Probing galactic growth with low surface brightness (LSB) data from deep imaging surveys

Elisabeth Sola

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE de Physique et Chimie-Physique (ED182)

Observatoire astronomique de Strasbourg (UMR 7550)

THÈSE présentée par :

Elisabeth SOLA

soutenue le : **07 septembre 2023**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Astrophysique

Sonder la croissance des galaxies avec les données de faible brillance de surface (LSB) de relevés d'imagerie profonde

THÈSE dirigée par :

M. DUC Pierre-Alain	Directeur de recherche, Observatoire astronomique de Strasbourg

RAPPORTEURS :

M. EMSELLEM Eric	Professeur, Observatoire européen austral (ESO)
M. KNAPEN Johan	Professeur, Institut d'Astrophysique des Canaries (IAC)

AUTRES MEMBRES DU JURY :

Mme DURRET Florence	Astronome, Institut d'Astrophysique de Paris
M. KAVIRAJ Sugata	Professeur, Université de Hertfordshire
Mme LANÇON Ariane	Professeure, Observatoire astronomique de Strasbourg

Publication list

2023 *Quantitative analyses of LSB tidal debris as a function of the host galaxy's properties and environment*
E. Sola; P.-A. Duc; and collaborators
In preparation

2022 *Characterization of low surface brightness structures in annotated deep images*
E. Sola; P.-A. Duc; F. Richards; A. Paiement; M. Urbano; J. Klehammer; M. Bílek; J.-C. Cuillandre; S. Gwyn; A. McConnachie
Astronomy & Astrophysics, Volume 662, id.A124, 29 pp.

2022 *Origin of the differences in rotational support among early-type galaxies: The case of galaxies outside clusters*
M. Bílek; P.-A. Duc; **E. Sola**
Astronomy & Astrophysics, Volume 672, id.A27, 26 pp.

2022 *Preparing for low surface brightness science with the Vera C. Rubin Observatory: Characterization of tidal features from mock images*
G. Martin et al (incl. **E. Sola**)
Monthly Notices of the Royal Astronomical Society, Volume 513, Issue 1, pp.1459-148

2020 *Learnable Gabor modulated complex-valued networks for orientation robustness*
F. Richards; A. Paiement; X. Xie; **E. Sola** ; P.-A. Duc
eprint arXiv:2011.11734

Remerciements

À mes parents, et à mes amis.

Je voudrais commencer par te remercier infiniment Pierre-Alain pour m'avoir donné la chance d'effectuer cette thèse à l'Observatoire astronomique de Strasbourg, pour m'avoir guidée et avoir cru en moi. Tu as toujours été très présent, disponible et à l'écoute malgré la quantité de travail considérable que tu as en tant que Directeur. Merci pour tous ces échanges, conseils et discussions. J'ai eu la chance de passer trois années de thèse formidables et tu y as très fortement contribué.

Je tiens ensuite à remercier Eric Emsellem, Johan Knapen, Florence Durret, Ariane Lançon et Sugata Kaviraj pour avoir accepté de consituer mon jury de thèse, pour leurs retours et le temps qu'ils m'accordent. Un grand merci à Rodrigo Ibata et Florence Durret pour avoir été dans mon comité de suivi et m'avoir apporté leurs conseils et soutien.

Durant ma thèse, j'ai eu l'occasion de travailler et d'échanger avec de nombreuses personnes dans différentes collaborations. Je souhaite en particulier remercier Felix Richards et sa directrice de thèse Adeline Paiement, qui ont développé le serveur d'annotations et sans qui mon sujet de thèse n'aurait pas été possible, pour leur temps, leur aide et pour nos réunions régulières. Merci également aux membres des collaborations MATLAS et CFIS avec qui j'ai pu échanger, notamment Francine Marleau, Oliver Müller, Michal Bílek, Melina Poulain, Nick Heester, Ivana Ebrova, Jean-Charles Cuillandre ainsi que Pierre Galois et Julie Klehammer que j'ai pu co-encadrer durant leurs stages.

Evidemment, j'adresse un grand merci aux personnes avec qui j'ai travaillé plus étroitement. Merci Olivier Marchal, pour avoir fait de notre bureau le meilleur de la coupole, pour nos délires, nos blagues de haute qualité et pour le temps passé ensemble, ainsi que pour l'aide sur le serveur d'annotation. Merci Mathias Urbano pour avoir été d'abord un stagiaire au top puis un super co-doctorant. Merci pour tes jeux de mots formidables (je n'ai pas trouvé d'acronyme suffisamment bien ici mais je compte sur bien sur la PAD'awan team pour en trouver un), ta bonne humeur à l'obs et en mission, et pour ton aide pour les annotations et l'article.

De manière plus générale, je remercie tout le personnel de l'Observatoire pour le temps passé auprès d'eux, pour les discussions et l'aide qu'ils ont pu m'apporter. Un grand merci tout particulièrement à Nicolas Martin, pour m'avoir intégrée à ton groupe, pour m'avoir fait rencontrer de nombreuses personnes, pour les évènements en-dehors de l'Observatoire. Merci à Anke (que je reverrais bientôt), Zhen, Rapha, Rodrigo. Merci à Véronique et Sandrine pour votre bienveillance, votre aide et disponibilité.

Je tiens à remercier très sincèrement toute l'équipe du Jardin des Sciences et notamment Milène Wendling pour m'avoir donné la possibilité de faire une mission de vulgarisation pendant deux années. J'ai beaucoup appris auprès de vous, j'ai pu faire de

nombreuses activités différentes, rencontrer des publics variés, me faire des contacts dans le monde de la vulgarisation, participer à différents évènements sans parler du travail sur le nouveau planétarium. Cette mission m'a vraiment tenu à coeur et je compte poursuivre la vulgarisation autant que possible dans la suite de ma carrière. Merci pour votre accueil, votre bonne humeur et toutes les occasions que vous m'avez offertes.

Merci à tous les doctorants et doctorantes qui ont partagé mes années, Jakub, Ivan, Julien, Mathias, Mei, Srikanth, Thibault, Yacine, et en particulier un grand merci à PA, Thibaut, Wassim et Lucie. Merci pour les repas du midi ensemble, les bars à jeux, les restaus, les karaokes et plus globalement pour le temps passé ensemble. On n'oublie pas que l'important c'est l'essentiel.

Qu'aurait été ma thèse sans vous, Emilie et Amandine ? Je n'arrive même pas à me le représenter. Sans vous la thèse aurait été bien moins drôle, sans votre soutien dans les galères et les bons moments, les repas du midi, les discussions et pauses que je n'ai prises que bien trop tard, pour les délires, les voyages, les moments où on râlait, ceux de réconfort, ceux de rigolade et ceux plus tristes. Même si je n'étais pas dans le même bureau (heureusement, comment vous aurais-je supportées sinon?!) vous êtes indissociables de ces trois superbes années. Merci les filles. Je suis contente qu'on continue toutes les trois en astro, mais triste car les postdocs seront différents sans vous. Pas le choix, on se revoit bientôt parce qu'on a des revanches sur Isaac à faire, des visites de Lyon, des balades sur Outer Wilds, du JdR, des concerts, des bouilloires à tester... On se revoit bientôt, dans le même labo et le même bureau évidemment, j'ai hâte d'entendre vos accents incroyables.

Je tiens maintenant à remercier toutes les personnes qui me sont chères, qui m'ont entourée, qui ont cru en moi, qui m'ont aidée et soutenue tout au long de mon parcours. C'est en grande partie grâce à vous que j'en suis là aujourd'hui et que j'ai pu réaliser le rêve que j'avais étant petite. Evidemment, ma famile. Papa, Maman, Cathy, Alex (sans oublier Comète, Cheasly et Châtaigne), merci pour tout, les mots ne suffisent pas et vous le savez. Merci. Merci Papy, Mamies (Mad et Nicole) et Tata et Olivier pour votre soutien et votre aide.

Je souhaite remercier ensuite mes amis, avec qui j'ai partagé tellement de choses. Merci Rémi et Gaëlle, pour votre soutien et votre écoute depuis un bon moment maintenant. Merci à mes amis de Strasbourg, et notamment à Laure et Quentin. Qu'est-ce que je serais devenue sans la bande de TPS ? Merci pour le temps à l'école et en-dehors, les repas, les délires, les photos, les vacances, les Nouvel an et évidemment les jeux (dédicace au Schluck'n'Spiel où on a passé tellement de week-ends que c'est un peu notre $2^{\text{ème}}$ maison). Merci d'être là, sur le discord et en vrai, vous êtes si importants pour moi: Guillaume, Julien, Maxime, Samuel, Solenn et Adrian. Adrian, je n'ai pas non plus les mots, je suis tellement chanceuse et heureuse de t'avoir à mes côtés. Merci pour ton amour, ton soutien sans faille, ta présence, ta bonne humeur, tes conseils, pour m'avoir aidée et soutenue dans mes nombreux moments de doute et d'angoisse.

À ma famille et à mes amis, encore une fois: merci. Je vous aime.

Abstract

Probing galactic growth with low surface brightness (LSB) data from deep imaging surveys

According to hierarchical models of galaxy formation, galaxies grow through galaxy mergers and accretion of gas. Such interactions leave tidal debris, such as tidal tails, streams or shells, as well as extended stellar haloes. This debris is the fossil record of the past assembly history of a galaxy, as the morphology and properties of this debris are a function of the type of merger that occurred. Detecting and characterizing this collisional debris is essential to better understand and constrain the evolution processes of galaxies. However, its detection is made complicated by its low surface brightness (LSB). Dedicated observing strategies and data reduction pipelines are necessary to disclose LSB features in deep images.

In order to constrain models of galactic evolution, large quantitative analyses of these faint tidal features are needed. The goal is to compare them to the predictions of numerical simulations, to put constraints on these models. It is necessary to probe galaxies spanning large ranges of mass and environment (field, group, cluster) to have a broad picture of galactic formation. Several studies performed qualitative analyses and produced a census of tidal debris around galaxies, but quantitative analyses were restricted to limited samples. The aim of this PhD thesis is to perform a quantitative study of LSB tidal features for a large sample of galaxies in various environments.

To that end, we used a new dedicated annotation tool and associated database that answers this need. It enables collaborators to precisely delineate the shape of LSB structures superimposed on deep images, and from it extract a wealth of information about these features. We developed analysis tools to obtain geometrical properties, surface brightness, colours, but also luminosity and stellar mass of tidal features and extended stellar haloes. We provided a first characterization of individual LSB structures around a sample of 475 nearby massive galaxies observed in deep images from four surveys carried out at the Canada-France-Hawaii Telescope (CFHT): MATLAS, CFIS, NGVS and VESTIGE.

Then, we focussed on the impact of the host galaxy mass and environment on galactic growth as seen through LSB features. We found that more massive galaxies host more tidal debris, which is also more luminous, as well as more luminous stellar haloes in the case of elliptical galaxies. The trends we saw are consistent with the hierarchical paradigm: more massive galaxies have undergone more mergers, hence have more debris and brighter haloes. The cluster environment affects spiral galaxies (through ram-pressure stripping) and tidal features. The galaxy mass seems to have a more prominent impact than the environment on tidal features and haloes.

We discussed the limitations of our study but also the future prospects. The comparison between observations and simulations using the annotation tool is an ongoing and very promising prospect. The use of deep learning algorithms, trained on our annotation database, is also promising to automatically detect tidal features in future large surveys.

