background is characterized with an uncertainty greater than 1/10 000. The current best practices (e.g. Ji et al. 2018) allow characterization of the sky background down to four parts in one million (which, in theory, would enable tidal features to be analysed down to $\sim$34.5 mag arcsec^{-2}), meaning limiting surface brightnesses greater than 31 mag arcsec^{-2} are achievable if LSST operates, at least in theory.

For comparison, Kniazev et al. (2004) measure an SDSS *r*-band limiting surface brightness[8] of 26.2 mag arcsec^{-2}, and the IAC Stripe82 Legacy Project are able to achieve an *r*-band limiting surface brightness of 28.7 mag arcsec^{-2} in the SDSS Stripe82 calibration area (Román & Trujillo 2017). Additionally, many contemporary wide-area surveys (e.g. Gwyn 2012; Aihara et al. 2018; Dey et al. 2019) now reach limiting surface brightnesses in excess of 28 mag arcsec^{-2}. Some of the deepest imaging currently available corresponds to targeted ground based campaigns (e.g. Fry et al. 1999; Cappellari et al. 2011; Abraham & van Dokkum 2014; Ferreras et al. 2014; Duc et al. 2015; Iodice et al. 2016; Merritt et al. 2016; Trujillo & Fliri 2016; Mihos et al. 2017; Spavone et al. 2017; Iodice et al. 2019; Montes et al. 2021; Ragusa et al. 2021) which can achieve limiting surface brightness of around 30 mag arcsec^{-2} (typically requiring much longer integration times and/or with relatively limited spatial resolution and field of view compared to the Rubin Observatory).

We consider multiple noise levels, which are calculated assuming a range of limiting surface brightnesses (3σ in a 10 arcsec $\times$ 10 arcsec box). These are converted to a per-pixel 1σ variance by rearranging the equation found in Román et al. (2020, Appendix A) (for simplicity, we neglect the zero-point):

$$\sigma_{\rm noise} = \frac{10^{-0.4\,\mu_r^{\rm lim}(n\sigma,\Omega\times\Omega)}\,{\rm pix}\,\Omega}{n}, \tag{3}$$

where Ω is the size of one of the sides of the box in arcseconds, pix is the pixel scale in arcseconds per pixel, n is the number of Gaussian standard deviations, and $\mu_r^{\rm lim}$ is the *r*-band limiting surface brightness, in this case $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec})$. Finally, we perform a 5 $\times$ 5 re-binning of the images to an angular scale of 1 arcsec. As the FWHM of the PSF is also around 1 arcsec, this represents the maximum binning we can perform before the images start to lose fidelity.

For the reasons outlined above, we do not target any specific prediction for image depth but instead opt to explore a range of values for $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec})$. We choose values between 28 and 31 mag arcsec^{-2}, which roughly encompass expected depths from a single exposure and close to the upper end of predictions of the depth of a full 10-yr stacked exposure (825 visits).

For each limiting surface brightness, we produce images assuming different redshifts. While the largest telescopes are capable of producing sufficiently deep imaging of local galaxies ($z < 0.01$) with minutes to a few hours of integration time (e.g. Sand et al. 2009; Trujillo et al. 2021), the Rubin Observatory will be capable of collecting much larger samples at higher redshifts ($0.05 < z < 0.1$) thanks to the large survey area of LSST. Therefore, we do not consider the very local Universe, instead picking a range of redshifts starting at $z = 0.05$ and extending to high redshift ($z = 0.8$), with cosmological redshift and surface brightness dimming taken into account, in order

[5]The PSF FWHM is always larger than the smoothing length in regions of interest (i.e. for the galaxy and dense tidal features as defined in Section 2.4).
[6]Measured to 289 arcsec and extrapolated to 420 arcsec based on a power-law fit.
[7]See Mihos (2019) for a review of recent advances and challenges in deep imaging.
[8]Note that estimates of the limiting surface brightness are not always directly comparable as they can vary depending on the exact methodology used.

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

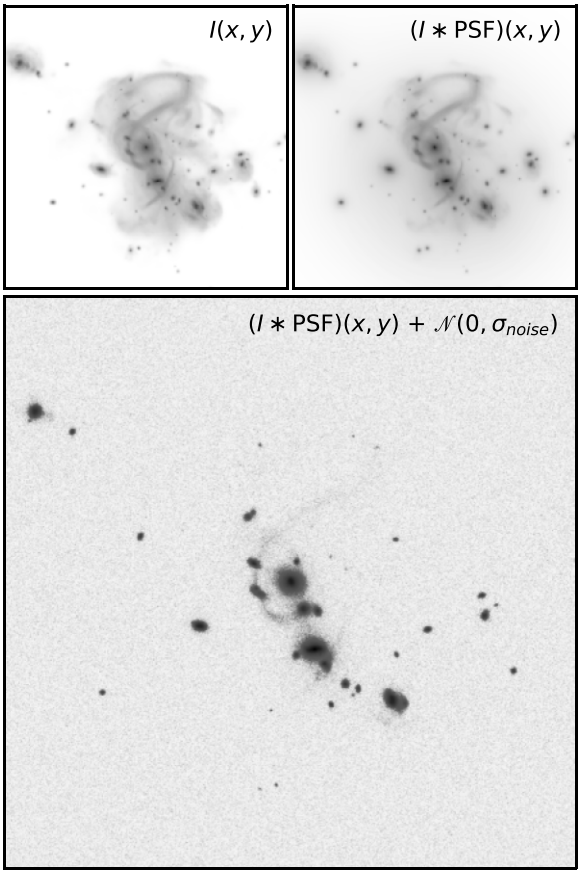


Figure 4. Plot demonstrating the steps taken to produce the mock images used for classification. As in Fig. 3, white corresponds to surface brightnesses fainter than ∼35 mag arcsec^{-2}. Top left-hand panel: *r*-band mock image of a galaxy at $z = 0.2$. Top right-hand panel: the image is convolved with the HSC PSF derived by Montes et al. (2021). Bottom: Gaussian random noise is added to the convolved image.

Table 1. Range of parameters used to generate mock observations: *a*, limiting *r*-band surface brightness for a 10 arcsec × 10 arcsec box with bracketed values showing the equivalent for a Chabrier (2003) IMF; *b*, redshift corresponding to viewing distance; *c*, axis of projection; *d*, point spread function.

Parameter	Values
$\mu_r^{\rm lim}(3\sigma, 10\,\mathrm{arcsec} \times 10\,\mathrm{arcsec})$ [a]	[28(27.43), 29(28.43), 30(29.43), 31(30.43)]
z [b]	[0.05,0.1,0.2,0.4,0.8]
π [c]	[*xy*, *xz*, *yz*]
PSF [d]	Montes et al. (2021)

to match the capabilities of the Rubin Observatory and test its ability to resolve tidal features in the more distant Universe.

Fig. 4 illustrates the process of producing a single mock observation in the *r*-band. Moving clockwise from the top left-hand panel we show the original *r*-band mock image created as described in Section 2.3.1, the same image convolved with the PSF and finally, with Gaussian random noise added.

Table 1 shows the full range of parameters used to produce mock observations, which add up to a total of 60 different variations per object. Values for the limiting surface brightness in brackets indicate the equivalent if we choose a Chabrier (2003) IMF[9] instead of a Salpeter (1955) IMF. We also produce an extra set of mock images for $\mu_r^{\rm lim} = 35$ mag arcsec^{-2} and $z = 0.05$ which we use as a 'ground truth' for the other mock observations. This value is informed by the stellar mass resolution of the simulation, since we do not expect tidal features to contain enough particles to produce sufficient signal-to-noise at stellar mass surface densities equivalent to ∼35 mag arcsec^{-2}. Visual inspection confirms that we do not visually detect additional structures in the diffuse light beyond 35 mag arcsec^{-2}. Given the finite resolution of the simulation and the hierarchical nature of galaxy assembly, it is likely that additional tidal features would become visible with finer resolution (see also Appendix A).

2.4 Measuring tidal features

We perform a separate measurement of the galaxy tidal features based on the distribution of particles in the simulation. In order to do this, we separately define tidal features or tidal debris as any star particles within the (1 Mpc)^3 volume that are not detected as part of an object or substructure by ADAPTAHOP (see Section 2.2). We define tidal features, as opposed to more diffuse tidal debris based on a $k = 5$ nearest neighbour stellar mass weighted density[10] threshold.

In order to determine the optimum density threshold we first assume that tidal features have a higher spatial frequency on average than the diffuse component, so that increasing the density threshold until the high frequency component is minimized should allow us effectively segregate features above and below a certain spatial frequency threshold. Based on the measurements of Sola et al. (2022), we choose a threshold value of 50 kpc, which is a little larger than the average width of the largest tidal tail measured. To determine the density threshold corresponding to this spatial frequency, we calculate the 2D Fourier transform of images produced at decreasing density thresholds until the power at frequencies smaller than 50 kpc approaches a minimum value (essentially applying a high-pass filter as in e.g. Popesso et al. 2012), arriving at an optimum density of $\rho_\star > 500\ \mathrm{M}_\odot\ \mathrm{kpc}^{-3}$.

Using these definitions we identify the following regions:

(i) *structure* (**S**) – members of any structure or substructure found by ADAPTAHOP

(ii) *host* (**H**) – members of the structure identified as the host galaxy

(iii) *tidal debris* (**T**) – not members of any structure or substructure and where $\rho_\star < 500\ \mathrm{M}_\odot\ \mathrm{kpc}^{-3}$

(iv) *dense tidal features* ($\mathbf{T_d}$) – not members of any structure or substructure and where $\rho_\star > 500\ \mathrm{M}_\odot\ \mathrm{kpc}^{-3}$

Using these definitions, we can again calculate star particle fluxes as described in Section 2.3.1 and create smoothed images using only the star particles identified as making up tidal features or tidal debris. Fig. 5 shows an example of this process: the surface brightness map created using all star particles is shown in the top left and the tidal feature and tidal debris surface brightness maps (panels c and d) are produced from particles that are identified in panel (b) in purple (dense tidal features; c) and red (extended tidal debris; d).

Our definition of tidal debris includes the diffuse material and debris in the outskirts of satellite galaxies. Their contribution to the

[9] A Kroupa (2002) IMF also yields very similar results.

[10] $\rho_\star = \sum_{i=1}^{k} m_{\star,i} \frac{3}{4\pi d_k^3}$

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

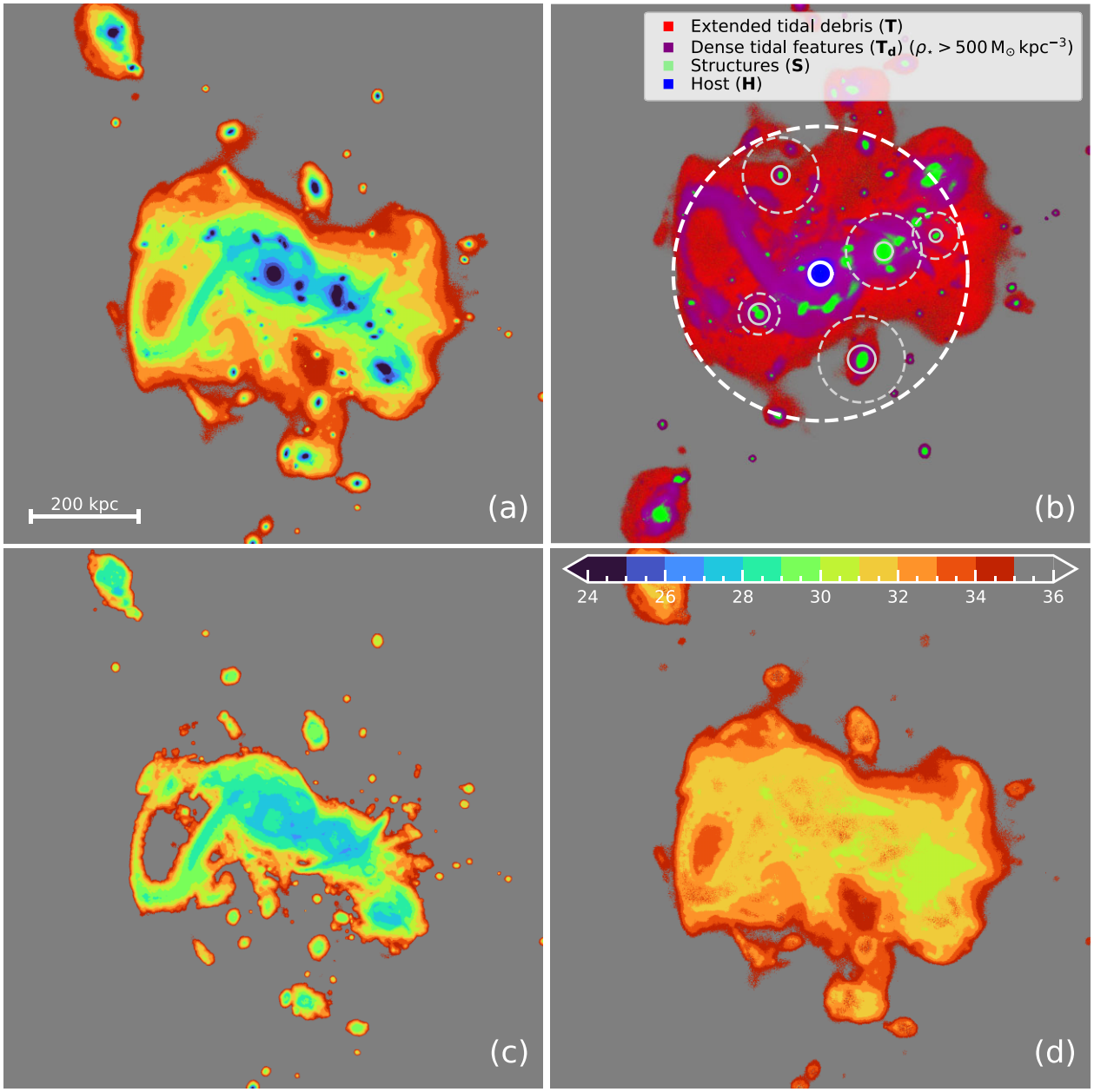


Figure 5. (a) *r*-band surface brightness map of one of the same mock fields as shown in Fig. 3. (b) A plot showing the same field with colours corresponding to tidal debris (T, red), dense tidal features ($\mathbf{T_d}$, purple), structures identified by ADAPTAHOP (S, green) and the central host galaxy (H, blue) with an equivalent stretch to the first image. The solid and dashed white circles indicate seven $R_{\rm eff}$ and $R_{\rm vir}$ of the host, respectively. The grey circles indicate the same for the most massive satellites. (c) *r*-band surface brightness map created from particles identified as dense tidal features, where $\rho_\star > 500$ M$_\odot$ kpc^{-3} (typical surface brightnesses of $\mu_r \sim 28 - 32$ mag arcsec^{-2}). (d): surface brightness map created from particles identified as tidal debris where, $\rho_\star < 500$ M$_\odot$ kpc^{-3} (typical surface brightnesses of $\mu_r \gtrsim 32$ mag arcsec^{-2}). Image scale in proper kpc is indicated by the scale bar in the bottom left of panel (a) and the colour bar at the top of panel (d) indicates the scale of the surface brightness maps in mag arcsec^{-2}. Pixels fainter than 35 mag arcsec^{-2} or brighter than 24 mag arcsec^{-2} are coloured grey and dark blue, respectively.

total tidal flux is typically small – for the most massive 20th percentile of haloes, we find that, of the total tidal flux found within 1 $R_{\rm vir}$ of the host galaxy, a median / mean of 7 / 8 cent and 24 / 25 per cent is found within 5 $R_{\rm eff}$ and 10 $R_{\rm eff}$ of its satellite galaxies, respectively. For any individual halo this value never exceeds 15 per cent or 50 per cent for 5 $R_{\rm eff}$ and 10 $R_{\rm eff}$, respectively (larger values come from systems with an ongoing major or minor merger where the host and one of its minor companions are close). The proportion of tidal flux contained in satellites decreases further for less massive haloes.

2.5 Merger tree

We construct merger trees for each galaxy according to the method of Tweed et al. (2009) based on the stellar particles of galaxies identified using ADAPTAHOP. The time resolution of the merger trees is ∼15 Myr, enabling us to track in detail the main progenitors (the object in the chain of the most massive progenitors at each snapshot) and mass assembly of each galaxy. We follow the stellar mass evolution and stellar mass accretion history of the host galaxy,

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

which in turn allows us to determine the merger history and ex-situ mass fraction of the galaxy.

We identify stars as ex-situ by iterating along the main branch of the merger tree. For each snapshot, star particles which were formed after the previous snapshot and which are identified by ADAPTAHOP as members of the host galaxy in the current snapshot are marked as in-situ, then at the final snapshot $t_{\rm max}$, any star particles identified by ADAPTAHOP as members of the host galaxy which are not marked as *in-situ* are considered to have formed ex-situ. The ex-situ mass fraction, $f_{\rm exsitu}$, between the current time, $t_{\rm max}$, and some previous time, $t_{\rm min}$, can then be defined for each host galaxy as follows:

$$f_{\rm exsitu}(t_{\rm max}, t_{\rm min}) = \frac{\sum\{m_\star \,|\, t_{\rm min} < t_{\rm birth} < t_{\rm max} \ \wedge \ {\rm ex-situ}\}}{\sum\{m_\star \,|\, t_{\rm min} < t_{\rm birth} < t_{\rm max}\}}, \quad (4)$$

where $m_\star$ is the set of particles all identified by ADAPTAHOP at $t_{\rm max}$, $\{m_\star \,|\, t_{\rm min} < t_{\rm birth} < t_{\rm max}\}$ is the subset of these particles with formation times, $t_{\rm birth}$, between $t_{\rm min}$ and $t_{\rm max}$ and similarly, $\{m_\star \,|\, t_{\rm min} < t_{\rm birth} < t_{\rm max} \ \wedge \ {\rm ex-situ}\}$ is the subset of all these particles formed between $t_{\rm min}$ and $t_{\rm max}$ which were formed ex-situ. Using this parameter, we are able to quantify how the visibility of tidal features correlates with the accretion history of each system. Throughout the rest of this paper we adopt a value of $t_{\rm min}$ equal to the earliest time that the main progenitor can be traced so that $f_{\rm exsitu}$ encompasses the entire lifetime of the object. As we discuss in Appendix C, increasing or reducing $t_{\rm min}$ does not have a statistically significant effect on either the correlation between $f_{\rm exsitu}$ and halo mass or the correlation between $f_{\rm exsitu}$ and tidal flux fraction, $f_{\rm tidal}$ (see Section 3.1.2).

Our definition of $f_{\rm exsitu}$ includes only stars that at are identified as part of the host galaxy by ADAPTAHOP (i.e. parts of panel (b) of Fig. 5 colour coded in blue), so the set of particles used to calculate the ex-situ mass and tidal flux are mostly mutually exclusive. $f_{\rm exsitu}$ should be considered as a measure of the ex-situ mass fraction within the central galaxy itself rather than of the entire system including the extended envelope.

