





# Characterizing tidal features around galaxies in cosmological simulations

A. Khalid <sup>1,2★</sup>, S. Brough <sup>1,2</sup>, G. Martin <sup>3,4,5</sup>, L. C. Kimmig,<sup>6</sup> C. D. P. Lagos,<sup>2,7,8</sup> R.-S. Remus<sup>6</sup>  
and C. Martinez-Lombilla <sup>1,2,9</sup>

<sup>1</sup>*School of Physics, University of New South Wales, Sydney, NSW 2052, Australia*

<sup>2</sup>*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia*

<sup>3</sup>*School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD, UK*

<sup>4</sup>*Korea Astronomy and Space Science Institute, 776 Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Korea*

<sup>5</sup>*Steward Observatory, University of Arizona, 933 N. Cherry Ave, Tucson, AZ 85719, USA*

<sup>6</sup>*Fakultät für Physik, Universitäts-Sternwarte München, LMU München, Scheinerstr 1, D-81679 München, Germany*

<sup>7</sup>*International Centre for Radio Astronomy Research, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*

<sup>8</sup>*Cosmic Dawn Center (DAWN), Denmark*

<sup>9</sup>*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia*

Accepted 2024 April 16. Received 2024 April 15; in original form 2023 September 29

## ABSTRACT

Tidal features provide signatures of recent mergers and offer a unique insight into the assembly history of galaxies. The Vera C. Rubin Observatory’s Legacy Survey of Space and Time (LSST) will enable an unprecedentedly large survey of tidal features around millions of galaxies. To decipher the contributions of mergers to galaxy evolution it will be necessary to compare the observed tidal features with theoretical predictions. Therefore, we use cosmological hydrodynamical simulations NEWHORIZON, EAGLE, ILLUSTRITNG, and MAGNETICUM to produce LSST-like mock images of  $z \sim 0$  galaxies ( $z \sim 0.2$  for NEWHORIZON) with  $M_{\star, 30 \text{ pkpc}} \geq 10^{9.5} M_{\odot}$ . We perform a visual classification to identify tidal features and classify their morphology. We find broadly good agreement between the simulations regarding their overall tidal feature fractions:  $f_{\text{NEWHORIZON}} = 0.40 \pm 0.06$ ,  $f_{\text{EAGLE}} = 0.37 \pm 0.01$ ,  $f_{\text{TNG}} = 0.32 \pm 0.01$ , and  $f_{\text{MAGNETICUM}} = 0.32 \pm 0.01$ , and their specific tidal feature fractions. Furthermore, we find excellent agreement regarding the trends of tidal feature fraction with stellar and halo mass. All simulations agree in predicting that the majority of central galaxies of groups and clusters exhibit at least one tidal feature, while the satellite members rarely show such features. This agreement suggests that gravity is the primary driver of the occurrence of visually identifiable tidal features in cosmological simulations, rather than subgrid physics or hydrodynamics. All predictions can be verified directly with LSST observations.

**Key words:** galaxies: clusters: general – galaxies: evolution – galaxies: groups: general – galaxies: interactions – galaxies: structure.

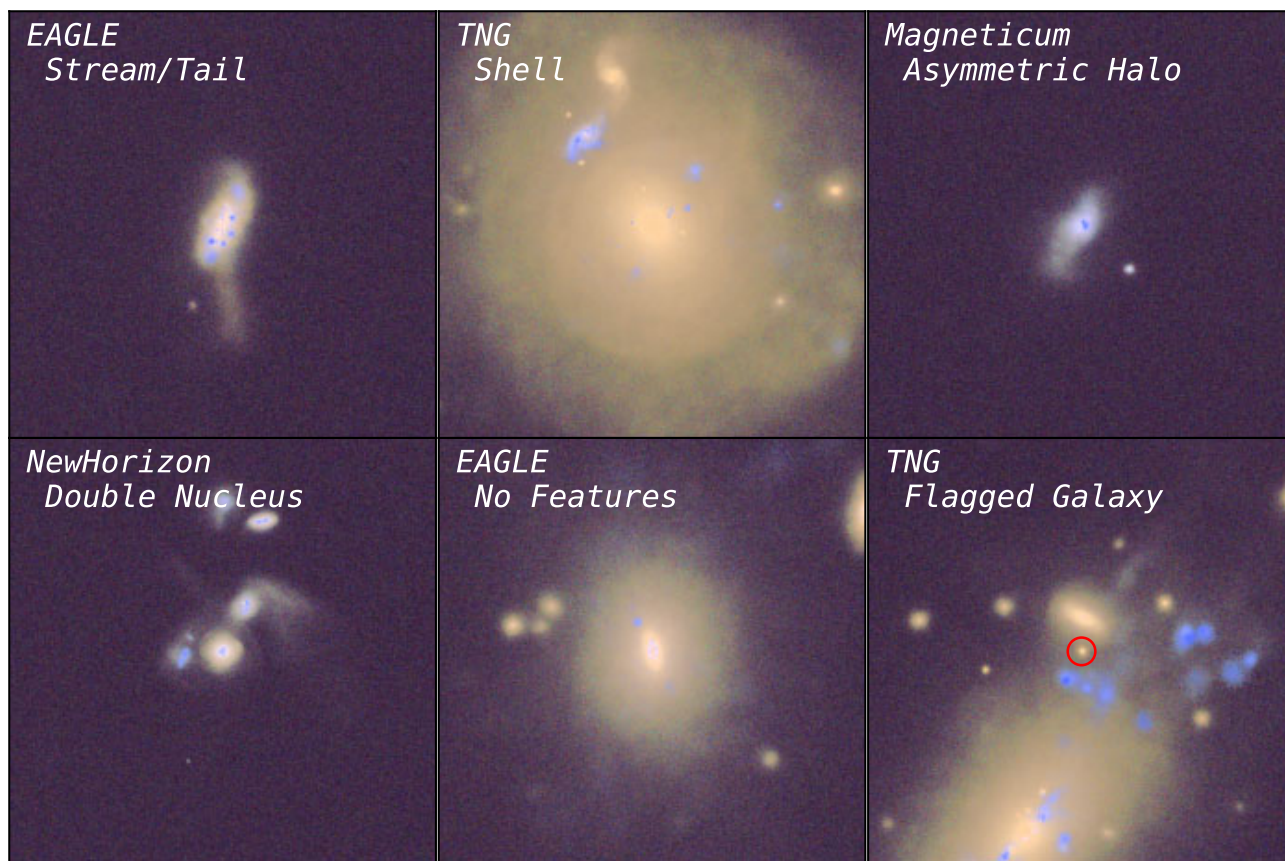
## 1 INTRODUCTION

In the ‘hierarchical structure formation’ model of the Universe (e.g. Press & Schechter 1974; Fall & Efstathiou 1980; Ryden & Gunn 1987; van den Bosch 2002; Agertz, Teyssier & Moore 2011), mergers play an important role in the evolution of galaxies, transforming galaxies through a variety of non-secular processes e.g. merger-driven star formation, merger-enhanced active galactic nuclei (AGNs) activity, dynamical evolution, and stellar mass accretion (e.g. Dubois et al. 2016; Martin et al. 2018, 2021; Davison et al. 2020; Remus & Forbes 2022; Cannarozzo et al. 2023). However, the intricacies of this picture, regarding the detailed merger statistics of galaxies, remain unresolved. There remain untested theoretical predictions from cosmological simulations. For example, Naab & Ostriker (2009) and Martin et al. (2018) predicted that low-mass mergers are the major contributor to galaxy morphological evolution since  $z \sim 1$ . There are also detailed predictions regarding the

environmental dependence of galaxy interaction and merger rates from cosmological  $N$ -body simulations (e.g. Gnedin 2003; Mihos 2003) and semi-analytic models (Jian, Lin & Chiueh 2012), where the galaxy encounter rates increase with increasing halo mass to a peak at halo masses corresponding to galaxy groups before dropping off at halo masses above that, corresponding to large galaxy clusters. To test these predictions we need to better understand galaxy mergers.

There are several ways to measure galaxy mergers observationally. These include detecting visible signatures of ongoing and past mergers via the presence of tidal features and selecting close pairs of galaxies that are likely to merge in the near future. Tidal features are diffuse non-uniform regions of stars that extend out from a galaxy, signatures of preceding, and concluded mergers in the forms of ‘tails’, ‘streams’, ‘asymmetric haloes’, ‘double nuclei’, and ‘shells’ (examples in Fig. 1). Observations and numerical simulations have found these features to have lifetimes of  $\sim 3$  Gyr (Ji, Peirani & Yi 2014; Mancillas et al. 2019; Yoon & Lim 2020; Huang & Fan 2022), making tidal features crucial probes for a galaxy’s recent merger history. Tidal feature detection requires very deep images to pick up signatures of ongoing and past mergers (e.g. Atkinson,

\* E-mail: [aman.khalid@unsw.edu.au](mailto:aman.khalid@unsw.edu.au)



**Figure 1.** Examples of *gri*-colour mock images from each simulation in our sample. Each of the tidal feature morphologies is illustrated from left to right, top to bottom: stream/tail, shell, asymmetric halo, double nuclei, and featureless. The galaxy marked by a red circle in the bottom right image is an example of one that we removed from the sample. While there is enough stellar mass within the 30 pkpc aperture for the object to be in our sample, the object itself is too dense and compact to be an accurate representation of a galaxy of  $M_{*,30 \text{ pkpc}} \geq 10^{9.5} M_{\odot}$ . Each image is 124 pkpc  $\times$  124 pkpc ( $\sim 2400$  pixels  $\times$  2400 pixels).

Abraham & Ferguson 2013; Kado-Fong et al. 2018; Martin et al. 2022). Close pairs are detected most accurately using spectroscopic observations to measure robust 3D positions for each galaxy (e.g. Robotham et al. 2014; Banks et al. 2021). Close pair detection is important in understanding the role mergers play in driving galaxy evolution; however, this method does have limitations. It cannot detect the presence of a secondary galaxy if that galaxy has been ripped apart by tidal forces; absorbed into a host galaxy; or the secondary galaxy is of too low mass to detect via spectroscopy (e.g. Lotz et al. 2011; Desmons et al. 2023b). Studying the tidal features of a galaxy provides insight into these aspects of the merging process.

The formation pathways of tidal features have been studied using *N*-body and hydrodynamic simulations to obtain insight into the merger histories of observed galaxies that host tidal features. These simulations have shown that tidal features are tracers of galaxy interactions and mergers. They probe detailed properties of these interactions (e.g. dynamics, mass ratios, and orbits). They have established that tail-like structures (top left panel Fig. 1) are formed from high angular momentum passages between similar mass galaxies (common in major mergers), while streams form from almost circularly infalling lower mass satellite galaxies (minor mergers; e.g. Toomre & Toomre 1972; Hendel & Johnston 2015; Karademir et al. 2019).

Various formation scenarios have been proposed to explain the formation of shells, the scenario most supported by analytical and numerical models is formation predominantly through radial mergers

(e.g. Amorisco 2015; Hendel & Johnston 2015; Pop et al. 2018; Karademir et al. 2019; Valenzuela & Remus 2023), resulting in overdensities of stripped stars accumulating at the apocentres of their orbits. Simulations have linked the properties of shells to the merger dynamics and mass ratios of the progenitor galaxies. *N*-body studies such as Hernquist & Spergel (1992) have found that shells formed from major mergers tend to be preferentially aligned with the major axis of the host galaxy. Pop et al. (2018) found that most of the  $z = 0$  shell population of the hydrodynamic-cosmological simulation ILLUSTRIS were formed predominantly through major mergers, suggesting that due to dynamical friction, pairs of more massive galaxies could probe a greater range of impact parameters (deviating further from a purely radial infall) and still form shells when compared to more minor mergers. These simulations provide strong evidence for shell formation occurring through mergers with more radial trajectories than those that form tails and streams.

The understanding of tidal feature formation from simulations is used to interpret observational studies of the characteristics of tidal features such as their colours, dynamics, appearance, and the environment in which the tidal feature host resides. These are then used to constrain the properties of the interaction that caused them. For example Foster et al. (2014) and Martínez-Delgado et al. (2021) used spectroscopic and photometric data, respectively, in conjunction with numerical simulations, to study the formation of tidal features around the NGC 4651 (Foster et al. 2014) and M104 (Martínez-Delgado et al. 2021) galaxies and constrain their recent merger

**Table 1.** Summary of the properties of the four cosmological hydrodynamical simulations. From left to right, the columns are the simulation name, matter density, dark energy density, Hubble constant, simulation box side length, stellar mass resolution, the redshift corresponding to the simulation time-step at which the particle data used were recorded, the number of galaxies in the simulations with  $M_{\star, 30 \text{ pkpc}} \geq 10^{9.5} M_{\odot}$ , the number of galaxies classified, number of compact point-like galaxies discarded, and final sample size for analysis.

Simulation	$\Omega_m$	$\Omega_{\Lambda}$	$H_0$ [km s <sup>-1</sup> Mpc <sup>-1</sup> ]	$V_{\text{box}}$ [cMpc <sup>3</sup> ]	$m_{\star}$ [ $M_{\odot}$ ]	$z$	$N_{\text{gals}}$	$N_{\text{classified}}$	$N_{\text{out}}$	$N_{\text{sample}}$
NEWHORIZON	0.272	0.728	70.4	16 <sup>3</sup>	$1.3 \times 10^4$	0.260, 0.263	105	62	0	62
EAGLE REFL0100N1504	0.307	0.693	67.8	100 <sup>3</sup>	$1.8 \times 10^6$	0.05	7273	2000	22	1978
TNG L75N1820TNG	0.309	0.691	67.7	111 <sup>3</sup>	$1.4 \times 10^6$	0.05	12 418	1907	81	1826
MAGNETICUM PATHFINDER BOX4-UHR	0.272	0.728	70.4	68 <sup>3</sup>	$2.6 \times 10^6$	0.05	4587	2000	10	1990

histories. Foster et al. (2014) were able to infer the recent passage of a satellite galaxy through the disc of NGC 4651 from their modelling of an observed stellar stream. Martínez-Delgado et al. (2021) were able to estimate the time around which the major wet merger forming M104 occurred. These particular examples illustrate the strength of tidal features in reconstructing merger histories of observed galaxies and hint at the potential of using tidal features to analyse larger statistical populations of galaxies to provide a detailed understanding of the merger process in our Universe.

One of the largest observational studies of tidal features to date, Kado-Fong et al. (2018), analysed the tidal features in  $0.05 < z < 0.45$  galaxies in the HSC-SSP (Hyper Suprime-Cam Subaru Strategic Program) Wide layer. They found evidence suggesting that shells are formed predominantly through minor mergers.

Cosmological hydrodynamical simulations enable the study of large statistically significant samples of galaxies to predict the dominant pathways of merger signatures in the Lambda cold dark matter (ΛCDM) model of hierarchical structure formation. Observations provide us with snapshots of our Universe at various lookback times, giving us the ground truth to which we can compare our models. Studying both observations and simulations enables us to understand the merger histories that are inferred from observations.

With the upcoming Vera C. Rubin Observatory’s Legacy Survey of Space and Time (LSST; Ivezić et al. 2019; Robertson et al. 2019; Brough et al. 2020) it will be possible to study tidal features around millions of galaxies (Martin et al. 2022), allowing for the most robust statistical survey of tidal features to date. To fully capitalize on this unprecedented wealth of data we will need robust predictions of the properties of these features from current hydrodynamic-cosmological simulations. In this study, we will characterize the tidal features around galaxies in the four hydrodynamic-cosmological simulations (described in Section 2.1), NEWHORIZON (Dubois et al. 2021), ILLUSTRISTNG (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Springel et al. 2018), EAGLE (Crain et al. 2015; Schaye et al. 2015), and MAGNETICUM (Teklu et al. 2015), in an observationally motivated approach that facilitates future comparison to LSST. We do this by visually classifying the tidal features present in LSST-like mock images, produced following Martin et al. (2022). We study four simulations to probe whether the different subgrid physics models applied to each simulation result in different merger pathways and whether this presents differently in the tidal features of galaxies. We compare the limited observations made to date to test how well the simulations are able to replicate them.

Our sample selection and production of mock images are described in Sections 2.2 and 2.3, respectively. Our methodology for visual classification, including the tidal feature morphologies considered, is given in Section 2.4 and we describe the methodology for calculating the errors on our tidal feature fractions in Section 2.5. We present a statistical analysis of our visual classifications, considering the

occurrence rates of tidal features in each simulation in Section 3 and then present how the fraction of galaxies exhibiting tidal features changes with both stellar mass (in Section 3.1) and halo mass (in Section 3.2). We then discuss our results with respect to past analyses of simulations and observational results in Sections 4.1 and 4.2, respectively. The implications of our visual classification method are addressed and discussed in Section 4.3. In Section 4.4 we consider the implications of our results on the occurrence and observability of tidal features across the cosmological hydrodynamical simulations that we study. We discuss the relationship between tidal feature occurrence in the hydrodynamic-cosmological simulations we study and the environment in which the galaxies reside in Section 4.5. In Section 4.6 we explore the predictions for LSST regarding tidal feature fractions and their trends with stellar and halo mass based on the hydrodynamic-cosmological simulations studied here. We draw our conclusion in Section 5.

We use the native cosmology from each simulation for calculating the distances between particles and creating our mock images; these are given in Table 1. For distances we use a ‘c’ prefix to denote comoving coordinates and a ‘p’ prefix to denote proper coordinates, e.g. cMpc is comoving megaparsecs and pMpc is proper megaparsecs.

## 2 SIMULATIONS AND METHODS

### 2.1 Simulations

To investigate the characteristics of tidal features in cosmological hydrodynamical simulations we explored four different simulations. These simulations evolve gas, stars, and dark matter under gravity and hydrodynamics along with a range of models for subgrid processes to simulate the realistic hierarchical structure formation of galaxies.

For our study we selected four simulations: NEWHORIZON, EAGLE REFL0100N1504, ILLUSTRISTNG L75N1820TNG, MAGNETICUM PATHFINDER BOX4-UHR (from here on, NEWHORIZON, EAGLE, TNG, and MAGNETICUM). These four simulations, summarized in Table 1, allow us to probe tidal features across a range of different resolutions, galaxy environments, and subgrid physics models. As we approach the advent of large observational surveys of galaxies within the sensitivity to observe tidal features, it is important to characterize what each of these simulations predicts concerning these features. TNG and EAGLE probe a similar volume and therefore the same range of environments (isolated galaxies, groups, and clusters). The MAGNETICUM simulation probes an intermediate-sized box about one-third the volume of the two larger simulations and has a slightly lower stellar mass resolution. NEWHORIZON probes a much smaller volume with a much higher stellar mass resolution, allowing for more detailed tidal structures to be resolved by probing a more limited sample of galaxies.

In this section we introduce each of the simulations, describing their mass resolution, volume, and how they were calibrated. We discuss how galaxy structure, stellar mass, and halo mass are measured in Section 2.1.1 and the simulations' stellar mass–halo mass (SMHM) relations in Section 2.1.2. We focus on these elements of the simulations as they are the salient differences for this study, directly impacting the number of stellar particles available to form tidal features and the halo masses that are covered by the simulations. The simulation subgrid models for star formation, stellar feedback, and black hole feedback, which could indirectly contribute to the factors above and therefore the tidal feature occurrence, are described in Appendix A.

The NEWHORIZON (Dubois et al. 2021) simulation is a zoom-in simulation from the parent HORIZON-AGN simulation (Dubois et al. 2014), which adopts the WMAP-7 cosmology (Komatsu et al. 2011). NEWHORIZON employs a spherical volume of  $16^3$  cMpc<sup>3</sup> with varying dark matter resolution. The initial 10 cMpc radius has  $m_{\text{DM}} = 1.2 \times 10^6 M_{\odot}$ ; this high-resolution patch is embedded in spheres of decreasing mass resolution of  $10^7$ ,  $8 \times 10^7$ , and  $6 \times 10^8 M_{\odot}$  corresponding to spheres of radius 10.6, 11.7, and 13.9 cMpc. The remaining volume is resolved at  $5 \times 10^9 M_{\odot}$  (Martin et al. 2022). The stellar mass resolution is  $m_{\star} = 1.3 \times 10^4 M_{\odot}$ . Each stellar particle represents a simple stellar population with a Chabrier initial mass function (IMF; Chabrier 2003) in the range 0.1–150  $M_{\odot}$ .

HORIZON-AGN is calibrated to match the black hole density and the scaling relations between black hole mass and galaxy properties (albeit at a lower spatial resolution to NEWHORIZON). The NEWHORIZON simulation also reproduces the observed galaxy surface brightness as a function of stellar mass (Dubois et al. 2021), without requiring any calibration for surface brightness.

Table 1 shows that NEWHORIZON has stellar particles approximately two orders of magnitude less massive than the other three simulations allowing it to resolve tidal features from higher mass ratio mergers. However, NEWHORIZON only covers field and group environments and does not include galaxy clusters of significant mass (Dubois et al. 2021).