Résumé

Sonder la croissance des galaxies avec les données de faible brillance de surface (LSB) de relevés d'imagerie profonde

Selon les modèles hiérarchiques de formation des galaxies, les galaxies se développent par le biais de fusions galactiques et d'accrétion de gaz intergalactique. Ces interactions laissent des débris de marée, tels que des queues de marée, des courants stellaires ou des coquilles, ainsi que des halos stellaires étendus. Ces débris sont les témoins fossiles de l'histoire de l'assemblage d'une galaxie, car leur morphologie et leurs propriétés dépendent du type de fusion qui s'est produit. La détection et la caractérisation de ces débris de collision sont essentielles pour mieux comprendre et contraindre les processus d'évolution des galaxies. Cependant, leur détection est rendue compliquée par leur faible brillance de surface (LSB). Des stratégies d'observation dédiées et des chaînes de réduction de données sont nécessaires pour révéler les structures LSB dans les images profondes.

Afin de contraindre les modèles d'évolution galactique, de grandes analyses quantitatives de ces débris de marée sont nécessaires. L'objectif est ensuite de les comparer aux prédictions des simulations numériques, afin de contraindre ces modèles. Il est nécessaire de sonder des galaxies couvrant de larges gammes de masse et d'environnements (champ, groupe, amas) pour avoir une image globale de la formation galactique. Plusieurs études ont effectué des analyses qualitatives et un recensement des débris de marée, mais les analyses quantitatives ont été limitées à des échantillons restreints. Le but de cette thèse est de réaliser une étude quantitative des stuctures de marée LSB pour un large échantillon de galaxies dans différents environnements.

Pour cela, nous avons utilisé un nouvel outil d'annotation dédié, et une base de données associée, qui répondent à ce besoin. Il permet aux collaborateurs de délimiter précisément la forme des structures LSB superposées aux images profondes, et d'en extraire une multitude d'informations sur ces débris. Nous avons développé des outils d'analyse pour obtenir les propriétés géométriques, la brillance de surface, les couleurs, mais aussi la luminosité et la masse stellaire des structures de marée et des halos stellaires étendus. Nous avons fourni une première caractérisation des structures LSB individuelles autour d'un échantillon de 475 galaxies massives et proches observées dans des images profondes provenant de quatre relevés effectués au Télescope Canada-France-Hawaï (CFHT) : MATLAS, CFIS, NGVS et VESTIGE.

Ensuite, nous nous sommes concentrés sur l'impact de la masse et de l'environnement de la galaxie hôte sur la croissance galactique à travers les débris LSB. Nous avons constaté que les galaxies plus massives hébergent plus de débris de marée, et des débris plus lumineux, ainsi que des halos stellaires plus lumineux pour les galaxies elliptiques. Les tendances observées sont cohérentes avec le paradigme hiérarchique : les galaxies plus massives ont subi plus de fusions, d'où un plus grand nombre de débris et des halos plus lumineux. L'environnement des amas influe sur les galaxies spirales (par le biais de la pression dynamique) et sur les débris de marée. La masse des galaxies semble avoir un impact plus important que l'environnement sur les structures de marée et les halos.

Nous avons discuté des limites de notre étude mais aussi des perspectives futures. La comparaison entre les observations et les simulations à l'aide de l'outil d'annotation est une perspective en cours et très prometteuse. L'utilisation d'algorithmes d'apprentissage profond, entraînés sur notre base de données d'annotations, est également prometteuse pour détecter automatiquement les débris des marées dans les grands relevés du ciel à venir.

Contents

List of Figures

Glossary

ADU	*Analog-Digital Unit*
CADC	*Canadian Astronomy Data Centre*
CDS	*Centre de Données astronomiques de Strasbourg*
CSS	*Cascading Style Sheets*
CFHT	*Canada-France-Hawaii Telescope*
CFIS	*Canada-France Imaging Survey*
DL	*Deep Learning*
ETG	*Early Type Galaxy*
FITS	*Flexible Image Transport System*
HiPS	*Hierarchical Progressive Surveys*
HTML	*HyperText Markup Language*
MOC	*Multi-Order Coverage map*
ML	*Machine Learning*
LSB	*Low Surface Brightness*
LTG	*Late Type Galaxy*
MATLAS	*Mass Assembly of early-Type GaLAxies with their fine Structures*
NGVS	*Next Generation Virgo Cluster Survey*
PSF	*Point Spread Function*
VESTIGE	*Virgo Environmental Survey Tracing Ionised Gas Emission*

Chapter 1 General introduction

1.1 Galactic evolution

1.1.1 Galaxies

Galaxies are conglomerates of billions of stars, gas and dust held together by the gravitational force, and likely enclosed in a dark matter halo. They exhibit a large variety of properties, from their morphologies, colours, luminosities, masses, kinematics, ages or metallicities. Understanding how these objects form and evolve through time remains one of the most fundamental cosmological questions.

The discovery that galaxies are actually extragalactic systems only dates back a century ago. In the XVIII$^{\text{th}}$ century, the French astronomer Charles Messier published a catalogue of diffuse fuzzy objects whose nature was not known but that should not be mistaken for comets. Messier himself had confused one of these diffuse objects (later known as the Crab Nebula) with Halley's comet. In the 1920's the nature of these diffuse objects, called nebulae, was still highly controversial. Some astronomers such as Shapley defended the idea that they were part of the Milky Way, which was seen as being the entire Universe. On the contrary, other astronomers such as Curtis argued in favor of 'island universes', with these nebulae being separate and distant external galaxies. The Great Debate between Shapley and Curtis was held in 1920, and reported in Shapley & Curtis (1921) which summarizes the arguments from both sides. A few years later, Hubble pointed a telescope towards the Andromeda Nebula and found a Cepheid variable star. Using the period-luminosity relation of Cepheid stars, discovered in 1912 by Henrietta Leavitt (Leavitt & Pickering 1912), Hubble determined the distance of Andromeda to be at least 8 times larger than the furthest star in the Milky Way, hence proving the existence of other galaxies and putting an end to the Great Debate. His results on the distances to Andromeda and other galaxies were published in Hubble (1925, 1929b,a). Nowadays, billions of galaxies are known in the observable Universe.

As galaxies are found to display a wide range of properties, several schemes were developed with the aim of classifying them. In particular, Hubble (1926) proposed a classification based on their morphology, also known as the Hubble Tuning Fork. Other classifications were proposed such as the one by De Vaucouleurs that extended Hubble's scheme to include the presence of structural components such as bars, spiral arms or rings (de Vaucouleurs 1959). It introduced the T-type, a numerical value indicating the class to which a given galaxy belongs. The Hubble classification is still widely used and it divides galaxies into two main groups: elliptical/lenticular, spiral/barred spiral. Some galaxies do not fit in this tuning fork diagram, such as irregular galaxies, dwarf galaxies (which are much smaller galaxies containing millions to several billions stars) or mergers. An illustration of the classification scheme is presented in Figure 1.1.

Elliptical galaxies, shown on the left part of the diagram, have spheroidal shapes and relatively smooth, featureless light distributions. They are mainly supported by the random motion of their stars. They are named 'E_n' were n is their degree of apparent ellipticity on the sky, defined as $n = 10 \times (1 - \frac{b}{a})$ where a (resp b) is the length of the semi-major (resp semi-minor) axis of the ellipse. Hence, nearly circular galaxies are E0 while more flattened ones are E7. Spiral galaxies have a flattened disc, a spiral structure and a bulge (i.e. a central concentration of stars). Some spirals show a bar-like structure and are called barred spirals (noted SB), while regular spirals without bar are denoted S. An additional letter, from a to c, precises the opening of the spiral pattern and brightness of the bulge. Spirals show star formation activity and are mostly supported by rotation. In between the morphology and properties of ellipticals and spirals, lenticular (or S0) galaxies have a prominent bulge and disc but without any visible spiral structure and without clear star formation activity. Finally, irregular (Irr) galaxies do not fall in one of the previous categories, because they do not have an obvious regular structure (disc or ellipsoid).

Elliptical and lenticular galaxies are often referred to as Early-Type Galaxies (ETGs), while spirals and irregulars are grouped in Late-Type Galaxies (LTGs). Contrary to what the names and the shape of Hubble's tuning fork may suggest, the classification does not imply an evolution from early to late-type galaxies. On the contrary, our current models

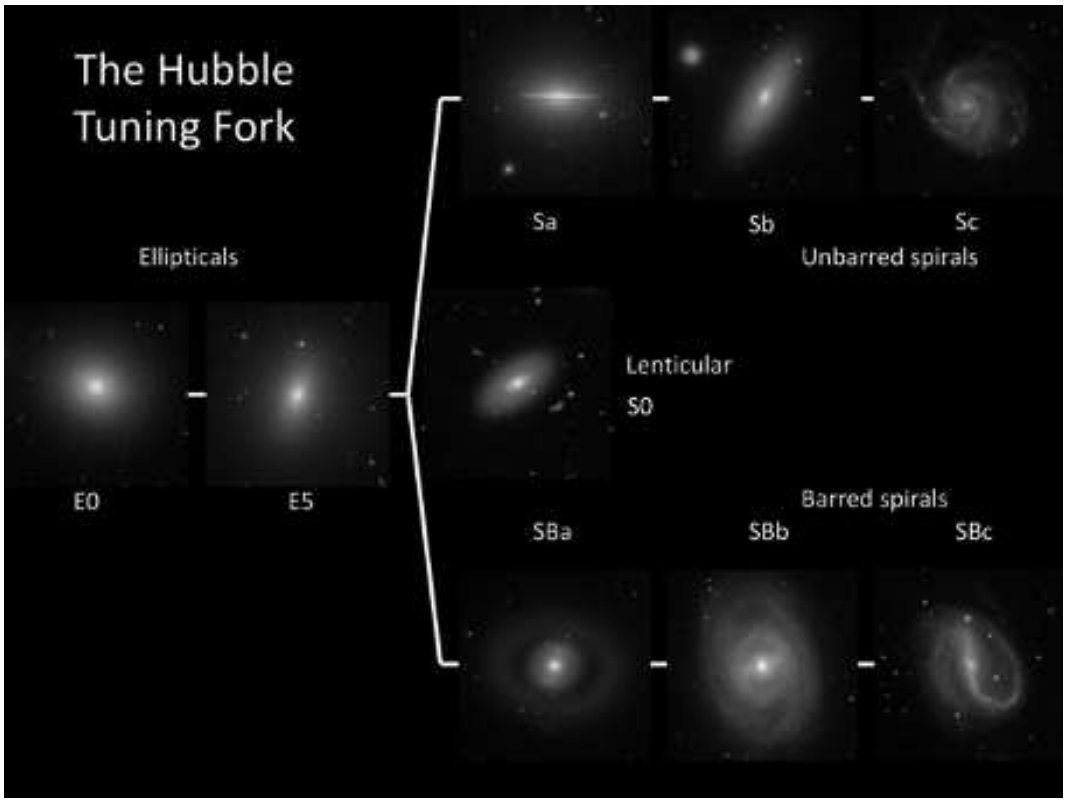

Figure 1.1: Illustration of the Hubble classification scheme with galaxy images from the Sloan Digital Sky Survey (Masters et al. 2019).

of galaxy formation and evolution suggest that galaxies form as spirals and evolve to ellipticals through accretion of gas and mergers.