2.6 Visual classification

Except for images deemed too noisy to effectively classify,[11] which were rejected based on visual inspection, we perform visual classifications for all of the permutations of each object (totalling ∼8000 unique images). Images were shared among 45 expert classifiers so that each permutation was classified separately by at least two people. A subset of ∼600 images were classified five times each in order to more robustly measure variation between classifiers. To maximize reproducibility, all classifiers were asked to follow a set of detailed instructions that included prototypical examples of each class of tidal feature and asked to classify around 300 individual images. For each permutation, classifiers viewed a set of three greyscale surface brightness maps with a fixed noise level but with a maximum stretch set to 27, 29, and 33 mag arcsec^{-2} and an arcsinh stretched g, r, i false colour image which were cropped to a physical size of 100 kpc × 100 kpc. They were asked to count the number of instances that they identified certain categories of tidal feature with various signatures of mergers and interactions considered – interacting galaxies with double nuclei, merger remnants, bridges, tidal tails, stellar streams, shells and plumes (e.g. Lofthouse et al. 2017; Bílek et al. 2020). Classifiers were asked to make their determinations according to the following criteria, taking into account the surrounding context and their physical interpretation of the image:

(i) *Stellar streams* – A stream of stars that can appear to trace an ellipse, spiral or straight line depending on the angle at which they are viewed.

(ii) *Tidal tails* – Differing from stellar streams in that a tidal tail must originate from material ejected from the host galaxy or a massive companion. They are therefore likely to be associated with interactions between similar mass galaxies and consist of material that has been unbound from a disrupted galaxy rather than gradually stripped.

(iii) *Asymmetric stellar haloes* – Diffuse, low surface brightness features in the outskirts of the galaxy that do not have a well-defined structure like stellar streams or objects where the overall structure of the diffuse stellar halo is clearly not symmetric.

(iv) *Shells* – radial structures consisting of concentric arcs or ring-like structures that do not cross one another.

(v) *Tidal bridges* – a tidal tail that connects two interacting galaxies. Bridges are formed from the material that gets dragged out during high mass ratio mergers, rather than material has been gradually stripped away over many orbits.

(vi) *Merger remnants* – A strongly morphologically disturbed galaxy where the merging objects have recently coalesced. May also exhibit tidal tails but there should be no indication of a second galaxy.

(vii) *Double nuclei* – visibly merging / interacting with both objects still clearly separated. There must be some sign of an interaction taking place (i.e. not only close pairs).

Examples of each of these categories are shown in Fig. 6.

In comparison to simpler categorization schemes, such as separating galaxies into elliptical and spiral morphologies (e.g. Lintott et al. 2008; Uzeirbegovic, Geach & Kaviraj 2020), a higher level of domain knowledge is required to perform detailed classifications of galaxy tidal features. This stems from the fact that correctly interpreting tidal features can depend on context and is often reliant on a physical understanding of the interactions taking place. For example, stellar streams can form shell-like morphologies (Foster et al. 2014) or tidal tails may appear similar to linear streams when observed edge-on. Even for expert classifiers, characterizing tidal features can still be quite uncertain, especially at high redshifts or at low limiting surface brightness, which both significantly alter the appearance of tidal features. The reliability of visual classifications is discussed in Section 4.1.

3 PROPERTIES AND DETECTABILITY OF THE TIDAL FEATURES AND DIFFUSE LIGHT AROUND GALAXIES

3.1 Quantifying the distribution and strength of tidal features

In this section, we consider how light is distributed around galaxies and their extended envelopes. We explore the distribution of light as a function of surface brightness in the main body of galaxies and in their extended envelopes as well as the total fraction of light that makes up different regions of the galaxy. In Appendix D, we show that the mass accretion histories of the galaxies in our sample are sufficiently stochastic that any average evolution between $z = 0.2$ and $z = 0.8$ is largely washed out. Based on this finding, we treat each instance of the same galaxy across the four snapshots considered as

[11] All objects had sufficient signal-to-noise to make classification possible except at $z = 0.8$, where only half of objects at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 30 - 31$ mag arcsec^{-2} were classifiable and almost no objects at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 28 - 29$ mag arcsec^{-2} were classifiable.

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

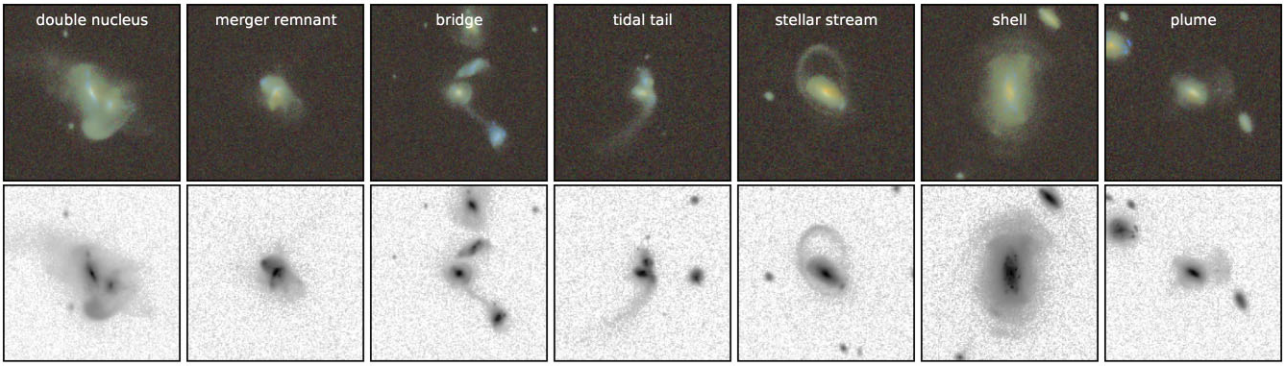


Figure 6. Example 100 Mpc ×100 Mpc thumbnail images for each category. For each object, we show the arcsinh stretched *g*, *r*, *i* false colour image (top row) and greyscale surface brightness map (bottom row).

independent objects. If we restrict our analysis only to galaxies at a single snapshot rather than include their progenitors at different redshifts, we do not see a notable difference in our results.

3.1.1 *Distribution of light in extended structures*

We first discuss how tidal feature flux is distributed, spatially and as a function of surface brightness.

The top left-hand and right-hand panels of Fig. 7 show the distribution of flux as a function of the galaxy's 2D effective radius and of surface brightness respectively. In both cases, galaxies are observed at a redshift of $z = 0.05$ and an angular resolution of 0.2 arcsec. The median fraction of cumulative flux contained within pixels brighter than 35 mag arcsec^{-2} and $R < R_{\rm vir}$ is shown separately for the host galaxy (**H**), dense tidal features ($\mathbf{T_d}$), and extended tidal debris (**T**), indicated by thick blue, purple and red lines, respectively. Within the host galaxy itself, the majority of pixels lie at low surface brightnesses (50 per cent fainter than 27 mag arcsec^{-2}) and small radii ($R < 5\ R_{\rm eff}$). Median profiles for the dense tidal features are plotted for three different mass ranges with $M_\star > M^\star$ ($M_\star > 10^{10.8}\,{\rm M}_\odot$), Milky Way mass galaxies ($10^{10.25}{\rm M}_\odot < M_\star < 10^{10.75}{\rm M}_\odot$) and low-mass galaxies ($M_\star < 10^{10}{\rm M}_\odot$) indicated by dashed, dash-dotted and dotted purple lines, respectively.

Typically, a majority of the flux from dense tidal features is found at smaller radii than extended tidal debris, with 50 per cent of flux contained within 7 $R_{\rm eff}$ compared with 10 $R_{\rm eff}$ for extended tidal debris. There is significant scatter in the cumulative flux distributions for both dense tidal features and extended tidal debris, meaning that, in many cases, a majority of the tidal flux lies at a large separation from the central galaxy. Typically, close to 100 per cent of tidal flux is contained within 25 $R_{\rm eff}$ or ∼0.6 $R_{\rm vir}$ (where $R_{\rm eff} \sim 4$ kpc for $M^\star$ galaxies on average), but a fairly substantial proportion lies beyond this in some cases. In particular, more massive galaxies tend to have tidal features whose flux extends further into the halo, with an average of 40 per cent of the total stellar halo flux lies beyond 25 $R_{\rm eff}$ in the most massive galaxies ($M_\star > 10^{11}{\rm M}_\odot$). This is significantly further into the stellar halo than many contemporary studies are typically able to probe. For example Merritt et al. (2016) and Trujillo et al. (2021) measure galaxy surface brightness profiles out to 15–20 $R_{\rm eff}$. While this is partially limited by the depth of imaging available (LSST will be similarly limited), at increasingly large radii, measurements of the surface brightness profile are increasingly likely to be contaminated by nearby bright objects. Note that studies which do probe deeper into the stellar halo (e.g. Buitrago et al. 2017; Borlaff et al. 2019) do appear to detect a larger fraction of flux outside of the main galaxy.

Without very coarse binning, and for the majority of galaxies in our sample ($M_\star \sim M^\star$), surface brightness limits significantly fainter than 32 mag arcsec^{-2} would be needed to recover a significant fraction of the flux found in extended diffuse light, which, in integrated light, is beyond the capabilities of any contemporary or forthcoming instrument including the Rubin Observatory. However, as we discuss in Section 3.3, the brighter parts of denser tidal features (and therefore much of the total light) will be detectable in many cases.

Finally, the bottom panel of Fig. 7 shows the joint distribution of the cumulative flux fraction for dense tidal features as a function of projected radius and surface brightness (i.e. the fraction of total flux in dense tidal features that is contained in pixels that are both within a given radius and brighter than a given surface brightness). Again, the majority of the total flux in dense tidal features resides within relatively small radii and at surface brightnesses that are in reach of LSST. For example, on average 50 per cent of flux lies within 10 $R_{\rm eff}$ and in pixels brighter than 31 mag arcsec^{-2} with a 1σ scatter of $\pm^{28}_{41}$ percentage points, and 76 per cent lies within 15 $R_{\rm eff}$ and $\mu_r = 32$ mag arcsec^{-2} with a 1σ scatter of $\pm^{18}_{26}$ percentage points. As these values indicate, there is significant scatter in the flux fraction, which appears particularly large at the expected LSST surface brightness limits (30–31 mag arcsec^{-2}).

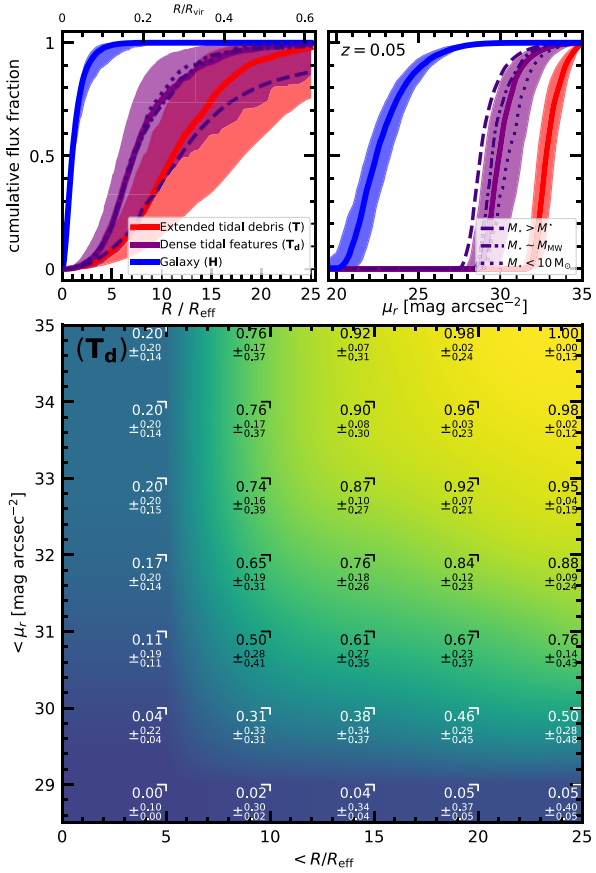
3.1.2 *Proportion of galaxy flux in tidal features*

In this section, we consider the amount of flux found in dense tidal features and extended tidal debris compared to the host galaxy – the *tidal flux fraction*, $f_{\rm tidal}$ – defined as:

$$f_{\rm tidal} = \frac{F_{\mathbf{T}} + F_{\mathbf{T_d}}}{F_{\rm tot}}, \tag{5}$$

where $F_{\mathbf{T_d}}$ and $F_{\mathbf{T}}$ are the total flux within a 3D radius greater than 5 $R_{\rm eff}$ and smaller than 1 $R_{\rm vir}$ found in dense tidal features and extended tidal debris respectively and $F_{\rm tot}$ is the total flux from every particle within 1 $R_{\rm vir}$ of the centre of the host galaxy. A minimum radius of 5 $R_{\rm eff}$ is sufficient to avoid any components of the galactic disc that may have been missed by ADAPTAHOP (typically recently formed resolved clusters of stars that are rejected for falling below the minimum particle number threshold) but not large enough to exclude a significant contribution from the stellar halo (e.g. Abadi, Navarro & Steinmetz 2006; Pillepich, Madau & Mayer 2015) outside of the central few effective radii (however in almost all cases all particles within 5 $R_{\rm eff}$ are associated to a structure or substructure of the host galaxy and the contribution by other particles not associated with any structure or sub-structure within 5 $R_{\rm eff}$ is negligible).

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023



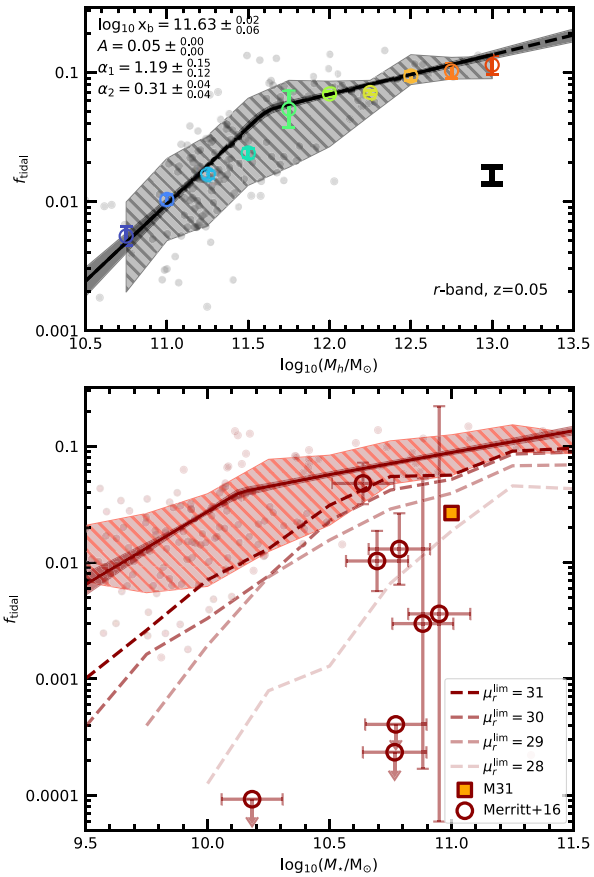


Figure 7. Top left-hand panel: The median cumulative fraction of flux as a function of projected radius ($\mu_r < 35$, $R < R_{\rm vir}$) for the host galaxy (**H**, blue), dense tidal features ($\mathbf{T_d}$, purple), and extended tidal debris (**T**, red). The coloured filled regions indicate the central 68th percentile (1σ) of the distribution. The top x-axis gives an indication of the approximate equivalent value of $R/R_{\rm vir}$ corresponding to the value of $R/R_{\rm eff}$ (based on median measured ratio for $R_{\rm vir}$ to $R_{\rm eff}$ of 1: 40.) shown on the bottom x-axis. Top right-hand panel: The median cumulative fraction of flux ($\mu_r < 35$, $R < R_{\rm vir}$) as a function of surface brightness. For both panels, median dense tidal feature profiles for three different mass ranges are shown with corresponding line styles indicated in the legend of the top right-hand panel. Bottom: colour plot showing the joint distribution of the cumulative flux fraction for dense tidal features as a function of projected radius and surface brightness. Values shown indicate the median and 1σ scatter of the cumulative flux fraction where brackets to the top right of each value indicate the maximum radius and surface brightness that the cumulative flux fraction is calculated within. Typically, flux from extended tidal debris is distributed towards larger radii and lower surface brightnesses than the dense tidal features, although there is significant scatter in the distribution of both components. In some cases, a non-negligible fraction of flux lies even beyond 25 $R_{\rm eff}$.

Figure 8. Top: Scatter plot showing the tidal flux fraction ($f_{\rm tidal}$, equation 5) in the r-band as a function of halo mass. The black line and grey shaded region show a broken power-law fit to the grey points representing individual NEWHORIZON galaxies and the associated 1σ uncertainty obtained from 100 000 bootstraps. The larger shaded and hatched region indicates the 1σ scatter of the grey points. The median $f_{\rm tidal}$ and associated errors in halo mass bins are indicated with coloured errorbars. The black error bar indicates the mean fractional variation in $f_{\rm tidal}$ due to orientation (∼30 per cent). Bottom: Similarly to the top panel, scatter points show $f_{\rm tidal}$ in the r-band as a function of stellar mass and a fit to $f_{\rm tidal}$ as a function of stellar mass is shown as a red line. The red open circles with error bars indicate the stellar halo mass fraction and associated 1σ uncertainties as a function of stellar mass from Merritt et al. (2016), with an additional data point for M31 (Courteau et al. 2011) plotted as a filled orange square. The dashed red lines indicate the median $f_{\rm tidal}$ that is recovered at different r-band limiting surface brightnesses (3σ, 10 arcsec × 10 arcsec). As expected, $f_{\rm tidal}$ increases as a function of halo mass, however the normalization and scatter of the relationship changes considerably with limiting surface brightness.

We calculate $f_{\rm tidal}$ in each LSST band for each object, fitting a broken power law with the form

$$f_{\rm tidal}(M) = \begin{cases} A[\log_{10}(M/M_b)]^{\alpha_1} & M < M_b, \\ A[\log_{10}(M/M_b)]^{\alpha_2} & M > M_b, \end{cases} \quad (6)$$

where M is the galaxy mass (stellar mass or halo mass), M_b is the mass at the break point, A is the amplitude at M_b, and α_1 and α_2 are power law indices before and after M_b, respectively.