The EAGLE (Evolution and Assembly of GaLaxies and their Environments) project is a large set of cosmological hydrodynamical simulations. EAGLE contains simulations of cubic volume  $12^3$ ,  $25^3$ ,  $50^3$ , and  $100^3$  cMpc<sup>3</sup>. The data which will be used in this study are from the reference model REFLO100N1504 which has a volume of  $100^3$  cMpc<sup>3</sup>. The reference model is the simulation run with standard parameters and physics as described in Schaye et al. (2015) and Crain et al. (2015). The simulation uses the cosmological parameters advocated by the Planck 2013 results (Planck Collaboration XVI 2014). The dark matter particle mass is  $m_{\text{DM}} = 9.70 \times 10^6 M_{\odot}$  and the stellar particle mass is  $m_{\star} = 1.81 \times 10^6 M_{\odot}$ .

The simulation was calibrated to match the galaxy stellar mass function (GSMF) and the relationship between black hole mass and galaxy stellar mass at  $z \sim 0$  (details in Schaye et al. 2015).

ILLUSTRISTNG is a suite of simulations, with the following box volumes:  $51.7^3$ ,  $110.7^3$ , and  $302.6^3$  cMpc<sup>3</sup>. We use simulation L75N1820TNG, which models the physics of dark matter and baryons in a  $110.7^3$  cMpc<sup>3</sup> box. This simulation adopts the  $\Lambda$ CDM model fit by the Planck 2015 results (Ade et al. 2016). The dark and baryonic particle resolutions for TNG100-1 are  $m_{\text{DM}} = 7.5 \times 10^6 M_{\odot}$  and  $m_{\star} = 1.4 \times 10^6 M_{\odot}$ .

TNG is calibrated to match the star formation rate density, the GSMF, the stellar-to-halo mass relation, the total gas mass contained within the virial radius of massive groups, the stellar mass–stellar size, and the black hole–galaxy mass relations all at  $z = 0$ , in addition

to the overall shape of the cosmic star formation rate density at  $z \lesssim 10$  (Pillepich et al. 2018b).

MAGNETICUM PATHFINDER simulations are a suite of cosmological hydrodynamical simulations, ranging in box size from  $25.6^3$  to  $3818^3$  cMpc<sup>3</sup>. We use the BOX4-UHR simulation, which models the physics of dark matter and baryons in a  $68^3$  cMpc<sup>3</sup> box. All simulations are performed with an updated version of the smooth particle hydrodynamic (SPH) code GADGET-3 (Springel et al. 2005b). We use the Box4-uhf run, which adopts a WMAP-7 fit cosmology (Komatsu et al. 2011). The mass resolution for stellar particles in this run is  $m_{\star} = 2.6 \times 10^6 M_{\odot}$  and  $m_{\text{DM}} = 3.6 \times 10^7 M_{\odot}$  for dark matter particles.

The MAGNETICUM simulations are calibrated to match the intracluster gas content of galaxy clusters, rather than matching any characteristics with individual galaxies.

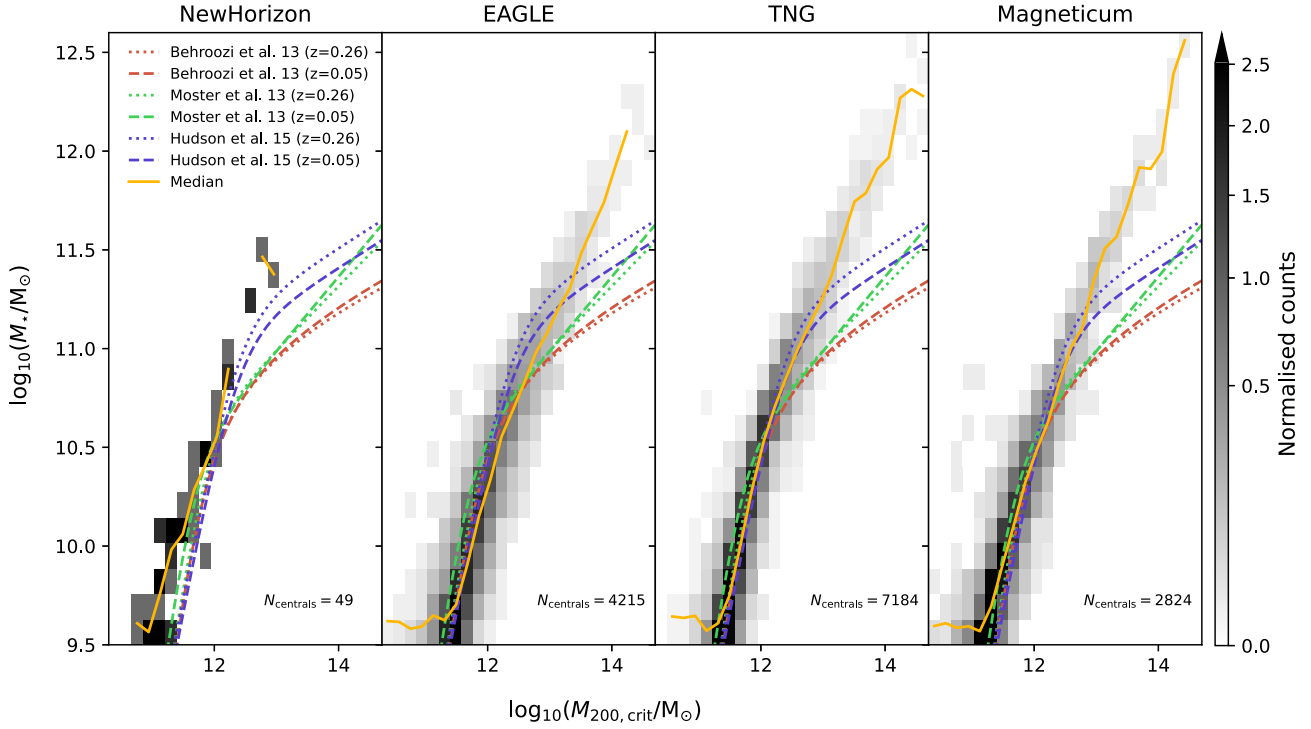
### 2.1.1 Identifying structure and measuring stellar mass

To study the properties of galaxies and their tidal features the particles in the simulations must be grouped into central galaxies and their satellites. For TNG, EAGLE, and MAGNETICUM, the galaxy and halo structure is computed using the baryonic version of the SUBFIND code (Springel; Yoshida & White 2001; Dolag et al. 2009). The versions of structure finder used across these simulations are similar but not identical. In contrast, NEWHORIZON utilizes ADAPTAHOP for finding haloes and subhaloes (Tweed et al. 2009). It is important to note that the stellar masses shown in Figs 2 and 3 are computed directly from each simulation's native structure finders. They are computed by summing the stellar masses of the particles assigned to each subhalo.

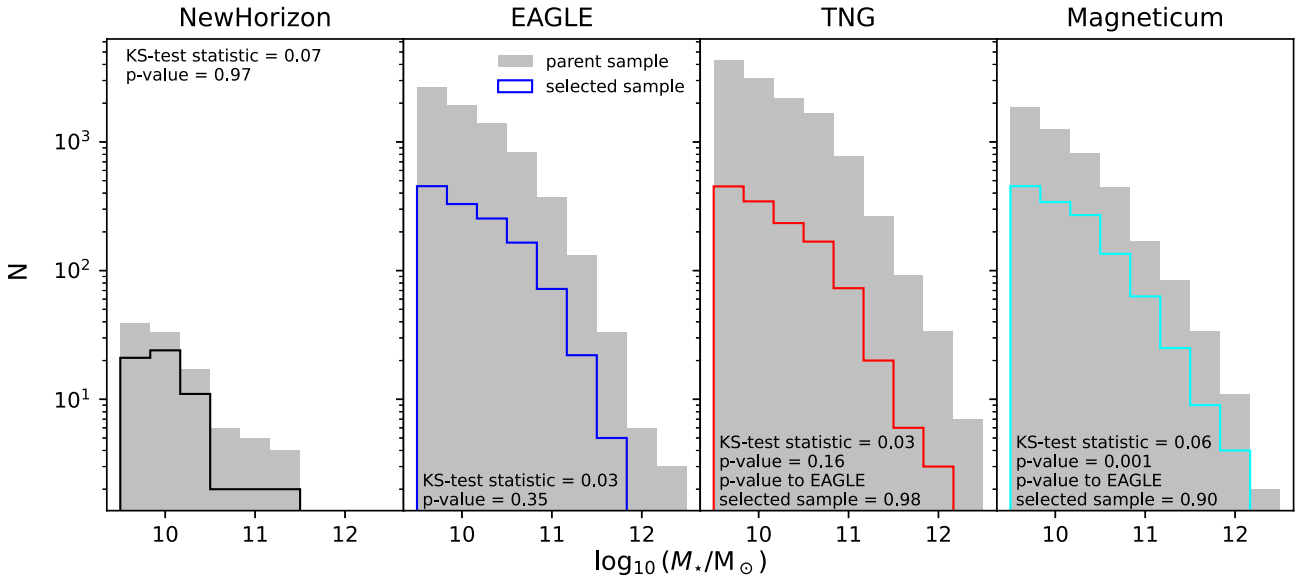
In general, SUBFIND and ADAPTAHOP differ in how they assign particles to haloes and subhaloes: SUBFIND identifies substructure based on its local overdensity and gravitational boundedness. In its initial step SUBFIND uses a Friends-of-Friends (FOF) algorithm on dark matter particles to distribute them into haloes; the baryons are then assigned to the halo of the nearest dark matter particle (if any). Substructures within each FOF halo are identified by searching for local density peaks, considering all particle types and resolution elements. Subhaloes are identified using contours of isodensity that pass through limiting saddle points in the density field. Particles/resolution elements not gravitationally bound to the subhalo to which they have been assigned are removed by applying an iterative unbinding procedure. Finally, all particles and resolution elements that are unassigned to a subhalo are assigned to the central subhalo (the largest subhalo within a halo). One notable limitation of the SUBFIND method is that any resolution element/particle lying beyond the limiting isodensity contour is ignored, even if they are gravitationally bound to a subhalo (e.g. Muldrew, Pearce & Power 2011; Cañas et al. 2019).

ADAPTAHOP identifies substructure based on topology and does not carry out any unbinding. Particles are assigned to groups defined by peaks in the density field that are linked to other groups by saddle points in the density field. The substructure is then assigned by repeating this process with an increasing density threshold, resulting in smaller groups within groups. ADAPTAHOP defines haloes as a group of groups that exceeds 160 times the mean dark matter density. Groups within each halo are regrouped such that each subhalo has a mass smaller than the host halo.

The differences between the algorithms are non-trivial and will lead to differences in the stellar masses of galaxies. ADAPTAHOP tends to identify fewer structures than FOF, i.e. several FOF haloes are included in one ADAPTAHOP halo; these differences can in-turn



**Figure 2.** The galaxy SMHM relation for all central galaxies within our stellar mass range for each of our simulations; the sample sizes are given in the bottom right corner. The relation is plotted using 25 equally spaced bins for  $9.5 \leq \log_{10}(M_*/M_\odot) \leq 12.6$  and  $10.3 \leq \log_{10}(M_{200, \text{crit}}/M_\odot) \leq 14.7$ . The dashed and dotted lines show observationally derived relations from the literature for redshifts comparable to the simulation snapshots; the dotted lines correspond to  $z = 0.26$  (NEWHORIZON snapshot redshift) and the dashed lines correspond to  $z = 0.05$  (EAGLE, TNG, and MAGNETICUM snapshot redshifts). The median relation for central galaxies in each simulation is illustrated with an amber line. On average we see that the relations for EAGLE, TNG, and MAGNETICUM are similar; however, NEWHORIZON is missing high-mass haloes and contains more stars for a given halo mass than the other simulations. Note for this figure we present the ADAPTAHOP (for NEWHORIZON) and SUBFIND (for EAGLE, TNG, and MAGNETICUM) based galaxy stellar masses ( $M_*$ ) for the galaxies. The counts have been normalized such that the volume under the histogram integrates to 1.



**Figure 3.** Stellar mass distributions of parent (grey) and selected samples (NEWHORIZON – black, EAGLE – blue, TNG – red, and MAGNETICUM – cyan) for each simulation. The  $M_* \geq 10^{9.5} M_\odot$  range explored by our study is well sampled by the parent distribution for TNG, EAGLE, and MAGNETICUM, but the smaller volume NEWHORIZON simulation has fewer galaxies at the high-mass end. We include the results of the KS tests comparing the sample to the parent distributions. EAGLE roughly samples its parent distribution and TNG and MAGNETICUM have been sampled to match EAGLE. The KS test results show that TNG remains similar to its parent, whereas MAGNETICUM no longer resembles the parent distribution. Note for this figure we present the ADAPTAHOP (for NEWHORIZON) and SUBFIND (for EAGLE, TNG, and MAGNETICUM) based galaxy stellar masses ( $M_*$ ) for the parent and subsampled galaxies.

influence how SUBFIND assigns subhaloes compared to ADAPTAHOP (Tweed et al. 2009). We only use galaxy stellar masses derived from the simulation’s native structure finding algorithm in Figs 2 and 3, where we compare our visually classified sample’s stellar mass distribution to the parent stellar mass distribution in the simulation. In light of the differences between the algorithms we state that while the stellar mass distributions of both parent and child samples should not be compared trivially across the simulations, we feel comfortable in using the distributions to check how well our classified samples probe the stellar mass distributions of the simulations above a stellar mass of  $10^{9.5} M_{\odot}$ . We also use the subhalo centres of potential as defined by SUBFIND and ADAPTAHOP for the centres of our LSST-like mock images. Additionally, we rely on the halo assignment when loading particles to produce mock images as we only use particles assigned to the same halo; for high halo masses (ADAPTAHOP halo mass  $\geq 10^{13} M_{\odot}$ ), FOF tends to assign  $\gtrsim 2$  times as many haloes to a given collection of particles as ADAPTAHOP (Tweed et al. 2009). We do not probe such high halo masses with NEWHORIZON therefore the differences in halo assignment will not have a significant impact.

From Section 3.1 onwards, when comparing the occurrence of tidal features as a function of stellar mass, to limit systematic biases between ADAPTAHOP and SUBFIND derived masses we use the 30 pkpc spherical aperture,  $M_{*, 30\text{pkpc}}$ , to measure the stellar masses. As discussed in Schaye et al. (2015), the 30 pkpc spherical aperture is comparable to the 2D Petrosian aperture often used in observational studies, making it a good choice for our aim to analyse the data in a manner directly comparable to observations. The spherical aperture measurement of stellar masses requires only the stellar particles assigned to a subhalo/galaxy within a 30 pkpc sphere centred on the subhalo’s centre of potential to be included for the galaxy’s stellar mass.

In Section 3.2, for our analysis of tidal feature fractions as a function of halo mass, we rely on the halo assignment from each simulation’s native structure finding algorithms to define main haloes, assign galaxies to main haloes, and the consequent measurement of  $M_{200, \text{crit}}$  for the main halo.  $M_{200, \text{crit}}$  is the total mass, including dark matter, gas, and stars within the radius where the average density is 200 times the critical density of the universe. For TNG, EAGLE, and MAGNETICUM, which rely on SUBFIND, the calculation of  $M_{200, \text{crit}}$  only relies on FOF for the identification of the halo’s centre; the calculation of its mass is independent of the halo assignment and includes all particles within the volume. In contrast, NEWHORIZON’s measurement of  $M_{200, \text{crit}}$  exclusively uses the resolution elements assigned to a particular halo by ADAPTAHOP. While each subhalo is assigned to an FOF identified host halo in SUBFIND, ADAPTAHOP’s use of different density thresholds to separate halo and subhalo structure and the application of minimum density of 160 times the mean dark matter density for a structure containing substructures to be identified as a group leads to some subhaloes in NEWHORIZON not being assigned to any parent haloes.

### 2.1.2 SMHM relation

The number of stars involved in interactions between haloes determines the number of stars available to form tidal features. Therefore, it is important to consider the efficiency of star formation at a given halo mass for each of the simulations.

In Fig. 2 we show the SMHM distributions for each simulation, displaying all central galaxies within our stellar mass cut (sample sizes given by  $N_{\text{centrals}}$  in Fig. 2). For comparison, we provide the observationally constrained empirical SMHM relations from Behroozi, Wechsler & Conroy (2013, constrained by stellar mass function, cosmic star formation rate, and specific star formation

rate) and Moster, Naab & White (2013, constrained by the stellar mass function) and the weak lensing derived relation from Hudson et al. (2015) and give the median SMHM relation for each of the simulations. We note that the median relation for each simulation appears to flatten for  $M_{200, \text{crit}} \lesssim 10^{11} M_{\odot}$ . This is driven by our lower stellar mass cut meaning that in this regime we only sample high-mass haloes due to not including the scatter to lower stellar masses (e.g. Teklu et al. 2017).

NEWHORIZON qualitatively reproduces the shape of the observational relations for the halo mass range of the simulation. Relative to the empirical relations NEWHORIZON does deviate to higher stellar masses below the knee of the SMHM relation. This is at least partly due to the smaller simulation volume (Martin et al. 2022), resulting in underrepresentation of group and cluster environments where star formation tends to be less efficient (e.g. Garrison-Kimmel et al. 2019; Samuel et al. 2022). Another contributing factor to the higher star formation efficiency of NEWHORIZON could be the subgrid physics implemented (described in Appendix A). We also note that recent empirical studies of the low-mass end of the SMHM relation, constrained by the star formation rates of Local Group dwarf galaxies, have shown significant scatter to higher stellar masses for  $M_{200, \text{crit}} \lesssim 10^{10} M_{\odot}$ , which is qualitatively similar to what is observed for NEWHORIZON (O’Leary et al. 2023).

EAGLE, TNG, and MAGNETICUM fall within the scatter of the observationally constrained relations for  $10^{11} M_{\odot} \lesssim M_{200, \text{crit}} \lesssim 10^{12} M_{\odot}$ . For  $M_{200, \text{crit}} \gtrsim 10^{13.5} M_{\odot}$  for EAGLE and  $M_{200, \text{crit}} \gtrsim 10^{12.5} M_{\odot}$  for TNG and MAGNETICUM we see that the centrals are sitting at systematically higher stellar masses than the observational relations. A factor contributing to this difference could be the differences between stellar mass measurements in simulations and observations. For example, Remus & Forbes (2022) found better agreement between the cosmological simulations EAGLE, MAGNETICUM, and TNG-300 and observational measurements of the SMHM relation (Kravtsov, Vikhlinin & Meshcheryakov 2018) when the light from the central galaxy and the intracluster light (ICL) around it was included in the observations. This deviation between simulations, observations, and models has been known for a while. Kravtsov et al. (2018) showed that the observed SMHM relation deviates from the model predictions if the brightest cluster galaxy is measured out to large radii and their ICL is included. This leads in those cases to larger observed stellar masses at a given halo mass than predicted by Behroozi et al. (2013) and Moster et al. (2013). This is found to be true for many simulations, where the SMHM relation from models and observations that exclude the ICL is usually reproduced if aperture cuts are applied to the simulations, but results in larger stellar masses if the stellar masses from the halo finders are used, that is, the ICL and residuals from satellite galaxies are included in the stellar mass measurement of the brightest cluster galaxy (see for example Remus & Forbes 2022 for MAGNETICUM, Pillepich et al. 2018b for ILLUSTRITNG, Contreras-Santos et al. 2024 for THE THREE HUNDRED SIMULATIONS, and Schaye et al. 2023 for FLAMINGOS). Brough et al. (2024) recently demonstrated, by applying observational methods to separate brightest cluster galaxies from their ICL to mock images of simulations from MAGNETICUM, EAGLE, ILLUSTRITNG, and HORIZON-AGN, that the best agreement is found for a separation by a given aperture. This provides observational support for the previously artificial aperture choice that we use to measure stellar masses.

There are small offsets between the SMHM relation for EAGLE, TNG, and MAGNETICUM, albeit smaller than the offset with NEWHORIZON. EAGLE is shifted to lower stellar masses compared to the other simulations and MAGNETICUM shifted to marginally higher

stellar masses than TNG at the low and high end of the relation. A detailed comparison between the models was carried out by Remus & Forbes (2022). They suggest that the likely cause of this is the star formation efficiencies of low-mass haloes, where MAGNETICUM is observed to convert more gas into galaxies. As a result of its SMHM relation, we expect NewHorizon to bring in more stars into mergers than the other three simulations for mergers containing haloes of a given size.

## 2.2 Sample selection

We select all simulated galaxies with  $M_{\star, 30 \text{ pkpc}} \geq 10^{9.5} M_{\odot}$ . We make no distinction between centrals and satellites in our sampling, allowing the underlying distributions of centrals/satellites to carry over to our samples. At a stellar mass of  $10^{9.5} M_{\odot}$  MAGNETICUM resolves these galaxies with  $\sim 1200$  stellar particles, EAGLE with  $\sim 1750$ , TNG with  $\sim 2250$ , and NEWHORIZON with  $\sim 24\,300$ . With any lower stellar mass threshold, there are fewer particles representing galaxies and our mock images will struggle to resolve significant tidal distortions to the stellar components of the galaxies. We also have few observational results to compare to below this stellar mass.

We sampled our galaxies from similar redshift snapshots for TNG, EAGLE, and MAGNETICUM. However, for NEWHORIZON we used the two most recent available snapshots, taking care not to sample galaxies twice (Table 1). The number of galaxies with  $M_{\star, 30 \text{ pkpc}} \geq 10^{9.5} M_{\odot}$  in each simulation is given in Table 1.