It is important to mention here that although the Hubble tuning fork is widely used, there are other classification schemes that extend it or differ from it. For instance, van den Bergh (1976) proposed a trident classification scheme with three branches after the elliptical: the 'normal' branch of Sa, Sb, Sc; but also 'anemic spirals' Aa, Ab, Ab parallel to the normal branch (smooth-armed spirals with small amounts of gas, intermediate between normal spirals and S0s); and a third branch of S0a, S0b, S0c parallel to the normal one that resemble spirals without dust. Other classifications could be based on colours rather than morphology, as ellipticals are mostly red with old stars and spirals mostly blue with young stars. Yet, a limitation arises from the fact that there exist some galaxies intermediate between spiral and S0s with red colours (Dressler et al. 1999; Poggianti et al. 1999; Goto et al. 2003a). A different approach is to focus on internal properties of galaxies such as kinematics. Indeed, some ellipticals rotate slowly (slow rotators, e.g., Bertola & Capaccioli 1975; Illingworth 1977) while others show a significant rotation pattern (fast rotators, e.g., Davies & Birkinshaw 1986; Franx et al. 1989; Jedrzejewski & Schechter 1989). Emsellem et al. (2007); Cappellari et al. (2007, 2011b) revisited the classification scheme using fast and slow rotators. An illustration of their scheme is presented in Figure 1.2.

1.1.2 Hierarchical paradigm

The modern standard model of cosmology is the Λ Cold Dark Matter (ΛCDM) model, introduced by Ostriker & Steinhardt (1995). It relies on the assumption that the Uni-

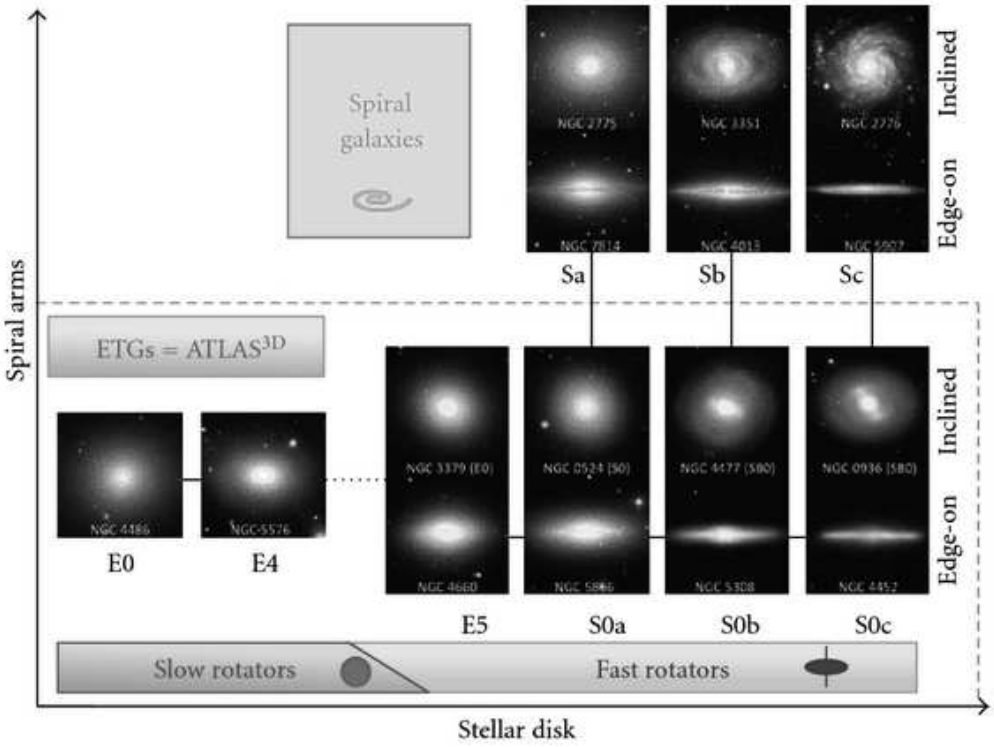

Figure 1.2: Figure 1 from Cappellari et al. (2011b). Proposed classification scheme, based on internal kinematics (separation between fast and slow rotators among ETGs) that extends the Hubble tuning fork.

verse is homogeneous and isotropic (also known as the cosmological principle) and that it contains cold non-collisional dark matter. Λ was introduced to represent the cosmological constant, which later on was associated to dark energy, to explain the accelerated expansion of the Universe. In this parametrization of the Big Bang cosmological model, structures grow hierarchically. The primordial density fluctuations are generated during the rapid expansion known as inflation and will then collapse under their own gravity to form the first dark matter haloes. The baryonic matter (i.e. gas) will fall into these dark matter potential wells, and it will itself condensate and collapse once cold enough to form the first stars and galaxies (White & Rees 1978; Blumenthal et al. 1984). These galaxies will then grow through star formation, successive mergers, accretion of gas and of smaller systems to form the present-day galaxies. Their dark matter haloes will merge as well. This hierarchical, or 'bottom-up', formation scenario of galaxies is a continuous and on-going process.

This model has replaced the original view of monolithic collapse (or 'top-down') (Eggen et al. 1962; Larson 1975) where galaxies would form coherently from a single giant gas cloud in the early Universe that would rapidly collapse gravitationally, fragment and form stars. The halo stars and globular clusters would be formed while the cloud was still condensing, and the other parts of the galaxy would have been formed at the same time. Since then, new observations have ruled out this scenario, such as the spread of age or metallicity distribution in globular clusters that should not be present if they had all formed at the same time (e.g., Forbes et al. 1997); or the fact that high-redshift galaxies appear to be different from present-day ones, in terms of size for instance as high redshift

galaxies are more compact (e.g., Daddi et al. 2005; Toft et al. 2007; van Dokkum et al. 2008).

Therefore, in the hierarchical paradigm, galaxies acquire their present-day characteristics through successive mergers, accretion of smaller satellite systems and of gas. Determining which of these processes dominates depends on the mass and redshift of the galaxy (e.g., White & Rees 1978; White & Frenk 1991; Kauffmann et al. 1993; Cole et al. 2000; Baugh et al. 2003; Somerville & Primack 1998; Dekel et al. 2009; Stringer & Benson 2007; Genel et al. 2010; L'Huillier et al. 2012). These studies rely on numerical models, either on semi-analytical models (in which a set of simplified physical rules is used for the baryonic matter, and where the halo formation and merger histories are derived using Monte Carlo methods); or on hydrodynamical simulations (in which the evolution of dark and baryonic matter is computed using the fundamental equations of hydrodynamics and gravity). The process of gas accretion, which is highly involved in the formation of clumpy high-redshift and disc galaxies (e.g., Daddi et al. 2007; Förster Schreiber et al. 2009; Agertz et al. 2009), also depends on galaxy mass (e.g., Kereš et al. 2005; Ocvirk et al. 2008). For massive galaxies, the growth seems to be dominated by mergers (e.g., Murali et al. 2002), as the accreted cold gas cannot pass through the shock-heated galactic halo (Afruni et al. 2019).

Gas accretion plays an important role in galactic evolution. Star formation occurs in galaxies at all redshifts. Without fresh incomes of gas, galaxies would rapidly exhaust their gas reservoir, from less than a Gyr at $z > 1$ (e.g., Genzel et al. 2010; Daddi et al. 2010) to less than a few Gyr at lower redshifts (e.g., Bigiel et al. 2011; Kennicutt & Evans 2012). In addition, the number of low-metallicity G-dwarf stars observed in the Milky Way is lower than the one predicted by closed-box models of chemical evolution (in which galaxies form from a closed box of gas): this discrepancy can be explained by a continuous accretion of relatively low-metallicity gas (e.g., Larson 1972; Tinsley 1981; Fenner & Gibson 2003; Greener et al. 2021). Therefore, gas accretion is needed through cosmic time to balance the gas consumption for star formation and explain the observed stellar metallicity distributions (see also e.g., Kereš et al. 2005; Sancisi et al. 2008; Putman 2017). There are three main sources of gas accretion: the intergalactic medium (IGM), satellites and recycled feedback gas (e.g., Katz et al. 2003; Bland-Hawthorn et al. 2007; Joung et al. 2012). Numerical simulations predict two main modes of accretion: hot and cold modes (e.g., Kereš et al. 2005; Dekel & Birnboim 2006; Kereš et al. 2009). In the hot mode, a part of the inflowing gas is shock heated to a temperature close to the virial temperature of the halo, then a fraction of this gas is able to cool in the central regions through radiative cooling (e.g., Rees & Ostriker 1977; Silk 1977; White & Rees 1978; White & Frenk 1991). A smooth accretion of intergalactic gas, or cold mode, dominates at high redshifts and in low-density environments (Kereš et al. 2005; Ocvirk et al. 2008; Kereš et al. 2009; Benson & Bower 2010).

The evolutionary pathway of galaxies is also a function of their morphological type. Ellipticals were long thought to form according to two scenarios: a simple monolithic collapse, or only by mergers. In the monolithic collapse scenario, ellipticals form quickly from the dissipationless collapse of a gas cloud, and stars form simultaneously with the assembly of the final galaxy (Eggen et al. 1962; Larson 1975) On the contrary, in the pure merger scenario (Toomre 1977), ellipticals result from the merging of two or more galaxies, so star formation does not depend on the assembly of the final galaxy. However, individually these two scenarios cannot properly reproduce the observations (for instance, the monolithic collapse scenario implies the formation of ellipticals should have taken

place at $z > 20$ which is incompatible with the fact that there are no massive extended ETG in the early Universe yet discovered), therefore a combination of both is needed to explain galaxy formation (see Mo et al. 2010). Current models predict that ETGs are most likely formed in a two-phase scenario (Oser et al. 2010; Naab et al. 2014). The first phase, occurring before $z = 2$, is dominated by smooth accretion of gas and of gas-rich mergers, with many stars formed inside the galaxy (i.e. in-situ star formation) (Thomas et al. 1999; Kereš et al. 2005; Dekel et al. 2009). At the end of this first phase (around $z = 2$) most galaxies reach the end of active star formation (although more massive galaxies can still form stars at later redshifts or on the contrary be quenched at earlier times, e.g., McDermid et al. 2015; Forrest et al. 2020; Tacchella et al. 2022). Yet, observational data of galaxies at a redshift of around two require a second phase to explain the formation of ETGs. Indeed, galaxies at $z \sim 2$ appear to be 2-4 times more compact, less massive and more disky (Daddi et al. 2005; Trujillo et al. 2006; Toft et al. 2007; Zirm et al. 2007; Cimatti et al. 2008; Buitrago et al. 2008; Damjanov et al. 2009; van Dokkum et al. 2010; Szomoru et al. 2012) than their $z = 0$ counterparts of similar mass. This considerable growth in mass and size could be explained primarily by the accretion of smaller systems, i.e. minor mergers (e.g., Naab et al. 2009; Trujillo et al. 2011; Bluck et al. 2012). This is the second phase of ETG formation, i.e. the accretion of stars formed in other galaxies (i.e. ex-situ star formation) at lower redshifts ($0 < z < 2$) (e.g., Meza et al. 2003; Naab et al. 2007; Feldmann et al. 2008; Lackner et al. 2012). Some galaxies can form out of this two-phase scenario: more rarely, some ETGs can result from the merger of two spirals (e.g., Toomre 1977; Barnes 1988). An additional layer of complexity arises when we focus on the different types of ETGs. For instance, some of them are characterized by a regular rotation pattern (fast rotators), while others do not show any sign of rotation or they display complex kinematic features (slow rotators) (e.g., Emsellem et al. 2007, 2011; Krajnović et al. 2011; Cappellari 2016). Different assembly histories are expected for these two classes of objects. Numerical simulations (e.g., Naab et al. 2014; Penoyre et al. 2017; Lagos et al. 2018) suggest that ellipticals originally form as fast rotators, and then are later transformed into slow rotators through mergers, typically wet minor mergers (Bílek et al. 2022).