Fig. 8 shows the distribution of $f_{\rm tidal}$ in the r-band as a function of halo mass (top panel) and stellar mass (bottom panel) as black or red coloured points. Solid black and red lines and smaller shaded regions show the broken power-law fit to the grey points and the associated 1σ uncertainty obtained from 100 000 bootstraps. The larger hatched and shaded regions indicate the central 68th percentile (1σ) of the distribution of the grey and red points. In the top panel, we also show the median $f_{\rm tidal}$ and associated errors for a number of overlapping 0.5 dex wide mass bins as coloured errorbars. While there is considerable scatter as a result of the stochastic nature of galaxy accretion histories, the fraction of flux in tidal features increases towards higher masses on average for both stellar and halo mass. This is consistent with both observational and theoretical studies, which show that the merger rates and $f_{\rm exsitu}$ of galaxies are larger for more massive galaxies (e.g. Stewart et al. 2008; Rodriguez-Gomez et al. 2015; Martin et al. 2019, 2021) (as we show

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

later in Section 3.2, $f_{\rm exsitu}$ and $f_{\rm tidal}$ are correlated in our simulations). The fraction, $f_{\rm tidal}$, increases from less than 1 per cent for $M_h = 10^{10.5}$ M$_\odot$ up to around 10 per cent at $M_h \sim 10^{13}$ M$_\odot$. We observe a break in the relation at a halo mass of $10^{11.6\pm^{0.03}_{0.06}}$ M$_\odot$ or a stellar mass of $10^{10.1\pm^{0.01}_{0.05}}$ M$_\odot$, which corresponds with the crossover mass at which elliptical galaxies begin to dominate and mergers become the dominant process driving the evolution of galaxies (e.g. Huertas-Company et al. 2010; Robotham et al. 2014; Thanjavur et al. 2016). One possible explanation is that the bulge itself is a remnant of past interactions (e.g. Martin et al. 2018a; Park et al. 2019) (but see also Gargiulo et al. 2019). Since more massive early-type galaxies typically formed their bulge at earlier times (Martin et al. 2018b) and typically exhibit fewer tidal features the earlier they formed (Yoon & Lim 2020), it is expected that the relationship between galaxy mass and $f_{\rm tidal}$ should weaken as spheroidal component of galaxies begin to dominate. It should be noted that while tidal features generally trace relatively recent events in a galaxy's accretion history, the extent that tidal features trace mass assembly or accretion history is complicated by the fact that the different types of tidal feature fade over different time-scales and their flux may become more difficult to detect. Therefore, $f_{\rm tidal}$ can also be sensitive to a range of factors beyond bulk accretion history.

Based on an extrapolation of our fit (dashed line), we would expect to find over 20 per cent of flux in tidal features for the most massive galaxies ($M_\star > 10^{12}$M$_\odot$). We also note that, if we consider the fraction of stellar mass found in tidal features instead of flux, we obtain very similar results. It should be noted that we do not see any significant change in our results if we consider other LSST photometric bands (u, g, i, z or y) except that the low mass slope becomes slightly shallower towards redder bands so that $\alpha_1 = 1.33\pm^{0.22}_{0.15}$ in the u-band falling to $\alpha_1 = 1.13\pm^{0.14}_{0.13}$ in the y-band.

The black error bar in Fig. 8 indicates the mean fractional variation in $f_{\rm tidal}$ due to orientation, which accounts for a variation of around 30 per cent. This variation is driven by changes in the effective gas geometry (e.g. Calzetti 2001), which act primarily to change the amount of dust attenuation and therefore the integrated flux of the galaxy. As dust column densities are much lower in the outskirts of galaxies, the integrated tidal flux in not similarly affected. As we discuss later, the fraction of observed tidal debris is much more dependent on the geometry of the tidal features as they may move above or below the limiting surface brightness depending on the angle at which they are viewed.

Additionally, we compare our results with those of Merritt et al. (2016). The red open circles with error bars or limits in the bottom panel indicate the stellar halo mass fraction and associated 1σ uncertainties as a function of stellar mass (corresponding to the red top x-axis). The value for M31 (Courteau et al. 2011) is plotted as a filled orange square and re-scaled by Merritt et al. (2016) to be consistent with their own definition. The quantity calculated by Merritt et al. (2016) is derived in a similar way to $f_{\rm tidal}$. In both cases any flux within 5 $R_{\rm eff}$ is ignored and an attempt is made to remove any contaminating flux from the galaxy itself (see Section 3.2 of Merritt et al. (2016) for a description of the method). Merritt et al. (2016) only account for flux from the main galaxy out to 7 $R_{\rm eff}$, while we find that a large fraction of tidal flux lies beyond 7 $R_{\rm eff}$ (see Fig. 7) and that the galaxy itself does not account for a significant fraction of flux by 7 $R_{\rm eff}$. Discrepancies could arise if the light of the galaxy within 7 $R_{\rm eff}$ is not accurately subtracted. Therefore, we expect these two methods to be somewhat comparable, but systematic differences likely exist. Also note that local galaxy samples are dominated by galaxies with late-type morphology and there is a considerable difference in values measured in these galaxies and similar mass early-type galaxies at higher redshifts (e.g. Buitrago et al. 2017).

While a naïve comparison with the Merritt et al. (2016) data points suggests a very significant level of disagreement with our simulated galaxies in both the normalization and level of scatter in the data points, it is important to account for the fact that some fraction of the total flux will always be missed due to the finite limiting surface brightness of the observed data. Merritt et al. (2016) achieve a limiting surface brightness of up to $\mu_g^{\rm lim}(3\sigma, 60\ {\rm arcsec} \times 60\ {\rm arcsec}) = 29.8$ mag arcsec^{-2} which we convert to a limiting surface brightness of around $\mu_r^{\rm lim}(3\sigma,\ 10\ {\rm arcsec} \times 10\ {\rm arcsec}) \sim 28$ mag arcsec^{-2} based on Appendix A of Román et al. (2020) and assuming a difference of 0.5 mag between the g and r bands. If we account only for detected flux (following the same method described later in Section 3.3) we find that the normalization of $f_{\rm tidal}$ falls as we move to lower limiting surface brightness (indicated by red dashed lines). Scatter also increases with shallower surface brightness limits. The standard deviation for approximately Milky Way mass galaxies ($10^{10.25}\,{\rm M}_\odot < M_\star < 10^{10.75}\,{\rm M}_\odot$) increases from 0.4 dex at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 31$ mag arcsec^{-2} to 1.3 dex at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 28$ mag arcsec^{-2}. This brings the data points from Merritt et al. (2016) into closer agreement with the simulated data, however the increase in scatter is comparable to the change in normalization (note that other studies (e.g. Monachesi et al. 2016; Harmsen et al. 2017) show a similar amount scatter and individual measurements (e.g. Carollo et al. 2010; Courteau et al. 2011; Deason, Belokurov & Sanders 2019; Smercina et al. 2020) span a similar range of values to Merritt et al. (2016). Again, we cannot be sure of the systematic differences between our methodology and that of Merritt et al. (2016) (especially as we do not take azimuthal averages as they have done). However, this result shows that while the normalization of $f_{\rm tidal}$ is expected to fall at brighter limiting surface brightness, the scatter is also expected to increase significantly, indicating that the uncertainty in the observations may currently be too large to make any valid comparison to theory.

In Fig. 9, we consider the fraction of flux that is found in dense tidal features versus more extended tidal debris, $F_{\mathbf{T}_d}/(F_{\mathbf{T}_d} + F_{\mathbf{T}})$, as a function of halo mass. Again, coloured error bars indicate the median and 1σ error for overlapping 0.5 dex wide mass bins and the grey filled region indicates the central 68th percentile (1σ) of the distribution of the grey points. The red dotted line shows the same relation as a function of stellar mass with the scale shown on the red top axes. For any given halo mass or stellar mass, around 80 per cent of the extended light of a galaxy is found in dense, generally higher surface brightness tidal features. The panel to the right of the main plot shows the probability density function (PDF) for $F_{\mathbf{T}_d}/(F_{\mathbf{T}_d} + F_{\mathbf{T}})$ marginalized over halo mass. As the PDF shows, galaxies are rarely found with more than 30 per cent of their extended flux in extended tidal debris rather than dense tidal features (<5 per cent chance). Additionally, we do not find any galaxies in which the amount of flux in extended tidal debris outweighs that found in dense tidal features. This argues that observational studies which recover much of the coherent tidal features will likely not be missing significant amounts of the accreted mass. With this in mind, missing flux due to finite resolution may, in some cases, be a similarly important consideration in measuring halo flux (in addition to the PSF and scattered light, as argued by e.g. Abraham & van Dokkum 2014), since smaller resolution elements allow tidal features to be analysed at higher surface brightness compared with the diffuse component of the halo (see also the discussion in Trujillo et al. 2021).

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

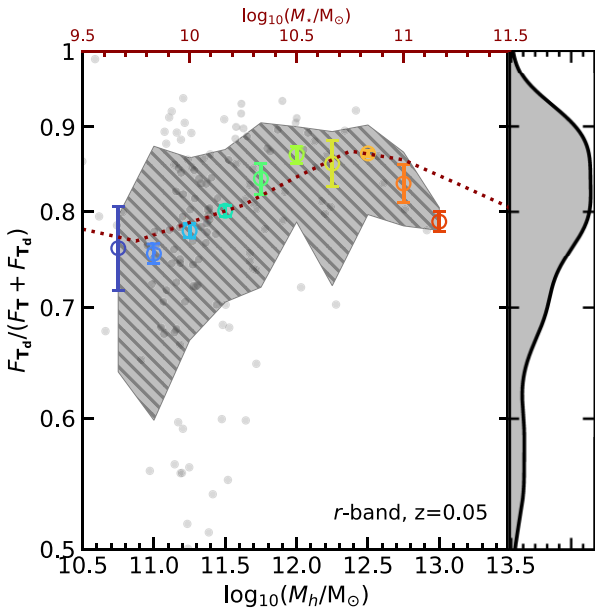


Figure 9. Scatter plot showing the fraction of tidal flux that is found in dense ($\rho_\star > 500$ M$_\odot$ kpc^{-3}) tidal features ($F_{\mathbf{T}_d}/(F_{\mathbf{T}_d} + F_{\mathbf{T}})$) versus the total flux in all tidal features) as a function of halo mass. Coloured error bars show the median value in overlapping mass bins with its 1σ uncertainty determined from 100 000 bootstraps. The dotted red line indicates the same median trend with stellar mass (errors are similar). The hatched and shaded grey region indicates the 1σ scatter in the grey points. The histogram on the y-axis shows the distribution of the dense $f_{\rm tidal}$ marginalized over mass. The fraction of flux in dense tidal features appears to increase by a small amount with halo mass (with large scatter), but there is some indication that it begins to fall again around $L^\star$.

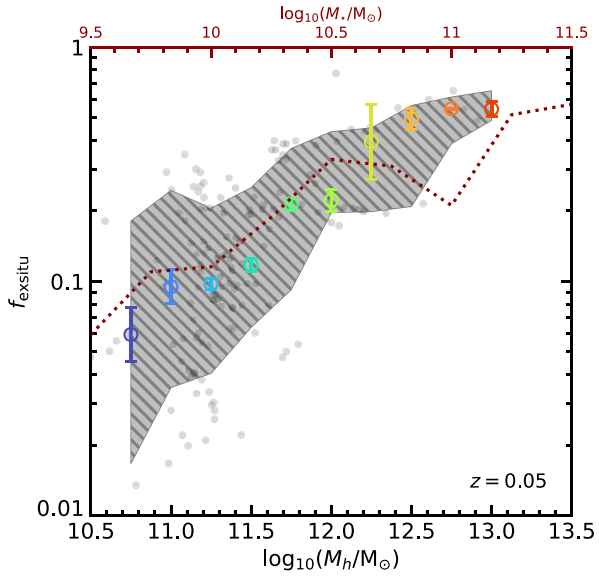


Figure 10. Scatter plot showing ex-situ mass fraction ($f_{\rm exsitu}$, equation 4) as a function of halo mass. The shaded and hatched region indicates the 1σ scatter of the grey points. Coloured error bars show the median value in overlapping halo mass bins with 1σ error bars determined from 100 000 bootstraps. The dotted red line indicates the same median trend with stellar mass (errors are similar). We observe a clear and relatively strong correlation between galaxy mass and accreted mass.

There is evidence of only a small amount of evolution in the fraction of extended light found in dense tidal features, increasing towards intermediate halo masses and then declining (although it is difficult to say definitively, given the relatively small sample of galaxies at higher masses). However, over the full mass range, dense tidal features are responsible for almost all of the flux. A lack of any significant evolution as a function of halo mass suggests that the nature of the tidal features does not change over this mass range or perhaps that the time-scales over which cohesive dense tidal features persist is not strongly affected by the host mass.[12] We will return to this topic in Section 4.1.

3.2 Tidal features and accreted mass

Fig. 10 shows the dependence of the ex-situ mass fraction on halo mass with respect to the total host stellar mass. The shaded region indicates the central 68th percentile (1σ) of the distribution of the grey points. We also show the median ex-situ mass fracton ($f_{\rm exsitu}$; equation 4) and associated errors for a number of overlapping 0.5 dex wide mass bins as coloured errorbars. Finally the red dotted line shows the same relation as a function of stellar mass (scale shown on the red top axes). As many other theoretical and observational studies (Purcell, Bullock & Zentner 2007; Oser et al. 2010; Cooper et al. 2013; Dubois et al. 2013; D'Souza et al. 2014; Lee et al. 2015; Rodriguez-Gomez et al. 2016; Harmsen et al. 2017; Pillepich et al. 2018; Spavone et al. 2018; Tacchella et al. 2019; Davison et al. 2020; Spavone et al. 2020; Martin et al. 2021) also predict, $f_{\rm exsitu}$ increases on average towards larger masses as a function of both halo and stellar mass. At lower masses ($M_h < 10^{11}$M$_\odot$) less than 10 per cent of stellar mass is formed ex-situ on average, rising to around half at the highest masses shown. We see that $f_{\rm exsitu}$ follows a similar trend with halo and stellar mass as $f_{\rm tidal}$, with both increasing towards higher masses (Fig. 8).

Finally, in Fig. 11, we plot $f_{\rm tidal}$ against $f_{\rm exsitu}$. The shaded region indicates the central 68th percentile (1σ) of the distribution of the grey points. We also show the median $f_{\rm exsitu}$ and associated errors for a number of overlapping bins in $f_{\rm tidal}$. The large overlapping error bars indicate the variation in $f_{\rm exsitu}$ over a 2 and 4 Gyr time-scale, respectively. While it is clear that $f_{\rm exsitu}$ correlates with $f_{\rm tidal}$ and halo mass, there appears to be a similar variance as a function of both variables.

In order to investigate whether accretion history ($f_{\rm exsitu}$) has a measurable influence in the strength of tidal features beyond the existing correlation of $f_{\rm exsitu}$ with halo mass, we calculate the partial distance correlation coefficient (Székely et al. 2014) between $f_{\rm tidal}$ and $f_{\rm exsitu}$, controlling for the halo mass, $\Re^*(\log_{10} f_{\rm tidal}, \log_{10} f_{\rm exsitu}; \log_{10} M_{\rm h})$, and between the halo mass and $f_{\rm exsitu}$, controlling for $f_{\rm tidal}$, $\Re^*(\log_{10} M_{\rm h}, \log_{10} f_{\rm exsitu}; \log_{10} f_{\rm tidal})$. The partial correlation coefficients and associated 1σ uncertainties are 0.28 ± 0.04 and 0.16 ± 0.04, respectively (with full correlation coefficients of 0.61 and 0.54, respectively), indicating a stronger association between $f_{\rm exsitu}$ and $f_{\rm tidal}$. This suggests that the tidal mass fraction may be a better predictor of $f_{\rm exsitu}$. The statistical significance

[12]Note that Pillepich et al. (2018) show the ratio of ICL to all diffuse mass increases in more massive haloes and our own examination of mock images of massive central galaxies in the Horizon-AGN simulation (Dubois et al. 2014) shows a relative dearth of distinct tidal features compared with lower mass galaxies in NewHorizon. However it is difficult to make any definitive statement due to the much more limited resolution of Illustris TNG300 and Horizon-AGN.

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

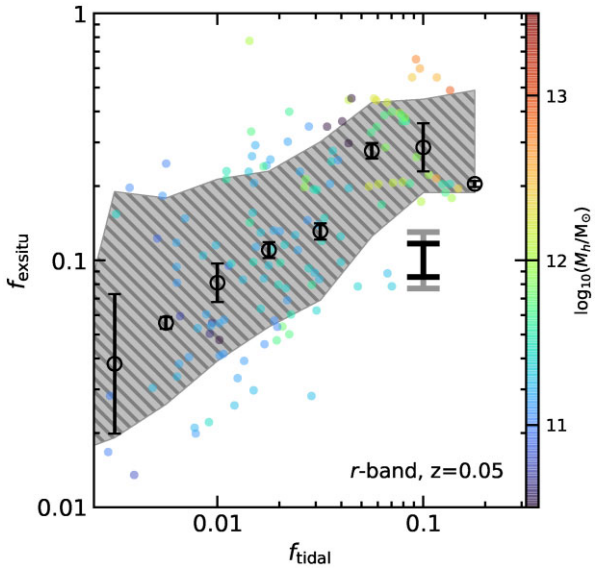


Figure 11. Scatter plot showing ex-situ mass fraction ($f_{\rm exsitu}$, equation 4) as a function of tidal flux fraction ($f_{\rm tidal}$, equation 5) in the r-band colour coded by halo mass. Black open circles with error bars indicate the median value in overlapping bins, with 1σ error bars determined from 100 000 bootstraps. The large overlapping black and grey error bars to the right indicate the mean fractional variation in ex-situ mass fraction over timescales of 2 and 4 Gyr (fractional errors of ∼30 per cent and ∼50 per cent, respectively). We observe a similarly strong correlation between $f_{\rm exsitu}$ and $f_{\rm tidal}$ as we do with $f_{\rm exsitu}$ and galaxy mass. As the coloured points indicate, both $f_{\rm exsitu}$ and $f_{\rm tidal}$ are correlated with mass.

of this result (that there is a stronger correlation between $f_{\rm exsitu}$ and $f_{\rm tidal}$ than $f_{\rm exsitu}$ and halo mass) stands at 2.2σ, giving a relatively weak indication that different accretion histories have a measurable impact on the amount of flux in the stellar halo. This result appears to be consistent with the idea that much of the tidal flux is contributed by recent mergers with a small number of relatively massive progenitors (e.g. Bullock & Johnston 2005; Purcell et al. 2007; Cooper et al. 2010). However, the total fraction of accreted mass alone is a relatively weak predictor of the final structure of the stellar halo. For example Rey & Starkenburg (2021) show that, even in galaxies with the same ex-situ mass fraction, the shape of the stellar halo is strongly sensitive to accretion history.

3.3 Tidal feature detection

We also consider the fraction of tidal features that we expect to be detected at different limiting surface brightnesses. Because it may, in some cases, be possible to detect contiguous structures by eye, even if they are fainter than the surface brightness limit, we adopt a definition for detected structures based on the connections between pixels that are 1σ above the noise level in images produced from particles that are part of dense tidal features only (so that contribution from well phase-mixed material is first removed i.e. Fig. 5, panel c). We describe the procedure and present an example of the procedure performed on the same object for a range of limiting surface brightnesses in Appendix E. Qualitatively, our results do not change if we only consider pixels that are brighter than a given limiting surface brightness.