From the total sample of galaxies with  $M_{\star, 30 \text{ pkpc}} \geq 10^{9.5} M_{\odot}$ , we randomly subsample 2000 galaxies to visually classify from each of MAGNETICUM and EAGLE and 1907 galaxies from TNG. This is done so that the visual classification can occur over a reasonable timeframe. Within these galaxies, we note the presence of several galaxies which seemed to be very compact, point-like, and have a stellar mass far higher than one would expect for such a compact object (see bottom right of Fig. 1). We discarded these compact, point-like objects from our sample; the number discarded is given under  $N_{\text{out}}$  in Table 1. Our final sample sizes are 62 for NEWHORIZON, 1978 for EAGLE, 1826 for TNG, and 1990 for MAGNETICUM.

To study how tidal features depend on stellar mass in each of the simulations it is necessary to match the stellar mass distributions between the classified simulation samples to avoid a mass-biased comparison. We subsample from the classified TNG and MAGNETICUM samples to match the stellar mass distribution of the parent EAGLE sample, which was calibrated to match the local GSMF (Furlong et al. 2015). Matching to EAGLE's stellar mass distribution therefore makes our results more directly comparable with observational samples of tidal features. However, for NEWHORIZON, we do not subsample, leaving the sample as it is as we do not wish to reduce the sample size any further.

To construct our stellar mass-matched sample, we fit the probability densities for the simulation parent stellar mass distributions using Gaussian kernel density estimation (KDE). Due to the unavailability of  $M_{\star, 30 \text{ pkpc}}$  measurements for the parent samples of TNG and MAGNETICUM we use SUBFIND-derived  $M_{\star}$  values to fit the KDE for the parent stellar mass distributions from EAGLE, TNG, and MAGNETICUM. Using these KDEs we infer the likelihoods of sampling galaxies with particular  $M_{\star, 30 \text{ pkpc}}$  from each of the simulations. We then use these likelihoods to construct weights with which we sample the classified TNG and MAGNETICUM galaxies to match the parent distribution of EAGLE. Our samples are constructed using the following steps:

- (i) We draw, without replacement,  $N$  galaxies from EAGLE.

- (ii) We draw, without replacement, a sample of  $N$  galaxies from TNG and MAGNETICUM, weighting the samples so that the likelihood of a galaxy being drawn is the same as it would be if we were sampling from the EAGLE parent stellar mass distribution.

- (iii) We redraw  $N$  galaxies from TNG and MAGNETICUM until the distributions reasonably match that of EAGLE, by requiring that a KS test (Hodges 1958) yield a  $p$ -value  $> 0.9$ .

- (iv) The largest sample that could be subsampled from TNG and MAGNETICUM while satisfying condition (iii) is  $N = 1300$ .

The exact weights we used for subsampling from our classified TNG and MAGNETICUM galaxies are as follows:

$$W_{\text{TNG}}(M_{\star, 30 \text{ pkpc}}) = P_{\text{EAG}}(M_{\star, 30 \text{ pkpc}}) / P_{\text{TNG}}(M_{\star, 30 \text{ pkpc}}) \quad (1)$$

$$W_{\text{MAG}}(M_{\star, 30 \text{ pkpc}}) = P_{\text{EAG}}(M_{\star, 30 \text{ pkpc}}) / P_{\text{MAG}}(M_{\star, 30 \text{ pkpc}}), \quad (2)$$

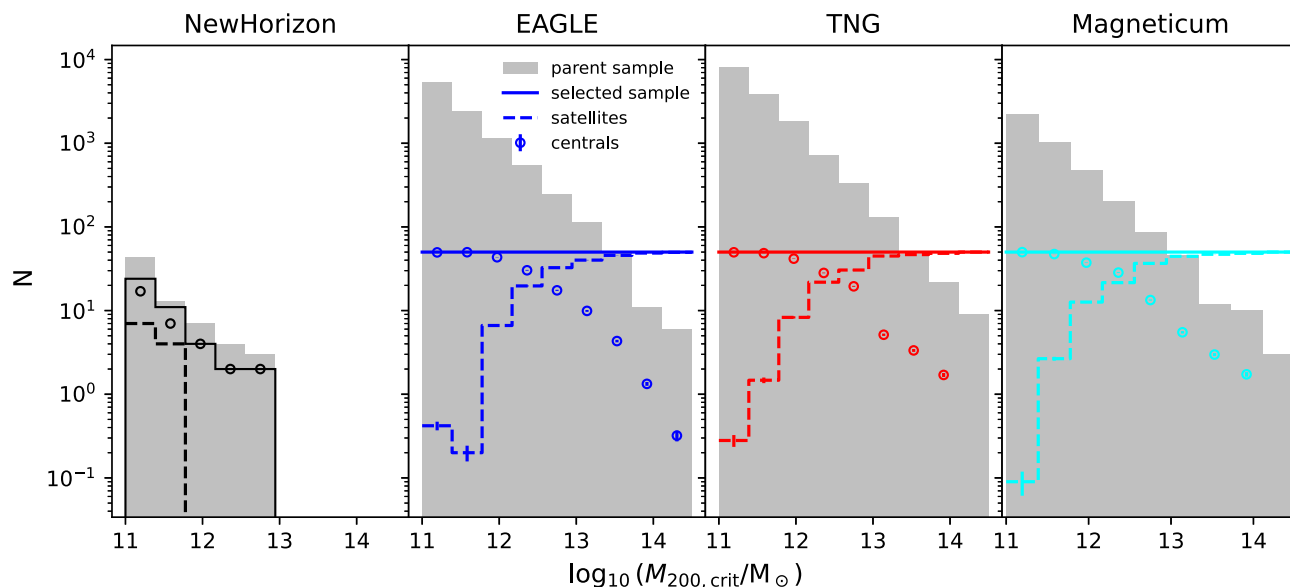
where  $W_{\text{TNG}}(M_{\star, 30 \text{ pkpc}})$  and  $W_{\text{MAG}}(M_{\star, 30 \text{ pkpc}})$  are the weights assigned to a TNG and MAGNETICUM galaxy, respectively, based on their  $M_{\star, 30 \text{ pkpc}}$ .  $P_{\text{EAG}}(M_{\star, 30 \text{ pkpc}})$ ,  $P_{\text{TNG}}(M_{\star, 30 \text{ pkpc}})$ , and  $P_{\text{MAG}}(M_{\star, 30 \text{ pkpc}})$  are the probability density functions from which we can calculate the likelihoods of galaxies with a particular  $M_{\star, 30 \text{ pkpc}}$  being sampled from the parent distribution of the respective simulations.

Fig. 3 shows the stellar mass distributions of the parent samples for each of the simulations in grey histograms and the distributions of our selected sample in coloured histograms, black for NEWHORIZON, blue for EAGLE, red for TNG, and cyan for MAGNETICUM. TNG and MAGNETICUM are mass-matched to EAGLE. The KS-test  $p$ -values for the matching of the  $M_{\star, 30 \text{ pkpc}}$  are 0.98 for EAGLE and TNG and 0.90 for EAGLE and MAGNETICUM.

We also compare our classified samples with the parent distributions from which they were sampled. This is done using a KS-test to compare their distributions in structure finder-defined stellar mass ( $M_{\star}$ ). The high  $p$ -value of 0.35 for our EAGLE sample is expected as we have not attempted to adjust the sampling with respect to its parent distribution; we see TNG ( $p = 0.16$ ) and MAGNETICUM ( $p = 0.001$ ) are less like their parent distributions as they have been sampled to match the 1300 galaxies sampled from EAGLE. We note that the conclusions we draw are not qualitatively changed by stellar mass-matching the samples.

For our halo mass-matched comparison we construct a flat distribution of log halo masses, from  $11 \leq \log_{10}(M_{200, \text{crit}}/M_{\odot}) \leq 14.5$  for EAGLE, TNG, and MAGNETICUM. Due to the small sample size of NEWHORIZON we do not attempt to perform any subsampling for the simulation. There is greater degeneracy among the parent halo masses of the galaxies as many galaxies can belong to the same halo. This makes it more difficult to match the populations with the weighted sampling method described above, so we opt for a Monte Carlo-based approach. To perform this we subsample our classified sample 100 times. We choose a linear bin spacing in  $\log_{10}(M_{200, \text{crit}}/M_{\odot})$ , such that our tidal feature fraction with halo mass measurement would have a sufficient number of galaxies in each bin while having a large enough number of bins to resolve trends with halo mass. We found that a subsample of 450 galaxies per simulation, divided across nine linearly spaced bins in log halo mass was appropriate for our purposes.

The leftmost panel in Fig. 4 shows our NEWHORIZON sample with respect to halo mass with a solid line; we note that NEWHORIZON does not include higher mass haloes, equivalent to galaxy group and cluster mass systems, due to the small volume of this simulation. Furthermore, four of the galaxies in our NEWHORIZON sample had no parent halo assigned, reducing our sample to 58 galaxies for this



**Figure 4.** The solid lines show the total number of galaxies sampled in each of the halo mass bins for NEWHORIZON and the mean number sampled in each of the 100 Monte Carlo iterations for EAGLE, TNG, and MAGNETICUM. The solid lines illustrate that our Monte Carlo samples are flat in log halo mass by construction. We subsample 50 galaxies in each of nine evenly spaced bins in the range  $11 \leq \log_{10}(M_{200, \text{crit}}/M_{\odot}) \leq 14.5$ . We also show the mean number of satellite galaxies per iteration using dashed lines and the mean number of centrals per iteration using circular points. We similarly depict the distributions of satellite and central galaxies for NEWHORIZON. The grey histograms show the parent distribution of log halo mass. The error bars give the standard error on the mean for each of the measurements obtained from the Monte Carlo iterations.

part of the analysis. We also consider the number of central and satellite galaxies in our haloes. The central galaxy is the subhalo consisting of the largest number of particles within a parent halo by the structure finder, the other galaxies in the halo are defined to be satellites. In Fig. 4 the dashed line shows the number of satellite galaxies in each halo mass bin and the open circles show the number of central galaxies. The three rightmost panels in Fig. 4 illustrate the flat sampling of galaxies with respect to their parent halo mass for EAGLE, TNG, and MAGNETICUM that we apply where analysing the halo mass and central and satellite galaxy properties of our sample. The solid lines in each of these panels show the 50 galaxies sampled in each of the halo mass bins. The mean number of satellites in each bin is shown using a dashed line and the mean number of centrals is shown using open circles. We see that the number of satellites in each bin increases with increasing halo mass, whereas the number of centrals decreases with halo mass. This is expected given that field galaxies are classified as centrals, while group and clusters contain many satellite galaxies orbiting around one central galaxy.

### 2.3 Mock images

Following the methods of Martin et al. (2022) we produce mock images that match the  $0.2 \text{ arcsec pixel}^{-1}$  spatial resolution and expected 10-yr surface brightness limits of LSST:  $\mu_g \sim 30.3 \text{ mag arcsec}^{-2}$ ,  $\mu_r \sim 30.3 \text{ mag arcsec}^{-2}$ , and  $\mu_i \sim 29.7 \text{ mag arcsec}^{-2}$  (P. Yoachim; private communication,  $3\sigma$ ,  $10 \text{ arcsec} \times 10 \text{ arcsec}$ ). We produce colour images using the methods of Lupton et al. (2004). The mock images in our study are placed at  $z = 0.025$ . The distance this corresponds to varies with the simulations adopted cosmology but is on average  $\sim 105 \text{ pMpc}$ . The mock  $g$ ,  $r$ , and  $i$  images are made to be squares of  $248 \text{ pkpc} \times 248 \text{ pkpc} \simeq 480 \text{ arcsec} \times 480 \text{ arcsec} = 2400 \text{ pixels} \times 2400 \text{ pixels}$ .

Examples of mock-observed images are shown in Fig. 1. Described briefly the construction of mock images involves:

(i) Extracting all the stellar particles associated with a group (FOF or ADAPTAHOP identified halo, depending on the simulation) within a 1 Mpc cube of the simulated galaxy and calculating spectral energy distributions (SEDs) for each stellar particle using simple stellar population models (Bruzual & Charlot 2003) extrapolated to the age and metallicity of each stellar particle. Dust attenuation is accounted for in Martin et al.’s (2022) method but is not included in the mock images created here. The brightness for each of the stellar particles is determined by the redshift of the galaxy and the convolution between the SED and the transmission function for the relevant LSST filters (Olivier, Seppala & Gilmore 2008). Using this method we produce apparent magnitudes for the LSST  $g$ ,  $r$ , and  $i$  filters.

(ii) Smoothing the particle distribution in regions where there is a potential for undersampling in the final image. This smoothing is necessary to remove unrealistic variations in flux between pixels in the mock image. This is done by subdividing stellar particles into smaller mass particles with positions distributed normally around the original particle, with a standard deviation equal to the distance to the fifth nearest neighbour of the stellar particle being smoothed (Martin et al. 2022, see also Merritt et al. 2020).

(iii) Produce mock images by collapsing the cube along one of its axes with a grid size of  $0.2 \text{ arcsec} \times 0.2 \text{ arcsec}$  (the LSST pixel size; Ivezić et al. 2019). The apparent magnitudes of the light from each of the pixels are measured accounting for the dimming due to the redshift ( $z = 0.025$ ).

(iv) To model the observational effects of seeing, the image is then convolved with the  $g$ -band Hyper-Suprime Cam point spread function (Montes et al. 2021) rescaled to match the LSST pixel size.

(v) To match observational background noise, Gaussian noise is added using the empirical relation between the standard deviation of the background noise and the limiting surface brightness, given in Román, Trujillo & Montes (2020).

$$\sigma_{\text{noise}} = \frac{10^{-0.4\mu_{\text{band}}^{\text{lim}}(n\sigma, \Omega \times \Omega)} \text{pix}\Omega}{n}, \quad (3)$$



where  $\Omega$  is the size of one side of the box over which the surface brightness limit is computed,  $=10$  arcsec,  $\text{pix}$  is the pixel scale in arcseconds per pixel,  $=0.2$  arcsec  $\text{pixel}^{-1}$ ,  $n$  is the number of Gaussian standard deviations for the surface brightness limits,  $=3$ , and  $\mu_{\text{band}}^{\text{lim}}$  is the limiting surface brightness for a particular photometric band. This model for noise intends to include all possible spatially invariant sources of noise, including instrumental noise, unresolved background sources, scattered light, etc. without explicitly modelling them. However, we note that this model does not include spatially varying sources of noise, e.g. sky gradients, diffraction spikes, imperfect sky subtraction, or galactic cirrus.

Martin et al. (in preparation) have shown using Sersic fits to mock images and galaxies in the HSC-SSP observed Cosmos field that this technique for producing mock images does not induce any measurable biases in galaxy structure measured from the mock images. They found that the variance between simulated and observed galaxies is mostly driven by the differences between the galaxy evolution models.

## 2.4 Tidal feature classification

We carry out our classification using single  $g$ ,  $r$ , and  $i$  band images along with the combined colour image, to give ourselves the best chance of detecting and correctly classifying these faint features. The images were viewed at a fixed range of contrast and brightness as shown in Fig. 1. The classification is carried out by the lead author. To classify the tidal features we chose to follow a classification scheme similar to that used in Břlek et al. (2020) and Desmons et al. (2023b). The categories include:

- (i) **Streams/Tails:** Prominent, elongated structures orbiting or expelled from the host galaxy. These usually have similar colours to the host galaxy.
- (ii) **Shells:** Concentric radial arcs or ring-like structures around a galaxy.
- (iii) **Plumes or asymmetric stellar haloes:** Diffuse features in the outskirts of the host galaxy, lacking well-defined structures like stellar streams or tails.
- (iv) **Double nuclei:** Two clearly separated galaxies within the mock image where merging is evident through the presence of tidal features.

In this scheme galaxies are allowed to host more than one feature type. A feature is allocated to the galaxy to which it appears to connect to, e.g. a tail/stream is allocated to the galaxy from which it stems, a shell is associated with the galaxy on which it is centred and a asymmetric halo is associated with the galaxy possessing the halo. In cases where features can be associated with multiple objects, e.g. a tail/stream between two galaxies (a tidal bridge) or an asymmetric halo enveloping a double nucleus, the tidal feature is counted for both objects. Following Desmons et al. (2023b) we have merged the categories of streams and tails as they are easily mistaken for one another at the surface brightness depth of LSST (Martin et al. 2022). We also use a system of confidence levels in our classification detailed in Table 2. This system is similar to that followed by Atkinson et al. (2013), but we group their confidence levels 3 and 4 which are represented as confidence level 3 here. Example features are shown in Fig. 1.

## 2.5 Calculating the errors on our counts and fractions

We estimate the binomial errors on the tidal feature fractions presented here using  $1\sigma \simeq 0.683$  confidence levels computed using a Bayesian beta distribution generator for binomial confidence intervals described in detail in Cameron (2011). This approach

**Table 2.** Descriptions of the classification confidence levels used for our analysis.

Confidence Level ( <i>conf.</i> )	Description
0	No tidal feature detected.
1	Hint of tidal feature detected, classification difficult.
2	Even chance of correct classification of tidal feature presence and/or morphology.
3	High likelihood of the tidal feature being present and morphology being obvious.

is robust against small and intermediate sample sizes, which is particularly important due to the small size of the NEWHORIZON sample.

## 3 RESULTS

We present the results of our classification in Fig. 5 and Table 3. Fig. 5 illustrates the fractions of galaxies exhibiting each kind of tidal feature for the four simulations: EAGLE in blue, TNG in red, NEWHORIZON in grey, and MAGNETICUM in cyan, for each of the confidence levels given in Table 2. The error bars show the  $1\sigma$  binomial errors on the fractions. We find excellent agreement between the four simulations at each confidence level. The differences between the tidal feature fractions are much smaller than the change in the fractions with feature morphology or with confidence level. The simulations all agree on the relative frequencies of tidal features, namely, asymmetries in the halo are the most common feature followed by double nuclei, streams/tails, and lastly shells.

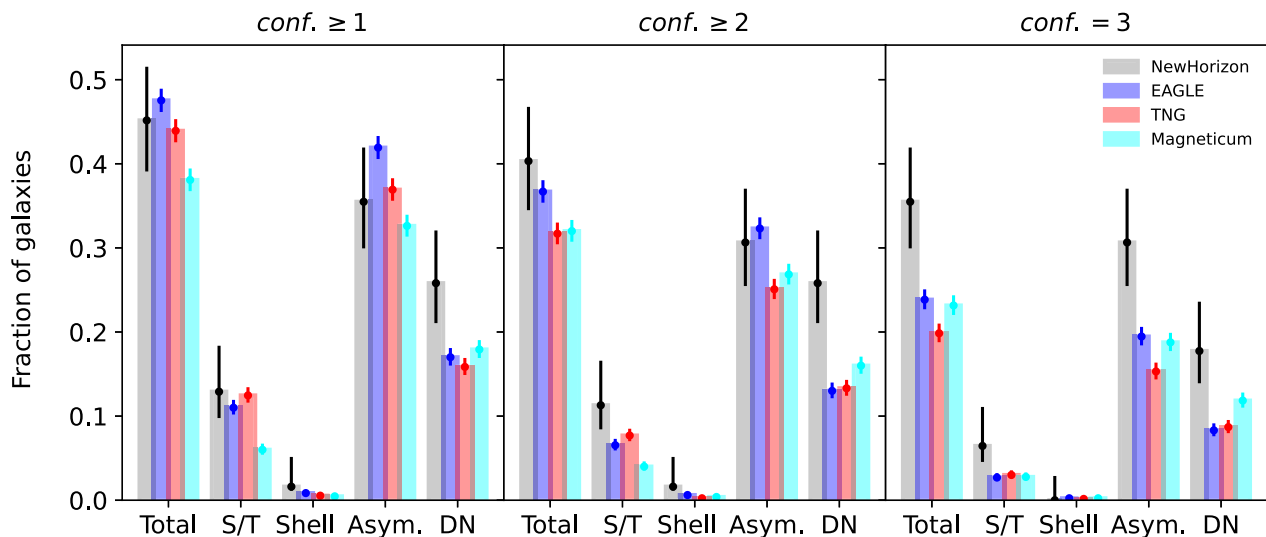
As expected, the fraction of galaxies exhibiting tidal features decreases with an increasing classification confidence threshold (*conf.*). The agreement between MAGNETICUM, TNG, and EAGLE gets stronger with the increasing confidence threshold, indicating that as we look at tidal features that are obvious/well resolved in the three simulations we have an increasing agreement.

We observe the inverse trend with NEWHORIZON, while still agreeing well with the frequencies of tidal features found in the other simulations. For *conf.*  $\geq 2$ , NEWHORIZON systematically sits at higher fractions of tidal features than any of the other simulations. Table 3 shows that between *conf.*  $\geq 1$  and *conf.*  $= 3$ , NEWHORIZON's tidal feature fraction drops by 0.1, whereas in EAGLE and TNG, it drops by 0.24 and in MAGNETICUM the fraction drops by 0.15. The smaller change in tidal feature fractions when moving from the lowest to the highest confidence level shows that tidal features detected in NEWHORIZON tend to be higher confidence than the other simulations.