The formation scenario of LTGs is likely to be smoother than for ETGs. Disc galaxies are thought to originate from the collapse of dark matter haloes and baryons inside it: during the hierarchical clustering of dark matter haloes, tidal torques will lead to a rotation of the system. Later, when the baryons cool and dissipate their gravitational energy, they concentrate within the dark matter halo and settle in a rapidly rotating disc (e.g., Dalcanton et al. 1997; Samland & Gerhard 2003; Silk 2003). Later, spirals grow through the accretion of cold intergalactic gas (e.g., Sancisi et al. 2008; Bílek et al. 2022) whose fresh infall compensates the gas already used for star formation, with a small contribution of minor mergers.

At the faint end of the mass function, dwarf galaxies are also assembled hierarchically in the ΛCDM model. Dwarf galaxies formed from the initial density perturbations that collapsed, and later were preserved from merging with larger and larger haloes. (e.g., Dekel & Silk 1986; Kravtsov 2010; Belokurov 2013). Some dwarf galaxies can also result from tidal interactions as hinted by observations and numerical simulations (e.g., Schweizer 1982; Mirabel et al. 1992; Barnes & Hernquist 1992; Elmegreen et al. 1993; Duc & Mirabel 1994; Okazaki & Taniguchi 2000). The comparison between the observed properties of dwarfs and numerical simulations provide fundamental tests of hierarchical models. For instance, Klypin et al. (1999) showed that there was a discrepancy between the number

of small-size dark matter clumps and the observed lower number of dwarf galaxies around the Milky Way. Other properties such as morphology, spatial distribution, presence of a nucleus or HI content are also essential in order to build a comprehensive view of galaxy formation and evolution across large mass ranges (e.g., Moore et al. 1999; Bode et al. 2001; McConnachie et al. 2009) as well as to put constraints on the nature of dark matter (e.g., Koposov et al. 2009; Peñarrubia et al. 2010; Nadler et al. 2019).

1.1.3 The role of the environment

The environment in which galaxies reside also has a strong impact on their properties and morphologies, which has been confirmed both observationally and with numerical simulations (e.g., Oemler 1974; Goto et al. 2003b). Dressler (1980) showed that there is a relation between the morphological type of a galaxy and the environment in which it resides. This is known as the morphology-density relation: ETGs are mostly found in higher-density environments such as groups or clusters, while LTGs are more often found in lower-density environments such as the field or the outskirts of groups and clusters. There is an additional redshift dependence, as the fraction of blue galaxies in clusters increases with redshift (Butcher & Oemler 1984).

Galaxies that enter clusters can be affected by multiple cluster processes that can drastically change their evolution and quench them. Ram-pressure stripping (e.g., Gunn & Gott 1972; Hester 2006; Boselli et al. 2022) removes the gas of galaxies as they move through the hot intracluster medium. Strangulation (e.g., Larson et al. 1980; Kawata & Mulchaey 2008), also known as starvation, is another effect where the hot gas of the galaxy is stripped. In that case, the galaxy is cut off from the supply of cold gas, as its hot gas that could have cooled down is no longer available, and it will continue to form stars until its gas reservoir is exhausted. Furthermore, due to the higher density in clusters, the probability of having galaxy interactions is higher. Because of the high typical velocity dispersion in clusters (around 1000 $\mathrm{km\,s^{-1}}$, Girardi et al. 1993), there may be multiple high-speed encounters (e.g., Moore et al. 1996, 1998), also known as galaxy harassment. During these numerous fly-bys, the internal energy of stars and gas in galaxies is increased, leading to heated, less-bound material. This material can be ejected in the intracluster medium and contribute to form the diffuse intracluster light. Although the high velocities of encounters reduce the probability that two galaxies will merge to form a single new galaxy (e.g., Toomre & Toomre 1972), the dynamical friction exerted by the cluster potential on infalling galaxies results in a loss of energy and momentum, which makes the galaxy fall onto the centre of the cluster where it can merge with the bright central galaxy (e.g., Webb et al. 2015). As galaxies mostly infall clusters in groups, the group dynamics and pre-processing inside it are also important factors to consider (e.g., Fujita 2004; Bianconi et al. 2018). Additional interactions with the gravitational potential of the cluster (e.g., Byrd & Valtonen 1990; Bekki 1999) also contribute to change galactic morphology. The combination of these mechanisms plays an important role in shaping galaxies, that are different in clusters and in the field (e.g., Mihos 2003; Berrier et al. 2008). The location of the galaxy inside the cluster itself is important: for instance, Janowiecki et al. (2010) and Adams et al. (2012) found a dependence between the presence of some stellar substructures around a galaxy and its distance to the center of the cluster; and Hatch et al. (2017) found increased fractions of red galaxies in the highest-density regions (groups) of a protocluster compared to the intergroup galaxies.

The aforementioned mechanisms, referred to as environmental quenching, are possible explanations for the transformation of morphologies in clusters. Yet, internal processes

within the galaxy itself can also trigger important changes, in clusters as well as in the field. In particular, supernovae (SN) and active galactic nuclei (AGN) feedbacks contribute to eject gas from the galaxy and to radiate energy that can be transmitted to the interstellar medium. This heated gas lowers or prevents star formation from occurring afterwards. Supernova feedback (e.g., Yorke et al. 1989; Hopkins et al. 2018) is more efficient in low mass galaxies, while AGN feedback (e.g., Silk & Rees 1998; Croton et al. 2006) is more efficient in gas depletion than SN feedback for more massive galaxies. These internal processes are known as mass quenching.

In addition to all these processes, galaxies continue to experience a slow and steady evolution due to internal factors: this is referred to as secular evolution (e.g., Kormendy 1979; Kormendy & Kennicutt 2004; Combes 2008; Falcón-Barroso & Knapen 2013). Energy, angular momentum and mass inside a galaxy are slowly rearranged due to collective interaction of instabilities like bars, spiral arms, rings, oval discs, thick disc and triaxial dark halos. Secular evolution can for instance lead to the formation of pseudo-bulges in galaxies.

1.1.4 Galaxy mergers and collisional debris

1.1.4.1 Mergers

One of the most important aspects of the hierarchical paradigm are interactions, collisions and mergers between galaxies. These events are crucial in shaping galaxies, and they occur at both high and low redshifts. Merger events can be divided in several classes as a function of the number of galaxies involved, their masses and gas content. There can be two (binary merger) or more (multiple merger) galaxies involved in the process. If the galaxies have relatively similar masses (i.e. with a mass ratio equal or larger than 1:4), the merger is said to be major. If galaxies have a mass ratio below 1:10, the merger is called minor. In between, intermediate-mass mergers have mass ratios between 1:4-1:10. Then, if the merger involves gas-rich galaxies, it is said to be wet, or dry on the contrary when galaxies are gas-poor. It is important to briefly remind here the terminology. Tidally interacting galaxies refers to systems that are strongly affected by the gravitational field of the other galaxy, that will get closer and closer up to the point that they will eventually merge and leave a single merger remnant (or post-merger galaxy). Colliding galaxies is a broader term that refers either to an encounter with another galaxy without a merger at the end, called flyby, or to interacting galaxies that will eventually merge.

The morphology of galaxies, kinematics, chemical composition, or stellar age are affected differently by mergers (e.g., Bílek et al. 2022; Yoon et al. 2023). For instance, kinematic features such as kinematic distinct cores, kinematic twists or counter-rotating cores are known to be potentially caused by mergers (e.g., Balcells & Quinn 1990; Bois et al. 2011; Schulze et al. 2017). The angular momentum is also affected (e.g., Cox et al. 2006; Emsellem et al. 2007; Hoffman et al. 2010; Lagos et al. 2017). The metallicity gradient is another modified property, as it is increased in case of minor mergers, where old and metal-poor newly accreted stars stay in the external parts of the galaxy; and it is flattened in case of major mergers where the stellar populations are mixed (e.g., Di Matteo et al. 2009; Amorisco 2016). Episodes of intense star formation can follow a merger event (e.g., Larson & Tinsley 1978; Hernquist 1989; Springel & Hernquist 2005), as gravitational instabilities induced by the coalescence of two galaxies can drive large fractions of the gas towards the inner galactic region. Minor mergers seem to be responsible for the mass and size growth of ETGs at $z < 1$ (e.g., Kaviraj et al. 2009; Hilz et al. 2012; Oogi &

Habe 2013; Lee et al. 2013; Nevin et al. 2023) as they occur more frequently than major mergers.

1.1.4.2 Tidal features

Interactions and collisional events contribute to modify the morphology of galaxies by their stripping stellar material. Indeed, the gravitational forces faced by a spatially extended galaxy differ on one side or on the other: the side facing the companion galaxy will be more attracted than the opposite side (e.g., Holmberg 1941; Zwicky 1953, 1956; Zasov 1968). The same mechanism is responsible for the tides on Earth with the Moon and the Sun. The stars and gas facing the companion may be stripped out of the galaxy, and form stellar structures called tidal features that extend around it (see the reviews by Schweizer 1998; Duc & Renaud 2013). For elliptical galaxies, dominated by the random motion of the stars, the net effect of such differential forces does not trigger a very clear deformation of a region of the galaxy, making these galactic tides more difficult to detect. On the other hand, when the system is dominated by ordered rotation like in spirals, the deformation is applied similarly to all the stars in a given region, resulting in a much more pronounced effect (see also Figure 1.5 reproduced from Duc & Renaud 2013, where in the case of the merger with a spiral tidal tails are produced, but none are generated in the case of an elliptical.). Figure 1.3 illustrates the different merger stages of interacting galaxies. It starts with the two galaxies approaching (panels 1-3) and by doing so they feel the differential forces of galactic tides that strip their stellar material and form long tidal features which are clearly visible. As they come even closer, their centres progressively merge together (panels 4-5), leaving a remnant with a completely disturbed morphology and tidal features (panel 6).