We define the total flux in detected structures within the original image (before noise is added) as a fraction of the total flux in pixels that are brighter than 35 mag arcsec^{-2} and which are within a radius

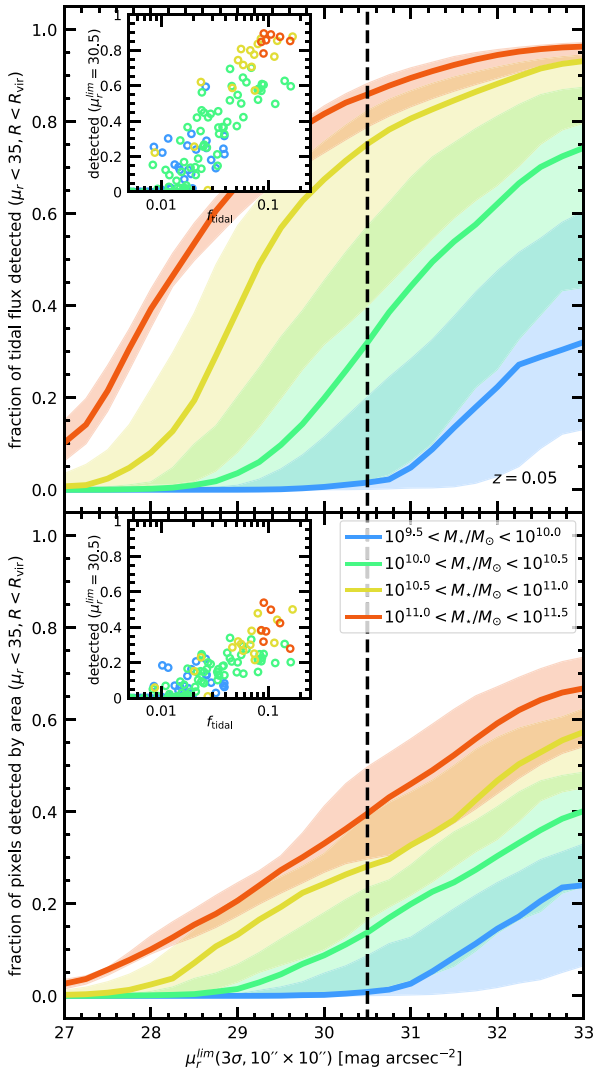


Figure 12. Top: the fraction of tidal flux found in detected structures as a function of r-band limiting surface brightness. Coloured lines indicate the median fraction of flux in detected structures for different stellar mass bins, and coloured regions indicate the central 68th percentile (1σ). The inset plot indicates the fraction of tidal flux detected per galaxy as a function of their tidal flux fraction for a limiting surface brightness of 30.5 mag arcsec^{-2}, and the colour of each point indicates which mass the galaxy is in. Bottom: the fraction of pixels detected as a function of r-band limiting surface brightness. The coloured lines indicate the median fraction of pixels detected in different stellar mass bins, and coloured regions indicate the central 68th percentile (1σ) of the distribution. The inset plot indicates the fraction of pixels detected per galaxy as a function of their tidal flux fraction for a limiting surface brightness of 30.5 mag arcsec^{-2}. The colour of each point indicates which mass bin the galaxy is in. An interactive version of the inset plots showing multiple surface brightness limits can be found at garrethmartin.github.io/files/frac_recovered.html.

of 1 $R_{\rm vir}$. We define the detected area similarly by counting the total area of the detected mask as a fraction of the total area of pixels that are brighter than 35 mag arcsec^{-2} and which are within a radius of 1 $R_{\rm vir}$.

In Fig. 12, we show the fraction of flux (top) and the fraction of area (bottom) that is detected in different mass bins as a function of r-band limiting surface brightness (3σ, 10 arcsec × 10 arcsec) at $z = 0.05$. We consider only pixels with a surface brightness brighter

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

than 35 mag arcsec^{-2} and which are within 1 $R_{\rm vir}$ of the centre of the galaxy. More massive galaxies tend to have tidal features that are more easily detectable. For example, based on a predicted 10-yr depth of $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) \approx 30.5$ mag arcsec^{-2} (Laine et al. 2018), we can expect to detect a non-negligible fraction ($\gtrsim 25$ per cent) of the total area and most of the flux ($\gtrsim 80$ per cent) that makes up the dense tidal features of galaxies more massive than the Milky Way ($M_\star > 10^{10.5}$ M$_\odot$). Even assuming a shallower final depth of 29.5 mag arcsec^{-2}, we still expect to detect more than 60 per cent of flux in tidal features for the same mass range.

A significant fraction of the mass associated with tidal features is detected at lower masses. The inset plots, which show the fraction of detected flux and area as a function of $f_{\rm tidal}$, indicate that this is largely a result of the fact that more massive galaxies tend to have stronger tidal features on average. As Fig. 8 shows, there is a significant spread in $f_{\rm tidal}$ such that lower mass galaxies can exhibit fairly strong tidal features. An interactive online supplement to this plot (garrethmartin.github.io/files/frac_recovered.html) shows a version of the two inset plots for different surface brightness limits. We see that, for very weak tidal features (low values of $f_{\rm tidal}$), a majority of flux is still not detected even for very high limiting surface brightnesses (e.g. ∼30 per cent of flux detected for $f_{\rm tidal} = 0.01$ for a limiting surface brightness of 33 mag arcsec^{-2}). Therefore, in the nearby Universe ($z < 0.05$), we can expect the LSST (or any similarly deep survey) to find a significant fraction of the mass associated with tidal features around intermediate and high mass galaxies ($M_\star > 10^{10.5}$M$_\odot$). Most tidal features found around galaxies in the low mass regime, however, are likely to remain inaccessible at least in the near future. As we highlight later in Section 4.2, this means that more massive galaxies exhibit tidal features that are both more frequent and stronger.

Finally, we explore how our ability to detect tidal features around galaxies declines with redshift. Fig. 13 shows the same detected flux and detected area fractions as a function of the redshift at which each object is observed for approximately Milky Way mass galaxies and for a range of limiting surface brightnesses, which are indicated in the legend. Towards higher redshifts, the fraction of detected flux declines quite sharply such that we do not expect to be able to detect tidal features around any Milky Way mass galaxy after a redshift of $z = 0.4$, even at 10-yr LSST depth.

An interactive online supplement to this plot (garrethmartin.github.io/files/completeness.html) shows how the detected flux and detected area fraction evolve with redshift for different stellar mass bins and limiting surface brightnesses.

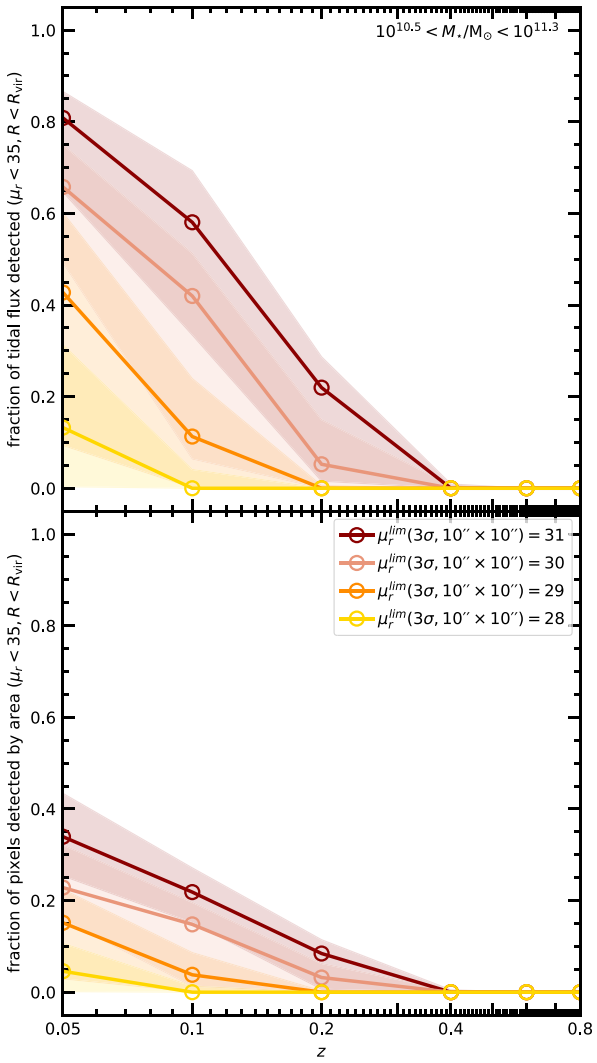


Figure 13. Top the fraction of tidal flux found in detected structures as a function of redshift in ∼$M^\star$ galaxies for different r-band limiting surface brightnesses. The coloured lines indicate the median fraction of flux in detected structures for a given redshift and coloured regions indicate the central 68th percentile (1σ) of the distribution. Bottom: the fraction of pixels detected as a function of redshift for different r-band limiting surface brightness. The coloured lines indicate the medial fraction of pixels detected in different stellar mass bins and coloured regions indicate the central 68th percentile (1σ) of the distribution. An interactive version of this plot showing multiple limiting surface brightnesses and stellar mass bins can be found at garrethmartin.github.io/files/completeness.html.

4 CHARACTERIZATION OF TIDAL FEATURES BY HUMAN CLASSIFICATION

4.1 Visual classification of mock images

As discussed in Section 2.6, all images were classified independently by at least two people and a set of 600 images were independently classified by five people. Each of our classifiers has experience in the field of LSB science and therefore has a good level of domain knowledge and many are also experts in the morphological classification of galaxies.

4.1.1 Collated feedback

Classifiers were also asked to provide feedback commenting on their experience classifying the mock images. Classifiers were asked to evaluate their confidence in the reliability and reproducibility of their classifications, whether there were any particular categories that they found difficult to classify and whether they employed a particular strategy or methodology when performing the classifications. We list our conclusions based on classifier responses below:

(i) Some classifiers were more conservative in their classifications than others, particularly in classifying mergers or fainter tidal features. Classifiers made what they felt was the most reasonable interpretations, but most felt that these choices were subjective and therefore liable to change between classifiers.

(ii) Some classifiers were more confident in the reproducibility of their classification than others. Broadly, classifiers felt

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

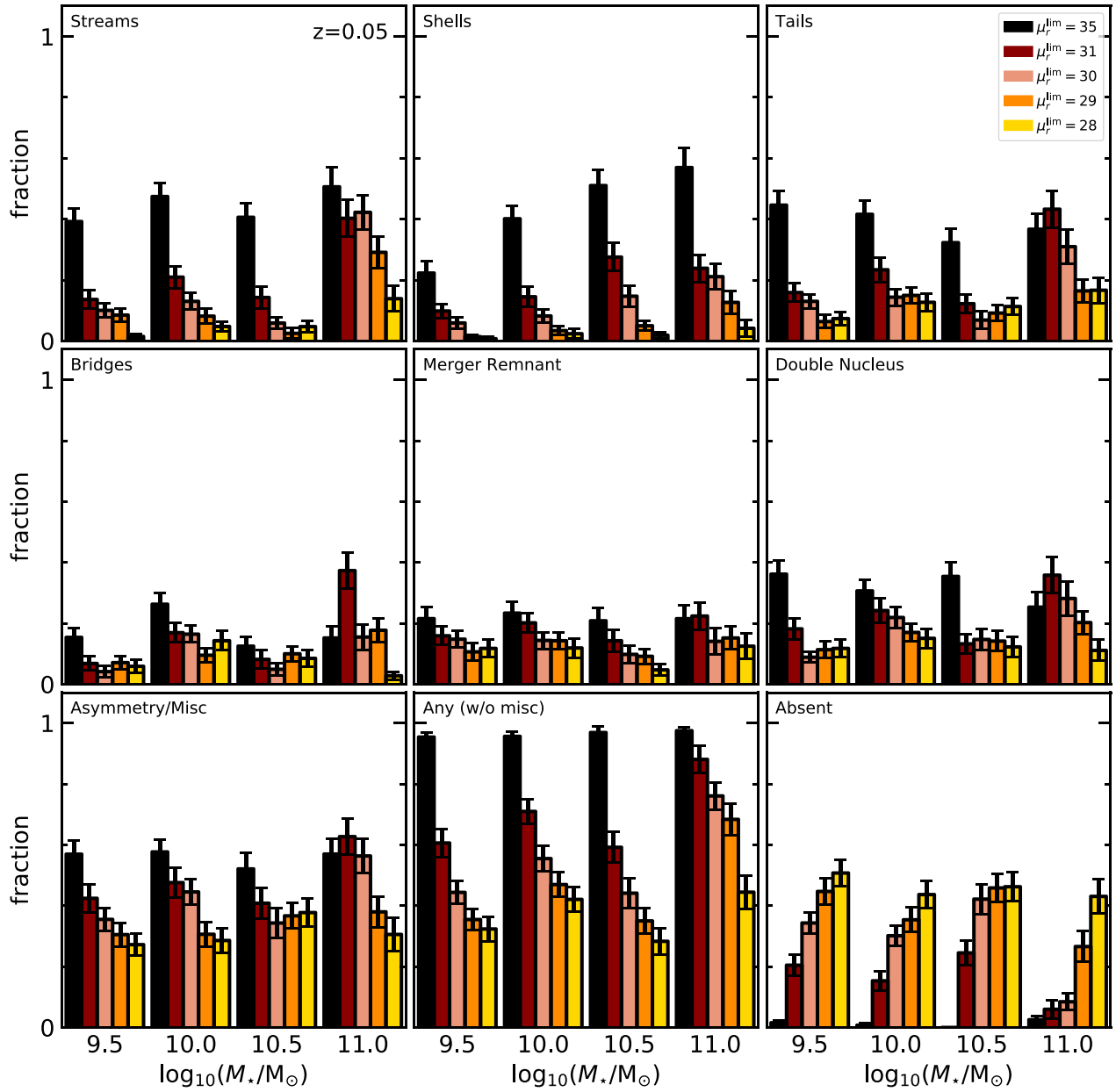


Figure 14. Histograms indicating the prevalence of different classes of tidal feature as a function of stellar mass and limiting surface brightness. The bars indicate the fraction of galaxies that exhibit at least one instance of a given class of tidal feature in each mass bin. The four mass bins have a width of 0.5 dex and run from 9.25 $M_\odot$ to 11.25 $M_\odot$, and errorbars for each bin are determined by bootstrap. At the highest limiting surface brightness, almost all galaxies exhibit coherent tidal features regardless of their mass. We observe significantly different behaviour in the prevalence of tidal features as a function of limiting surface brightness and stellar mass across different classes.

that they were able to identify the presence of tidal features very reliably, but felt that, in some cases, detailed characterization and distinguishing between similar categories of tidal feature was difficult (e.g. differentiating tidal tails, stellar streams, plumes, and shells from one another) and that precisely determining the frequency of these features in each image was also difficult.

(iii) Classifiers generally felt that their classifications became more consistent as they classified more objects. In some cases, where classifiers went back through and repeated their classifications, they ended up revising some of their original classification and particularly the earliest images that they classified.

(iv) In the case of faint or poorly resolved tidal features, it was often difficult to place these features into a specific category. Classifiers generally placed such features into miscellaneous category in these cases.

(v) In the absence of 3D kinematic information, ambiguities arise in the classification of certain features. Excluding tidal features within the field of view that were not associated with any interaction with the host galaxy added a degree of complexity to the classification as making this determination could be difficult, for example in the case where there is significant overlap between two objects but not clear tidal disturbance. For the deepest / highest spatial resolution images, classifications tended to become more difficult

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

as the complexity of the morphology and environment of the tidal features became more apparent.

(vi) Counting certain categories of tidal feature, like the number of shells, plumes, or asymmetric features was not always simple. In a single halo, it is much easier to identify multiple distinct plumes than in a group environment, where it becomes much more difficult due to multiple overlapping features.

(vii) For more distant (poorly resolved) objects there was a feeling among classifiers that they may have been susceptible to seeing asymmetries in the stellar halo that were not present.

(viii) At low spatial resolution, distinguishing merging systems from close pairs is challenging due to a lack of resolved tidal features. In general, classifiers found the concept of double nuclei systems to be quite uncertain as there was a large degree of subjectivity in determining if two nuclei are close, share a common envelope, are just the result of projection, etc, which becomes more difficult for more distant objects.

Despite the high level of expertise of our classifiers, many of them still found the exercise challenging. Some of the difficulties raised by the classifiers could be alleviated by designing the study differently, but there are also factors that are more difficult or impossible to address. For example it is doubtful whether it is possible to produce entirely consistent classifications between classifiers due to the subjective nature of many of the decisions that classifiers are required to make. Additionally, classifiers likely have differing notions of what exactly constitutes a given class of tidal feature, a lack of any standard definitions likely compounds this in addition to making any comparisons between classifiers or between studies necessarily qualitative. Since it is impossible to completely standardize the classification process, this implies that there will always be differences between human classifiers.

4.1.2 *Census of tidal features by class*

In this section, we consider the frequency at which different classes of tidal feature were identified by our human classifiers. Fig. 14 shows the fraction of galaxies at $z = 0.05$ in which different classes of tidal feature were identified for different stellar masses and limiting surface brightnesses.

We see similar behaviour in the prevalence of shells as seen in observational studies like Bílek et al. (2020),[13] and the decline in the prevalence of shells with stellar mass remains across different limiting surface brightness (with differing normalization). Interestingly, while we see similar behaviour for tails and streams at lower limiting surface brightnesses, at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2}, there is little difference in their prevalence across mass bins, indicating that while these features are present at similar levels across the mass range we consider, they are typically fainter and more difficult to detect in lower mass galaxies.

In the highest mass bin, streams and tails are both detected at similar rates regardless of limiting surface brightness for $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) \geq 30$ mag arcsec^{-2} indicating that the majority of streams and tidal tails are at least this bright in these more massive galaxies. Meanwhile, merger remnants and double nuclei are quite reliably identified regardless of limiting surface brightness and occur at relatively similar levels across stellar mass.

[13]It should be noted that the sample of Bílek et al. (2020) is more strongly dominated by elliptical galaxies and that we have not attempted to replicate their definitions for different tidal feature classes.

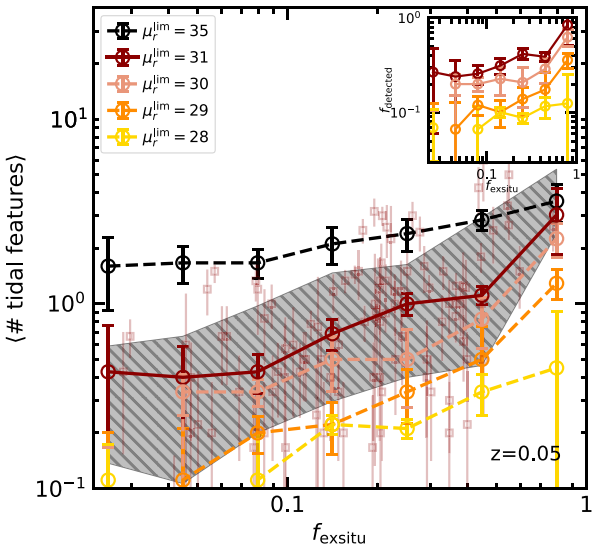


Figure 15. The average number of tidal features identified by classifiers as a function of galaxy ex-situ mass fraction. The solid red line indicates the average number of distinct tidal features identified per galaxy (excluding the merger, double nuclei, and asymmetry / misc category) at $z = 0.05$ for $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 31$ mag arcsec^{-2} as a function of $f_{\rm exsitu}$, with error bars determined by bootstrap. The coloured dashed lines show the same for different limiting surface brightnesses indicated in the legend. The light red squares with errorbars indicate the number of tidal features counted in individual galaxies for $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 31$ mag arcsec^{-2} with errors determined by the standard deviation across multiple classifiers and different projections. The hatched region indicates the central 68th percentile (1σ) spread for these points. The inset plot shows instead the number of tidal features recovered as a fraction of the average number of tidal features identified for $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2} in each $f_{\rm exsitu}$ bin. We observe fairly weak evolution and large scatter in the number of tidal features identified with $f_{\rm exsitu}$, especially at higher limiting surface brightness.