The qualitative threshold of confidence levels does not change the relative frequencies of tidal features within simulations; they act as a trade-off between the detectability of a tidal feature by visual classification and the accuracy of its classification. For this reason, we tread a middle path between detectability and accuracy for the remainder of the results and present the moderate confidence results (i.e. *conf.*  $\geq 2$ ), noting that the choice of confidence level does not qualitatively change the conclusions we draw.

At *conf.*  $\geq 2$  the total tidal feature fractions (given in Table 3) are  $0.40 \pm 0.06$  for NEWHORIZON,  $0.37 \pm 0.01$  for EAGLE,  $0.32 \pm 0.01$  for TNG, and  $0.32 \pm 0.01$  for MAGNETICUM.

We study the tidal feature fraction as a function of stellar and halo mass. In Section 3.1 we present the results of our visual classification in a mass-matched comparison of EAGLE, TNG, and MAGNETICUM and the entire sample of 62 galaxies from the higher resolution



**Figure 5.** The frequencies of the tidal feature morphologies for each simulation: NEWHORIZON in grey, EAGLE in blue, TNG in red, and MAGNETICUM in cyan. From left to right the panels show the results for each confidence threshold:  $conf. \geq 1$ ,  $conf. \geq 2$ , and  $conf. = 3$  (described in Table 2). The fractions are given for five feature categories: total tidal feature fraction (Total), stream/tail (S/T), shell, asymmetric halo (Asym.), and double nucleus (DN). The error bars give the  $1\sigma$  binomial confidence levels. The occurrence of features relative to one another is similar across each confidence level; the fractions of galaxies exhibiting features drop as the confidence level increases. Generally, the simulations are in good agreement with NEWHORIZON, tending to have a higher fraction tidal features than the other simulations.

**Table 3.** The tidal feature fractions for every feature type and classification confidence level ( $conf. \geq 1$ ,  $conf. \geq 2$ , and  $conf. = 3$ , described in Table 2). The five feature categories presented are total tidal features, stream/tail, shell, asymmetric halo, and double nucleus. The rightmost column gives the overall mean fraction across the four simulations. The uncertainties give the  $1\sigma$  binomial confidence intervals on each of the fractions. \* NEWHORIZON is excluded from the calculation of the mean shell fraction as there is only one feature in that sample.

Tidal features	Confidence level	NEWHORIZON	EAGLE	TNG	MAGNETICUM	Overall mean
Total tidal features	$conf. \geq 1$	$0.45 \pm 0.06$	$0.48 \pm 0.01$	$0.44 \pm 0.01$	$0.38 \pm 0.01$	$0.44 \pm 0.07$
	$conf. \geq 2$	$0.40 \pm 0.06$	$0.37 \pm 0.01$	$0.32 \pm 0.01$	$0.32 \pm 0.01$	$0.35^{+0.07}_{-0.06}$
	$conf. = 3$	$0.35 \pm 0.06$	$0.24 \pm 0.01$	$0.20 \pm 0.01$	$0.23 \pm 0.01$	$0.26^{+0.07}_{-0.06}$
Stream/Tail	$conf. \geq 1$	$0.13^{+0.05}_{-0.03}$	$0.110^{+0.009}_{-0.008}$	$0.124^{+0.01}_{-0.009}$	$0.060^{+0.007}_{-0.006}$	$0.11^{+0.06}_{-0.03}$
	$conf. \geq 2$	$0.11^{+0.05}_{-0.03}$	$0.065^{+0.008}_{-0.006}$	$0.077^{+0.006}_{-0.007}$	$0.040^{+0.006}_{-0.005}$	$0.07^{+0.05}_{-0.03}$
	$conf. = 3$	$0.06^{+0.05}_{-0.02}$	$0.027^{+0.005}_{-0.004}$	$0.03^{+0.005}_{-0.004}$	$0.028^{+0.005}_{-0.004}$	$0.04^{+0.05}_{-0.02}$
Shell	$conf. \geq 1$	$0.016^{+0.04}_{-0.005}$	$0.008^{+0.003}_{-0.002}$	$0.005^{+0.003}_{-0.001}$	$0.005^{+0.003}_{-0.001}$	$0.006^{+0.001}_{-0.0007}$ *
	$conf. \geq 2$	$0.016^{+0.04}_{-0.005}$	$0.006^{+0.003}_{-0.001}$	$0.0023^{+0.002}_{-0.0007}$	$0.004^{+0.003}_{-0.001}$	$0.0041^{+0.002}_{-0.0005}$ *
	$conf. = 3$	$0.0^{+0.03}$	$0.0023^{+0.002}_{-0.0007}$	$0.0015^{+0.002}_{-0.0005}$	$0.0023^{+0.002}_{-0.0007}$	$0.0020^{+0.001}_{-0.0004}$ *
Asymmetric halo	$conf. \geq 1$	$0.35 \pm 0.06$	$0.42 \pm 0.01$	$0.37 \pm 0.01$	$0.33 \pm 0.01$	$0.37^{+0.07}_{-0.06}$
	$conf. \geq 2$	$0.31^{+0.06}_{-0.05}$	$0.32 \pm 0.01$	$0.25 \pm 0.01$	$0.27 \pm 0.01$	$0.29^{+0.07}_{-0.06}$
	$conf. = 3$	$0.31^{+0.06}_{-0.05}$	$0.19 \pm 0.01$	$0.153^{+0.01}_{-0.009}$	$0.19 \pm 0.01$	$0.21^{+0.07}_{-0.05}$
Double nucleus	$conf. \geq 1$	$0.26^{+0.06}_{-0.05}$	$0.17 \pm 0.01$	$0.16 \pm 0.01$	$0.18 \pm 0.01$	$0.19^{+0.07}_{-0.05}$
	$conf. \geq 2$	$0.26^{+0.06}_{-0.05}$	$0.130^{+0.01}_{-0.009}$	$0.133^{+0.01}_{-0.009}$	$0.16 \pm 0.01$	$0.17^{+0.07}_{-0.05}$
	$conf. = 3$	$0.18^{+0.06}_{-0.04}$	$0.083^{+0.008}_{-0.007}$	$0.087^{+0.008}_{-0.007}$	$0.118^{+0.01}_{-0.008}$	$0.12^{+0.06}_{-0.04}$

NEWHORIZON simulation. In Section 3.2 we study the trends with halo mass, considering this as a proxy for galaxy environment (e.g. Yang et al. 2005).

### 3.1 Tidal features and stellar mass

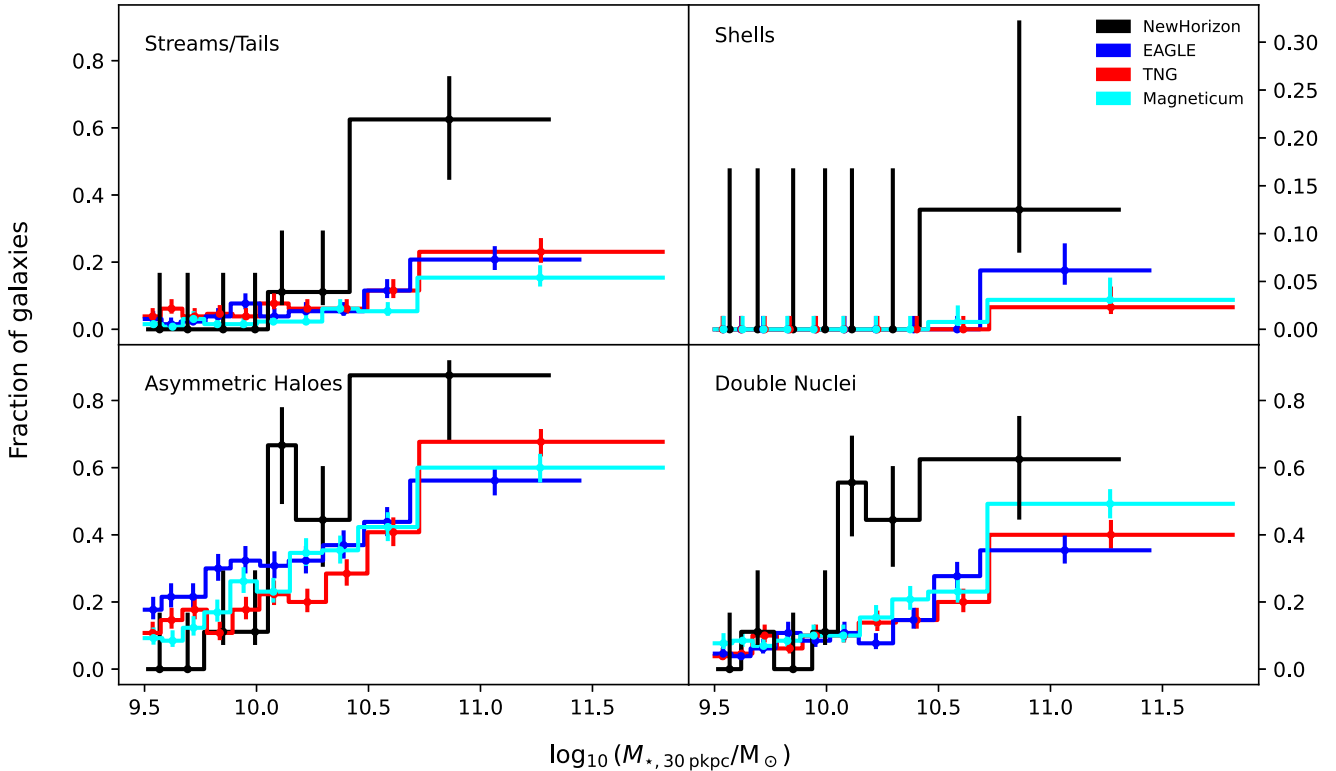
Table 4 shows the average stellar masses of the tidal feature hosts for every simulation at the moderate classification confidence level ( $conf.$

$\geq 2$ ). These are in good agreement across each of the simulations. The detected shells have the highest overall mean stellar masses across the simulations,  $M_{*, 30 \text{ pkpc}} \sim 10^{11.2} M_{\odot}$ , not including NEWHORIZON which only has one shell in the classified sample. The other tidal features appear to have very similar mean stellar masses.

Fig. 6 investigates the fraction of the galaxy population exhibiting tidal features above the LSST surface brightness limits as a function of stellar mass. We do this by constructing stellar mass bins in

**Table 4.** The mean galaxy stellar masses for each of the tidal feature categories, from top to bottom: all tidal features, stream/tail, shell, asymmetric halo, and double nucleus, for the four simulations, from left to right: NEWHORIZON, EAGLE, TNG, and MAGNETICUM. The final column shows the overall mean over the four simulations for the tidal feature host mass. The uncertainties show the standard deviations on the means. \* NEWHORIZON is excluded from the calculation of the average shell host mass as there is only one feature in that sample.

Mean masses [ $\log_{10}(M_{\star, 30 \text{ pkpc}}/M_{\odot})$ ]	NEWHORIZON	EAGLE	TNG	MAGNETICUM	Average
All tidal features	$10.3 \pm 0.8$	$10.3 \pm 0.6$	$10.4 \pm 0.7$	$10.4 \pm 0.9$	$10.3 \pm 0.4$
Stream/Tail	$10.9 \pm 0.4$	$10.6 \pm 0.4$	$10.7 \pm 0.5$	$10.7 \pm 0.4$	$10.7 \pm 0.2$
Shell	10.4	$11.0 \pm 0.1$	$11.1 \pm 0.2$	$11.1 \pm 0.4$	$11.1 \pm 0.1^*$
Asymmetric halo	$10.5 \pm 0.6$	$10.5 \pm 0.6$	$10.6 \pm 0.6$	$10.6 \pm 0.6$	$10.6 \pm 0.3$
Double nucleus	$10.6 \pm 0.6$	$10.6 \pm 0.5$	$10.7 \pm 0.6$	$10.8 \pm 0.7$	$10.7 \pm 0.3$



**Figure 6.** Distributions of the fraction of galaxies exhibiting at least one instance of a tidal feature over the range of stellar masses. From left to right, top to bottom we have the fractions of galaxies hosting: streams/tails, shells, asymmetric haloes, and a double nucleus. The fractions for NEWHORIZON are shown in black, EAGLE in blue, TNG in red, and MAGNETICUM in cyan. The error bars indicate the  $1\sigma$  binomial errors. For each simulation, each mass bin contains a similar number of galaxies,  $N_{\text{bin}, \text{TNG}} = N_{\text{bin}, \text{EAGLE}} = N_{\text{bin}, \text{MAGNETICUM}} = 130$  and  $N_{\text{bin}, \text{NEWHORIZON}} \simeq 9$ . The distributions of tidal features as a function of stellar mass agree well across all simulations, and exceptionally well between the stellar mass-matched samples from TNG, EAGLE, and MAGNETICUM.

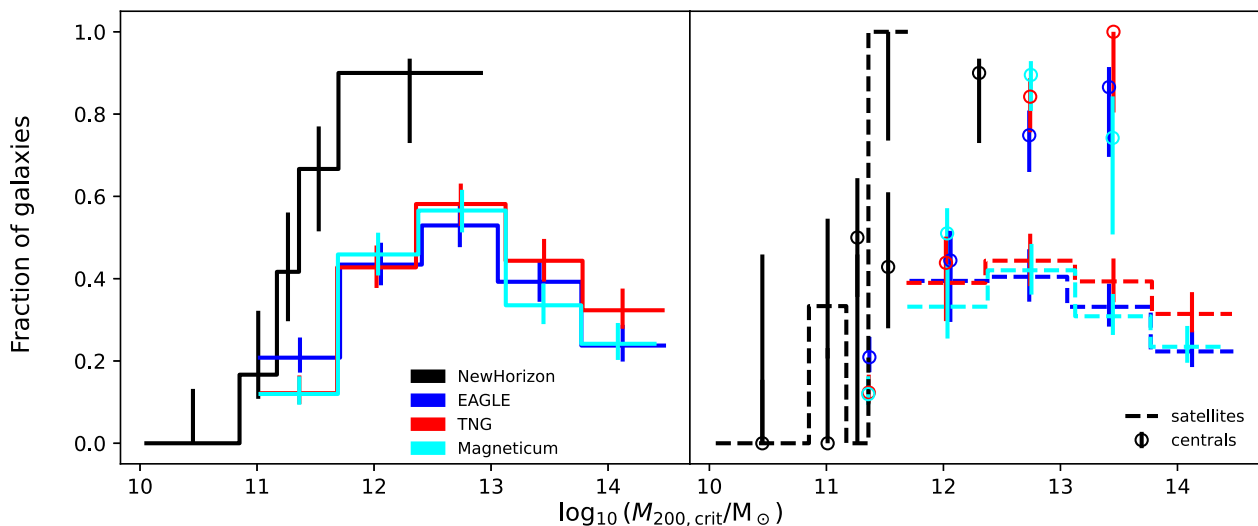
equal numbers of galaxies. For the stellar mass-matched samples of EAGLE, MAGNETICUM, and TNG we have 10 bins containing 130 galaxies each. For NEWHORIZON the 62 galaxies are distributed across seven bins of approximately nine galaxies. We again see excellent agreement between the mass-matched simulations and broadly good agreement for NEWHORIZON, though we note that NEWHORIZON has systematically higher feature fractions, particularly for  $M_{\star} > 10^{10} M_{\odot}$ .

The top left panel of Fig. 6 displays the stream/tail fraction–stellar mass distribution of galaxies for each simulation. All simulations exhibit a trend of increasing stream/tail fraction with increasing stellar mass, with very strong agreement between TNG, EAGLE, and MAGNETICUM. NEWHORIZON exhibits the same trend but has systematically higher fractions than the three other simulations. However, it is important to keep in mind the much smaller sam-

ple size (and therefore larger uncertainties) and higher resolution of NEWHORIZON when considering the significance of these offsets.

The top right panel shows the shell fraction–stellar mass distributions. There is a notable increase in the fraction of galaxies exhibiting shells at  $M_{\star, 30 \text{ pkpc}} \gtrsim 10^{11} M_{\odot}$  for TNG, EAGLE, and MAGNETICUM. The singular shell in NEWHORIZON does not provide sufficient information to determine any trends.

The asymmetric halo fraction–stellar mass distribution for each simulation is plotted in the bottom left panel of Fig. 6. Again we see excellent agreement between the simulations and a trend of increasing asymmetric halo fraction with increasing stellar mass. The rate of increase in tidal feature fraction with stellar mass is the largest here (from  $\sim 0.2$  to  $\sim 0.6$ ), indicating a strong dependence on stellar mass.



**Figure 7.** Distributions of the total tidal feature fractions as a function of halo mass. The left panel shows the total tidal feature fractions for each of our simulations and the right panel shows the total tidal feature fractions for the central and satellite galaxy populations. The results for NEWHORIZON are given in black, EAGLE in blue, TNG in red, and MAGNETICUM in cyan. The NEWHORIZON data are given for the sample of 58 classified galaxies with assigned parent haloes and each point represents the feature fraction for a halo mass bin containing  $\simeq 12$  galaxies. The points for EAGLE, TNG, and MAGNETICUM represent the mean of 100 Monte Carlo iterations, where each iteration subsampled 450 galaxies to the flat log halo mass distributions depicted in Fig. 4. Each of the bins for EAGLE, TNG, and MAGNETICUM contain 90 galaxies; these galaxies are a varying mix of central and satellite galaxies as can be seen in Fig. 4. The error bars show the mean  $1\sigma$  binomial errors.

Lastly, the bottom right panel of Fig. 6 shows the double nucleus fraction for each of the stellar mass bins. We see a sharply increasing double nucleus fraction with increasing stellar mass for TNG, EAGLE, and MAGNETICUM, and excellent agreement between the three. For NEWHORIZON, there is a sharp increase with increasing stellar mass and an indication of a higher fraction in the highest mass bins.

### 3.2 Tidal features and halo mass

The left panel of Fig. 7 shows our total tidal feature fractions as a function of halo mass for each of our simulations. We again find excellent agreement between our simulations, with EAGLE, TNG, and MAGNETICUM following the same trends and having consistent values for tidal feature fractions in each halo mass bin. For these simulations the fractions of galaxies hosting tidal features increase from halo masses of  $10^{11}$  to  $\sim 10^{12.7} M_{\odot}$  but decreases with increasing halo mass for halo masses above this. NEWHORIZON follows a similar trend for the halo masses it covers, with systematically higher fractions.

By comparing the location of the peak in tidal feature fractions with halo mass in the left panel of Fig. 7 to the proportions of central and satellite galaxies in our sample in Fig. 4, we see the peak coincides with the transition from a central-dominated population to a satellite-dominated population. Therefore, we further analyse this trend by comparing the total tidal feature fractions for central and satellite galaxies as a function of the parent halo mass in the right panel of Fig. 7. We see again similarly excellent agreement between the simulations, with EAGLE, TNG, and MAGNETICUM coinciding in their fractions for both central and satellite galaxies and NEWHORIZON following similar trends for centrals although at a systematically higher fraction. The fraction of central galaxies exhibiting tidal features increases with increasing halo mass. We note

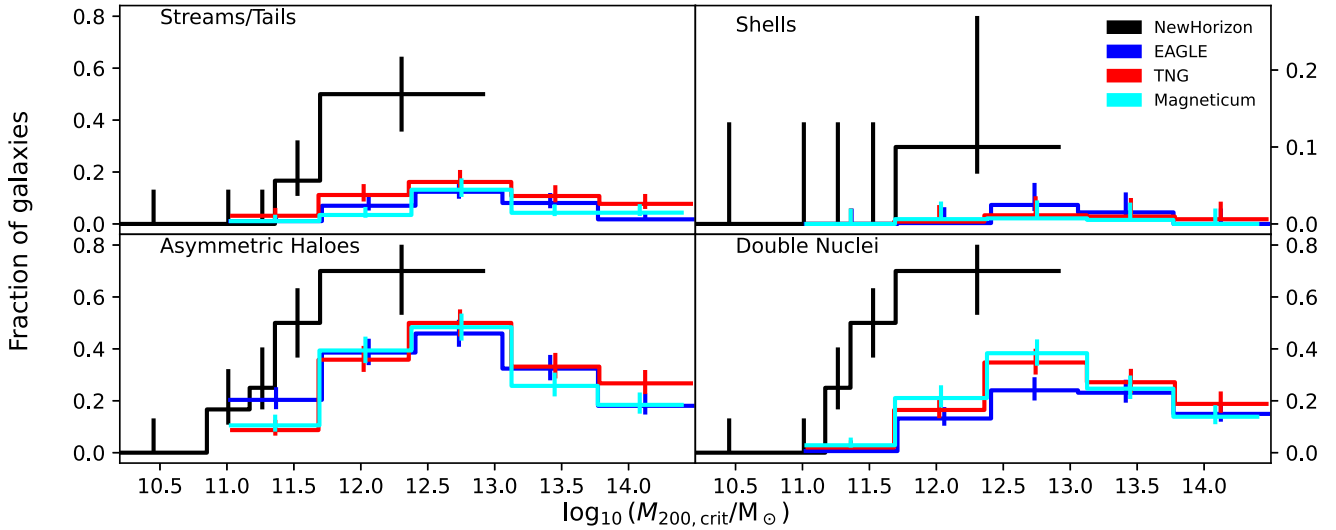
that there is substantial noise in the measurement in the  $\sim 10^{13.5} M_{\odot}$  halo mass bin as it contains on average  $\lesssim 10$  galaxies. The fractions of central galaxies exhibiting tidal features increase to fractions of  $0.86^{+0.05}_{-0.2}$ ,  $1.0_{-0.2}$ , and  $0.73^{+0.1}_{-0.2}$  by  $M_{200,crit} \gtrsim 10^{13} M_{\odot}$  for EAGLE, TNG, and MAGNETICUM. This indicates that at cluster-like halo masses a majority of central galaxies exhibit visually identifiable tidal features.