Tidal features were discovered in the late 1920's thanks to the development of large telescopes and photographic plates. Arp (1966) compiled an Atlas of Peculiar Galaxies that includes galaxies with a disturbed morphology, different from the classical Hubble tuning fork, as well as faint extensions around galaxies. Although these tidal features had been observed in many other surveys (e.g., Duncan 1923; Keenan 1935; Vorontsov-Velyaminov 1962; Zwicky 1963), their tidal origin remained a subject of controversy until the 1970's. Some argued that they arose from supernovae activity that would eject material outside the galaxy (Zwicky 1962); others considered the effect of jets (Ambartsumian 1961; Arp 1972); while other works proposed magnetic fields to explain the shape of tails (Burbidge & Burbidge 1963; Vorontsov-Vel'Yaminov 1965). Holmberg (1941) carried out an experiment that can be considered as the first N-body simulation to investigate whether these features could be explained by gravitational forces. To overcome the impossible manual numerical integration, Holmberg (1941) replaced gravitation by light: with photocells and an ingenious system of 37 light bulbs representing a galaxy, he simulated gravity and observed how two galaxies would come closer to each other as well as the type of perturbation produced. With the advent of computers, numerical simulations were developed and used to answer the same question. Toomre & Toomre (1972) showed that under certain conditions, the encounter between two disc galaxies could reproduce the long tidal tails in the observations (see Figure 1.4). Following these first numerical experiments, other studies were conducted numerically and confirmed the tidal origin of several observed diffuse structures around galaxies (e.g., Eneev et al. 1973; Lauberts 1974; Keenan & Innanen 1975). Numerical simulations have been conducted ever since in order to investigate the formation of these features and their link with the evolution of galaxies. Two examples of numerical simulations of an interaction between two galaxies of similar

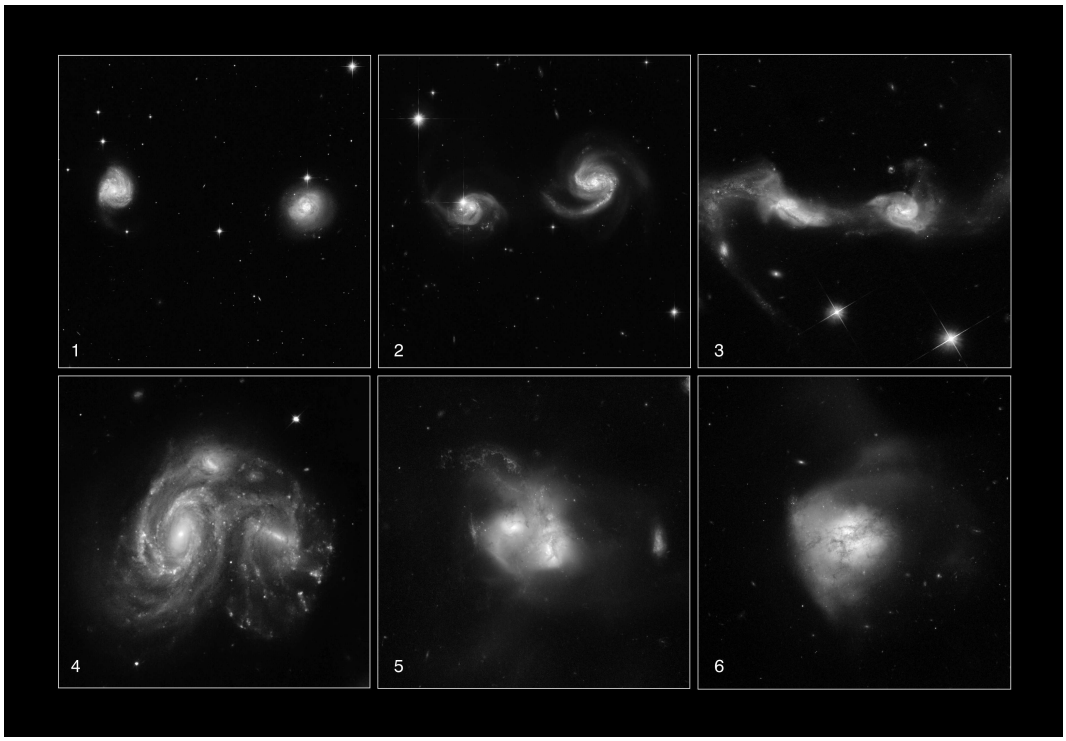

Figure 1.3: Illustration of the merger stages of interacting galaxies, taken by the Hubble space telescope[1]. Credits: NASA, ESA, the Hubble Heritage Team (STScI/AURA)-ESA/Hubble Collaboration and A. Evans, K. Noll, and J. Westphal . This sequence of images, taken from the Hubble atlas of interacting galaxies, represents 6 snapshots of galaxies that are involved in a merger process at different stages. *Panel 1*: The two galaxies are approaching, some deformations start to become visible. *Panel 2*: Galaxies are close enough so that the gravitational interaction is strong enough to disturb the morphology of the galaxies and start to create tidal features. *Panel 3*: Tidal features are clearly visible, they are long and extended and may form a 'bridge' between galaxies (i.e. two tidal features merged together) *Panel 4*: Galaxies are coming even closer (here two massive galaxies are actually merging with a third smaller one, as an example of a multiple merger). *Panel 5*: The centres of both galaxies are still visible but are starting to merge, the original shape and morphology of the galaxies are destroyed. *Panel 6*: The remnant of the merger, with a completely disturbed morphology and diffuse tidal features around it. There can be bursts of star formation triggered by the compression of gas during the merger.

mass can be seen in Figures 1.4 and 1.5.

Not all tidal features are alike. Numerical simulations indicate that their properties and morphologies highly depend on the type of the merger that occurred and on its impact parameters, such as the mass ratio between the galaxies, their inclinations or difference of velocities (e.g., Hibbard & Mihos 1995; Johnston et al. 2008; Peirani et al. 2010; Hendel & Johnston 2015; Amorisco 2015; Pop et al. 2018; Mancillas et al. 2019; Karademir et al. 2019). For instance, tidal tails are large elongated structures generated during the merger of two equal-mass galaxies (e.g., Toomre & Toomre 1972; Barnes 1992; Hopkins et al. 2008). Typical examples of tails are the famous wide and elongated structures around the Antennae galaxies (NGC4038/NGC4039) (Arp 1966). Simulations indicate they remain visible for about 2 Gyr (Mancillas et al. 2019). As tidal tails contain stars and gas,

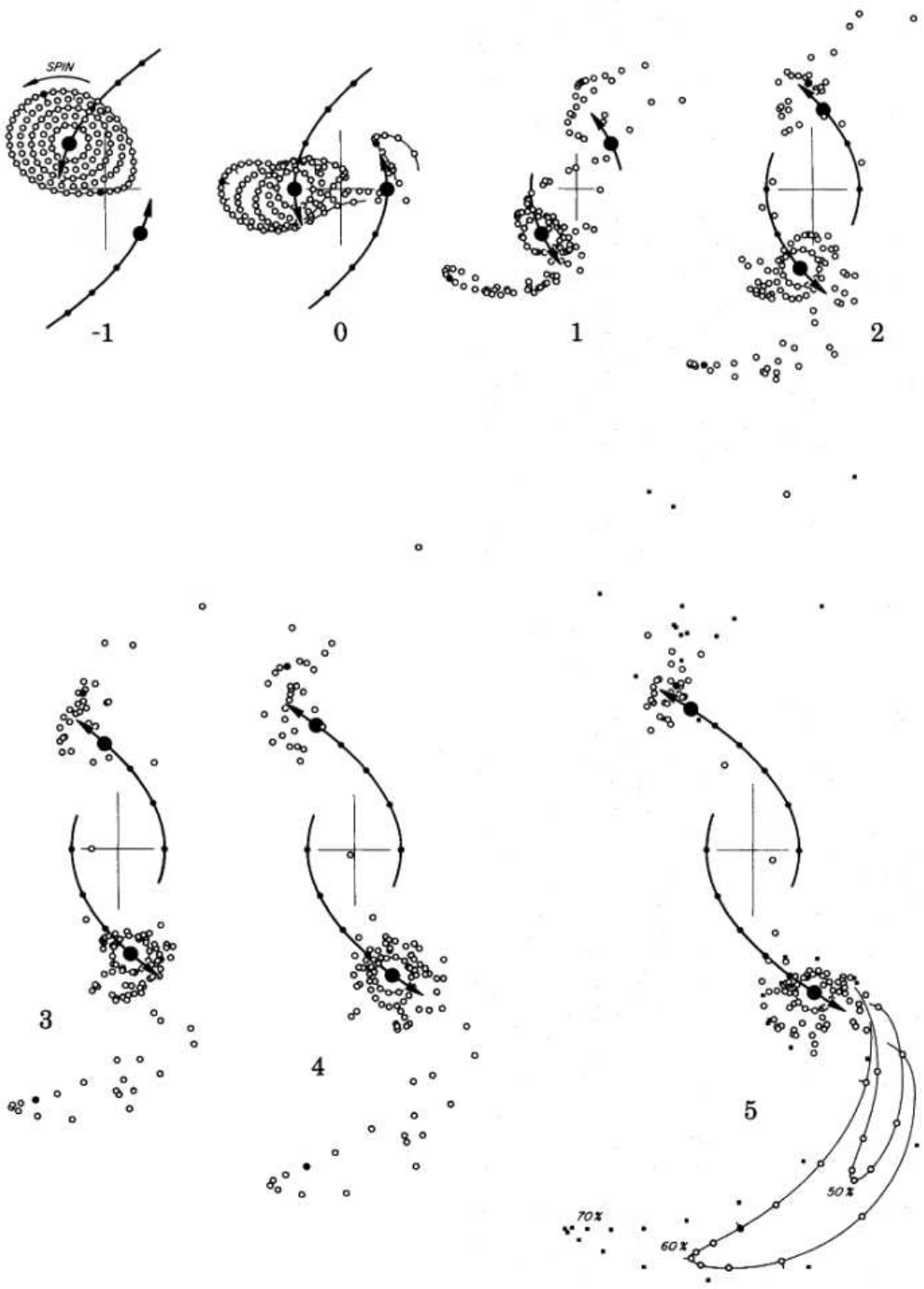

Figure 1.4: Figure 2 from Toomre & Toomre (1972): simulation of a flat direct parabolic passage of a companion of equal mass. The bulk of the two galaxies are represented by the large black dots and test particles with the small open circles. One can see the influence of the companion galaxy that triggers tides and strips material from the primary galaxy.

there can be regions of star formation due to the increase of gas pressure during the merger: star clusters and even tidal dwarf galaxies can be formed in tails (e.g., Elmegreen et al. 1993; Schweizer 2006; Duc 2012). Stellar streams are another type of tidal feature

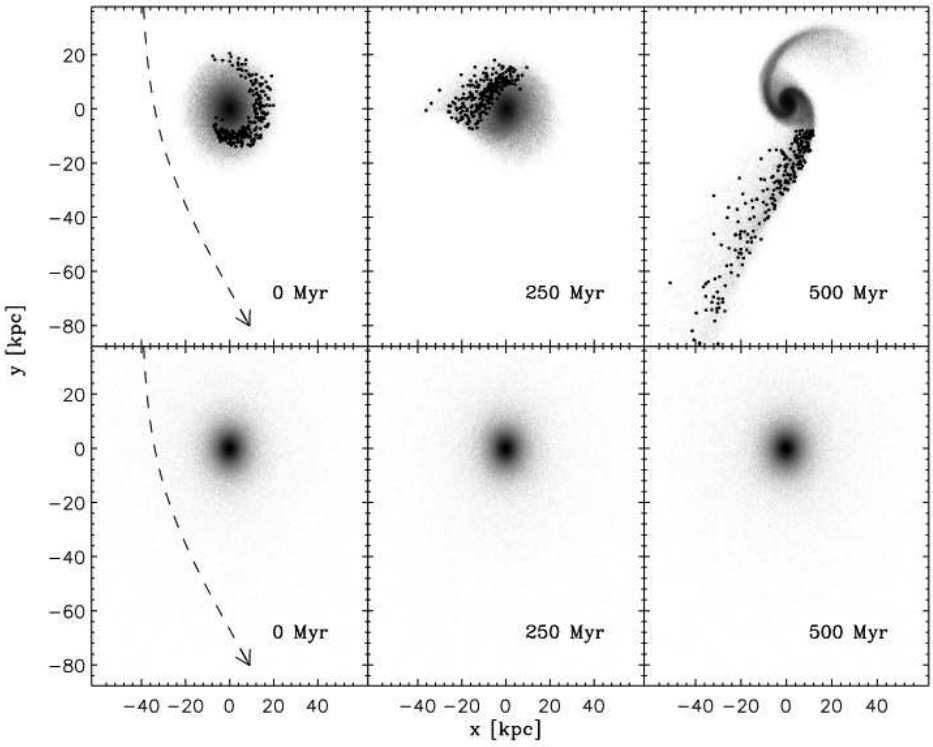

Figure 1.5: Figure 3 from Duc & Renaud (2013): N-body simulation of an coplanar encounter between a galaxy composed of a disc, bulge and dark matter halo, and a point mass particle (with a mass ratio of 1). This experiment applies to any mass element, and can in principle be extended to both the gaseous and stellar components of a galaxy. *Top panel*: Interaction between a disc galaxy seen face-on and a point-mass particle of similar mass, before the interaction (left), at pericentre (middle) and after (right). The black dots tag a subset of particles that are situated in one of the tidal tails at t = 500 Myr. *Bottom panel*: Same simulation but for an elliptical galaxy.