In all, close to 100 per cent of galaxies exhibit some kind of distinct tidal feature (i.e. not just asymmetries) at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2} regardless of mass. However, this number falls fairly significantly as more realistic limiting surface brightnesses are considered. This result is in broad agreement with Vera-Casanova et al. (2021), who show that ∼90 per cent of their Aurgia models show clear LSB features at 31 mag arcsec^{-2} for a sample of host galaxies with an average stellar mass of $10^{10.8}$ M$_\odot$ (compare with the most massive bin of the 'Any' panel of Fig. 14).

4.2 Visual biases

4.2.1 *Effect of redshift and limiting surface brightness*

In this section, we study how the number of tidal features identified changes as a function of limiting surface brightness and redshift. Fig. 15 shows the average number of distinct tidal features identified per galaxy (excluding mergers, double nuclei, and miscellaneous asymmetries) at $z = 0.05$ as a function of galaxy ex-situ mass for overlapping logarithmic bins of width 0.4 dex. Different coloured lines indicate our results based on mock images with different limiting surface brightnesses. We find a similar result to that shown in Fig. 11 – as $f_{\rm exsitu}$ increases the average number of tidal features identified also tends to increase. The red error bars

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

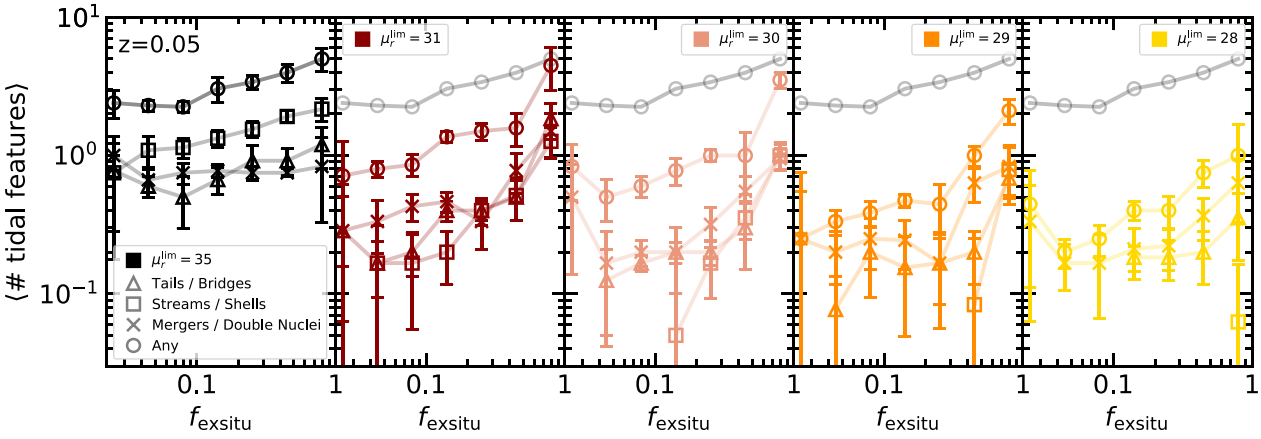


Figure 16. The average number of tidal features identified by classifiers as a function of galaxy ex-situ mass fraction at $z = 0.05$ split by limiting surface brightness and tidal feature class. Tidal tails and bridges are indicated by triangular markers, tidal streams, and shells by squares, mergers, or double nuclei by crosses, and all tidal features (not including the miscellaneous category) by circles. Different panels correspond to different limiting surface brightnesses with a line for the 'Any' category at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2} plotted in grey for comparison in each panel. Error bars show 1σ uncertainties obtained from bootstrapping.

indicate the number of tidal features detected in individual galaxies (for $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 31$ mag arcsec^{-2}), with errors derived from the standard deviation in three different projections with each projection classified independently by at least two classifiers. There exists a fairly wide spread, which can be seen in the hatched region indicating the 1σ dispersion of the points.

If we use the $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2} line as a proxy for the true number of tidal features, we see that the average galaxy has at least one identified tidal feature regardless of the ex-situ mass fraction and that the average number of tidal features identified increases only modestly with $f_{\rm exsitu}$, although the trend does appear to strengthen for large $f_{\rm exsitu}$. Considering the large variation over time of the individual galaxy merger histories seen in Fig. D2, this fairly weak correspondence is perhaps not surprising. It is also true that, while there is only a weak dependence in the average number of tidal features with redshift, tidal features in galaxies with low $f_{\rm exsitu}$ are typically weaker so that, at brighter limiting surface brightnesses, the trend strengthens.

The inset plot shows the number of tidal features identified as a fraction of the average number identified at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2} in each bin. At all limiting surface brightnesses shown, a greater fraction of tidal features are identified for galaxies with higher ex-situ mass fractions. Considering that high mass or high $f_{\rm exsitu}$ galaxies tend to exhibit stronger tidal features (e.g. Fig. 12), we can expect a greater fraction of tidal features to be bright enough to be detected at brighter limiting surface brightnesses. It is worth acknowledging that this reflects a possible observational bias – that the tidal features present in lower mass haloes have fewer tidal features, but these tidal features are also likely to be weaker – so that they are more likely to go undetected.

Fig. 16 again shows the average number of distinct tidal features identified per galaxy as a function of ex-situ stellar mass. Tidal features are broken down into three categories: tidal tails and bridges, streams and shells, and merger remnants or double nuclei denoted by open triangles, squares, and crosses, respectively. We observe markedly different behaviour in the trends across $f_{\rm exsitu}$ and limiting surface brightness for different classes of tidal feature. Streams and shells are the best tracer of ex-situ mass, with tails, bridges, mergers, and double nuclei occurring with roughly constant frequency across the range of ex-situ masses shown.

Being more numerous and longer lasting than other classes of tidal feature (Greco et al. 2018b), shells and streams are expected to better sample the average accretion history of the galaxy. However, at the limiting surface brightness achievable by the Rubin Observatory, streams and shells are the least frequently identified class of tidal feature (e.g. Fig. 14). Since prevalence of shells and streams appears to decline strongly towards fainter limiting surface brightnesses ($\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) \leq 31$ mag arcsec^{-2}), especially compared with other classes of tidal feature such as tidal tails, this indicates a decline in tidal feature strength at smaller ex-situ masses, rather than being a direct tracer of the frequency of accretion events. This is a natural consequence of the fact that features such as streams and shells are typically formed in more unequal mass ratio mergers with lower mass satellites.

Below $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 31$ mag arcsec^{-2}, there is almost no evolution in the normalization of the trend with $f_{\rm exsitu}$ for both the tails/bridges and mergers/double nuclei classes, with a stronger trend emerging at high $f_{\rm exsitu}$. Since both the frequency and strength of tidal features changes with $f_{\rm exsitu}$ it is difficult to disentangle the effect, but the shallower relation seen for the faintest limiting surface brightness suggests that it is driven by these tidal features being generally fainter.

Fig. 17 shows the average number of distinct tidal features identified per galaxy (excluding mergers, double nuclei, and miscellaneous asymmetries) as a function of redshift. Different coloured lines correspond to different limiting surface brightnesses and the black dashed line indicates the average number of tidal features identified for $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2} at $z = 0.05$. The average number of tidal features identified per galaxy falls rapidly with redshift so that around ten times fewer tidal features are identified when the same galaxies are viewed at $z = 0.8$ compared with $z = 0.05$. This decline is principally a consequence of cosmological dimming and the PSF or pixel scale blurring features as apparent size of objects decreases. At fainter limiting surface brightnesses, enough light could be scattered from the central galaxy to its extended tidal features to obscure them, especially as they move further into the core of the PSF towards higher redshifts While this is likely not a concern at limiting surface brightnesses achievable by the Rubin Observatory, the effect can become more important in deeper imaging (see Appendix B for further details).

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

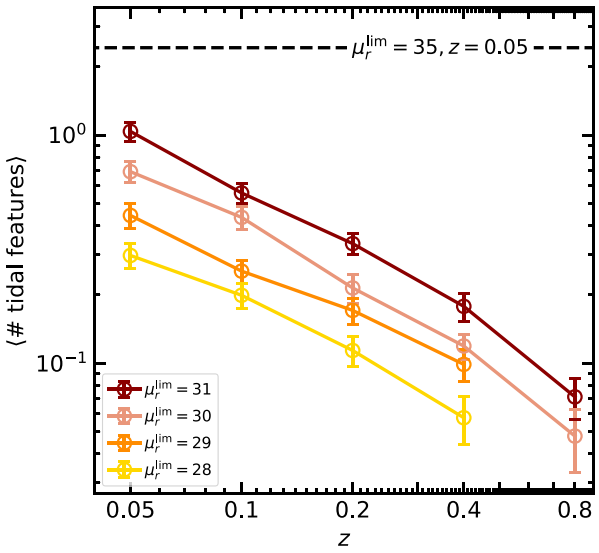


Figure 17. The average number of tidal features identified by classifiers as a function of redshift. The solid coloured lines indicate the average number of distinct tidal features identified per galaxy (excluding the asymmetry / misc category) as a function of redshift and for the different limiting surface brightnesses indicated in the legend. The dashed black line indicates the average number of tidal features identified for $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2} at $z = 0.05$. Lines do not extend to $z = 0.8$ for the two brightest limiting surface brightnesses because low signal-to-noise makes visual classification too difficult for a majority of objects.

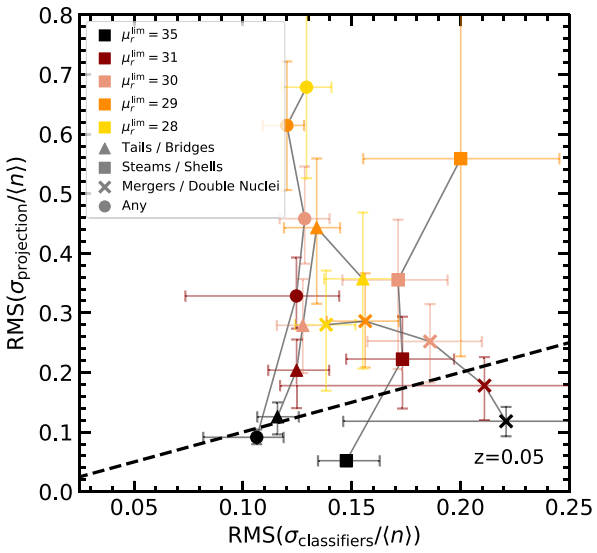


Figure 18. Evolution of the RMS fractional $\sigma_{\rm projection}$ and $\sigma_{\rm classifiers}$ and as a function of limiting surface brightness at $z = 0.05$. Tidal tails and bridges are indicated by triangular markers, tidal streams and shells by squares, mergers or double nuclei by crosses and all tidal features (not including the miscellaneous category) by circles. Colours show the limiting surface brightness as indicated in the legend and error bars show 1σ uncertainties obtained from bootstrapping. No marker is plotted for shells at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 28$ mag arcsec^{-2} because almost no objects display shells. While fainter limiting surface brightnesses typically improve the accuracy of classifications, the opposite is true in the case of the merger/double nuclei category.

4.2.2 *Effect of projection*

In this section, we consider the difference in classifications made in three different projections 90 degrees apart (*xy*, *xz*, *yz*) comparing the scatter in the number of features identified against the scatter between individual classifiers. We measure the standard deviation of the number of features identified for each galaxy in two ways: In the first case, we measure the standard deviation across classifiers, $\sigma_{\rm classifiers}$, treating each projection of the same object independently. In the second case, we measure the standard deviation across projections, $\sigma_{\rm projection}$, using the mean number of classifications across all classifiers for each projection. Any objects where no features of a given class were identified by any of the classifiers are not considered. We obtain the fractional standard deviation by dividing by the average number of features in each class and then take the RMS of this value over all galaxies.

Fig. 18 shows the evolution of these quantities as a function of limiting surface brightness and for different classes of tidal feature at $z = 0.05$. Tidal tails and bridges are indicated by triangular markers, tidal streams and shells by squares, mergers or double nuclei by crosses, and all tidal features (not including the miscellaneous category) by circles. Colours indicate the limiting surface brightness and error bars show 1σ uncertainties obtained from bootstrapping. At the brightest limiting surface brightness, the typical scatter in classifications is larger for $\sigma_{\rm classifiers}$, but towards fainter limiting surface brightnesses, $\sigma_{\rm projection}$ quickly becomes larger, while $\sigma_{\rm classifiers}$ does not change very significantly. In other words, at sufficient depth, disagreement between classifiers arising purely from subjective disagreement on the classification of identical images is the dominant source of uncertainty, but uncertainties arising from projection begin to dominate in shallower imaging. At expected LSST surface brightness limits (30–31 mag arcsec^{-2}), the uncertainty arising from projection is dominant for all but the mergers / double nuclei categories.

Most robust to changing the limiting surface brightness is the mergers and double nuclei category, which only sees a modest increase in uncertainty towards brighter limiting surface brightness. This is likely because features associated with merger remnants and double nuclei are typically bright so are robustly detected regardless of image depth. Interestingly, the variability of $\sigma_{\rm classifiers}$ for mergers and double nuclei is the largest of any category (while still quite modest) with an opposite trend to any of the other classes of tidal feature. This increase in uncertainty with fainter limiting surface brightness is also corroborated in the feedback given by classifiers, who mentioned that very deep imaging made certain features more difficult to classify. The visual appearance of lower surface brightness features, which tend to be significantly more extended, can change significantly through different projections leading to significant variation. In the case of higher surface brightness features, which are already clear at relatively bright limiting surface brightness (and less variable with projection), deeper imaging acts only to increase the prevalence of confounding sources which leads to an increase in the variance between classifiers at fainter limiting surface brightness.

Based on both qualitative evidence from classifiers themselves and our quantitative analysis, we see that there are inherent uncertainties which cannot be completely removed. While there is typically improvement with deeper imaging, in some circumstances, classifications actually become less robust. Of course, a solution to this problem, which would be effective at least up to the surface brightness of the features themselves, would be to apply brighter surface brightness cuts than the actual limiting surface brightness

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

of the data. However, if we note the difference in normalization between the $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2} and 31 mag arcsec^{-2} panels of Fig. 16, it is evident that some merging systems / double nuclei are only revealed at very faint limiting surface brightnesses meaning some fraction of systems would be missed in this case. Also note that we do not model foreground and background objects or other astrophysical contaminants in our mock images, the inclusion of which would likely further reduce the agreement between classifiers.

5 SUMMARY

In this paper, we have performed a comprehensive theoretical investigation of the extended diffuse light around galaxies and galaxy groups down to low stellar mass densities and explored the reliability of human classifications under different observational biases. Our sample consists of 37 unique objects from the NEWHORIZON simulation whose progenitors we select at $z = 0.2$, $z = 0.4$, 0.6, and 0.8 giving a total of 148 objects across four different redshifts with stellar masses $10^{9.5} < M_\star/{\rm M}_\odot < 10^{11.5}$. Our main findings based on automated techniques and human visual classification are as follows:

(i) Distribution of tidal flux:

(1) A large fraction of tidal flux is expected to be detectable at LSST 10-yr depth. Assuming the LSST pipeline is suitably optimized, 50 per cent of the total flux from substructure identifiable as distinct tidal features is detectable with a limiting surface brightness of $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 30.5$ mag arcsec^{-2}. 90 per cent of the pixels that make up tidal features are brighter than $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 32$ mag arcsec^{-2} by area. However, almost all of the more diffuse light around galaxies (which makes up around 25 per cent of the total light in tidal features) will remain undetectable at a limiting surface brightness of $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) =$ 30−31 mag arcsec^{-2} outside of very coarse binning.

(2) Much of the tidal flux in galaxies is found at large radii. While 50 per cent of tidal flux is contained within 7 $R_{\rm eff}$ of the galaxy centre on average, close to 100 per cent of flux is only reached by 25 $R_{\rm eff}$ or ∼0.6 $R_{\rm vir}$.

(3) The amount of tidal flux detected is strongly dependent on limiting surface brightness. At brighter limiting surface brightnesses the normalization of the relation between mass and tidal flux decreases so that the average $f_{\rm tidal}$ for an MW mass galaxy ($M_\star \approx 10^{10.5}$ M$_\odot$) decreases from ∼5 per cent at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 31$ mag arcsec^{-2} to only a fraction of a per cent at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) =$ 28 mag arcsec^{-2}, while the scatter increases from 0.4 dex at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 31$ mag arcsec^{-2} to 1.3 dex at $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 28$ mag arcsec^{-2}.

(4) At predicted LSST limiting 10-yr depth, a majority (75 per cent) of tidal flux in MW mass galaxies ($M_\star > 10^{10.5}$ M$_\odot$) is detectable at $z = 0.05$. At low masses ($M_\star < 10^{10}$ M$_\odot$), almost no galaxies are expected to exhibit visible tidal features. This is driven by the fact that tidal features are less frequent, but also generally weaker in lower mass galaxies. Even if a shallower final depth of 29.5 mag arcsec^{-2} is assumed, we still expect the Rubin Observatory to detect more than 60 per cent of flux in tidal features for galaxies of MW mass or greater.

(5) Similarly, tidal features become significantly more difficult to detect at higher redshifts so that we would not expect to routinely identify any tidal features around MW mass galaxies beyond$z = 0.2$. While cosmological dimming is the primary driver, smearing of tidal features as they move into the core of the PSF may also play a role, particularly if very faint limiting surface brightnesses are considered. In this case, diffraction-limited, space-based observatories such as Roman (Robertson et al. 2019) and Euclid (Borlaff et al. 2022) offer an important complement to the Rubin Observatory.

(6) ex-situ mass fraction correlates with galaxy mass and tidal flux fraction. Partial correlation coefficients indicate a more favourable correlation with ex-situ mass (at a significance 2.2σ) giving some indication that accretion history drives the tidal flux fraction beyond the simple correlation with mass. We observe a break in the relation between tidal flux and stellar mass at $10^{10.1\pm^{0.01}_{0.05}}$ M$_\odot$, corresponding to the crossover mass at which mergers are thought to become the dominant process driving galaxy evolution.

(ii) Reliability of human classification:

(1) Galaxies in the NEWHORIZON simulation exhibit a range of analogues to observed tidal features. While data at sufficient depth are relatively scant and exact comparisons are difficult, the NEWHORIZON simulation produces tidal features whose frequencies evolve with stellar mass in a way that is comparable to trends seen in available observational data.

(2) At very faint limiting surface brightnesses ($\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2}), expert classifiers were able to identify specific tidal features in close to 100 per cent of galaxies ($M_\star > 10^{9.5}$ M$_\odot$). Certain features, like merger remnants were identified at roughly the same frequency regardless of limiting surface brightness, while the detection of shells was found to be much more sensitive to image depth.

(3) A greater number of tidal features were identified in galaxies with high ex-situ mass fractions. When compared with the number of tidal features identified for a limiting surface brightness of $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 35$ mag arcsec^{-2}, a greater fraction of tidal features are detected in galaxies with higher ex-situ mass fractions due to the fact that they are typically also brighter. This reflects a possible observational bias since the tidal features present in galaxies with a smaller number of tidal features are also likely to be weaker and are, therefore, more likely to go undetected.