The satellite populations for EAGLE, TNG, and MAGNETICUM show a declining trend in tidal feature fraction with increasing halo mass. There are only 11 satellites in NEWHORIZON so it is difficult to comment on any trends with halo mass, but for the halo mass bin centred around  $10^{11.3} M_{\odot}$  we find a significantly higher fraction than the closest halo mass bins for the other three simulations, in line with the higher tidal feature fractions we find for NEWHORIZON throughout.

Fig. 8 shows our specific tidal feature fractions as a function of halo mass. The top left panel shows the stream/tail fraction as a function of halo mass. We find good agreement between the mean distributions for EAGLE, TNG, and MAGNETICUM. The fractions of galaxies with features appear to increase from a halo mass of  $10^{11} M_{\odot}$  to  $\sim 10^{12.7} M_{\odot}$ , from a fraction of  $\sim 0$  to  $\sim 0.15$ – $0.2$ , before decreasing again down to a fraction of  $\sim 0$ – $0.05$ . NEWHORIZON follows a similar shape for the range of halo masses it covers; however, with systematically higher fractions.

The top right panel of Fig. 8 shows the shell fraction as a function of halo mass. While there are very few shells in the sample, there is a small peak in the shell fraction at a halo mass  $\sim 10^{12.7} M_{\odot}$ . The single shell identified in the NEWHORIZON simulation does not allow for any meaningful comment on any potential trend. The peak is at a similar halo mass as for the other tidal features.

The fraction of asymmetric haloes with halo mass is presented in the bottom right panel of Fig. 8. The shape resembles that seen for streams/tails and shells, albeit the peak is more clearly resolved due



**Figure 8.** Distributions of tidal feature fractions as a function of halo mass. From left to right, top to bottom, we have the stream/tail fraction, shell fraction, asymmetric halo fraction, and double nucleus fraction. The results for NEWHORIZON are given in black, EAGLE in blue, TNG in red, and MAGNETICUM in cyan. The NEWHORIZON data are given for the sample of 58 classified galaxies with assigned parent haloes and each point represents the feature fraction for a halo mass bin containing  $\simeq 12$  galaxies. The points for EAGLE, TNG, and MAGNETICUM represent the mean of 100 Monte Carlo iterations, where each iteration subsampled 450 galaxies to the flat log halo mass distributions depicted in Fig. 4. Each of the bins for EAGLE, TNG, and MAGNETICUM contain 90 galaxies. The error bars show the mean  $1\sigma$  binomial errors. We see excellent agreement between the four simulations plotted.

to the larger number of asymmetric haloes. The peak is at a similar halo mass ( $\sim 10^{12.7} M_{\odot}$ ) to the peak for streams/tails and shells, and the fraction is larger in size (maximum asymmetric halo fraction  $\sim 0.5$ ). NEWHORIZON appears to again have a systematically higher fraction for each halo mass, while following the same trend for the halo masses it covers.

Lastly, the bottom right panel presents the double nucleus fraction as a function of halo mass. We again see a steady increase, peak, and decrease in feature fraction, similar to the other tidal feature morphologies. The halo mass peak is consistent with the peak found for the other features. NEWHORIZON follows a similar trend for the halo masses it covers with a systematically higher double nuclei fraction.

Generally, we see a trend of increasing tidal feature fraction with halo mass to  $M_{200, \text{crit}} \sim 10^{12.7} M_{\odot}$ , after which the fraction appears to decrease. NEWHORIZON appears to resemble the distributions of the other simulations for the range of halo masses it covers, despite the lack of mass matching and much smaller sample size.

Fig. 9 shows the specific tidal feature fractions for the central and satellite galaxy populations. For each of the feature types, we see that the tidal feature fractions of central galaxies tend to increase with increasing halo mass for each simulation, with the exception of the stream/tail fraction in the halo mass bin centred around  $\sim 10^{13.5} M_{\odot}$  for EAGLE and MAGNETICUM. This could be a consequence of the small number statistics of central galaxies in this bin; on average there are only  $10.21^{+0.04}_{-0.03}$  and  $5.54^{+0.04}_{-0.02}$  centrals in that halo mass bin for EAGLE and MAGNETICUM, respectively. We find that shells appear almost exclusively in central galaxies in all the simulations.

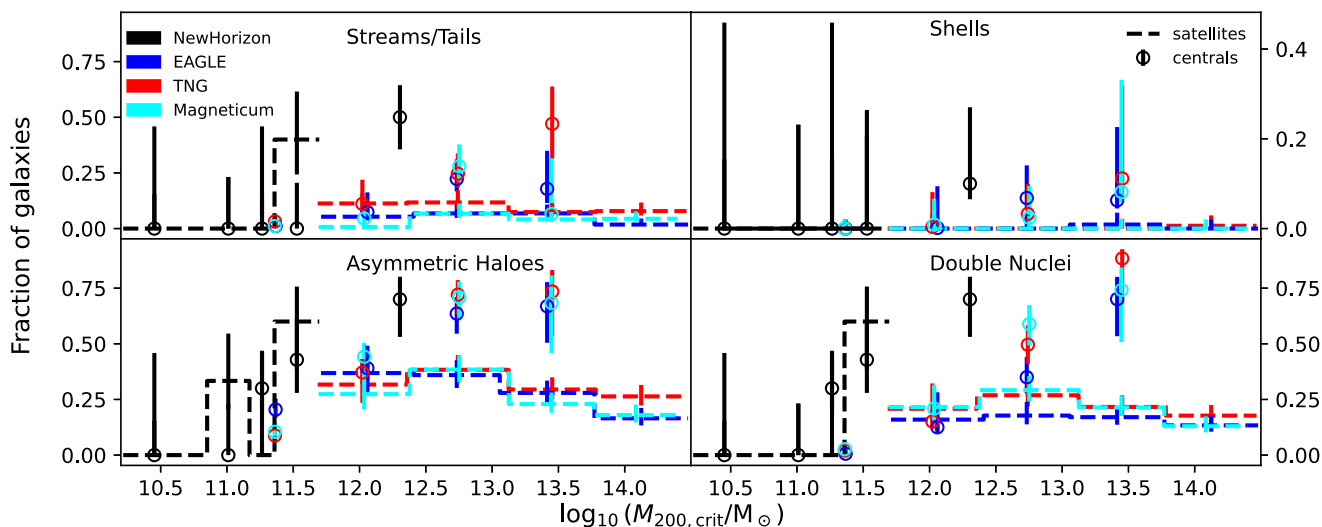
For tails/streams in Fig. 9, satellite galaxies appear to follow a flat distribution with halo mass. For asymmetric haloes we find that there is some evidence for a small peak in the distribution at  $M_{200, \text{crit}} \sim 10^{12.7} M_{\odot}$ . We see similar evidence for a peak in the distributions of double nuclei fractions in the satellite populations of TNG and MAGNETICUM. However, the distribution of the fraction of satellites exhibiting double nuclei from EAGLE remains flat with respect to halo mass.

## 4 DISCUSSION

We have presented a visual classification of the tidal features around galaxies in four cosmological hydrodynamical simulations: NEWHORIZON, EAGLE, TNG, and MAGNETICUM. Our visual classification finds broad agreement between the four simulations regarding the tidal feature fractions and their behaviour with stellar mass and halo mass while showing systematically higher tidal feature fractions for NEWHORIZON in the moderate and highest confidence levels. In Section 4.1 we compare our results to previous studies of tidal features in simulations and in Section 4.2 we compare them to observational results to ensure that they are consistent with previous work. In Section 4.3, we discuss the impact of our visual classification method, namely having the classifications done by an individual rather than a group. In Section 4.4, we discuss the possible implications of our results on the occurrence of tidal features in cosmological hydrodynamical simulations and in Section 4.5 we explore the potential of an environmental dependence for tidal feature occurrence in our results. We go on to give testable predictions for LSST in Section 4.6.

### 4.1 Comparison with other simulated analyses

We check whether our results are consistent with previous analyses of tidal features around galaxies in simulations. Martin et al. (2022) performed a study of visually classified tidal features using NEWHORIZON across a similar range of stellar masses to our study. Their original sample consisted of 37 galaxies from the  $z = 0.2$  time-step and their progenitors at  $z = 0.4, 0.6,$  and  $0.8$ , leading to a sample of 148 galaxies. They had 45 astronomers visually classify the images and varied galaxy orientations relative to the observer and the distances at which the object was placed relative to the observer in order to gather  $\sim 8000$  images, 600 of which were classified by up to five different classifiers. They applied an extended classification scheme, including bridges and merger remnants in addition to the classifications we use here. Comparing our results for the fractions of

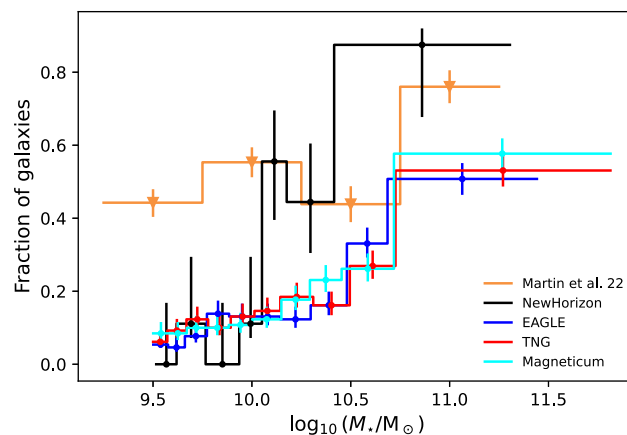


**Figure 9.** Distribution of tidal feature fractions for satellite and central galaxies as a function of halo mass. From left to right, top to bottom, we have the stream/tail, shell, asymmetric halo, and double nucleus fractions. The results for NEWHORIZON are given in black, EAGLE in blue, TNG in red, and MAGNETICUM in cyan. The NEWHORIZON data are given for the sample of 58 classified galaxies with assigned parent haloes. The dashed lines show the feature fractions for satellite galaxies and the circles show the feature fractions for centrals for each halo mass bin containing  $\simeq 10$  galaxies. For the Monte Carlo sampled distributions the points for centrals and satellites show the means of 100 iterations, each bin contains 90 galaxies. Each bin contains a varying mix of central and satellite galaxies as can be seen in Fig. 4. The error bars for NEWHORIZON shows the  $1\sigma$  binomial errors and the error bars for EAGLE, TNG, and MAGNETICUM data show the mean  $1\sigma$  binomial errors.

galaxies exhibiting at least one instance of a tidal feature morphology (Fig. 6), we see qualitatively similar trends of the fraction of galaxies exhibiting streams/tails, double nuclei, and asymmetric haloes all increasing with increasing stellar mass.

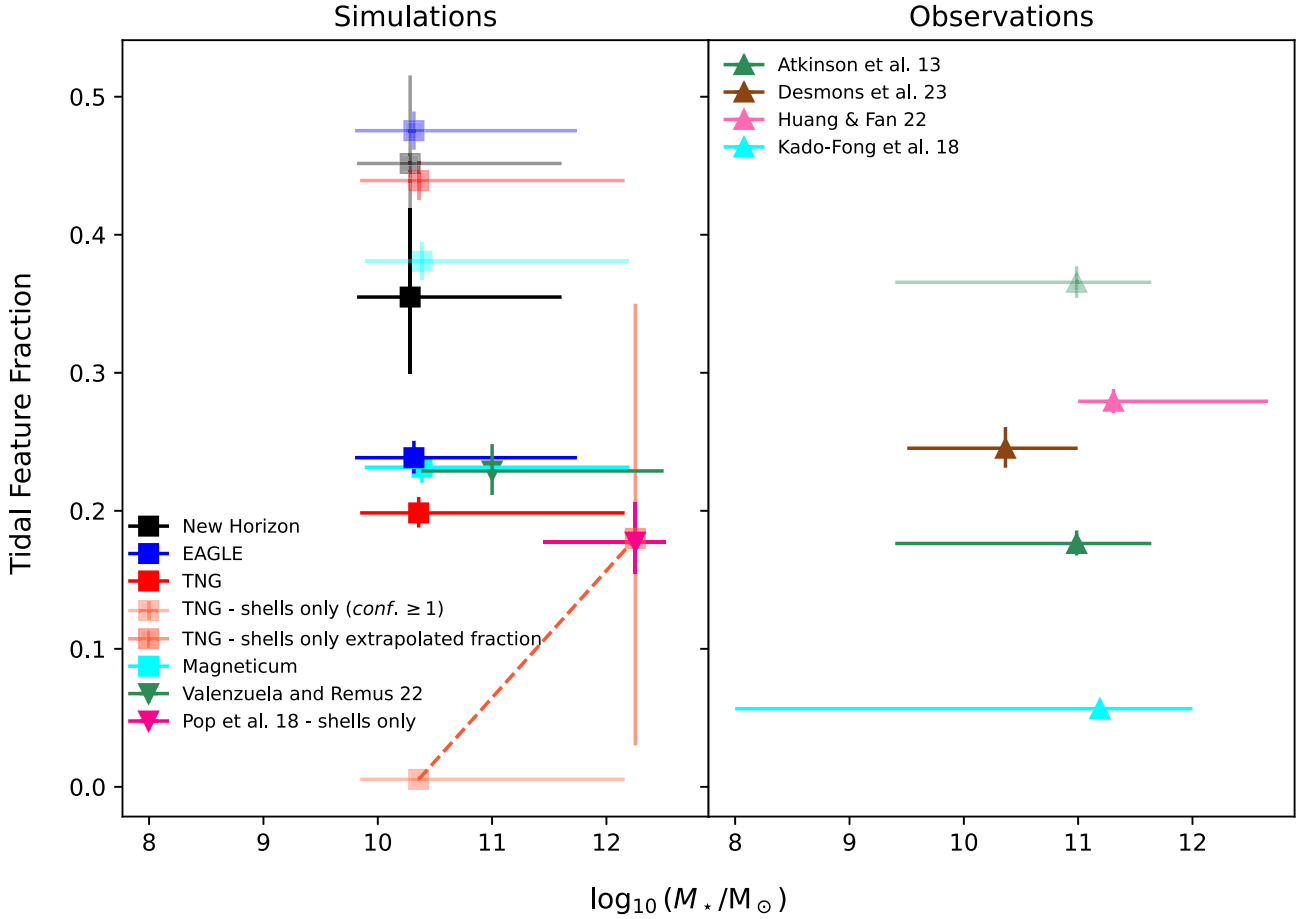
The most easily comparable measure between the two studies is the total tidal feature fraction. We compare to the Martin et al. (2022) results for all tidal features, excluding asymmetric haloes to align with their approach, in Fig. 10. The results are depicted similarly to Fig. 6, with identical binning and including only classifications with  $conf. \geq 2$ . Our NEWHORIZON results agree well for stellar masses  $\geq 10^{10} M_{\odot}$ . For stellar masses below  $10^{10} M_{\odot}$ , we report lower tidal feature fractions, with our fractions being lower by  $\sim 0.4$ . For galaxies with  $10^{9.5} M_{\odot} \leq M_{*} \leq 10^{10} M_{\odot}$ , Martin et al. (2022) sample seven galaxies from the  $z = 0.2$  time-step, excluding the remaining galaxies due to their haloes being contaminated by lower resolution dark matter particles. This contamination is exclusively a concern for zoom-in simulations and therefore only impacts NEWHORIZON in our sample. We include these galaxies, sampling 32 galaxies in this range to maximize our number statistics for NEWHORIZON. This difference in sampling is the likely cause of the differences in the measured fraction.

We further check for consistency by comparing our results to those from Valenzuela & Remus (2023). They visually classified the presence of streams, tails, and shells in the same MAGNETICUM simulation box we use for this study. Valenzuela & Remus (2023) performed their visual classification on the 3D stellar component of MAGNETICUM galaxies with  $M_{*} \geq 2.4 \times 10^{10} M_{\odot}$  and virial mass  $M_{\text{vir}} \geq 7.1 \times 10^{11} M_{\odot}$  ( $M_{\text{vir}} \approx M_{200, \text{crit}}$ ). The most direct comparison to our results is using the total tidal feature fractions. In Fig. 11, we show the fraction of galaxies exhibiting tidal features in each of our samples, for  $conf. \geq 1$  and  $conf. = 3$  classifications. The confidence levels increase with the opacity of the point. The left panel compares the simulated tidal feature fractions. The stellar mass error bars for each of the fractions give the range of stellar masses studied (using the  $M_{*, 30\text{pkpc}}$  masses for our results), while



**Figure 10.** Distribution with stellar mass of the total tidal feature fraction (excluding asymmetric haloes) presented with results from Martin et al. (2022). Stellar mass is measured by  $M_{*, 30 \text{ pkpc}}$  for our sample and ADAPTAHOP stellar masses for Martin et al. (2022). NEWHORIZON is shown in black, EAGLE in blue, TNG in red, and MAGNETICUM in cyan. The Martin et al. (2022) results are shown in orange. We compute the fractions using classifications with  $conf. \geq 2$ . The error bars indicate the  $1\sigma$  binomial uncertainty for our results and the  $1\sigma$  uncertainty obtained from 100 000 bootstraps for Martin et al. (2022). For each simulation, each mass bin contains a similar number of galaxies,  $N_{\text{bin,NEWHORIZON}} \simeq 9$  and  $N_{\text{bin,EAGLE}} = N_{\text{bin,TNG}} = N_{\text{bin,MAGNETICUM}} = 130$ . Martin et al. (2022) bin their galaxies using stellar mass bins of equal width.

the point itself gives the fraction and the mean stellar mass where the measurement is available, and the mid-point of the stellar mass range where the mean is unavailable. Valenzuela & Remus (2023) measured a tidal feature fraction of  $0.23 \pm 0.02$ , which agrees well with our tidal feature fraction for MAGNETICUM and other simulations at  $conf. = 3$ . Limiting ourselves to just tails, streams, and shells within the stellar mass range covered by Valenzuela & Remus



**Figure 11.** Total fraction of galaxies exhibiting tidal features from various observations and simulations studies compared to the results presented here. On the left, we display our own simulation results and simulation results from the literature. On the right, we give the results of observational studies. The error bars along the stellar mass axis show the range of stellar masses for each measurement, with the marker situated on the mean where the sample mean stellar mass is provided (where unavailable the mid-point of the stellar mass range is used instead). Our  $conf. \geq 1$  results are presented with a transparent marker and our  $conf. = 3$  results are presented with the solid marker. We use a similar opacity to indicate the different confidence level derived fractions from observations. The dashed line links our TNG  $conf. \geq 1$  shell fraction with an estimated fraction when linearly extrapolated to higher stellar masses. We find our tidal feature fractions to be in the expected range for observational results and simulation results.

(2023), we find a feature fraction of  $0.17 \pm 0.02$  at  $conf. \geq 1$  and  $0.09^{+0.02}_{-0.01}$  for  $conf. = 3$ . Therefore, our fractions sit a factor of  $\sim 1.4$ – $2.5$  times below the fractions for the same MAGNETICUM simulation box. Valenzuela & Remus (2023) also find an increasing fraction of galaxies exhibiting tidal features with stellar mass, qualitatively agreeing with our predictions. The differences in our fractions could be due to us visually detecting and classifying tidal features from mock images while Valenzuela & Remus (2023) classified based on unprojected 3D data. Mock images are subject to projection effects and the surface brightness limit we have applied which Kado-Fong et al. (2018), Pop et al. (2018), and Martin et al. (2022) suggest have a significant impact on the detectability of tidal features. Therefore, lower fractions from these projected mock images are expected.

Pop et al. (2018) studied shell formation in an analogous run used from the previous generation ILLUSTRIS simulation which has a similar resolution to that used in this study (Genel et al. 2014; Vogelsberger et al. 2014). They performed their analysis on 220 of the most massive galaxies in the ILLUSTRIS simulation ( $M_* \gtrsim 10^{11} M_\odot$ ) and visually detected the presence of shells using stellar surface density maps across three orthogonal projections. Their shell fraction of  $0.18^{+0.03}_{-0.02}$  is presented in the left panel of Fig. 11. Furthermore,

Pop et al. (2018) predict that in an observational survey (i.e. with only one projection available) they would measure a shell fraction of  $0.14 \pm 0.03$ . For a direct comparison, we take our most optimistic shell fraction ( $conf. \geq 1$ ) for our stellar mass-matched TNG sample. This gives a substantially lower fraction of shells than Pop et al. (2018),  $0.005^{+0.003}_{-0.001}$  (Table 3). From Fig. 11, we can see that our studies cover different mass ranges. Pop et al. (2018) probe galaxies that are much higher in stellar mass than ours. As shell fractions increase with stellar mass, we find it likely that the different stellar mass ranges of the samples contribute to the different tidal feature fractions. To test the impact of the different stellar mass ranges, we extrapolate our shell fraction to the mid-point of Pop et al.’s (2018) stellar mass range,  $10^{12.3} M_\odot$ . While our  $conf. \geq 2$  results plotted in Fig. 6 only have measured shell fractions in the highest stellar mass bin, if we include  $conf. \geq 1$  classifications we obtain measurements in the last two stellar mass bins of  $0.008^{+0.02}_{-0.002}$  and  $0.05^{+0.03}_{-0.01}$ . An approximate linear gradient of  $0.1 \pm 0.1$  can be obtained using the mean galaxy stellar masses in these two bins. Using this as an estimate for the rate of shell fraction increase with stellar mass for our sample, we can estimate what our total shell fraction would be if our stellar masses are extrapolated from the mean mass in our highest stellar

mass bin to the mid-point of Pop et al. (2018) stellar mass range. This gives a shell fraction  $0.18^{+0.17}_{-0.15}$ , which is consistent with the fraction measured by Pop et al. (2018), suggesting that the differences are mainly due to the stellar mass ranges of the respective samples.