generated during a minor merger, i.e. by the tidal disruption of a lower-mass companion (e.g., Bullock & Johnston 2005; Johnston et al. 2008; Belokurov et al. 2006; Martínez-Delgado et al. 2010). They globally are thinner than tails and can be visible for 2-3 Gyr. Shells are arc-shaped features, often concentric. There can be one to many shells around a given galaxy. They are likely formed during intermediate-mass radial mergers (e.g., Quinn 1984; Wilkinson et al. 1987; Prieur 1990; Ebrova 2013; Duc et al. 2015; Pop et al. 2018), and can remain visible for about 3-4 Gyr. Examples of tails, streams and shells are visible in Figure 1.6. There can also be other types of tidal debris, such as broad fan, rings (e.g., Bekki 1998; Mosenkov et al. 2022) or disturbed external isophotes that are signs of more or less strong levels of interactions or mergers. As the stars that compose them will eventually die and not be replaced, tidal features will phase-mix over time and become fainter and fainter, up to the point where they will not be visible anymore as a well-defined structure. These features will contribute to form an extended and very faint stellar halo around the galaxy, which remains visible for a longer time than individual features. Therefore, as the different classes of features are produced by different types of mergers, their identification and classification give valuable information about the mass assembly history of the host galaxy during the past few gigayears (i.e. the late assembly) and about models of galactic evolution in general. Likewise, probing the extended stellar haloes enables to push this study an additional few gigayears back.

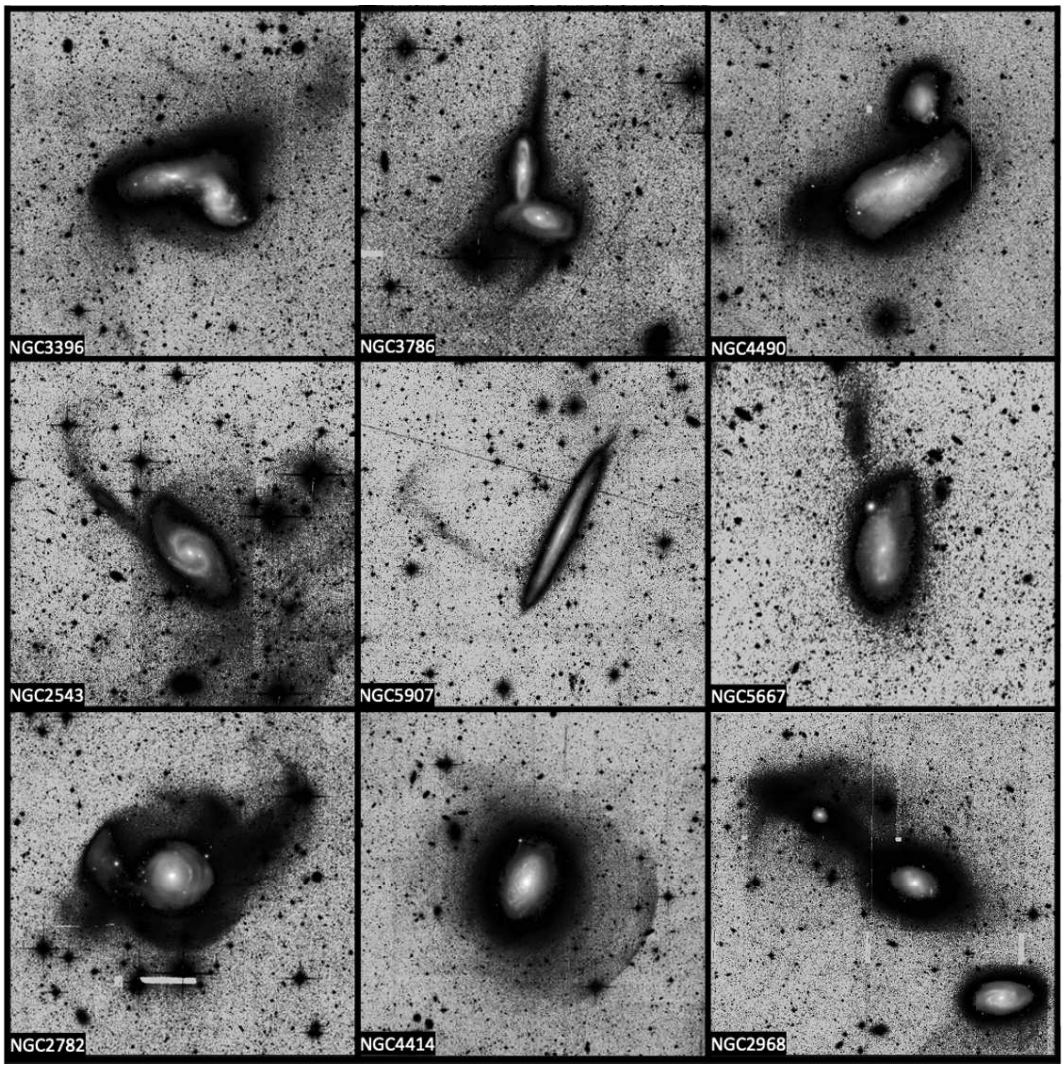

Figure 1.6: Figure 3 from Sola et al. (2022). Example of tidal features in CFIS r-band images displayed with a asinh scale. A composite colour image from PanSTARRS DR1 shows the centre of the galaxy, while the extended dark structures correspond to stellar material in the extended stellar halo or tidal features. The first row shows tidal tails, the middle row streams and the bottom one shells.

Like the galaxies around which they can be found, tidal features are also affected by the environment. In the field, stripped debris remains loosely bound to the galaxy and can be visible for several Gyr before phase-mixing, if there is no external perturbation (e.g., Lotz et al. 2008; Ji et al. 2014; Mancillas et al. 2019). In clusters, it is the result of competing processes, that either tend to generate it (through galaxy harassment, Moore et al. 1996); interactions between galaxies in an infalling group (Mihos 2003) or with the cluster itself (Byrd & Valtonen 1990; Gnedin 2003), or erase it (through dynamical heating from the whole cluster, interactions with the intracluster medium or ram-pressure stripping, Mihos et al. (2005); Tal et al. (2009)). Determining which effect is dominant is a complicated task as they occur simultaneously. The diffuse debris then contributes to the intracluster light (ICL), made of stars that are not bound anymore to any galaxy (e.g., Rudick et al. 2006; Conroy et al. 2007; Montes 2019).

The comparison of tidal features between observations and simulations of galactic evolution models, where the formation history is known, enables to put constraints on

the orbit, mass ratio and accretion time of satellites, the importance of accretion versus in-situ star formation, the merger rates or the overall galaxy evolution (e.g., Johnston et al. 2008; Cooper et al. 2010; Hendel & Johnston 2015; Pop et al. 2018; Hood et al. 2018; Mancillas et al. 2019; Bílek et al. 2022). Tidal features are also useful probes of the underlying galactic potential. In particular, stellar streams approximately follow the orbit of their progenitor (Dehnen et al. 2004; Eyre & Binney 2011), which can be used to derive information about the gravitational potential and dark matter distribution (e.g., Hernquist & Quinn 1987; Johnston 1998; Ebrova 2013; Thomas et al. 2017; Nibauer et al. 2023). When the stars constituting the stream are individually resolved in real data, it is possible to use kinematics and positions in order to constrain the potential (e.g., Ibata et al. 2001b; Helmi 2004; Lux et al. 2013; Malhan & Ibata 2019; Koposov et al. 2023).

However, several points must be kept in mind when inferring the past assembly history of galaxies through the study of collisional features. First, it is only possible to probe the late assembly history, given the limited lifetime of tidal features (less than a few Gyr, e.g., Lotz et al. 2008; Ji et al. 2014; Mancillas et al. 2019). Then, not all interactions and mergers will produce prominent tidal features. For instance, due to the dominance of random motions in ellipticals, it is much more difficult to produce tidal tails around ETGs than around LTGs. This implies that the visible prominent tails rather trace gaseous (wet) mergers than dry ones (e.g., Duc & Renaud 2013). In addition, features can also be generated during non-merging flybys. Finally, the implications between the presence of tidal features and galactic assembly previously mentioned are valid in the hierarchical paradigm of the ΛCDM model. However, as this model still faces some challenges, other theories such as modified gravity were developed to try to overcome these issues. In particular, MOND (Milgrom 1983; Famaey & McGaugh 2012) proposes a modified gravity in the weak acceleration field regime as an alternative to dark matter in galaxies. From comparative numerical simulations (e.g., Nipoti et al. 2007; Tiret & Combes 2007; Combes & Tiret 2010), the dynamical friction which causes galaxies to merge is weaker in MOND than in ΛCDM; hence galaxies can be found in close encounters that will not lead to a merger. This means that the galaxies in close pairs observed today may simply be involved in non-merging flybys. In addition, tidal dwarf galaxies form more easily in MOND (Kroupa 2015). Therefore, tidal features can arise from non-merging flybys and from the disruption of tidal dwarf galaxies by tides, which is different from the hierarchical paradigm where tidal feature trace past or on-going mergers (Bílek et al. 2018, 2019). It is also important to make a distinction here between tidal features that originate from on-going interaction events and the ones resulting from past mergers. The former give information about the present and future history of the galaxy, i.e. we see the two galaxies that are involved in the interaction and we predict their merger. The latter are located around post-merger galaxies, i.e. there is only one merger remnant so the tidal debris gives clues about the type of collision that had already happened, hence the past assembly history. Finally, a distinction must be made between tidal features observed at low and high redshifts. At higher redshifts, the fraction of galaxies involved in a merger increases, in particular the fraction of major mergers increases up to $z = 1.5$ (see López-Sanjuan et al. 2009; Conselice 2009, and references therein), so one could expect to see more tidal features in high-redshift galaxies. However, due to the surface brightness dimming that affects all sources, only the bright enough features will be visible, i.e. they are mostly features originating from on-going mergers. As previously mentioned, these features give information about the future of the galaxy but not its past, contrary to features around post-merger galaxies. Besides, the gas dominance and clumpy nature of distant galaxies

may prevent the formation of well-defined tidal tails.