(4) Concurrence between classifiers generally improves with deeper imaging but morphologies can become more complex, introducing uncertainty in precise characterization. In particular, classifiers were less likely to concur with each other the presence of a merger remnant and double nuclei when viewing deeper images.

(5) Concurrence between classifiers is quite robust to different limiting surface brightnesses, but brighter limiting surface brightnesses produce much weaker agreement when classifications over different projections of the same object are compared. Typically different projections of the same object produce a larger scatter in classifications than the scatter between different classifiers viewing the same object in the same orientation.

Our findings, which are based on realistic Rubin Observatory mock images at the final LSST survey depth (30–31 mag arcsec^{-2}), indicate that the Rubin Observatory will be well situated to provide high quality observations of the tidal features surrounding galaxies. We expect the Rubin Observatory to open up a new region of discovery space by delivering sufficiently deep imaging down to intermediate

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

redshifts ($z < 0.2$) and stellar masses ($M_\star > 10^{10}\,\mathrm{M}_\odot$) to study these structures in detail.

ACKNOWLEDGEMENTS

GM thanks Peter Yoachim for sharing theoretical results for LSST surface brightness metrics and Yohan Dubois for fruitful discussion. JAB acknowledge financial support from CONICET through PIP 11220170100527CO grant. JLC acknowledges support from National Science Foundation (NSF) grant AST-1816196. CAC acknowledges support from the Science and Technology Research Council under grant ST/S006095/1 and LJMU. FAG acknowledges financial support from FONDECYT Regular 1211370 and from the Max Planck Society through a Partner Group grant. GG gratefully acknowledges support by the ANID BASAL projects ACE210002 and FB210003. RAJ acknowledges support from the Yonsei University Research Fund (Yonsei Frontier Lab. Young Researcher Supporting Program) of 2021 and from the Korean National Research Foundation (NRF-2020R1A2C3003769). SK acknowledges support from the STFC [ST/S00615X/1] and a Senior Research Fellowship from Worcester College Oxford. JHK acknowledges financial support from the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No 721463 to the SUNDIAL ITN network, from the State Research Agency (AEI-MCINN) of the Spanish Ministry of Science and Innovation under the grant 'The structure and evolution of galaxies and their central regions' with reference PID2019-105602GB-I00/10.13039/501100011033, and from IAC project P/300724, financed by the Ministry of Science and Innovation, through the State Budget and by the Canary Islands Department of Economy, Knowledge and Employment, through the Regional Budget of the Autonomous Community. DJP acknowledges funding from an Australian Research Council Discovery Program grant DP190102448. JR acknowledges support from the State Research Agency (AEI-MCINN) of the Spanish Ministry of Science and Innovation under the grant 'The structure and evolution of galaxies and their central regions' with reference PID2019-105602GB-I00/10.13039/501100011033. EAS thanks the LSSTC Data Science Fellowship Program, which is funded by LSSTC, NSF Cybertraining Grant #1829740, the Brinson Foundation, and the Moore Foundation. AEW acknowledges support from the STFC [ST/S00615X/1]. MB acknowledges support from the Science and Technology Facilities Council through grant number ST/N021702/1. FB acknowledges support grants PID2020-116188GA-I00 and PID2019-107427GB-C32 from The Spanish Ministry of Science and Innovation. FD acknowledges support from the STFC [ST/V506709/1]. RD gratefully acknowledges support by the ANID BASAL projects ACE210002 and FB210003.

This work has made use of the Horizon cluster on which the simulation was post-processed, hosted by the Institut d'Astrophysique de Paris. We warmly thank S. Rouberol for running it smoothly.

The authors thank the referee, Andrew Cooper, for a detailed and constructive report which helped improve the final paper.

DATA AVAILABILITY

The simulation data analysed in this paper were provided by the NEWHORIZON collaboration. The data will be shared on request to the corresponding author, with the permission of the NEWHORIZON collaboration or may be requested from https://new.horizon-simulation.org/data.html.

REFERENCES

Abadi M. G., Navarro J. F., Steinmetz M., 2006, MNRAS, 365, 747
Abraham R. G., van Dokkum P. G., 2014, PASP, 126, 55
Agertz O., Teyssier R., Moore B., 2011, MNRAS, 410, 1391
Aihara H. et al., 2018, PASJ, 70, S4
Aihara H. et al., 2019, PASJ, 71, 114
Akhlaghi M., Ichikawa T., 2015, ApJS, 220, 1
Aubert D., Pichon C., Colombi S., 2004, MNRAS, 352, 376
Barro G. et al., 2017, ApJ, 840, 47
Baumer M., Davis C. P., Roodman A., 2017, PASP, 129, 084502
Beaton R. L. et al., 2014, ApJ, 790, 117
Beck M. R. et al., 2018, MNRAS, 476, 5516
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Belokurov V., Erkal D., Deason A. J., Koposov S. E., De Angeli F., Evans D. W., Fraternali F., Mackey D., 2017, MNRAS, 466, 4711
Bílek M. et al., 2020, MNRAS, 498, 2138
Blumenthal K. A. et al., 2020, MNRAS, 492, 2075
Borlaff A. et al., 2019, A&A, 621, A133
Borlaff A. S. et al., 2022, A&A, 657, A92
Bournaud F., Duc P. A., Amram P., Combes F., Gach J. L., 2004, A&A, 425, 813
Bovy J., Bahmanyar A., Fritz T. K., Kallivayalil N., 2016, ApJ, 833, 31
Brough S. et al., 2020, preprint (arXiv:2001.11067)
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Buitrago F., Trujillo I., Curtis-Lake E., Montes M., Cooper A. P., Bruce V. A., Pérez-González P. G., Cirasuolo M., 2017, MNRAS, 466, 4888
Bullock J. S., Johnston K. V., 2005, ApJ, 635, 931
Calzetti D., 2001, PASP, 113, 1449
Camps P., Baes M., 2020, Astron. Comput., 31, 100381
Cappellari M. et al., 2011, MNRAS, 413, 813
Carollo D. et al., 2010, ApJ, 712, 692
Chabrier G., 2003, PASP, 115, 763
Chang C., Drlica-Wagner A., Kent S. M., Nord B., Wang D. M., Wang M. H. L. S., 2021, Astron. Comput., 36, 100474
Choi P. I., Guhathakurta P., Johnston K. V., 2002, AJ, 124, 310
Choi H., Yi S. K., Dubois Y., Kimm T., Devriendt J. E. G., Pichon C., 2018, ApJ, 114 , 856
Cibinel A. et al., 2019, MNRAS, 485, 5631
Conselice C. J., Bershady M. A., Jangren A., 2000, ApJ, 529, 886
Conselice C. J., Yang C., Bluck A. F. L., 2009, MNRAS, 394, 1956
Cooper A. P. et al., 2010, MNRAS, 406, 744
Cooper A. P., D'Souza R., Kauffmann G., Wang J., Boylan-Kolchin M., Guo Q., Frenk C. S., White S. D. M., 2013, MNRAS, 434, 3348
Courteau S., Widrow L. M., McDonald M., Guhathakurta P., Gilbert K. M., Zhu Y., Beaton R. L., Majewski S. R., 2011, ApJ, 739, 20
D'Souza R., Kauffman G., Wang J., Vegetti S., 2014, MNRAS, 443, 1433
Darg D. W. et al., 2010, MNRAS, 401, 1043
Davison T. A., Norris M. A., Pfeffer J. L., Davies J. J., Crain R. A., 2020, MNRAS, 497, 81
Deason A. J., Belokurov V., Sanders J. L., 2019, MNRAS, 490, 3426
Deeley S. et al., 2017, MNRAS, 467, 3934
Dekel A., Sari R., Ceverino D., 2009, ApJ, 703, 785
Dey A. et al., 2019, AJ, 157, 168
Diehl T., Dark Energy Survey Collaboration, 2012, Phys. Proc., 37, 1332
Draine B. T. et al., 2007, ApJ, 663, 866
Dubinski J., Mihos J. C., Hernquist L., 1996, ApJ, 462, 576
Dubinski J., Mihos J. C., Hernquist L., 1999, ApJ, 526, 607
Dubois Y., Gavazzi R., Peirani S., Silk J., 2013, MNRAS, 433, 3297
Dubois Y. et al., 2014, MNRAS, 444, 1453
Dubois Y. et al., 2021, A&A, 651, A109
Duc P.-A., Brinks E., Wink J. E., Mirabel I. F., 1997, A&A, 326, 537
Duc P.-A. et al., 2015, MNRAS, 446, 120
Dumas J., Newberg H. J., Niedzielski B., Susser A., Thompson J. M., Weiss J., Lewis K. M., 2015, ApJ, 811, 36
Elbaz D., Cesarsky C. J., 2003, Science, 300, 270

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189
Fensch J. et al., 2020, A&A, 644, A164
Ferreras I. et al., 2014, MNRAS, 444, 906
Fiacconi D., Feldmann R., Mayer L., 2015, MNRAS, 446, 1957
Fliri J., Trujillo I., 2016, MNRAS, 456, 1359
Foster C. et al., 2014, MNRAS, 442, 3544
Fry A. M., Morrison H. L., Harding P., Boroson T. A., 1999, AJ, 118, 1209
Gargiulo I. D. et al., 2019, MNRAS, 489, 5742
Garrison-Kimmel S. et al., 2019, MNRAS, 489, 4574
Gentile G., Famaey B., Combes F., Kroupa P., Zhao H. S., Tiret O., 2007, A&A, 472, L25
Gómez F. A. et al., 2017, MNRAS, 472, 3722
Graham A. W., 2002, ApJ, 568, L13
Graham A. W., Dullo B. T., Savorgnan G. A. D., 2015, ApJ, 804, 32
Greco J. P. et al., 2018a, PASJ, 70, S19
Greco J. P. et al., 2018b, ApJ, 857, 104
Grillmair C. J., Freeman K. C., Irwin M., Quinn P. J., 1995, AJ, 109, 2553
Gwyn S. D. J., 2012, AJ, 143, 38
Harmsen B., Monachesi A., Bell E. F., de Jong R. S., Bailin J., Radburn-Smith D. J., Holwerda B. W., 2017, MNRAS, 466, 1491
Hendel D., Johnston K. V., 2015, MNRAS, 454, 2472
Hendel D., Johnston K. V., Patra R. K., Sen B., 2019, MNRAS, 486, 3604
Hood C. E., Kannappan S. J., Stark D. V., Dell'Antonio I. P., Moffett A. J., Eckert K. D., Norris M. A., Hendel D., 2018, ApJ, 857, 144
Huertas-Company M., Aguerri J. A. L., Tresse L., Bolzonella M., Koekemoer A. M., Maier C., 2010, A&A, 515, A3
Ibata R., Thomas G., Famaey B., Malhan K., Martin N., Monari G., 2020, ApJ, 891, 161
Iodice E. et al., 2016, ApJ, 820, 42
Iodice E. et al., 2019, A&A, 623, A1
Ivezić Ž. et al., 2019, ApJ, 873, 111
Jackson R. A., Martin G., Kaviraj S., Laigle C., Devriendt J. E. G., Dubois Y., Pichon C., 2019, MNRAS, 489, 4679
Ji I., Hasan I., Schmidt S. J., Tyson J. A., 2018, PASP, 130, 084504
Johnston K. V., Sigurdsson S., Hernquist L., 1999, MNRAS, 302, 771
Johnston K. V., Sackett P. D., Bullock J. S., 2001, ApJ, 557, 137
Johnston K. V., Choi P. I., Guhathakurta P., 2002, AJ, 124, 127
Johnston K. V., Bullock J. S., Sharma S., Font A., Robertson B. E., Leitner S. N., 2008, ApJ, 689, 936
Jonsson P., 2006, MNRAS, 372, 2
Junais et al., 2020, A&A, 637, A21
Kado-Fong E. et al., 2018, ApJ, 866, 103
Kadowaki J., Zaritsky D., Donnerstein R. L., 2017, ApJ, 838, L21
Karunakaran A., Spekkens K., Zaritsky D., Donnerstein R. L., Kadowaki J., Dey A., 2020, ApJ, 902, 39
Kaviraj S., 2014, MNRAS, 440, 2944
Kaviraj S., 2020, preprint (arXiv:2001.01728)
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., Devriendt J., Dubois Y., Slyz A., Welker C., Pichon C., Peirani S., Le Borgne D., 2015, MNRAS, 452, 2845
Kaviraj S. et al., 2017, MNRAS, 467, 4739
Kaviraj S., Martin G., Silk J., 2019, MNRAS, 489, L12
Kawinwanichakij L. et al., 2017, ApJ, 847, 134
Kesden M., Kamionkowski M., 2006, Phys. Rev. D, 74, 083007
Kim T. et al., 2012, ApJ, 753, 43
Kniazev A. Y., Grebel E. K., Pustilnik S. A., Pramskij A. G., Kniazeva T. F., Prada F., Harbeck D., 2004, AJ, 127, 704
Komatsu E. et al., 2011, ApJS, 192, 18
Kroupa P., 2002, Science, 295, 82
Kuijken K. et al., 2002, The Messenger, 110, 15
Laine S. et al., 2018, preprint (arXiv:1812.04897)
Lambrides E. L. et al., 2021, ApJ, 919, 43
Lauer T. R., 1988, ApJ, 325, 49
Lee N. et al., 2015, ApJ, 801, 80
Lintott C. J. et al., 2008, MNRAS, 389, 1179
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
Lupton R., Blanton M. R., Fekete G., Hogg D. W., O'Mullane W., Szalay A., Wherry N., 2004, PASP, 116, 133
Malhan K., Valluri M., Freese K., 2021, MNRAS, 501, 179
Malin D. F., Carter D., 1983, ApJ, 274, 534
Mancillas B., Duc P.-A., Combes F., Bournaud F., Emsellem E., Martig M., Michel-Dansac L., 2019, A&A, 632, A122
Mantha K. B. et al., 2019, MNRAS, 486, 2643
Martin G., Kaviraj S., Devriendt J. E. G., Dubois Y., Laigle C., Pichon C., 2017, MNRAS, 472, L50
Martin G., Kaviraj S., Devriendt J. E. G., Dubois Y., Pichon C., 2018a, MNRAS, 480, 2266
Martin G., Kaviraj S., Devriendt J. E. G., Dubois Y., Pichon C., Laigle C., 2018b, MNRAS, 474, 3140
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
Martínez-Delgado D., Pohlen M., Gabany R. J., Majewski S. R., Peñarrubia J., Palma C., 2009, ApJ, 692, 955
Martínez-Delgado D. et al., 2012, ApJ, 748, L24
Martinez-Delgado D. et al., 2021, preprint (arXiv:2104.06071)
Merritt A., van Dokkum P., Danieli S., Abraham R., Zhang J., Karachentsev I. D., Makarova L. N., 2016, ApJ, 833, 168
Merritt A., Pillepich A., van Dokkum P., Nelson D., Hernquist L., Marinacci F., Vogelsberger M., 2020, MNRAS, 495, 4570
Mihos J. C., 2019, preprint (arXiv:1909.09456)
Mihos J. C., Hernquist L., 1996, ApJ, 464, 641
Mihos J. C., Harding P., Feldmeier J., Morrison H., 2005, ApJ, 631, L41
Mihos J. C. et al., 2015, ApJ, 809, L21
Mihos J. C., Harding P., Feldmeier J. J., Rudick C., Janowiecki S., Morrison H., Slater C., Watkins A., 2017, ApJ, 834, 16
Miville-Deschênes M. A., Duc P. A., Marleau F., Cuillandre J. C., Didelon P., Gwyn S., Karabal E., 2016, A&A, 593, A4
Miyazaki S. et al., 2002, PASJ, 54, 833
Miyazaki S. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 9
Moffat A. F. J., 1969, A&A, 3, 455
Monachesi A., Gómez F. A., Grand R. J. J., Kauffmann G., Marinacci F., Pakmor R., Springel V., Frenk C. S., 2016, MNRAS, 459, L46
Monachesi A. et al., 2019, MNRAS, 485, 2589
Montes M., 2019, preprint (arXiv:1912.01616)
Montes M., Infante-Sainz R., Madrigal-Aguado A., Román J., Monelli M., Borlaff A. S., Trujillo I., 2020, ApJ, 904, 114
Montes M., Brough S., Owers M. S., Santucci G., 2021, ApJ, 910, 45
Morales G., Martínez-Delgado D., Grebel E. K., Cooper A. P., Javanmardi B., Miskolczi A., 2018, A&A, 614, A143
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Müller O., Vudragović A., Bílek M., 2019, A&A, 632, L13
Naab T. et al., 2014, MNRAS, 444, 3357
Nelson D. et al., 2019, MNRAS, 490, 3234
Olivier S. S., Seppala L., Gilmore K., 2008, in Atad-Ettedgui E., Lemke D., eds, Proc. SPIE Conf. Ser. Vol. 7018, Advanced Optical and Mechanical Technologies in Telescopes and Instrumentation. SPIE, Bellingham, p. 70182G
Olsen K. P. et al., 2021, ApJ, 922, 88
Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, ApJ, 725, 2312
Park M.-J. et al., 2019, ApJ, 883, 25
Pawlik M. M., Wild V., Walcher C. J., Johansson P. H., Villforth C., Rowlands K., Mendez-Abreu J., Hewlett T., 2016, MNRAS, 456, 3032
Pawlowski M. S., 2021, Galaxies, 9, 66
Pearson S., Clark S. E., Demirjian A. J., Johnston K. V., Ness M. K., Starkenburg T. K., Williams B. F., Ibata R. A., 2022, ApJ, 926, 166