Fig. 8 shows the relationship between halo mass and tidal feature occurrence. It suggests that we may be able to test the rates of close encounters as a function of the environment predicted by  $N$ -body simulations (e.g. Gnedin 2003) using observations from LSST. The simulations predicted that the rate of close encounters per galaxy as a function of the relative velocities of the galaxies during the encounter has a peak at an encounter velocity  $\sim 500 \text{ km s}^{-1}$ . This corresponds to two galaxies moving at speeds  $\sim 250 \text{ km s}^{-1}$  relative to the centre of mass of the cluster. The halo mass–velocity dispersion relationship shows that  $\sigma \simeq 250 \text{ km s}^{-1}$  corresponds to a peak at  $M_{\text{vir}} \sim 10^{13} M_{\odot}$  (Elahi et al. 2018). We expect there to be a corresponding peak in the relationship we measure between tidal feature occurrence and halo mass. As tidal features are tracers of galaxy encounters, the fact that we see a peak at  $M_{200, \text{crit}} \sim 10^{12.7} M_{\odot}$  in Fig. 8 is an encouraging prediction from our simulations.

Jian et al. (2012) studied close pairs in the MILLENIUM  $N$ -body simulation and a suite of semi-analytic models. They found a peak in the relationship between the fractions of close pairs undergoing mergers and halo mass in the range  $10^{12}$ – $10^{13} M_{\odot}$ . This is in good agreement with our double nuclei panel in Fig. 8.

We find good agreement between the tidal feature fractions in our simulations, suggesting that tidal features are a genuine probe of mass assembly and may not be sensitive to processes being modelled by subgrid physics. Canas et al. (2020) and Proctor et al. (2023) studied intra-halo light (IHL) in HORIZON-AGN and EAGLE, respectively, and Brough et al. (2024) studied the IHL in HORIZON-AGN, HYDRANGEA, MAGNETICUM, and ILLUSTRISTNG. These works all found good agreement in IHL fractions despite the different subgrid physics models between simulations. As IHL is a similarly low surface brightness feature formed through mergers, the agreement between IHL fractions across these simulations suggests that gravitational physics is playing a dominant role in this probe of assembly history.

In general, we find an encouraging agreement with existing simulation results. The differences in the tidal feature fractions in our results and existing studies are likely explained by a combination of the differences in stellar mass ranges and the potential reduction of tidal feature detectability due to the surface brightness limits of the mock images.

## 4.2 Comparison with observational results

Observed tidal feature fractions measured by prior studies span a large range of values. Most of these differences are likely driven by the different surface brightness limits of the surveys, as tidal features become more visible with deeper imaging and at higher galaxy stellar masses (e.g. Kado-Fong et al. 2018; Martin et al. 2022). To ease comparison with the values measured here we have plotted the total tidal feature fractions measured in observational studies in the right panel of Fig. 11.

Atkinson et al. (2013) used observations from the wide component of the Canada–France–Hawaii Telescope Legacy Survey. They visually classified  $g'-r'-i'$  stacked images of 1781 galaxies with a redshift range of  $0.02 < z < 0.4$ . The limiting surface brightness of their data was  $\mu_{g',AB} = 27.7 \pm 0.5 \text{ mag arcsec}^{-2}$ . The classification scheme used by Atkinson et al. (2013) identifies streams, arms, linear features, miscellaneous structure, diffuse fans, and shells and has significant overlap with our own classification types. Our

streams/tails category is roughly analogous to their streams + arms + linear features, and our asymmetric halo category corresponds to their diffuse fans + miscellaneous structure. For their lowest confidence level Atkinson et al. (2013) measure a total tidal feature fraction of  $0.37 \pm 0.01$  and at their highest confidence level, they measure  $0.18 \pm 0.01$ . Their lowest confidence level sits below our  $\text{conf.} \geq 1$  fractions for NEWHORIZON, EAGLE, and TNG:  $0.45 \pm 0.06$ ,  $0.48 \pm 0.01$ , and  $0.44 \pm 0.01$ . However, their fraction is comparable to MAGNETICUM:  $0.38 \pm 0.01$ , which presents a tidal feature fraction slightly lower than the other simulations, likely due to its slightly lower resolution discussed in more detail in Section 4.4. Their highest confidence level tidal feature fraction sits significantly below our  $\text{conf.} = 3$  fractions for NEWHORIZON, EAGLE, TNG, and MAGNETICUM:  $0.35 \pm 0.06$ ,  $0.24 \pm 0.01$ ,  $0.20 \pm 0.01$ , and  $0.23 \pm 0.01$ . This is expected given Atkinson et al. (2013) examined data at a limiting surface brightness  $\sim 2.6 \text{ mag arcsec}^{-2}$  brighter than our mock images.

There have been many studies of tidal features made using HSC-SSP observations, such as the work of Kado-Fong et al. (2018), Huang & Fan (2022), and Desmons et al. (2023b). Kado-Fong et al. (2018) studied 20 000 galaxies across a redshift range of  $0.05 < z < 0.45$  using the first data release of the HSC-Wide catalogue (Bosch et al. 2018). They used an automated method, creating high spatial frequency images to detect tidal features and then visually classifying the morphologies of these features. They found a much lower total tidal feature fraction of  $0.056 \pm 0.002$ . They cover a much larger range of stellar masses than our sample and detect tidal features to surface brightness depths of  $\mu_i \sim 26.4 \text{ mag arcsec}^{-2}$ ,  $3.3 \text{ mag arcsec}^{-2}$  shallower than our mock images. With their stellar mass range probing down to  $10^8 M_{\odot}$  and redshift range reaching  $z \sim 0.45$ , many tidal features will fall below their detection limit resulting in their lower fraction.

Huang & Fan (2022) used the coadded images from the Wide layer of HSC-SSP third data release, which reaches a surface brightness limit of  $\mu_i \sim 28.5 \text{ mag arcsec}^{-2}$  ( $\sim 1.2 \text{ mag arcsec}^{-2}$  brighter than our mocks). Their final sample of galaxies consists of 2649 early-type galaxies without companions (i.e. the sample has no double nuclei) with stellar masses  $\geq 10^{11} M_{\odot}$ . They apply an automated procedure to model and remove the host galaxy and leave behind only the tidal features. Their total tidal feature fraction is  $0.28 \pm 0.01$ . Their fraction is slightly higher than our  $\text{conf.} = 3$  fractions for all of our simulations other than NEWHORIZON. Their fraction falls  $\sim 0.2$  below our  $\text{conf.} = 1$  fractions; this indicates that when we include all our galaxies that exhibit any trace of a tidal disturbance, even where it is difficult to determine its morphology, we visually detect far more than Huang & Fan (2022). This is likely due to the deeper surface brightness limits applied to our images.

Desmons et al. (2023b) used the HSC-SSP second data release UltraDeep region images, with a limiting surface brightness of  $\mu_g \sim 30.76$ ,  $\mu_r \sim 29.82$ , and  $\mu_i \sim 29.41 \text{ mag arcsec}^{-2}$  ( $3\sigma$ ;  $10 \text{ arcsec} \times 10 \text{ arcsec}$ ). They limited their sample to galaxies in the redshift range  $0.04 < z < 0.2$  and applied stellar mass limits of  $10^{9.5} M_{\odot} \leq M_{\star} \leq 10^{11} M_{\odot}$ , constructing a volume-limited sample of 852 galaxies. Their classification scheme is identical to ours, making this a particularly relevant comparison. They measure a total tidal feature fraction of  $0.23 \pm 0.02$ . This is in good agreement with our  $\text{conf.} = 3$  results for EAGLE, TNG, and MAGNETICUM and  $\sim 0.1$  lower in fraction than NEWHORIZON, while we measure significantly higher fractions when including all possible tidal features ( $\text{conf.} \geq 1$ ). Given their lower stellar mass range, it could be that the differences in fractions are driven by the contribution of simulated galaxies with  $M_{\star, 30 \text{ pkpc}} \geq 10^{11} M_{\odot}$  that tend to have higher tidal feature fractions



than lower stellar mass galaxies. Furthermore, all the galaxies in the Desmons et al. (2023b) sample are farther from the observer than in our sample. Galaxies at  $z = 0.2$  are further away than the galaxies in our mock images; more distant tidal features may be too dim to detect visually. Desmons et al. (2023b) found qualitatively similar trends with increasing tidal feature fractions as a function of stellar mass. They do detect a substantially higher fraction of galaxies exhibiting shells,  $0.02 \pm 0.01$ , higher than our mean  $\text{conf.} \geq 1$  shell fraction of  $0.006^{+0.001}_{-0.0007}$  (excluding NEWHORIZON as we only detect one shell in this simulation). This result is surprising as the Desmons et al. (2023b) sample is measured over lower stellar masses than ours. Given that shell fraction increases with increasing galaxy stellar mass we would expect our sample to return a higher shell fraction. Constraining our sample to  $M_{*, 30 \text{ pkpc}} \leq 10^{11} M_{\odot}$ , we find EAGLE, TNG, and MAGNETICUM to have mean  $\text{conf.} \geq 1$  shell host masses of  $\log_{10}(M_{*, 30 \text{ pkpc}}/M_{\odot}) = 10.8 \pm 0.2$ ,  $10.7 \pm 0.1$ , and  $10.8 \pm 0.1$ . These are consistent within  $2\sigma$  of the mean shell host stellar mass of  $\log_{10}(M_{*}/M_{\odot}) = 10.56 \pm 0.04$ , found by Desmons et al. (2023b).

The higher fractions of galaxies hosting shells in observations could be a result of factors relating to the properties of the host galaxies, the characteristics of the shells themselves and the methods we use to produce the mock images. The detectability of shells was examined in Martin et al. (2022) and Bazkiaei (2023), who found that it depends significantly on the degree of contrast against the background galaxy. The variation in detectability is driven more by the physical parameters of a galaxy (stellar mass, Sersic index, and galaxy size) than by the characteristics of the shells themselves (opening angle and shell width). Bazkiaei (2023) find that detection of shells is dependent on the stellar mass resolution of the simulations, which limits the amount of contrast a shell can achieve against the host galaxy in the inner radii close to the galaxy. Martin et al. (2022) found that within  $\sim 4 R_{\text{eff.}}$ , shells will not be detectable in NEWHORIZON mock images but could be detectable in LSST. We expect this region to extend to larger radii for the other simulations as they have lower stellar mass resolutions (Bazkiaei 2023).

It is possible that systematics introduced by the mock image construction (e.g. smoothing) could reduce the visibility of some shells, making shells more difficult to detect or potentially be misidentified as a different class (e.g. asymmetric halo). However, the agreement of our TNG shell fractions with the shell fractions measured directly from ILLUSTRIS simulation data by Pop et al. (2018) suggests that the mock image creation method does not significantly impact the visibility of shells. Direct comparison with LSST data could help clarify this tension between simulation and observed shell fractions.

### 4.3 Discussion of visual classification methodology

The majority of the visual classification was performed by the lead author. There was consultation on a subset of 31 NEWHORIZON galaxies to ensure good agreement on the visually classified morphological types of features. Other studies have already used visual classifications by groups of scientists and reported on the stability and convergence of such classification systems. Bílek et al. (2020) used a group of scientists of different experience levels to identify streams, shells, and tails on the MATLAS Survey data. They found that there is a significant scatter between classifiers and that this scatter tends to decrease with the experience of the classifiers. Martin et al. (2022) also found that with faint tidal features, there was decreased concurrence between classifiers on the nature of the

morphologies, increasing the scatter. The above suggests that using multiple classifiers helps to characterize the scatter in the morphology and detectability of the tidal feature and could increase the accuracy and reliability of visual detection and classification. However, we argue that having a single classifier makes for a similarly biased comparison across all of the simulations, as the biases of the single classifier are carried across the simulations as systematic errors, allowing for a fair consensus on whether the simulations agree with each other.

### 4.4 Comparison between the different simulations

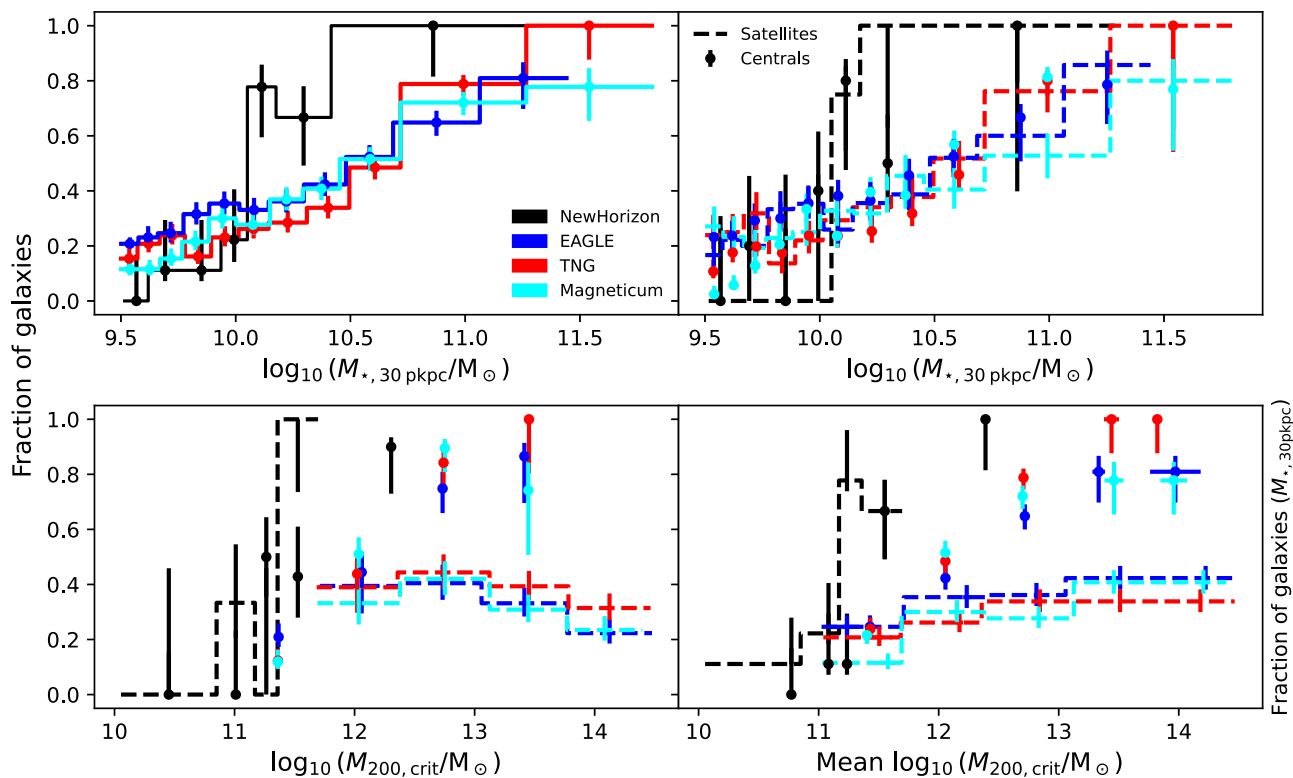
The broad agreement between the four different simulations analysed here, and the behaviour of these fractions with respect to stellar mass and halo mass suggest that the different subgrid physics and the different hydrodynamics schemes applied are not playing a dominant role in the observability of tidal features. The impact of classification confidence on the feature fractions is presented in Table 2 and Fig. 5. Importantly we see that the decrease in the fractions of galaxies exhibiting tidal features with an increasing confidence level is less significant in the NEWHORIZON simulation than in the other simulations. This implies that the classified tidal features are of a higher confidence and therefore easier to detect in NEWHORIZON than the other simulations. A contributing factor to the higher confidence tidal features in the NEWHORIZON simulation could be its approximately two orders of magnitude higher stellar mass resolution compared to the other simulations. For a given number of particles, NEWHORIZON would be able to resolve tidal features from disrupted galaxies with  $\gtrsim 100$  times lower stellar mass than the other simulations, which could result in the higher fractions (Appendix B). The slightly lower mass resolution for MAGNETICUM when compared to the other simulations could also be contributing to its lower fraction when including  $\text{conf.} \geq 1$  tidal features that are likely to include fainter features where the mass resolution could significantly impact its visibility when near a bright host (Appendix B).

The galaxy SMHM relation (Fig. 2) points to a different potential source of NEWHORIZON's higher tidal feature fractions as a function of stellar and halo mass. In this relation, we see that for a given parent halo mass NEWHORIZON has on average higher stellar mass galaxies populating the halo compared to the other simulations. This means that when galaxies undergo mergers, NEWHORIZON will tend to bring more stellar mass into the merger and therefore is likely to produce more visible tidal features for a galaxy of a particular stellar mass than the other simulations.

These results suggest that the simulations are in reasonable agreement with one another regarding the occurrence of visually detectable tidal features. The higher resolution and higher galaxy stellar masses for a given halo mass of NEWHORIZON could be driving the higher fractions we measure.

### 4.5 Tidal feature fractions and environment

The relationships between tidal feature fractions and halo mass, plotted in the left panel of Fig. 7, illustrate a peak in the occurrence of tidal features at  $M_{200, \text{crit}} \sim 10^{12.7} M_{\odot}$ . This peak coincides with the transition from a central-dominated population of galaxies to a satellite-dominated population. These observations together suggest that the occurrence of tidal features could depend on environment both measured as a function of halo mass and as location in the halo. In Fig. 12 we investigate whether the trends we measure with halo mass are indicative of an environmental influence on the occurrence



**Figure 12.** Investigating whether trends of tidal feature fraction with halo mass and as a function of central and satellite galaxy populations can be explained by the correlation between tidal feature fraction and galaxy stellar mass. The top left panel shows the tidal feature fraction–stellar mass relation shown in Fig. 10 with the inclusion of asymmetric haloes and with the final bin for EAGLE, TNG, and MAGNETICUM divided into two bins, linearly spaced in log-stellar mass. The top right panel uses identical binning to the top left, and plots the fractions of centrals and satellite galaxies exhibiting tidal features as a function of stellar mass. This illustrates that there are no differences in the relations followed by central or satellite galaxies across all simulations. The bottom left panel is identical to the right panel of Fig. 7 and shows the tidal feature fraction as a function of halo mass for central and satellite galaxies. The bottom right panel plots the tidal feature fraction estimated from the tidal feature fraction–stellar mass relation shown in the top left panel, for the populations of centrals and satellites as a function of mean halo mass. For central galaxies we find reasonable agreement between the bottom two panels, indicating that the trends seen in the tidal feature fraction–halo mass relation are a result of the relationship of the tidal feature fractions with stellar mass. However, satellites in the bottom right panel have fractions that systematically increase with increasing halo mass compared to the bottom left panel, where they increase and then fall as a function of increasing halo mass. This could indicate that there are factors other than stellar mass that drive tidal feature fractions for satellite galaxies. The uncertainties in the tidal feature fraction are given by the  $1\sigma$  binomial errors for the relations with stellar mass and the mean  $1\sigma$  binomial errors for the relations with halo mass. The uncertainties on the mean halo mass are given by standard deviation of the mean halo mass over the 100 Monte Carlo iterations.

of tidal features or whether the observed trends can be reproduced by accounting for the trends of tidal features with galaxy stellar mass.

We start in the top panel by investigating whether there is any separation in the tidal feature–stellar mass relation when dividing it into central and satellite galaxies. The top left panel is a version of Fig. 10, including asymmetric haloes and with the highest stellar mass bin in the EAGLE, TNG, and MAGNETICUM simulations, split into two bins linearly spaced in log-stellar mass to allow for the trends in the higher stellar mass ranges ( $M_{\star, 30 \text{ pkpc}} \gtrsim 5 \times 10^{10} M_{\odot}$ ) to be explored. The top right panel shows tidal feature fractions with galaxy stellar mass for central and satellite galaxies in each of the bins in the top left panel. We see for all the simulations that the tidal feature fraction with stellar mass does not depend significantly on whether the galaxy is a satellite or a central.