As the environment plays an important role in shaping galaxies, the investigation of tidal features as a function of the environment of the host galaxy could help distinguish between different origin scenarios (e.g., Mihos 2003; Adams et al. 2012; Venkatapathy et al. 2017). By extension, the study of the ICL in clusters enables to put constraints on the dynamical state of the cluster and on its hierarchical formation process (e.g., Rudick et al. 2009; Montes 2019). Therefore, to have a broad picture of galaxy evolution through tidal features, it is necessary to study them statistically around large and representative samples of galaxies, spanning various scales of mass and environment. This also requires to study tidal features originating from past mergers, not on-going ones, which limits the study to the nearby Universe where it is possible to see them (unlike the high-redshift Universe where only bright tidal features from on-going mergers can be recovered).

Detecting tidal features in the Local Group is possible through stellar count, even using ground-based missions, such as the Sloan Digital Sky Survey (SDSS) with which Belokurov et al. (2006) revealed streams and stellar structures in the Sagittarius dwarf spheroidal. The resolved stellar populations photometric Pan-Andromeda Archaeological Survey (PAndAS) (Martin et al. 2014) also studied this 'field of streams', as well as the surrounding of M31 where signs of galactic interactions were studied (McConnachie et al. 2009). Substructures in M31 halo were detected in other surveys, such as Ibata et al. (2001a), or Ferguson et al. (2002) with the Isaac Newton Telescope, or Gilbert et al. (2012) with the Spitzer Large Area Survey with Hyper-Suprime-Cam (SPLASH) survey. A wealth of new structures have been disclosed in the Milky Way by the Gaia mission (e.g., Malhan et al. 2018; Antoja et al. 2020). However, for systems with increasing distance from the Local Group, stars are less and less resolved individually so that the detection of tidal features relies mainly on the study of the combined diffuse light they emit. A limitation to these studies arises from the fact that these features are very faint, they are called low surface brightness (LSB) structures: they can only be disclosed using dedicated observing strategies, instruments and data processing pipelines.

1.2 Low surface brightness structures

1.2.1 The LSB Universe

1.2.1.1 (Re)-discovering the LSB Universe

In order to understand the broad picture of galaxy formation and evolution, it is important to study galaxies of various properties in many different environments across the entire sky. However, a difficulty in detecting galaxies is the fact that a huge fraction of them are too faint, i.e. their surface brightness is below the value of the sky background.

The idea that such very faint objects existed started to appear in the 1970's. With the technology and photographic plates available at that time, Freeman (1970) studied the photometry of 36 spiral galaxies and in particular their exponential disc, and concluded that they all had roughly the same central surface brightness of 21.65 mag arcsec^{-2} in the B-band. However, this value was explained to be a selection effect by Disney (1976) who analysed the consequences of the biases introduced by the surface brightness cut in surveys. Indeed, the technological means available at the time were limited by the sky brightness, around 21.8 mag arcsec^{-2} which is close to the value found by Freeman (1970), so objects fainter than this could not be seen, which biased the apparent size and luminosities of galaxies. These biases had already been hinted by Arp (1965) and were

later confirmed by other studies, in particular by McGaugh et al. (1995). At the same time, Sandage (1976) reported the detection of very faint 'nebulae' at high latitudes with a magnitude down to 25 mag arcsec^{-2} in the V-band, using wide-field photographic plates, that were actually galactic cirrus clouds. From that moment, the search for LSB galaxies intensified, either through the careful re-examination of already-existing photographic plates, or with new observations with photographic plates or CCD cameras.

For instance, Phillipps et al. (1987) studied the surface brightness of galaxies in the first complete homogeneous survey of the Fornax cluster, correcting for SB-dependent selection effects. They divided galaxies in two groups: the 'normal' ones, with a central SB peaking around 21.75 mag arcsec^{-2} in the B-band, and 'LSB' ones with a central SB lower than 22.5 mag arcsec^{-2} even if they were less numerous. Likewise, Impey et al. (1988) published a catalogue of 137 LSB galaxies in the Virgo cluster, among them 27 were not known before. The photographic plates reached a SB of 27 mag arcsec^{-2}. In addition, Bothun et al. (1987) accidentally discovered Malin 1, an extremely large and massive LSB disc galaxy. Later, Impey et al. (1996) provided a catalogue of around 700 galaxies in a survey aiming at detecting nearby LSB galaxies. To that end, they re-analysed scans of photographic plates and conducted follow-up observations at radio and optical wavelengths. With the advent of CCD cameras, new surveys were conducted with the aim of revealing the LSB Universe with this technology. Davies et al. (1990) gathered B and R-band photometry of LSB galaxies in the Fornax cluster using CCDs. Davies et al. (1988) warned about the importance of flat-fielding in the search for faint objects. During their observations for LSB galaxies in the direction of the cluster A1367, they experimented a new flat-fielding technique by pointing towards 'blank' regions. They actually found in one of their blank frame a new Malin 1-type galaxy. Afterwards, Schwartzenberg et al. (1995) carried out a deep CCD survey of 'blank', i.e., non-clusters, regions to search for LSB galaxies, and they detected 500 LSB objects.

The detection of LSB galaxies in deep images with CCD cameras in comparison to photographic plates however suffered technological limitations. Indeed, deep photographic plates could undergo an appropriate technique to enhance faint structures on fine-grain high-contrast emulsions (Malin 1978). This technique, in addition to very long exposure times, enabled to reach limiting SB of about 28 mag arcsec^{-2} (e.g Kormendy & Bahcall 1974). On the contrary with digital CCD cameras, such non-linear processes are not possible and the issue of the sky background subtraction remained a blocking factor for a few decades. In addition, other issues such as internal reflections of the light in the optics of the telescope and the CCD camera made the LSB study even more complicated. Another limitation of the CCD cameras is that for decades their field of view was much smaller than photographic plates, which prevented the imaging of the bright parts of nearby galaxies and could not reach out to the faint outskirts (e.g., Impey & Bothun 1997; Knapen & Trujillo 2017). For instance, Turner et al. (1993) and Schwartzenberg et al. (1995) used CCDs with a total field of view of 17×17 arcmin per frame.

It is worth noting that the term of 'deep' images completely depends on the techniques available at a given moment. It is also important to mention that the estimation of the surface brightness depth, i.e. the faintest SB value that can be reached, has no precise definition and can vary from one study to the other, so it must be kept in mind when comparing studies.

1.2.1.2 Modern observations

Only in the last decade was the LSB Universe re-discovered (see the review about ultra-deep imaging by Knapen & Trujillo (2017) and references therein). Thanks to progress in telescopes, CCD camera and detectors (both in terms of field of view and sensitivity), in combination with dedicated observing strategies and data reduction pipelines, the SB limit was lowered below 25 mag arcsec^{-2}, enabling statistical studies of LSB objects. As an illustration of what the LSB regime enables to discover, Figure 1.7 shows a comparison of four galaxies observed with deep and shallower images: LSB tidal features appear where there seemed to be none before. The search for LSB features can be carried out with different instruments, ranging from small-sized to professional ones. As previously mentioned, the measure of the surface brightness limit can vary from one survey to the other (e.g., Duc et al. 2015; Trujillo & Fliri 2016). Simply reporting the depths announced by the different teams is not correct as they cannot be compared. A common definition must be used, which is often computed as a fluctuation of 3σ using $10'' \times 10''$ boxes. We will use this definition here and report the values computed by Euclid Collaboration et al. (2022), in particular in their Figure 11, for a comparison of depths between several surveys.

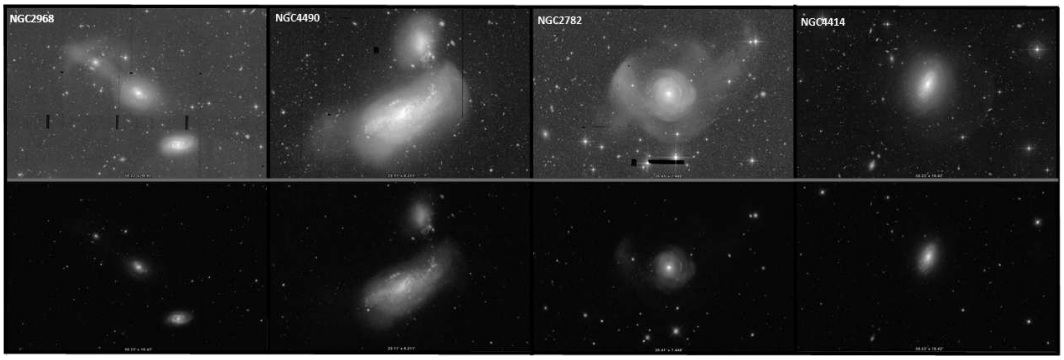

Figure 1.7: Comparison between deep images (*top row*) from CFIS and shallower images (*bottom row*) from SDSS for four different galaxies. LSB tidal features are clearly visible in the deeper image while they are not revealed in the shallower one.

If we focus on surveys aiming at disclosing tidal features, we can start by projects conducted with small-sized telescopes. In collaboration with amateur astronomers, Martínez-Delgado et al. (2010) used three small-sized (0.1-0.5 m) private telescopes located in dark sites to obtain deep images of a few nearby galaxies that were known to host stellar features. Likewise, Javanmardi et al. (2016) used small amateur telescopes to target nearby spiral galaxies and disclose tidal features as well as dwarf galaxies. van Dokkum et al. (2014) adopted another approach: they built the Dragonfly Telephoto Array from 48 (originally 8) telephoto lenses on a common mount coupled to commercial CCD cameras. This robotic system has the advantage of reducing the scattered light thanks to excellent coating of the lenses, it is comparable to a 1-m aperture refracting telescope (Abraham & van Dokkum 2014), with a depth of around 28 mag arcsec^{-2} ($3\sigma, 10'' \times 10''$ arcsec^2) for observations of the Coma cluster (Euclid Collaboration et al. 2022). Gilhuly et al. (2020) also used Dragonfly to image the halo of a galaxy down to 28.7 mag arcsec^{-2} in the r-band ($3\sigma, 10'' \times 10''$ arcsec^2). The Huntsman telescope (Spitler et al. 2019) follows the Dragonfly concept and consists of ten telephoto lenses.

Slightly larger telescopes were also used, such as the 0.6-m Burrell Schmidt telescope.

Mihos et al. (2015); Mihos (2016) obtained deep images with it of the Virgo cluster up to 28.5 mag arcsec^{-2} ($3\sigma, 10'' \times 10''$ arcsec^2) in the V-band and were able to study the diffuse intracluster light as well as tidal features in the cluster. Likewise, Watkins et al. (2015) studied the faint tidal features around M51 with the Burrell Schmidt telescope. Rich et al. (2019) used the 0.7-m Jeanne Rich telescope to study the haloes of 119 nearby galaxies (with a SB estimated around 28 mag arcsec^{-2} but not measured in $10'' \times 10''$ arcsec^2 boxes) for the HERON survey. Tal et al. (2009) imaged a complete sample of luminous elliptical galaxies with the 1-m SMARTS telescope at the Cerro Tololo Interamerican Observatory (CTIO) (with an estimated depth of 27.7 mag arcsec^{-2} in the V-band but not measured in $10'' \times 10''$ arcsec^2 boxes) and studied tidal features. In addition, Byun et al. (2018) used a dedicated observing strategy and reduction pipeline for KMTNet, a system comprising three 1.6-m identical telescopes, which enabled them to reach the LSB regime.