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

Peirani S., Crockett R. M., Geen S., Khochfar S., Kaviraj S., Silk J., 2010, MNRAS, 405, 2327
Pfleiderer J., 1963, Z. Astrophys., 58, 12
Pillepich A., Madau P., Mayer L., 2015, ApJ, 799, 184
Pillepich A. et al., 2018, MNRAS, 475, 648
Pontzen A., Tremmel M., Roth N., Peiris H. V., Saintonge A., Volonteri M., Quinn T., Governato F., 2017, MNRAS, 465, 547
Popesso P. et al., 2012, preprint (arXiv:1211.4257)
Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14
Press W. H., Schechter P., 1974, ApJ, 187, 425
Prole D. J., Davies J. I., Keenan O. C., Davies L. J. M., 2018, MNRAS, 478, 667
Purcell C. W., Bullock J. S., Zentner A. R., 2007, ApJ, 666, 20
Quinn P. J., 1984, ApJ, 279, 596
Ragusa R. et al., 2021, A&A, 651, A39
Reddish J. et al., 2022, MNRAS, 512, 160
Ren J. et al., 2020, MNRAS, 499, 3399
Renaud F., Famaey B., Kroupa P., 2016, MNRAS, 463, 3637
Rey M. P., Starkenburg T. K., 2022, MNRAS, 510, 4208
Rich R. M. et al., 2019, MNRAS, 490, 1539
Robertson B. E. et al., 2017, preprint (arXiv:1708.01617)
Robertson B. et al., 2019, Bull. Am. Astron. Soc., 51, 30
Robotham A. S. G. et al., 2014, MNRAS, 444, 3986
Rockosi C. M. et al., 2002, AJ, 124, 349
Rodriguez-Gomez V. et al., 2015, MNRAS, 449, 49
Rodriguez-Gomez V. et al., 2016, MNRAS, 458, 2371
Román J., Trujillo I., 2017, MNRAS, 468, 4039
Román J., Trujillo I., Montes M., 2020, A&A, 644, A42
Salpeter E. E., 1955, ApJ, 121, 161
Samuel J., Wetzel A., Santistevan I., Tollerud E., Moreno J., Boylan-Kolchin M., Bailin J., Pardasani B., 2022, preprint (arXiv:2203.07385)
Sand D. J., Olszewski E. W., Willman B., Zaritsky D., Seth A., Harris J., Piatek S., Saha A., 2009, ApJ, 704, 898
Sanderson R. E., Bertschinger E., 2010, ApJ, 725, 1652
Sanderson R. E., Helmi A., 2013, MNRAS, 435, 378
Schawinski K. et al., 2014, MNRAS, 440, 889
Schweizer F., 1982, ApJ, 252, 455
Seigar M. S., Graham A. W., Jerjen H., 2007, MNRAS, 378, 1575
Sérsic J. L., 1968, Atlas de Galaxias Australes. Observatorio Astronomico, Cordoba, Argentina
Shipp N. et al., 2018, ApJ, 862, 114
Slater C. T., Harding P., Mihos J. C., 2009, PASP, 121, 1267
Smercina A. et al., 2020, ApJ, 905, 60
Sola E. et al., 2022, preprint (arXiv:2203.03973)
Spavone M. et al., 2017, A&A, 603, A38
Spavone M. et al., 2018, ApJ, 864, 149
Spavone M. et al., 2020, A&A, 639, A14
Stewart K. R., Bullock J. S., Wechsler R. H., Maller A. H., Zentner A. R., 2008, ApJ, 683, 597
Székely G. J. et al., 2007, Ann. Stat., 35, 2769
Székely G. J. et al., 2014, Ann. Stat., 42, 2382
Tacchella S. et al., 2019, MNRAS, 487, 5416
Tal T., van Dokkum P. G., Nelan J., Bezanson R., 2009, AJ, 138, 1417
Tanoglidis D. et al., 2021a, preprint (arXiv:2109.08246)
Tanoglidis D., Ćiprijanović A., Drlica-Wagner A., 2021b, Astron. Comput., 35, 100469
Taranu D. S., Dubinski J., Yee H. K. C., 2013, ApJ, 778, 61
Teyssier R., 2002, A&A, 385, 337
Thanjavur K., Simard L., Bluck A. F. L., Mendel T., 2016, MNRAS, 459, 44
Toomre A., 1977, in Tinsley B. M., Larson D., Campbell R. B. G., eds, Evolution of Galaxies and Stellar Populations. Yale University Observatory, New Haven, p. 401
Toomre A., Toomre J., 1972, ApJ, 178, 623
Toro E., 1999, Riemann Solvers and Numerical Methods for Fluid Dynamics. Springer-Verlag, Berlin
Torrealba G. et al., 2018, MNRAS, 475, 5085
Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, MNRAS, 470, 1121
Trujillo I., Fliri J., 2016, ApJ, 823, 123
Trujillo I. et al., 2021, A&A, 654, A40
Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, A&A, 506, 647
Uzeirbegovic E., Geach J. E., Kaviraj S., 2020, MNRAS, 498, 4021
van den Bosch F. C., Abel T., Croft R. A. C., Hernquist L., White S. D. M., 2002, ApJ, 576, 21
van Dokkum P. et al., 2018, Nature, 555, 629
van Dokkum P. et al., 2019, ApJ, 883, L32
Varghese A., Ibata R., Lewis G. F., 2011, MNRAS, 417, 198
Vera-Casanova A. et al., 2021, preprint (arXiv:2105.06467)
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Walmsley M., Ferguson A. M. N., Mann R. G., Lintott C. J., 2019, MNRAS, 483, 2968
Wang J., Hammer F., Athanassoula E., Puech M., Yang Y., Flores H., 2012, A&A, 538, A121
Weingartner J. C., Draine B. T., 2001, ApJ, 548, 296
Welker C., Dubois Y., Devriendt J., Pichon C., Kaviraj S., Peirani S., 2017, MNRAS, 465, 1241
Yoon Y., Lim G., 2020, ApJ, 905, 154
Zaritsky D. et al., 2019, ApJS, 240, 1
Zaritsky D., Donnerstein R., Karunakaran A., Barbosa C. E., Dey A., Kadowaki J., Spekkens K., Zhang H., 2021, ApJS, 257, 60

APPENDIX A: EFFECTIVE RESOLUTION LIMIT FOR DETECTING SHELLS

The stellar particle mass resolution of a cosmological simulation puts limits on its ability to resolve structures. For instance, an image of a structure with a total stellar mass close to the stellar mass resolution of a simulation may not have sufficient contrast against the background of the host galaxy to be detectable, even if the particles that make up the structure are kinematically distinct from the galaxy. Given the stellar mass of a structure in a simulation, one can estimate the number of stellar particles in the structure and use this number to predict whether it is resolved (detectable) in a simulation with known stellar particle mass. In this section, we present the lower limits on detectable shells based on the NEWHORIZON stellar particle mass resolution (1.3×10^4 $M_\odot$).

Following Bazkiaei et al. (in preparation), we combine analytical profiles of shells based on Sanderson & Bertschinger (2010) and Sanderson & Helmi (2013) with Sérsic (1968) models in order to estimate the numerical limits of the simulation to resolve tidal features.

Table A1 specifies four host galaxy Sérsic models which we select to bracket a realistic range of parameter space and to be roughly representative of galaxies found in the four mass bins shown in Fig. 12. Shell models are generated according to equations (1) and (19) of Sanderson & Helmi (2013). Following the same notation used by Sanderson & Helmi (2013), we generate models for a range of characteristic widths, δ_r, galactocentric radii r_s, and opening angles, α, as well as for a range of stellar masses.

To find the stellar mass within a shell, we produce mass maps for every combination of galaxy and shell model matching the Rubin Observatory 0.2 arcsec pixel scale (corresponding to a physical size of 0.19, 0.37, 0.66, 1.08, 1.51 kpc per pixel for the lowest to highest redshift of this work). The maps are then re-binned to 1 arcsec to mimic the procedure we use to produce mock images from NEWHORIZON galaxies (see Section 2.3.3). Two types of re-binned 1 arcsec image created:

(i) 1 arcsec mass maps – we perform a simple re-binning of the original 0.2 arcsec maps to 1 arcsec – used to determine the stellar

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

Table A1. Summary of the parameters used to generate each model galaxy: *a*, model name; *b*, total stellar mass; *c*, Sérsic index; *d*, effective radius.

Model [a]	$M_*/M_\odot$ [b]	n [c]	$r_{\rm eff}$ [kpc] [d]
Galaxy 1	$10^{9.5}$	0.5	2
Galaxy 2	$10^{10.0}$	1.0	4
Galaxy 3	$10^{10.5}$	2.5	6
Galaxy 4	$10^{11.0}$	4.0	8

mass of each shell model and the region of each galaxy overlapped by the shell model.

(ii) PSF convolved surface brightness maps – the 0.2 arcsec mass maps are used to calculate the *r*-band surface brightnesses of the tidal feature models following the same procedure as detailed in Section 2.3.1 and assuming all stars making up the tidal features are born at $z = 2$ and have a metallicity of $Z = 0.1$. They are then convolved with the Hyper Suprime-Cam PSF (Montes et al. 2021) and re-binned to 1 arcsec.

Using the first set of mass maps, we calculate the signal-to-noise-ratio ($\rm SNR_{shell}$) for each shell as follows:

$$\rm SNR_{shell} = \frac{n_{shell}}{\sigma_{shell}} = \frac{n_{shell}}{\sqrt{2\, n_{galaxy} + n_{shell}}}, \quad (A1)$$

where $n_{\rm shell}$ and $n_{\rm galaxy}$ are the total number of star particles that comprise the arc of the shell (i.e. the brightest part) and galaxy, respectively, within region described by the arc of the shell model given the stellar mass resolution of the simulation. Shells are considered to be detected if $\rm SNR_{shell} > 5$.

Then, using the PSF convolved surface brightness maps, we calculate the average surface brightness across the arc of the shell in order to find the faintest detectable shell from among all the detected shell models.

The surface brightness of faintest detectable shells, based on the NewHorizon stellar particle mass resolution are presented in Fig. A1. This figure shows the numerical lower limits for the faintest *r*-band surface brightness at which shells around each of the model galaxies are detected as a function of galactocentric radius. Different galaxy models and shell model parameters are represented by different colours and line styles, respectively.

Even with its relatively high stellar mass resolution, we expect NewHorizon to struggle to resolve shells with sufficient contrast close to the central parts of galaxies. Around our most massive model galaxy in particular, signal-to-noise is not sufficient to detect any of our model shells with surface brightnesses fainter than 31 mag arcsec^{-2} within galactocentric radii smaller than 4.5 $R_{\rm eff}$. For less massive galaxy models, it is possible to detect faint shells at significantly smaller radii, however.

Overall, NewHorizon is expected to resolve most of the shells (and likely other types of tidal feature) which would be detectable at limiting surface brightnesses realistically achievable by the main LSST survey ($\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) \leq 31$ mag arcsec^{-2}) outside of the central few effective radii. More care needs to be taken in interpreting predictions at higher limiting surface brightnesses as observationally detectable features may not reach sufficient signal-to-noise to be resolved in the simulation out to relatively large radii ($r_s \gg 10\, R_{\rm eff}$).

Note that in our treatment we do not account for possible differences in the stellar populations of the galaxy and tidal features which result in differing mass-to-light ratios between the galaxy and shell. Tidal features may also be easier to detect in false colour images

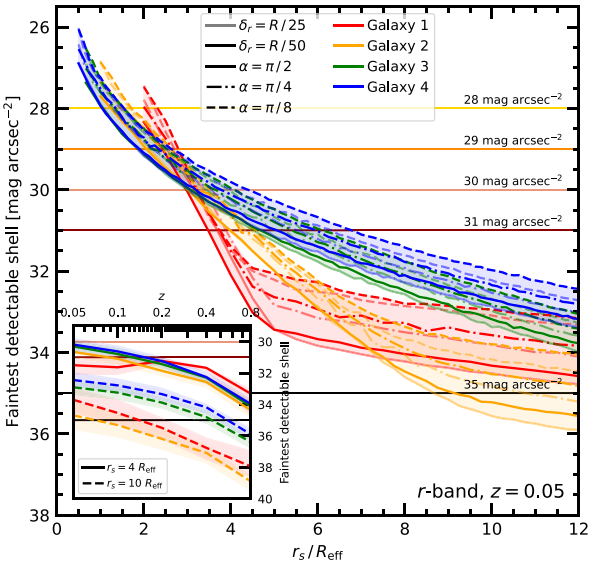


Figure A1. The main plot shows numerical lower limits for the average surface brightness of the faintest detectable shell as a function of galactocentric radius for different galaxy and shell models. Different coloured lines indicate different galaxy Sérsic models, while solid, dot-dashed, and dashed lines indicate different opening angles. Darker or lighter colours indicate different characteristic widths with all values shown in the legend. We expect any shells (with a given set of parameters and at a given radius) with surface brightnesses fainter than these lines to be undetectable in NewHorizon mock images, regardless of the image depth. For reference we also show coloured horizontal lines, which highlight the surface brightness limits of mock images used throughout this work. The inset plot shows the surface brightness of the faintest detectable shell averaged over the different parameter values of δ_r and α and measured at 4 $R_{\rm eff}$ (solid lines) and 10 $R_{\rm eff}$ (dashed lines) as a function of redshift. Line colours again indicate the Sérsic models while the shaded region indicate the range of values measured across the different shell parameters.

in cases where their colours differ enough from host galaxy, even if there is insufficient signal-to-noise in any single band. We cannot quantify how much this would improve our results, except to say that we do see this effect in at least some mock images. Without knowing the underlying spatial distribution of shells, it is difficult quantify the significance of the impact that an inability to detect shells close to the central galaxy has, however. On the other hand, depending on viewing angle, shells may be expected to be brighter at their maximum than other types of tidal feature since stars accumulate at the shell apocentre. However, since we consider an average signal-to-noise across the whole arc of the shell, we expect this effect to be lessened.

APPENDIX B: EFFECT OF THE PSF ON THE VISIBILITY OF TIDAL FEATURES

Fig. B1 illustrates the effect of the PSF on the median surface brightness profile of our galaxies. In the top right-hand panel we show a comparison of the measured PSF from Montes et al. (2021) and best fitting Moffat (1969) and Gaussian distributions. The wings of the PSF are not well described by either Gaussian or Moffat distributions beyond a few arcseconds and this becomes increasingly severe towards large radii. We therefore find that neither are appropriate choices for modelling the effect of the PSF in the faint outskirts of the galaxy. The remaining panels show median surface brightness

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

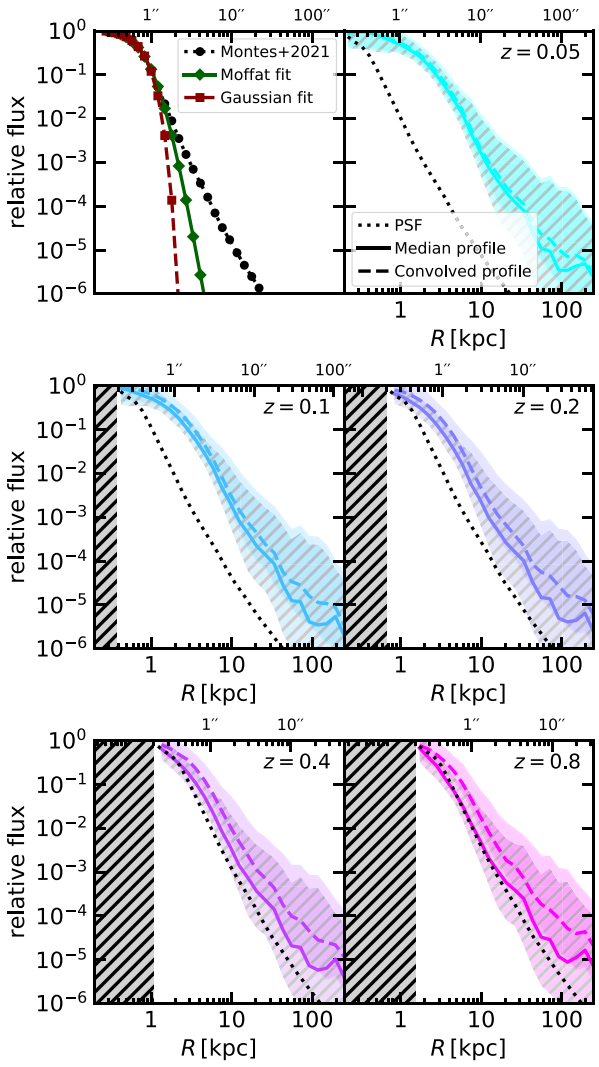


Figure B1. Top left-hand panel: Comparison of the measured 1D PSF from Montes et al. (2021) with a the best fit Moffat (1969) and Gaussian distributions. Other panels: surface brightness profiles compared to the PSF for different redshifts (indicated in the top right-hand corner of each panel). The solid lines show the median relative surface brightness profile of the raw mock images, dashed lines show the median relative surface brightness profiles after they have been convolved with the PSF, and the dotted black line shows the Montes et al. (2021) PSF. Coloured filled regions indicate the central 68th percentile (1σ) of profiles where the hatched region indicates the unconvolved profile. The x-axis limits are fixed in physical units between 0.2 and 250 kpc with the equivalent scale in arcseconds for each redshift shown on the top x-axis of each panel. The grey hatched region indicates radii smaller than the 0.2 arcsec Rubin Observatory pixel size.

profiles compared to the PSF for different redshifts. The solid lines show the median relative surface brightness profile of the raw mock images, dashed lines show the median relative surface brightness profiles after they have been convolved with the PSF, and the dotted black line indicates the PSF. The profiles and PSF in each panel are scaled to a fixed physical scale, with the angular scale indicated separately on the upper x-axis. Towards higher redshift, as the increasing angular scale means that fainter regions of the galaxy surface brightness profile move further into the core of the PSF, the convolved profiles start to depart significantly from the original profiles. As we neglect any possible contribution of scattered light from stars, this effect is entirely the result of smearing and the scattering of light from the bright core of the galaxy into the fainter outskirts. We therefore expect that, for very deep imaging, the PSF will have some impact on how well we are able to detect tidal features. For a noise level of $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 31$ mag arcsec^{-2}, we find that the visibility of tidal features is noticeably impacted by the PSF, becoming especially apparent after $z = 0.4$. However, we do not find that this effect is strong enough that previously visible tidal features routinely become invisible, particularly at typical limiting surface brightnesses accessible to the Rubin Observatory. As instruments improve further and it becomes possible to probe even deeper into the outskirts of galaxies, we can expect that this effect will become more important.

APPENDIX C: SELECTING THE EX-SITU MASS TIME INTERVAL

We try to choose a time interval that gives the tightest relation between ex-situ mass and halo mass, however it is not obvious what this interval should be. While longer time-scales probe more of the accretion history galaxy, they may not reflect the current state of the galaxy (for example, if the galaxy recently underwent a merger) and so at some point may begin to correlate poorly with the halo mass. Equally a time-scale that is too small will be effected more strongly by the stochasticity inherent in galaxy accretion histories.

By changing the value of $t_{\rm min}$ in equation (4), we vary the time over which we measure the ex-situ mass fraction. Here we consider the affect of adopting different time intervals Δt such that $t_{\rm min} = t_{\rm max} - \Delta t$. We then measure the distance correlation coefficient (Székely et al. 2007) between $f_{\rm exsitu}$ and M_h.

Figs C1 and C2 show how the correlation coefficient behaves for different values of Δt. Fig. C1 shows the correlation between halo mass and $f_{\rm exsitu}$ and Fig. C2 shows the correlation between $f_{\rm tidal}$ and $f_{\rm exsitu}$. Open circles with error bars indicate the value of the correlation coefficient and associated 1σ uncertainty as a function of Δt. In the

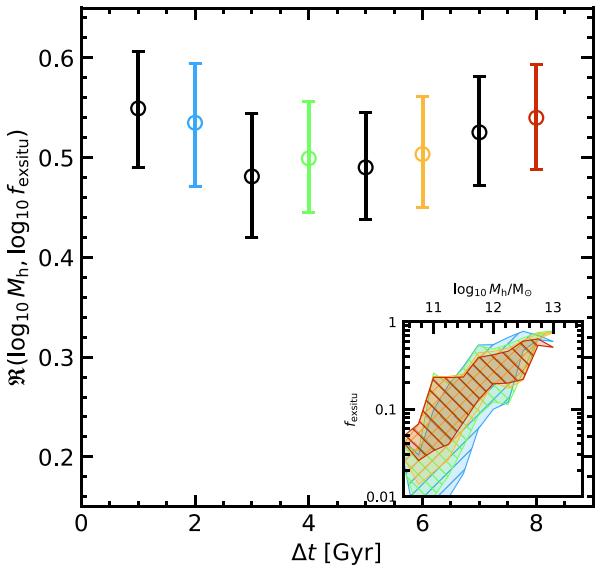


Figure C1. Correlation coefficient for different time intervals with 1σ uncertainties indicated by error bars. The inset plot shows a region enclosing the 1σ scatter for the M_h–$f_{\rm exsitu}$ relation for $\Delta t = [2, 4, 6, 8]$ where the colour of the region corresponds to the value Δt indicated by the coloured error bars in the main plot.