In the lower panels we explore whether the relationship observed with halo mass in Fig. 7 could be a result of the observed relationship with stellar mass. We repeat Fig. 7 in the bottom left panel for direct comparison. The bottom right panel shows the tidal feature fraction estimated using the mean stellar masses of the central and satellite galaxies in each of the bins in the bottom left panel. The trends seen

for central galaxies in the bottom left panel are reproduced well in the bottom right panel, suggesting that their tidal feature fractions are mostly driven by their stellar mass. The satellite tidal feature fractions inferred from their stellar mass systematically increase with increasing halo mass reaching 0.3–0.4 for  $M_{200, \text{crit}} \gtrsim 10^{13} M_{\odot}$ . However, the true satellite tidal feature fractions increase with increasing halo mass to  $M_{200, \text{crit}} \sim 10^{12.7} M_{\odot}$  and decrease with increasing halo mass beyond this mass. While this is within  $2\sigma$  of the uncertainties on these fractions, the systematic nature of the difference suggests there might be factors other than stellar mass driving the relationship between tidal feature fractions and halo mass for satellite galaxies. This might indicate an enhancement of mergers in satellites residing in haloes of  $M_{200, \text{crit}} \sim 10^{12.7} M_{\odot}$  and a reduction in mergers in satellites residing in haloes of  $M_{200, \text{crit}} \gtrsim 10^{13}$  (e.g. Jian et al. 2012). The work of Omori et al. (2023) provides another example of galaxy mergers rates exhibiting a dependence on environment. Omori et al. (2023) used a deep learning model to detect mergers in HSC-SSP, finding that on scales of 0.5 to 8 Mpc  $h^{-1}$ , mergers tended to occur in lower density environments, showing that galaxy environment is a significant factor in the occurrence of mergers. The

role of environment in the occurrence of galaxy mergers is further highlighted by Sureshkumar et al. (2024). Sureshkumar et al. (2024) studied the spatial clustering of merger and non-merger galaxies in the GAMA catalogue and found that galaxy mergers tend to occur in underdense environments on spatial scales greater than  $50 \text{ kpc h}^{-1}$ .

#### 4.6 Predictions for LSST

Given the broadly good agreement between the simulations regarding tidal feature fractions and their trends with stellar mass and halo mass, we can be confident in the simulations' predictive power for comparison with LSST. We note that NEWHORIZON's higher tidal feature fractions in some relations could potentially indicate that the higher mass resolution of this simulation allows it to resolve tidal features from lower stellar mass progenitors. We use a simple model to probe the limiting progenitor mass for tidal features that each simulation can detect in the mock images in Appendix B. In NEWHORIZON we can detect tidal features down to stellar masses of  $\sim 10^6 M_\odot$ , so long as they are not disrupted over an area  $\gtrsim 36 \text{ pkpc}^2$  and in the other simulations this limit is  $\sim 10^8 M_\odot$  with the tidal feature being disrupted over an area  $\sim 2500 \text{ pkpc}^2$  so long as the light from the host galaxy is at a similar surface brightness as the feature or fainter within the same region as the tidal feature. This could suggest that the predictions here offer a lower bound on the potential detections in LSST. However, the higher stellar mass for a given halo mass in NEWHORIZON's SMHM relation as well as its smaller volume could also contribute to the higher fractions. The fact that the three simulations of a similar resolution are in agreement regarding the tidal feature fractions and their trends with stellar and halo mass, indicates that the simulations are well converged and offer robust predictions down to the lowest stellar mass tidally disrupted galaxy that can be both resolved in the simulation and visibly detected in LSST (e.g. Martin et al. 2022).

Given that the simulations offer robust predictions, complete to their lowest stellar mass tidally disrupted galaxy that can be resolved in our mock images, they provide a lower bound on tidal feature fractions should LSST be able to probe tidal features below simulation resolution limits, which will depend on observational processing (Watkins et al. 2024). We inspect Figs 6 and 7 to examine what the simulations may predict with respect to visually detected tidal features in LSST. The increasing tidal feature fraction with stellar mass is commonly observed in both simulations and observations. This trend supports the observational evidence that the visual detectability of tidal features increases with the stellar mass of the host. A comparison of our trends of tidal feature fraction with stellar mass against future LSST observations will enable us to further test whether these predictions are consistent with the real Universe. We can also probe whether the tidal feature fraction relation with halo mass for satellite galaxies has any dependence on environment.

LSST will survey billions of galaxies. To take full advantage of this unprecedented data set we will need automated detection and classification of tidal features. There is ongoing work on automated detections of tidal features (e.g. Kado-Fong et al. 2018; Pearson et al. 2019; Desmons, Brough & Lanusse 2023a; Omori et al. 2023). Desmons et al. (2023a) have developed a self-supervised machine-learning method to detect tidal features in HSC-SSP galaxies. Further work to convert these detections into classifications of specific tidal feature morphologies would enable the measurement of trends of total and specific tidal feature fractions with stellar and halo mass over large samples for which visual classification would be prohibitively time-consuming.

## 5 CONCLUSIONS

We have presented the results of our visual classifications of tidal features around galaxies with  $10^{9.5} M_\odot \leq M_\star \lesssim 10^{12} M_\odot$  in mock images made to predicted LSST 10-yr surface brightness depths of  $\mu_g \sim 30.3 \text{ mag arcsec}^{-2}$ ,  $\mu_r \sim 30.3 \text{ mag arcsec}^{-2}$ , and  $\mu_i \sim 29.7 \text{ mag arcsec}^{-2}$  from four cosmological hydrodynamical simulations. We compared the tidal feature fractions and their behaviour as a function of stellar and halo mass. We draw the following conclusions from our results:

(i) Tidal feature fractions decrease with increasing classification confidence levels. This highlights the need to specify a confidence level of visual classification appropriate for the analysis. We present our results using  $\text{conf.} \geq 2$  (see Table 2).

(ii) The total tidal feature fractions are consistent between the simulations:  $f_{\text{NEWHORIZON}} = 0.40 \pm 0.06$ ,  $f_{\text{EAGLE}} = 0.37 \pm 0.01$ ,  $f_{\text{TNG}} = 0.32 \pm 0.01$ , and  $f_{\text{MAGNETICUM}} = 0.32 \pm 0.01$  (Table 3). This shows that the occurrence of visually identified tidal features in the cosmological hydrodynamical simulations may not be sensitive to the different subgrid physics applied by each.

(iii) The higher tidal feature fractions for NEWHORIZON as a function of stellar and halo mass may be driven by differences between its galaxy SMHM relation and that of the other simulations (Fig. 2) as well as its higher stellar mass resolution (Appendix B).

(iv) The impact of simulation stellar mass resolution is similar for all feature types, generally enhancing the likelihood of detecting a feature around a given galaxy.

(v) The specific tidal feature fractions are consistent across the four simulations (values for each simulation given in Table 3), with average specific feature fractions across the four simulations of  $f_{\text{stream/tail}} = 0.07_{-0.03}^{+0.05}$ ,  $f_{\text{shell}} = 0.0041_{-0.0007}^{+0.002}$ ,  $f_{\text{asym.}} = 0.29_{-0.005}^{+0.07}$ , and  $f_{\text{DN}} = 0.17_{-0.05}^{+0.07}$ .

(vi) Total tidal feature fractions across the simulations increase with increasing stellar mass from  $0.0^{+0.2}$ ,  $0.208_{-0.009}^{+0.02}$ ,  $0.15_{-0.03}^{+0.04}$ , and  $0.12_{-0.02}^{+0.03}$  to  $\sim 1_{-0.2}$ ,  $0.67 \pm 0.04$ ,  $0.81_{-0.04}^{+0.03}$ , and  $0.73 \pm 0.04$  for NEWHORIZON, EAGLE, TNG, and MAGNETICUM, respectively, between  $M_\star, 30 \text{ pkpc} \sim 10^{9.5} - 10^{11.3}$  (NEWHORIZON),  $10^{11.5}$  (EAGLE), and  $10^{11.8} M_\odot$  (TNG and MAGNETICUM) (Fig. 12), showing the impact of stellar mass on the merger frequency of galaxies.

(vii) Tidal feature fractions increase with increasing halo mass to a peak mass of  $M_{200, \text{crit}} \sim 10^{12.7} M_\odot$  before declining with increasing halo mass (Fig. 8).

(viii) Central galaxies exhibit increasing tidal feature fractions with increasing halo mass, a trend that can be accounted for by the relationship between tidal feature fraction and galaxy stellar mass. We expect over half of the central galaxies in haloes with masses  $\gtrsim M_{200, \text{crit}} 10^{12.7} M_\odot$  to have evidence of at least one tidal feature.

(ix) Satellite galaxies tend to exhibit a declining tidal feature fraction for  $M_{200, \text{crit}} \gtrsim 10^{13} M_\odot$ , and this trend cannot be fully accounted for by the stellar masses of the galaxies alone, suggesting a potential additional environmental dependence (Fig. 12).

(x) Comparison with observations indicates that our results are consistent with the tidal feature fractions that we would expect from LSST, complete to the tidal features with stellar masses  $\gtrsim 10^6 M_\odot$  for NEWHORIZON that are disrupted to an area  $\lesssim 36 \text{ pkpc}^2$  and  $\gtrsim 10^8 M_\odot$  for EAGLE, TNG, and MAGNETICUM that are detectable even if disrupted to an area of  $\sim 2500 \text{ pkpc}^2$ , provided the light from the host in the same region is of a similar surface brightness or fainter than the tidal feature (Appendix B). There is some indication from our comparisons with Desmons et al. (2023b) that shells might be detectable at lower host galaxy stellar masses in observations than in simulations.

We have produced predictions from cosmological simulations that are directly testable using LSST data, in particular exploration of trends of tidal feature fractions as a function of stellar and halo mass. Of particular interest will be using observed tidal features as a proxy for close encounters and mergers between galaxies, and determining how this fraction depends on halo mass to probe the galaxy assembly histories and the impact of the environment on them.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for their thoughtful report that has improved the paper. We thank Annalisa Pillepich for helpful discussions regarding ILLUSTRISTNG data. SB acknowledges funding support from the Australian Research Council through a Discovery Project DP190101943. This research includes computations using the computational cluster Katana supported by Research Technology Services at UNSW Sydney. This work was performed on the OzSTAR national facility at Swinburne University of Technology. The OzSTAR program receives funding in part from the Astronomy National Collaborative Research Infrastructure Strategy (NCRIS) allocation provided by the Australian Government, and from the Victorian Higher Education State Investment Fund (VH-ESIF) provided by the Victorian Government. The NEWHORIZON simulation was undertaken with HPC resources of CINES under the allocations c2016047637, A0020407637, and A0070402192 by Genci, KSC-2017-G2-0003 by KISTI, and as a ‘Grand Challenge’ project granted by GENCI on the AMD Rome extension of the Joliot Curie supercomputer at TGCC. A large data transfer was supported by KREONET which is managed and operated by KISTI. We acknowledge the Virgo Consortium for making their simulation data available. The EAGLE simulations were performed using the DiRAC-2 facility at Durham, managed by the ICC, and the PRACE facility Curie based in France at TGCC, CEA, Bruyères-le-Châtel. TNG100 was run on the HazelHen Cray XC40 system at the High Performance Computing Center Stuttgart as part of project GCS-ILLU of the Gauss Centre for Supercomputing (GCS). Ancillary and test runs of the IllustrisTNG project were also run on the Stampede supercomputer at TACC/XSEDE (allocation AST140063), at the Hydra and Draco supercomputers at the Max Planck Computing and Data Facility, and on the MIT/Harvard computing facilities supported by FAS and MIT MKI. The *Magneticum* simulations were performed at the Leibniz-Rechenzentrum with CPU time assigned to the Project *pr83li*. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2094 - 390783311. Parts of this research were supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

Software: ASTROPY (Astropy Collaboration 2013, 2018, 2022), NUMPY (Harris et al. 2020), PANDAS (McKinney 2010; The pandas development team 2023), SCIPY (Virtanen et al. 2020).

## DATA AVAILABILITY

NEWHORIZON data may be requested from <https://new.horizon-simulation.org/data.html>. EAGLE data used in this work are publicly available at <http://icc.dur.ac.uk/Eagle/>. The ILLUSTRISTNG data used in this work are publicly available at <http://www.tng-project.org>. MAGNETICUM data are partially available at <https://c2papcosmosim.uc.lrz.de/> (Ragagnin et al. 2017), with larger data sets on request.

## REFERENCES

- Ade P. A. R. et al., 2016, *A&A*, 594, A13  
 Agertz O., Teyssier R., Moore B., 2011, *MNRAS*, 410, 1391  
 Amorisco N. C., 2015, *MNRAS*, 450, 575  
 Arth A., Dolag K., Beck A. M., Petkova M., Lesch H., 2017, preprint (arXiv:1412.6533)  
 Astropy Collaboration, 2013, *A&A*, 558, A33  
 Astropy Collaboration, 2018, *AJ*, 156, 123  
 Astropy Collaboration, 2022, *ApJ*, 935, 167  
 Atkinson A. M., Abraham R. G., Ferguson A. M. N., 2013, *ApJ*, 765, 28  
 Banks K., Brough S., Holwerda B. W., Hopkins A. M., López-Sánchez Á. R., Philipps S., Pimblet K. A., Robotham A. S. G., 2021, *ApJ*, 921, 47  
 Bazkiaei A. E., 2023, *Thesis, Macquarie Univ.*,  
 Beck A. M. et al., 2016, *MNRAS*, 455, 2110  
 Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57  
 Bílek M. et al., 2020, *MNRAS*, 498, 2138  
 Booth C. M., Schaye J., 2009, *MNRAS*, 398, 53  
 Bosch J. et al., 2018, *PASP*, 70, S5  
 Brough S. et al., 2020, preprint (arXiv:2001.11067)  
 Brough S. et al., 2024, *MNRAS*, 528, 771  
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000  
 Cameron E., 2011, *Publ. Astron. Soc. Aust.*, 28, 128  
 Cañas R., Elahi P. J., Welker C., Lagos C. d. P., Power C., Dubois Y., Pichon C., 2019, *MNRAS*, 482, 2039  
 Canas R., del P. Lagos C., Elahi P. J., Power C., Welker C., Dubois Y., Pichon C., 2020, *MNRAS*, 494, 4314  
 Cannarozzo C. et al., 2023, *MNRAS*, 520, 5651  
 Cappellari M. et al., 2007, *MNRAS*, 379, 418  
 Chabrier G., 2003, *PASP*, 115, 763  
 Contreras-Santos A. et al., 2024, *A&A*, 683, A59  
 Crain R. A. et al., 2015, *MNRAS*, 450, 1937  
 Dalgarno A., McCray R. A., 1972, *ARA&A*, 10, 375  
 Dalla Vecchia C., Schaye J., 2008, *MNRAS*, 387, 1431  
 Dalla Vecchia C., Schaye J., 2012, *MNRAS*, 426, 140  
 Davison T. A., Norris M. A., Pfeffer J. L., Davies J. J., Crain R. A., 2020, *MNRAS*, 497, 81  
 de Zeeuw P. T. et al., 2002, *MNRAS*, 329, 513  
 Desmons A., Brough S., Lanusse F., 2023a, preprint (arXiv:2308.07962)  
 Desmons A., Brough S., Martínez-Lombilla C., De Propriis R., Holwerda B., López-Sánchez Á. R., 2023b, *MNRAS*, 523, 4381  
 Di Matteo T., Springel V., Hernquist L., 2005, *Nature*, 433, 604  
 Dolag K., Jubelgas M., Springel V., Borgani S., Rasia E., 2004, *ApJ*, 606, L97  
 Dolag K., Vazza F., Brunetti G., Tormen G., 2005, *MNRAS*, 364, 753  
 Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497  
 Dolag K., Mevius E., Remus R.-S., 2017, *Galaxies*, 5, 35  
 Donnert J., Dolag K., Brunetti G., Cassano R., 2013, *MNRAS*, 429, 3564  
 Dubois Y., Devriendt J., Slyz A., Teyssier R., 2010, *MNRAS*, 409, 985  
 Dubois Y. et al., 2014, *MNRAS*, 444, 1453  
 Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, *MNRAS*, 463, 3948  
 Dubois Y. et al., 2021, *A&A*, 651, A109  
 Elahi P. J., Power C., Lagos C. d. P., Poulton R., Robotham A. S. G., 2018, *MNRAS*, 477, 616  
 Emsellem E. et al., 2004, *MNRAS*, 352, 721  
 Fabjan D., Borgani S., Tornatore L., Saro A., Murante G., Dolag K., 2010, *MNRAS*, 401, 1670  
 Fall S., Efstathiou G., 1980, *MNRAS*, 193, 189  
 Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416  
 Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, *PASP*, 110, 761  
 Foster C. et al., 2014, *MNRAS*, 442, 3544  
 Furlong M. et al., 2015, *MNRAS*, 450, 4486  
 Garrison-Kimmel S. et al., 2019, *MNRAS*, 489, 4574  
 Genel S. et al., 2014, *MNRAS*, 445, 175  
 Gnedin O. Y., 2003, *ApJ*, 582, 141

- Haardt F., Madau P., 2001, preprint ([arXiv:astro-ph/0106018](https://arxiv.org/abs/astro-ph/0106018))
- Harris C. R. et al., 2020, *Nature*, 585, 357
- Hendel D., Johnston K. V., 2015, *MNRAS*, 454, 2472
- Hennebelle P., Chabrier G., 2011, *ApJ*, 743, L29
- Hernquist L., Spergel D. N., 1992, *ApJ*, 399, L117
- Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, *MNRAS*, 442, 2304
- Hodges J. L., 1958, *Arkiv för Matematik*, 3, 469
- Huang Q., Fan L., 2022, *ApJS*, 262, 39
- Hudson M. J. et al., 2015, *MNRAS*, 447, 298
- Ivezić Ž. et al., 2019, *ApJ*, 873, 111
- Ji I., Peirani S., Yi S. K., 2014, *A&A*, 566, A97
- Jian H.-Y., Lin L., Chiueh T., 2012, *ApJ*, 754, 26
- Kado-Fong E. et al., 2018, *ApJ*, 866, 103
- Karademir G. S., Remus R.-S., Burkert A., Dolag K., Hoffmann T. L., Moster B. P., Steinwandel U. P., Zhang J., 2019, *MNRAS*, 487, 318
- Katz N., Hernquist L., Weinberg D. H., 1992, *ApJ*, 399, L109
- Kennicutt R. C., Jr, 1998, *ApJ*, 498, 541
- Kimm T., Cen R., 2014, *ApJ*, 788, 121
- Kimm T., Cen R., Devriendt J., Dubois Y., Slyz A., 2015, *MNRAS*, 451, 2900
- Kimm T., Haehnelt M., Blaizot J., Katz H., Michel-Dansac L., Garel T., Rosdahl J., Teyssier R., 2018, *MNRAS*, 475, 4617
- Komatsu E. et al., 2011, *ApJS*, 192, 18
- Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, *Astron. Lett.*, 44, 8
- Lotz J. M., Jonsson P., Cox T. J., Croton D., Primack J. R., Somerville R. S., Stewart K., 2011, *ApJ*, 742, 103
- Lupton R., Blanton M. R., Fekete G., Hogg D. W., O'Mullane W., Szalay A., Wherry N., 2004, *PASP*, 116, 133
- Mancillas B., Duc P.-A., Combes F., Bournaud F., Emsellem E., Martig M., Michel-Dansac L., 2019, *A&A*, 632, A122
- Marinacci F. et al., 2018, *MNRAS*, 480, 5113
- Martin G., Kaviraj S., Devriendt J. E. G., Dubois Y., Pichon C., 2018, *MNRAS*, 480, 2266
- Martin G. et al., 2021, *MNRAS*, 500, 4937
- Martin G. et al., 2022, *MNRAS*, 513, 1459
- Martínez-Delgado D. et al., 2021, *MNRAS*, 506, 5030
- McKinney W., 2010, in van der Walt S., Millman J., eds, Proc. 9th Python in Science Conference. Austin, Texas, p. 56
- Mehrgan K., Thomas J., Saglia R., Parikh T., Neureiter B., Erwin P., Bender R., 2024, *ApJ*, 961, 127
- Merritt A. et al., 2020, *MNRAS*, 495, 4570
- Mihos C., 2003, preprint ([arXiv:astro-ph/0305512](https://arxiv.org/abs/astro-ph/0305512))
- Montes M., Brough S., Owers M. S., Santucci G., 2021, *ApJ*, 910, 45
- Moster B. P., Naab T., White S. D. M., 2013, *MNRAS*, 428, 3121
- Muldrew S. I., Pearce F. R., Power C., 2011, *MNRAS*, 410, 2617
- Naab T., Ostriker J. P., 2009, *ApJ*, 690, 1452
- Naiman J. P. et al., 2018, *MNRAS*, 477, 1206
- Nelson D. et al., 2018, *MNRAS*, 475, 624
- Nelson D. et al., 2019, *Comput. Astrophys. Cosmol.*, 6, 2
- O'Leary J. A., Steinwandel U. P., Moster B. P., Martin N., Naab T., 2023, *MNRAS*, 520, 897
- Olivier S. S., Seppala L., Gilmore K., 2008, in Atad-Ettdedgui E., Lemke D., eds, Proc. SPIE Conf. Ser. Vol. 7018, Advanced Optical and Mechanical Technologies in Telescopes and Instrumentation. SPIE, Bellingham, p. 70182G
- Omori K. C. et al., 2023, *A&A*, 679, A142
- Pearson W. J., Wang L., Trayford J. W., Petrillo C. E., van der Tak F. F. S., 2019, *A&A*, 626, A49
- Pillepich A. et al., 2018a, *MNRAS*, 473, 4077
- Pillepich A. et al., 2018b, *MNRAS*, 475, 648
- Planck Collaboration XVI, 2014, *A&A*, 571, A16
- Pop A.-R., Pillepich A., Amorisco N. C., Hernquist L., 2018, *MNRAS*, 480, 1715
- Press W. H., Schechter P., 1974, *ApJ*, 187, 425
- Proctor K. L., Lagos C. d. P., Ludlow A. D., Robotham A. S. G., 2023, *MNRAS*, 527, 2624
- Ragagnin A., Dolag K., Biffi V., Cadolle Bel M., Hammer N. J., Krukau A., Petkova M., Steinborn D., 2017, *Astron. Comput.*, 20, 52
- Remus R.-S., Forbes D. A., 2022, *ApJ*, 935, 37
- Robertson B. E. et al., 2019, *Nat. Rev. Phys.*, 1, 450
- Robotham A. S. G. et al., 2014, *MNRAS*, 444, 3986
- Román J., Trujillo I., Montes M., 2020, *A&A*, 644, A42
- Ryden B. S., Gunn J. E., 1987, *ApJ*, 318, 15
- Samuel J., Wetzel A., Santistevan I., Tollerud E., Moreno J., Boylan-Kolchin M., Bailin J., Pardasani B., 2022, *MNRAS*, 514, 5276
- Schaye J., 2004, *ApJ*, 609, 667
- Schaye J., Dalla Vecchia C., 2008, *MNRAS*, 383, 1210
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Schaye J. et al., 2023, *MNRAS*, 526, 4978
- Schmidt M., 1959, *ApJ*, 129, 243
- Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, *MNRAS*, 380, 877
- Spitzer L., 1962, *Physics of Fully Ionized Gases*, 2nd edn. Interscience Publishers, New York
- Springel V., 2010, *MNRAS*, 401, 791
- Springel V., Hernquist L., 2003, *MNRAS*, 339, 289
- Springel V., Yoshida N., White S. D. M., 2001, *New Astron.*, 6, 79
- Springel V., Di Matteo T., Hernquist L., 2005a, *MNRAS*, 361, 776
- Springel V. et al., 2005b, *Nature*, 435, 629
- Springel V. et al., 2018, *MNRAS*, 475, 676
- Sureshkumar U. et al., 2024, preprint ([arXiv:2402.18520](https://arxiv.org/abs/2402.18520))
- Sutherland R. S., Dopita M. A., 1993, *ApJS*, 88, 253
- Teklu A. F., Remus R.-S., Dolag K., Beck A. M., Burkert A., Schmidt A. S., Schulze F., Steinborn L. K., 2015, *ApJ*, 812, 29
- Teklu A. F., Remus R.-S., Dolag K., Burkert A., 2017, *MNRAS*, 472, 4769
- Teyssier R., 2002, *A&A*, 385, 337
- Teyssier R., Moore B., Martizzi D., Dubois Y., Mayer L., 2011, *MNRAS*, 414, 195
- The pandas development team, 2023, Pandas-Dev/Pandas: Pandas, Zenodo
- Toomre A., Toomre J., 1972, *ApJ*, 178, 623
- Trebitsch M. et al., 2021, *A&A*, 653, A154
- Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, *A&A*, 506, 647
- Valenzuela L. M., Remus R.-S., 2023, preprint ([arXiv:2308.08443](https://arxiv.org/abs/2308.08443))
- van den Bosch F. C., 2002, *MNRAS*, 331, 98
- Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, 436, 3031
- Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518
- Watkins A. E., Kaviraj S., Collins C. C., Knapen J. H., Kelvin L. S., Duc P.-A., Román J., Mihos J. C., 2024, *MNRAS*, 528, 4289
- Weinberger R. et al., 2017, *MNRAS*, 465, 3291
- Wiersma R. P. C., Schaye J., Smith B. D., 2009, *MNRAS*, 393, 99
- Willmer C. N. A., 2018, *ApJS*, 236, 47
- Yang X., Mo H. J., Jing Y. P., van den Bosch F. C., 2005, *MNRAS*, 358, 217
- Yoon Y., Lim G., 2020, *ApJ*, 905, 154

## APPENDIX A: OVERVIEW OF SUBGRID PHYSICS MODELS

### A1 NewHorizon

NEWHorizon uses the adaptive mesh refinement code (AMR) RAMSES (Teyssier 2002). Star formation is set to occur in regions with hydrogen gas number density higher than  $10 \text{ cm}^{-3}$ , following the Schmidt law:  $\dot{\rho}_* = \epsilon_* \rho_g / t_{\text{ff}}$ , which relates the star formation rate mass density,  $\dot{\rho}_*$ , to the gas mass density,  $\rho_g$ , local free fall time of the gas,  $t_{\text{ff}}$ , and lastly a varying star formation efficiency,  $\epsilon_*$  (Kimm et al. 2018; Trebitsch et al. 2021). The star formation efficiency is computed following the prescriptions of Hennebelle & Chabrier (2011), which account for the turbulence of the gas. H and He gas are modelled using an equilibrium chemistry model with a homogeneous UV background. This gas is allowed to cool

to  $\simeq 10^4$  K through collisional ionization, excitation, recombination, bremsstrahlung, and Compton cooling. This gas, if metal-enriched, is allowed to further cool down to 0.1 K using the rates from Dalgarno & McCray (1972) and Sutherland & Dopita (1993).

The impact of supernovae (SNe) on the gas and therefore star formation is another process below the resolution limit of cosmological simulations. NEWHORIZON models SNe as explosions releasing  $10^{51}$  erg of energy into the surrounding baryons, with a minimum of  $6 M_{\odot}$  needed in a star to go supernova (SN), they model a specific SNe frequency of  $0.03 M_{\odot}^{-1}$  using the mechanical SN feedback scheme described in Kimm & Cen (2014) and Kimm et al. (2015).

NEWHORIZON implements two modes of active galactic nuclei (AGNs) feedback, thermal and kinetic (radio and quasar), depending on the accretion rate of the supermassive black hole (following Dubois et al. 2010; Teyssier et al. 2011). AGN feedback can help remove gas from the galaxy and/or heat it up making it unsuitable for star formation.

### A2 EAGLE

EAGLE runs on the Lagrangian smooth particle hydrodynamic (SPH) code GADGET-3. Star formation is implemented stochastically, following the pressure law scheme of Schaye & Dalla Vecchia (2008), implementing the observed Kennicutt–Schmidt (Schmidt 1959; Kennicutt 1998) relation into the simulation. Star formation is set to occur in regions with a metallicity-dependent hydrogen mass density threshold, formulated by Schaye (2004), to account for the more efficient radiative transfer in metal-rich gas clouds. Stellar particles are assumed to follow a Chabrier IMF (Chabrier 2003), with stars spanning in the range  $0.1$ – $100 M_{\odot}$ . EAGLE accounts for stellar winds, radiation, and SNe feedback by implementing their collective feedback through thermal heating, distributing the energy produced at each time-step by a stellar particle to the neighbouring particles using the stochastic thermal feedback scheme of Dalla Vecchia & Schaye (2012).

The radiative cooling and heating rates of gas resolution elements at a given density, temperature, and redshift are computed with the software CLOUDY (Ferland et al. 1998). The gas is assumed to be optically thin, in an ionization equilibrium, and exposed to the cosmic microwave background and a spatially homogeneous UV/X-ray background that evolves over time (Haardt & Madau 2001; Wiersma, Schaye & Smith 2009). EAGLE implements one mode of AGN feedback through stochastic heating, following the prescriptions of Dalla Vecchia & Schaye (2008) and Booth & Schaye (2009).

### A3 IllustrisTNG

TNG used the moving mesh code AREPO (Springel 2010), which differs from AMR code (e.g. RAMSES) and SPH code (e.g. GADGET-3) in that it uses an unstructured mesh defined by a Voronoi tessellation of a set of discrete points. This alternative methodology was formulated to address issues with AMR and SPH, which impact their accuracy in particular situations (e.g. suppression of fluid instabilities in SPH and lack of Galilean invariance and overmixing present in AMR codes).

TNG models star formation stochastically, treating the star formation and pressurization of a multiphase interstellar medium following Springel & Hernquist (2003). In the model by Springel & Hernquist (2003), cold gas above a density threshold of  $0.1 \text{ H cm}^{-3}$  forms star particles following the empirically defined Kennicutt–Schmidt relation (Schmidt 1959; Kennicutt 1998) and the Chabrier IMF (Chabrier

2003). Under the Springel & Hernquist (2003) prescription, SNe pressurizes the gas and can lead to enhanced star formation. The stellar winds produced by star formation can carry and inject kinetic energy into the surrounding gas. They are modelled as described in Vogelsberger et al. (2013) and Pillepich et al. (2018a).

The radiative cooling of gas is modelled accounting for its metal enrichment (following Wiersma et al. 2009), and a time-evolving homogeneous UV background with self-shielding corrections in dense interstellar medium (following Katz, Hernquist & Weinberg 1992; Faucher-Giguère et al. 2009). Following the models of Vogelsberger et al. (2013), radiative cooling is further influenced by the nearby radiation fields of AGNs.

TNG implements subgrid AGN feedback with a radio and quasar mode. For accretion rates below 5 per cent of the Eddington limit, the radio-mode feedback is active, injecting bursty thermal energy into a  $\sim 50$  pc bubble displaced away from the host galaxy (Sijacki et al. 2007). For higher accretion rates the quasar mode is active, injecting thermal energy continuously into the adjacent gas (Di Matteo, Springel & Hernquist 2005; Springel, Di Matteo & Hernquist 2005a). The details of all subgrid physics can be found in Weinberger et al. (2017), Pillepich et al. (2018a), and Nelson et al. (2019).

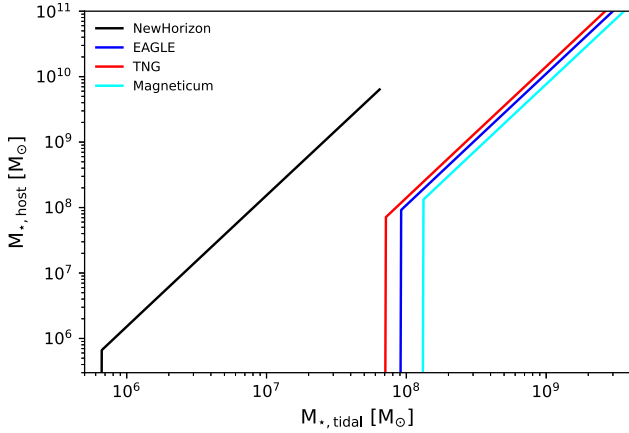
### A4 Magneticum

MAGNETICUM PATHFINDER simulations are a suite of cosmological hydrodynamical simulations, ranging in box size from  $25.6^3$  to  $3818^3 \text{ cMpc}^3$ . We use the BOX4-UHR simulation, which models the physics of dark matter and baryons in a  $68^3 \text{ cMpc}^3$  box. MAGNETICUM PATHFINDER has been performed with a modified version of the Lagrangian SPH code GADGET-3, with modifications on the viscosity and treatment of the kernel functions following Dolag et al. (2005), Donnert et al. (2013), and Beck et al. (2016). Star formation and the stellar wind-driven kinetic feedback are modelled following Springel & Hernquist (2003). Each star particle represents a stellar population following a Chabrier IMF (Chabrier 2003) and each gas particle can form up to four stars. Star formation and metal enrichment from SN feedback and asymptotic red giant branch stars is modelled following the Springel & Hernquist (2003) as well as the local metallicity-dependent processes (Wiersma et al. 2009; Dolag, Mevius & Remus 2017). The radiative cooling and heating rates of gas resolution elements at a given density, temperature, and redshift are computed with the software CLOUDY (Ferland et al. 1998). The radiative cooling accounts for a UV/X-ray background following the prescriptions of Haardt & Madau (2001) and Wiersma et al. (2009).

The subgrid physics treatment of supermassive black holes and their AGN feedback is implemented as described by Fabjan et al. (2010) and Hirschmann et al. (2014). The implemented black holes feedback scheme accounts for a transition from quasar to radio mode according to Sijacki et al. (2007). Note that the black holes in this simulation are not pinned to the potential minimum. Thermal conduction is implemented according to Dolag et al. (2004) but following Arth et al. (2017), with one-twentieth of the classical Spitzer value (Spitzer 1962). The details of all subgrid physics are described by Teklu et al. (2015) and Hirschmann et al. (2014).

## APPENDIX B: TIDAL FEATURE DETECTION LIMITS

The stellar mass resolutions of our simulations as well as the surface brightness limits of our mock images limits the minimum stellar masses of the tidal features we can resolve. Through some simple modelling we investigate the tidal feature detection limits for each



**Figure B1.** The maximum stellar mass from the host galaxy allowed in the same region as the tidal feature without making it undetectable as a function of the tidal feature’s stellar mass. For a tidal feature to be detectable it has  $\text{SNR} > 5$ .

of our simulations. Following Martin et al. (2022) we calculate the signal-to-noise ratio for a tidal feature against its background galaxy and take an  $\text{SNR}_{\text{tidal}} > 5$  as the range for a detectable tidal feature.

$$\text{SNR}_{\text{tidal}} = \frac{n_{\text{tidal}}}{\sigma_{\text{tidal}}} = \frac{n_{\text{tidal}}}{\sqrt{2n_{\text{host}} + n_{\text{tidal}}}} \quad (\text{B1})$$

Here,  $n_{\text{tidal}}$  is the number of particles that comprise a tidal feature, and  $n_{\text{host}}$  is the number of particles from the host in the same region as the tidal feature. Using this equation and the simulation mass resolutions (Table 1) we determined for each of our simulations, given a tidal feature of a specific stellar mass, what was the maximum amount of stellar mass from the host that could be in the same region without making the tidal feature undetectable. The results of this calculation are shown in Fig. B1, where we show the maximum host stellar mass contamination in the same region as the tidal feature as a function of tidal feature stellar mass. We can see that each simulation has a drop off corresponding to where its mass resolution is limiting its ability to resolve tidal features with  $\text{SNR} > 5$  without there being virtually no contamination from the host galaxy. From this calculation we can predict that NEWHORIZON is able to resolve tidal features that are  $\gtrsim 2$  orders of magnitude less massive than in the other three simulations. NEWHORIZON should also be able to resolve tidal features with much greater contamination from the host galaxy when compared to the other three simulations.

To investigate the minimum mass galaxy that can be tidally disrupted and detected in our mock images, we must construct a model for the surface brightnesses of tidal features stemming from different stellar mass hosts. We build a simple model assuming the following:

- (i) the mass-to-light ratio for the stellar particles is 1;

- (ii) the light from a tidal feature is homogeneously distributed across a square region.

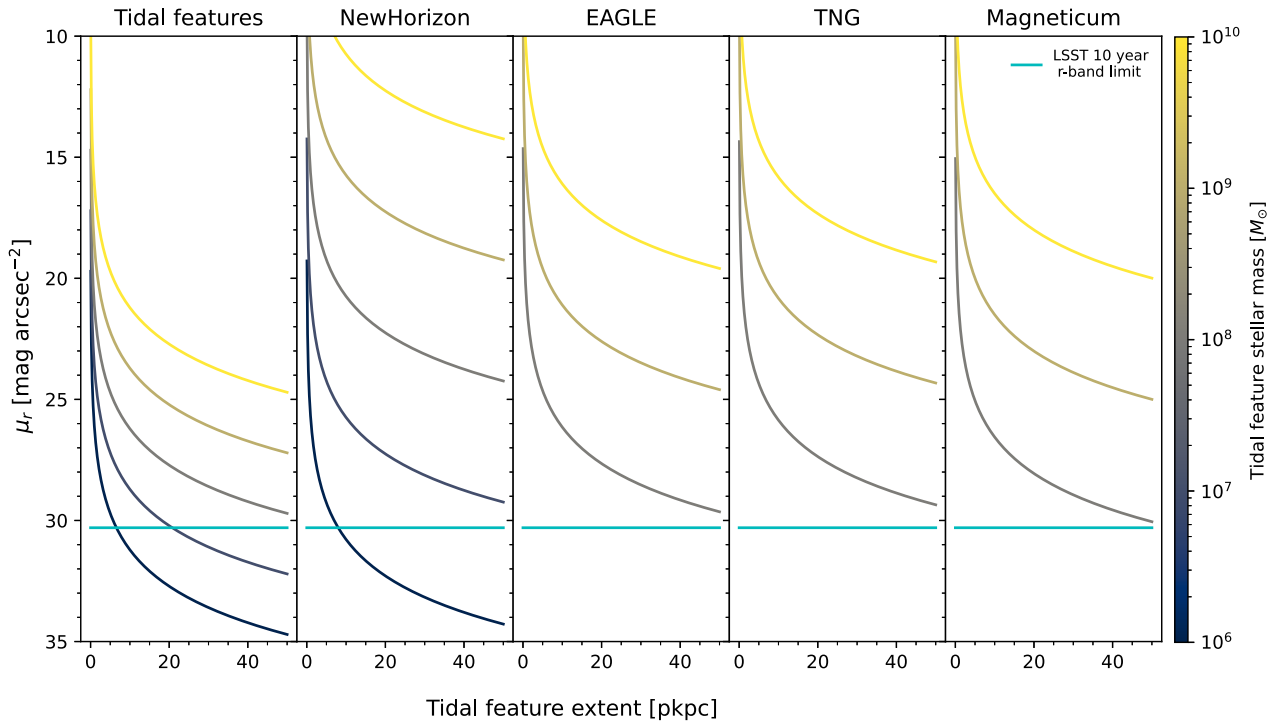
While the mass-to-light ratios vary as a function of galaxy morphology (e.g. de Zeeuw et al. 2002; Emsellem et al. 2004; Cappellari et al. 2007) and could vary significantly across a galaxy (e.g. Mehrgan et al. 2024), 1 is a physical mass-to-light ratio and eases computation substantially. The homogeneous distribution of light across a square area is a simplifying assumption that allows us to compute the surface brightness of tidal features without simulating features of varying light distributions and regions. If the mass-to-light ratio for the tidally disrupted galaxy is above 1 or the distribution of the light is inhomogeneous resulting in some regions that are brighter than the average brightness of the tidal feature, the resultant tidal features would be detectable despite having lower stellar mass than our estimates.

We convert from the physical surface brightness of the stellar particles in  $L_{\odot} \text{pc}^{-2}$  to LSST  $r$ -band surface brightness in  $\text{mag arcsec}^{-2}$  using the following relationship:

$$\mu_r = -2.5 \log_{10} (\text{SB} (L_{\odot} \text{pc}^{-2})) + M_{\odot,r} + 5 \log_{10} \left( \frac{648000}{10\pi} \right), \quad (\text{B2})$$

where we use the estimated absolute AB magnitude of the Sun in the LSST  $r$  band,  $M_{\odot,r} = 4.64$  (Willmer 2018).

The leftmost panel in Fig. B2 shows the surface brightness with tidal feature extent for tidal features of stellar masses  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$ , and  $10^{10} M_{\odot}$ . The surface brightness profiles in the remaining panels of Fig. B2 indicate the maximum allowed host brightness within the same area to still allow the tidal feature to be resolved against the host background contamination. We compute these assuming a minimum  $\text{SNR}_{\text{tidal}} > 5$  and again assuming the host light is distributed evenly in the background. Our results highlight that even tidal features from galaxies that are  $10^6 M_{\odot}$  are above the mock image surface brightness limits provided their radius remains within  $\sim 6$  pkpc. However, only NEWHORIZON has the sufficient resolution to resolve them so long as the light from the host galaxy in the same region as the tidal feature is of a similar brightness or fainter. The surface brightness profile for the maximum allowed host contamination indicates that these features are only detectable if there is a similar amount of light from the host in this region. From Fig. B1 we can estimate that NEWHORIZON will resolve a  $10^6 M_{\odot}$  feature within the survey detection limits if there is  $\lesssim 1.5 \times 10^6 M_{\odot}$  of stellar mass from the host within the same region, provided the feature is no more extended than  $\sim 6$  pkpc. For the larger cosmological simulations, we can detect tidal features of stellar mass  $\gtrsim 10^8 M_{\odot}$  even if they have been disrupted over extents  $\gtrsim 50$  pkpc, so long as the host stellar contamination is  $\lesssim 1.5 \times 10^8 M_{\odot}$  within the same region. As the tidal features become more massive than these lower limits, the maximum surface brightness of the host galaxy within the same region as the tidal feature increases for each of the simulations.



**Figure B2.** The  $r$ -band surface brightness of the model tidal features as a function of their radii and the corresponding maximum possible surface brightness of the light from the host galaxy within the tidal features area such that the tidal feature is still detectable for each of the simulations. The tidal features panel shows the surface brightness of each tidal feature as a function of the radius of the circle the light is spread over. The remaining panels show the maximum surface brightness of the light from the host that can be within the same region as the tidal feature without making it undetectable for each simulation. The colours correspond to the stellar masses of the tidal features,  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$ , and  $10^{10} M_{\odot}$ . The LSST 10 yr  $r$  band,  $3\sigma$ ,  $10 \text{ arcsec} \times 10 \text{ arcsec}$ , limiting surface brightness is shown in Rubin Turquoise.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.