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

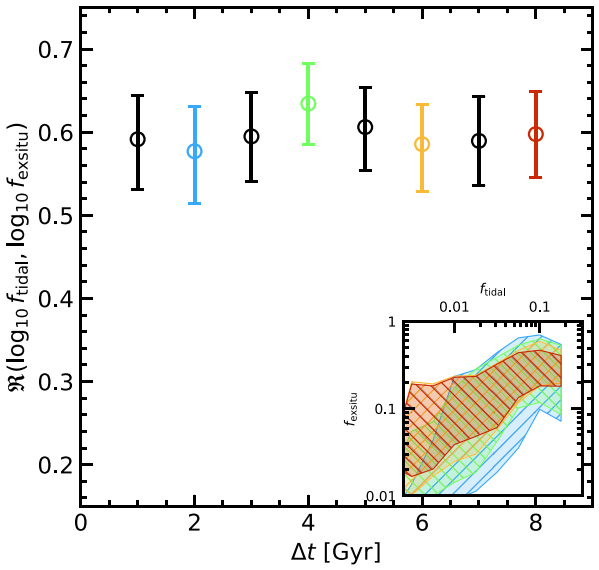


Figure C2. Correlation coefficient for different time intervals with 1σ uncertainties indicated by error bars. The inset plot shows a region enclosing the 1σ scatter for the M_h–$f_{\rm exsitu}$ relation for $\Delta t = [2, 4, 6, 8]$ where the colour of the region corresponds to the value Δt indicated by the coloured error bars in the main plot.

inset panel we plot filled and hatched regions which enclose the 1σ scatter of the M_h–$f_{\rm exsitu}$ relation for multiple values of Δt ($\Delta t = [2, 4, 6, 8]$). The colour of each region corresponds to the value Δt and is indicated by the four coloured error bars in the main plot.

We find that increasing the time-scale does not have a significant influence on the level of correlation in either case. We therefore adopt the maximum possible value of Δt for each galaxy (i.e. over the whole lifetime of the galaxy) as this better reflects the overall accretion history of the galaxy.

APPENDIX D: REDSHIFT EVOLUTION

We calculate the change in $f_{\rm exsitu}$ and $f_{\rm tidal}$ for our sample of 37 galaxies in the time interval between highest and lowest redshifts that we consider ($z = 0.8$ and $z = 0.2$) and define the growth rate of $f_{\rm exsitu}$, $\Gamma_{\rm exsitu}$; equation (D1), and the tidal mass fraction, $\Gamma_{\rm tidal}$; equation (D2), as follows:

$$\Gamma_{\rm exsitu} = \frac{f_{\rm exsitu}(z = 0.2, z = \infty) - f_{\rm exsitu}(z = 0.8, z = \infty)}{f_{\rm exsitu}(z = 0.2, z = \infty)\,\Delta t}, \quad \text{(D1)}$$

where $f_{\rm exsitu}$ is defined in equation (4) and Δt is the time between $z = 0.8$ and $z = 0.2$ (~4 Gyr).

$$\Gamma_{\rm tidal} = \frac{f_{\rm tidal}(z = 0.2) - f_{\rm tidal}(z = 0.8)}{f_{\rm tidal}(z = 0.2)\,\Delta t}, \quad \text{(D2)}$$

where $f_{\rm tidal}$ is defined in equation (5).

Fig. D1 shows the distribution $\Gamma_{\rm exsitu}$ (top panel) and $\Gamma_{\rm tidal}$ (bottom panel) both obtained from Gaussian kernel density estimates using 10 000 bootstraps. The numbers in red and green at the top of both panels indicate the percentage of galaxies whose growth rate is either negative or positive, respectively, with their associated 1σ errors. The standard deviation of the distribution of $\Gamma_{\rm exsitu}$ and $\Gamma_{\rm tidal}$ is 0.15 and 0.18 Gyr^{-1}, respectively, indicating relatively large swings in the growth rate (amounting to a greater than ~50 per cent change over the 4 Gyr between $z = 0.8$ and $z = 0.2$ in 40–50 per cent of the population). In both cases the growth rate in the redshift range

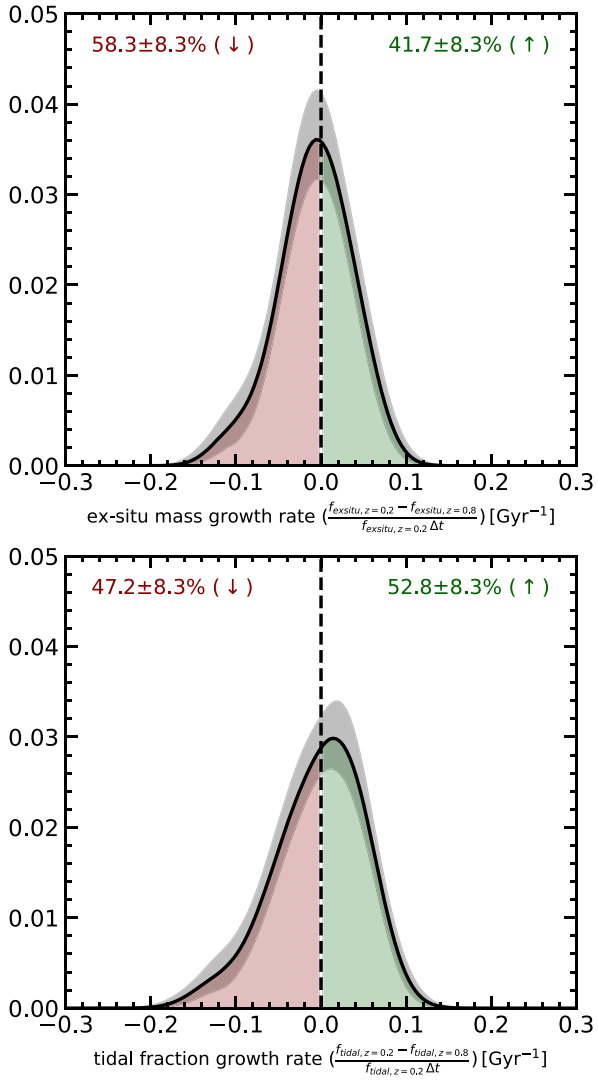


Figure D1. Top: the distribution of the $f_{\rm exsitu}$ growth rate ($\Gamma_{\rm exsitu}$) from Gaussian kernel density estimates from 10 000 bootstraps. Bottom: the distribution of the $f_{\rm tidal}$ growth rate ($\Gamma_{\rm tidal}$) from Gaussian kernel density estimates from 10 000 bootstraps. Both growth rates are calculated between $z = 0.8$ and $z = 0.2$. The grey region in both panels shows the 1σ uncertainty in the kernel density estimate.

between $z = 0.8$ and $z = 0.2$ is consistent with an equal number of galaxies having negative and positive growth rates.

Additionally, the median $\Gamma_{\rm exsitu}$ and the median $\Gamma_{\rm tidal}$ are both consistent with no average change (-0.0186 ± 0.0261 Gyr^{-1} for $\Gamma_{\rm exsitu}$ and -0.0230 ± 0.0652 Gyr^{-1} for the $\Gamma_{\rm tidal}$). This is also true if we consider the overall fractional change in the net $f_{\rm exsitu}$ and $f_{\rm tidal}$, which we calculate by taking the mean value weighted by the host galaxy masses (-0.0292 ± 0.0295 Gyr^{-1} for $\Gamma_{\rm exsitu}$ and -0.0187 ± 0.0246 Gyr^{-1} for the $\Gamma_{\rm tidal}$).

Finally, Fig. D2 shows the total number of mergers undergone by the galaxies in our sample, which we adopt as a rough proxy for the number of discrete units of mass entering the galaxy halo over time (i.e. objects that could be disrupted in the galaxy halo to form tidal features). The main plot shows the total number of mergers with mass ratio $R > 1$: 1000 that each galaxy has undergone as a function of their stellar mass and the inset plot shows individual tracks indicating the change in the merger rate history of each galaxy as a function of

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

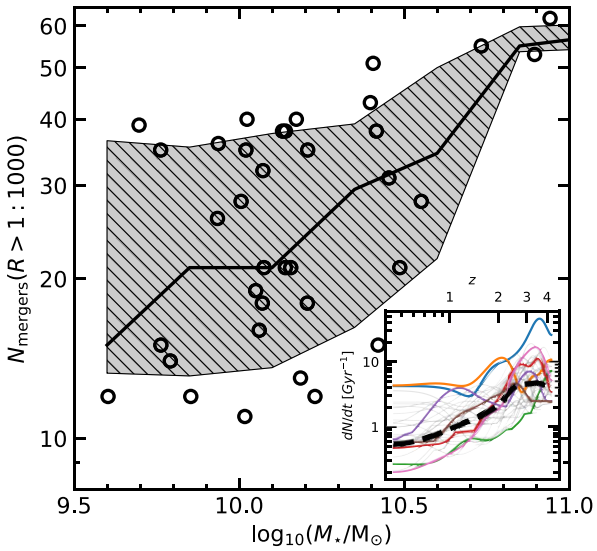


Figure D2. Number of mergers with mass ratio $R < 1{:}1000$ undergone per galaxy since $z = 5$ as a function of galaxy stellar mass. The hatched region encloses the 1σ scatter and the black line indicates the median. The inset plot shows the merger rate ($R < 1{:}1000$) per Gyr of individual galaxies as a function of redshift. The thick dashed line shows the evolution of the median merger rate of all the galaxies in the sample, thin grey lines show the merger rate history for each galaxy individually, with a small random subset highlighted with thicker coloured lines for clarity.

redshift for the same sample of galaxies. The merger rate history for a small subsample of galaxies is highlighted with thicker coloured lines, while the remaining galaxy merger histories are shown as thin grey lines. While we observe a clear average evolution in the merger rate, there is a very significant spread in both the total number of mergers that galaxies have undergone at fixed stellar mass as well as in the overall shape and normalization of galaxy merger histories. In the range of redshifts that we consider in this paper ($z = 0.2$ to $z = 0.8$), we see galaxy merger rates decreasing on average, but the variation in the merger histories between $z = 0.2$ and $z = 0.8$ is very large, with merger rates increasing significantly (fractional increase greater than 0.1) in ∼15 per cent of cases or remaining roughly flat (fractional change of less than 0.1) for a further ∼30 per cent of cases.

Together, these results indicate that, although there is a clear average evolution in galaxy accretion histories over cosmic time, the merger histories of individual galaxies are sufficiently stochastic that we do not expect to observe this trend in individual galaxies over the time-scale that we consider in this study.

APPENDIX E: TIDAL FEATURE DETECTION METHOD

Here, we describe our method for determining detected pixels which is used in the definition of the detection fraction used in Section 3.3. We take into account the fact that it is generally possible to detect contiguous structures by eye, even if they are made up of pixels that are mostly fainter than the surface brightness limit. This is because only relatively few detected pixels grouped close together are required for a contiguous structure to be recognized even if these detected pixels make up a small fraction of the total area of the visible tidal feature. Our aim is to identify these structures in a way

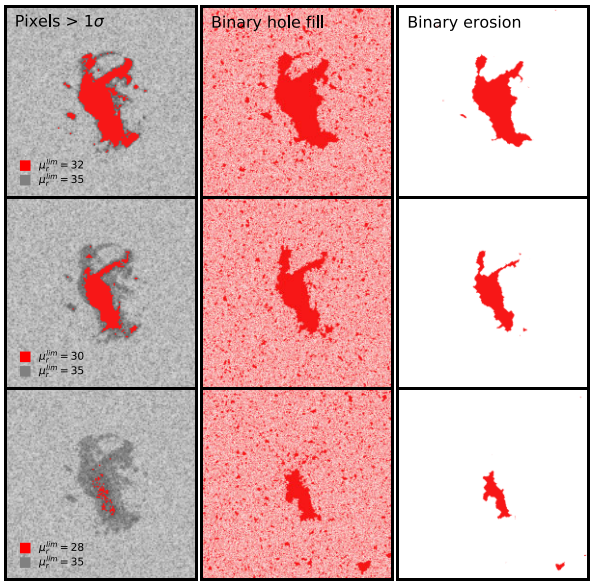


Figure E1. Example images showing the process of locating detected structures for limiting surface brightnesses of $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 32$ mag arcsec^{-2}, 30 mag arcsec^{-2}, and 28 mag arcsec^{-2}. Each panel of the first column shows a map of pixels that are at least 1σ brighter than the noise level for each given limiting surface brightness level (red) plotted over the same map for a limiting surface brightness of 35 mag arcsec^{-2} (grey). For this column only, the images are interpolated for illustrative purposes in order to average out the effect of noise, which otherwise makes it difficult to distinguish between the red and grey maps. The second column shows the binary mask resulting from a 3×3 element binary hole fill and the third column shows the final mask after applying a binary erosion.

that produces similar results to the human eye while rejecting regions of noise.

We adopt a definition for detected structures based on the connections between pixels that are 1σ above the noise level in images produced from particles that are part of dense tidal features only (e.g. Fig. 5, panel c). We proceed as follows:

(i) We first use the `binary_fill_holes` function implemented in SCIPY (Virtanen et al. 2020) using a cross-shaped 3×3 structuring element, which allows us to construct a mask consisting of every pixel lying within the boundary of a connected region (i.e. we fill any undetected pixels that are surrounded by detected pixels).

(ii) In order to remove small isolated structures which arise from spurious detections in the noise, we then perform a binary erosion on the mask with enough iterations that structures no longer appear in isolated regions of the image.

(iii) Any flux found in pixels that are within the mask is considered to be detected.

Fig. E1 shows our method performed on an example galaxy. Left-hand panels show detection maps consisting of pixels that are at least 1σ brighter the noise level for each given limiting surface brightness level (red) plotted over the same map for a limiting surface brightness of 35 mag arcsec^{-2} (grey). These images are interpolated (smoothed) in order to average out the effect of noise, which otherwise makes it difficult to distinguish between the red and grey maps (for illustrative purposes only). The second column shows the binary mask resulting from a 3×3 element binary hole fill and the third column shows the final mask after applying a binary erosion. Although the interpolated detection maps and final masks

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023

appear similar at fainter limiting surface brightnesses, the utility of this method becomes more apparent at limiting surface brightnesses closer to that of the tidal features. In this case the area of visible tidal features is significantly larger than the area of pixels that make up the detection map.

[1]*Korea Astronomy and Space Science Institute, 776 Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Korea*
[2]*Steward Observatory, University of Arizona, 933 N. Cherry Ave, Tucson, AZ 85719, USA*
[3]*Research Centre for Astronomy, Astrophysics & Astrophotonics, Macquarie University, Sydney, NSW 2109, Australia*
[4]*Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia*
[5]*Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D) , Australia*
[6]*INAF-Astronomical Observatory of Capodimonte, Salita Moiariello 16, I-80131 Naples, Italy*
[7]*Department of Astronomy, Case Western Reserve University, Cleveland, OH 44106, USA*
[8]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[9]*Instituto de Astronomía Teórica y Experimental, CONICET-UNC, Laprida. 854, X5000BGR Córdoba, Argentina*
[10]*Observatorio Astronómico de Córdoba, Universidad Nacional de Córdoba, Laprida 854, X5000BGR Córdoba, Argentina*
[11]*School of Physics, University of New South Wales, NSW 2052, Australia*
[12]*NSF's NOIRLab/Rubin Observatory Project Office, 950 North Cherry Avenue, Tucson, AZ 85719, USA*
[13]*Astrophysics Research Institute, Liverpool John Moores University, IC2, Liverpool Science Park, Liverpool L3 5RF, UK*
[14]*Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg (ObAS), UMR 7550, F-67000 Strasbourg, France*
[15]*Instituto de Investigación Multidisciplinar en Ciencia y Tecnología, Universidad de La Serena, Raúl Bitrán 1305, La Serena, Chile*
[16]*Departamento de Astronomía, Universidad de La Serena, Av. Juan Cisternas 1200 Norte, La Serena, Chile*
[17]*Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Macul, Santiago, Chile*
[18]*Instituto de Astronomía, Universidad Nacional Autónoma de México A.P. 70-264, 04510 México, D.F., México*
[19]*Department of Astronomy, Yonsei University Observatory, Yonsei University, Seoul 03722, Republic of Korea*
[20]*Centre for Astrophysics Research, Department of Physics, Astronomy and Mathematics, University of Hertfordshire, Hatfield AL10 9AB, UK*
[21]*Instituto de Astrofísica de Canarias, Vía Láctea S/N, E-38205 La Laguna, Spain*
[22]*Departamento de Astrofísica, Universidad de La Laguna, E-38206 La Laguna, Spain*
[23]*School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK*
[24]*Department of Physics, Lancaster University, Lancaster LA1 4YB, UK*
[25]*Department of Physics and Astronomy, UCLA, PAB 430 Portola Plaza, Los Angeles, CA 90095-1547, USA*
[26]*Department of Physics and Astronomy, University of California, Davis, One Shields Ave, Davis, CA 95616, USA*
[27]*Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy, Northwestern University, 1800 Sherman Ave, Evanston, IL 60201, USA*
[28]*Institut d'Astrophysique de Paris, Sorbonne Universités, UMPC Univ Paris 06 et CNRS, UMP 7095, 98 bis bd Arago, F-75014 Paris, France*
[29]*School of Physics, Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul 02455, Republic of Korea*
[30]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK*
[31]*IPAC, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA*
[32]*Departamento de Física Teórica, Atómica y Óptica, Universidad de Valladolid, E-47011 Valladolid, Spain*
[33]*Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Tapada da Ajuda, P-T1349-018 Lisbon, Portugal*
[34]*INAF - Catania Astrophysical Observatory, Via S. Sofia, 78, I-95123 Catania, Italy*
[35]*Departamento de Astronomía, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Concepción, Chile*
[36]*Universidad Internacional de Valencia. Carrer del Pintor Sorolla 21, E-46002 Valencia, Spain*
[37]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia*
[38]*Department of Physics and Astronomy, University of Louisville, Natural Science Building 102, Louisville, KY 40292, USA*
[39]*Astrophysics Research Centre, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa*
[40]*Department of Physics and Astronomy, The University of Alabama, Tuscaloosa, AL 35487, USA*
[41]*Minnesota State University, Mankato; Trafton North 141 Mankato, MN 56001, USA*
[42]*Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411007, India*
[43]*OzGrav-Swinburne, Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia*
[44]*Departamento de Física, Facultad de Ciencias, Universidad Nacional Autónoma de México, Ciudad Universitaria, CDMX 04510, México*
[45]*National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22033, USA*

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

Downloaded from https://academic.oup.com/mnras/article/513/1/1459/6566363 by CNRS user on 07 April 2023