Larger professional telescopes have been widely used in the search for LSB structures. In particular, the 3.6-m Canada-France Hawaii Telescope (CFHT), with its 36 CCD cameras and one degree field-of-view, has been extensively used by different surveys. The CFHT Legacy Survey was run during five years, covering a total of 155 square degrees (Cuillandre et al. 2012). The deep images were exploited with different goals, ranging from supernovae, gravitational lensing or galaxy morphology and evolution, including the search for tidal features (Bridge et al. 2010). Later, Ferrarese et al. (2012) conducted the Next Generation Virgo cluster Survey (NGVS) covering the 104 square degrees of the Virgo cluster. Duc et al. (2015) and Bílek et al. (2020) carried out follow-up observations of nearby massive elliptical galaxies with the Mass Assembly of early-type GaLAxies with their fine Structures (MATLAS) survey. The Canada-France imaging survey (CFIS) from the Ultraviolet Near Infrared Optical Northern Survey (UNIONS), will eventually cover 5,000 square degrees with LSB r-band images. Although not dedicated to the search of tidal structures, Boselli et al. (2018) also covered the entire Virgo cluster with the Virgo Environmental Survey Tracing Ionised Gas Emission (VESTIGE), in Hα+[NII], with additional deep r-band images that can be used to study LSB features. All these surveys enable to reach SB of about 28-29 mag arcsec^{-2} in several optical bands (e.g., Duc et al. 2015; Euclid Collaboration et al. 2022).

Another important telescope is the 2.5-m Sloan, with the Sloan Digital Sky Survey (SDSS) (York et al. 2000), and its multiple releases. Nair & Abraham (2010) visually inspected the morphology of thousands of galaxies (including tidal features) in SDSS DR4 images. Miskolczi et al. (2011) specifically searched for streams in SDSS DR7 images. However, a limitation of SDSS is the fact that it is relatively shallow, but the SDSS Stripe82 region was observed multiple times and stacked in order to reach a SB around 28.5 mag arcsec^{-2} ($3\sigma, 10'' \times 10''$ arcsec^2) in the r-band, which is 2 magnitudes deeper than single-pass SDSS images (Fliri & Trujillo 2016). As an example, Kaviraj (2010) studied the morphologies of ETGs in Stripe82, among which some disturbed morphologies are tidal features; while Peters et al. (2017) focussed on the extended stellar haloes of spiral galaxies in Stripe82 images.

Numerous surveys have been conducted using other modest to large telescopes. For instance, Iodice et al. (2016, 2017) and Venhola et al. (2017) used images from the 2.6-m VLT Survey Telescope (VST) to target the Fornax cluster, its intracluster light, the extended stellar haloes, tidal features and LSB galaxies (with a claimed depth of 31 mag arcsec^{-2} but not measured on $10'' \times 10''$ arcsec^2 boxes). Still focusing on the Fornax cluster, Muñoz et al. (2015) disclosed faint dwarf galaxies using the wide-field Dark Energy Camera (DECam) mounted on the 4.0-m Blanco telescope at CTIO as part of the Next

Generation Fornax Survey (NGFS). Hood et al. (2018) also used the DECam data to probe tidal features around nearby galaxies. The DECam Legacy Survey (DECaLS, Blum et al. 2016) reaches depths of around 28-29 mag arcsec^{-2} ($3\sigma, 10'' \times 10''$ arcsec^2) (Euclid Collaboration et al. 2022).

Even larger telescopes are used to open a window on the LSB Universe. The Subaru telescope is a 8.2-m telescope located at Mauna Kea like the CFHT, and with its program Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP, Aihara et al. 2018) it aims at covering 1400 square degrees in deep images. Kado-Fong et al. (2018), Alabi et al. (2020) or Jackson et al. (2021) all used data from the Subaru Telescope to study tidal features around galaxies. Montes et al. (2021) reached a SB limit of 30.9 mag arcsec^{-2} in the g-band ($3\sigma, 10'' \times 10''$ arcsec^2) with HSC-SSP data of a galaxy cluster. The Large Binocular Telescope (LBT), with its two 8.4-m mirrors, and its Smallest Scale of Hierarchy (SSH) survey (Annibali et al. 2020) targets stellar streams around nearby dwarf galaxies. Trujillo et al. (2021) also uses this telescope to carry out the LBT Imaging of Galactic Halos and Tidal Structures (LIGHTS) survey, focusing on stellar haloes and the low surface brightness population of satellites around nearby galaxies, reaching a depth of about 31 mag arcsec^{-2} ($3\sigma, 10'' \times 10''$ arcsec^2) in the V-band. Finally, Trujillo & Fliri (2016) aimed at reaching the theoretical SB limit of the 10.4-m Gran Telescopio de Canarias (GTC) and showed it was possible to reach around 31.5 mag arcsec^{-2} ($3\sigma, 10'' \times 10''$ arcsec^2) for an exposure time of 8h on a source.

The limit of the LSB Universe will soon be pushed further away with a new generation of instruments. There are several projects of LSB-dedicated space telescopes, such as the ARRAKIHS ESA's fast mission (PI. R.Guzmán) or the proposed MESSIER surveyor Valls-Gabaud & MESSIER Collaboration (2017). However, the most promising surveys are the upcoming Euclid and Rubin/LSST as they will cover thousands of square degrees in the sky with a depth sufficient to reveal LSB features around unprecedented samples of galaxies, enabling statistical studies (e.g., Euclid Collaboration et al. 2022; Martin et al. 2022). The Euclid VIS Wide Survey will supposedly reach a depth of 29.5 mag arcsec^{-2}, and 31.5 mag arcsec^{-2} for Euclid VIS Deep Survey ($3\sigma, 10'' \times 10''$ arcsec^2, Euclid Collaboration et al. 2022). The Rubin/LSST will reach a r-band depth lower than 30.5 mag arcsec^{-2} (in boxes of $10'' \times 10''$ arcsec^2)(Ivezić et al. 2019; Brough et al. 2020; Euclid Collaboration et al. 2022).

Therefore, the study of the LSB Universe can be conducted with instruments of various sizes, ranging from amateur telescopes to the largest professional ones. In spite of their differences in terms of optics, mount or diameter, they all face the same challenges to obtain deep images.

1.2.1.3 Challenges

In order to disclose the very faint LSB regime (>28 mag arcsec^{-2}), several challenges must be faced. They can be gathered in three broad classes: contributions from the instrument itself; contributions from the location where the observations were conducted; and contributions from astrophysical processes (see the reviews by Knapen & Trujillo 2017; Mihos 2019, and references therein).

We first focus on the contributions from the instrument itself. They are of several kinds. First, the light from bright objects, either stars or galaxy centres, can be reflected multiple times between the CCD camera and optical elements of the instrument. This triggers multiple extended 'ghost' haloes around the objects, that are concentric and more or less centred on the objects but that may extend over several arcminutes (e.g., Slater

et al. 2009; Duc et al. 2015). Some artificial haloes are presented on the left panel of Figure 1.8. The issue with this additional artificial light is the fact that they can overlap or hide LSB structures. It is possible to reduce these internal reflection by adding excellent coating on the CCD camera and the filters; but also to use a completely different telescope design such as Dragonfly (van Dokkum et al. 2014); or to adopt a particular observing strategy (Trujillo & Fliri 2016). An additional effect arises when the camera is composed of several CCDs: since each CCD is one detector, combining all the images into a single one where measurements can be performed requires precise calibration beforehand (e.g., Luppino et al. 1995; Zheng et al. 2022), that includes dealing with the gaps between CCDs (see the example on the middle panel of Figure 1.8). The shape of the instrument and more precisely the support structure called spider vanes will result in diffraction patterns that appear as strikes around bright objects, and that contribute to scatter light across the image. Another important point to consider is the impact of the Point-Spread Function (PSF) of the instrument. The PSF describes the response of an optical imaging system to a point source. If the optical system is perfect, without any aberration, then only the diffraction comes into play in the PSF. Yet, instruments do have aberrations and most of all when observing from Earth there are turbulences from the atmosphere, called seeing, that are the dominant contribution to the PSF. Solutions to overcome the seeing are to go into space (where there is no atmosphere hence no seeing) or to implement adaptive optics. The shape of the PSF also depends on the position on the image (e.g., Karabal et al. 2017). In all cases, the PSF has an important impact on the image as it contributes to scatter light. Indeed, the light from the object of interest as well as from nearby sources will be scattered by the PSF (by a convolution) and contribute to create a scattered light background across the entire image (e.g., Slater et al. 2009). For LSB science, a crucial part are the extended wings of the PSF, as they will contribute to add an excess of light towards the outer regions of galaxies and get mixed with the extended stellar halo (e.g., Sandin 2014; Trujillo & Fliri 2016; Karabal et al. 2017).

The second point to consider are the contributions from the location where the observations were carried out. Even for the darkest spot on Earth, far away from any source of light pollution, and on a portion of the sky that seems empty of any source, the sky has a SB value of around 22 $\mathrm{mag\,arcsec^{-2}}$ (Garstang 1989). This value is explained mostly by processes in the upper atmosphere, such as airglow which is a faint emission of light by the atmosphere. Indeed, during daylight the Sun photoionizes atoms which recombine during the night; there is the contribution of cosmic rays hitting the upper atmosphere and triggering luminescence; as well as chemiluminescence from the reaction of oxygen and nitrogen atoms with hydroxyl molecules. Using space-based telescopes enables to get rid of this effect: for instance, the Hubble Space Telescope (HST) can reach 1-2 mag deeper than ground-based telescopes. However, even in space, there is an additional contribution from zodiacal light, which is the light from the Sun that is scattered on dust on the ecliptic plane.

The last point to take into account are the contribution from astrophysical processes. In particular, Galactic cirrus is clouds of dust in the interstellar medium (ISM) from our Milky Way that absorb and scatter a part of the light from galaxies. Cirrus appears as complex diffuse regions on the sky with filamentary structures, and it is visible in the optical and infrared regime (where the dust radiates thermally). Some cirrus clouds are presented on the right panel of Figure 1.8. Cirrus features are present even at high Galactic latitudes (Planck Collaboration et al. 2016). An issue is that cirrus may be mistaken for tidal features or intracluster light (Davies et al. 2010; Cortese et al. 2010;

Duc et al. 2015; Miville-Deschênes et al. 2016). As it is very difficult to remove from deep images, cirrus is an important astrophysical contamination source. A smaller contribution comes from satellite trails (e.g., Cheselka 1999; Vandame 2001).

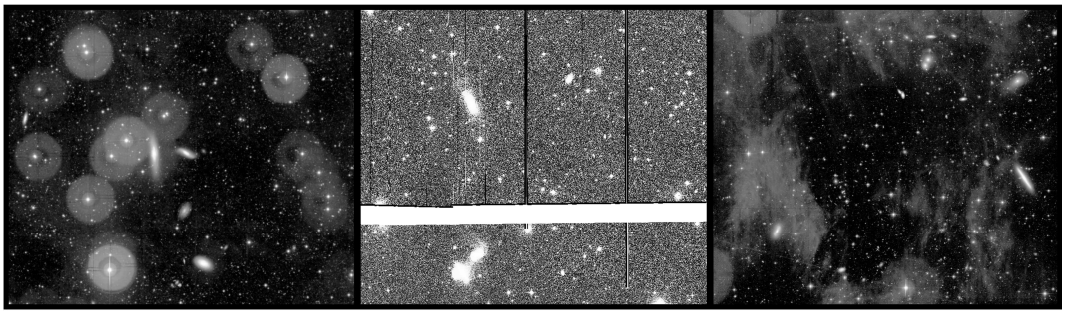

Figure 1.8: Illustration of some contamination sources that can be found in deep images. *Left*: Extended artificial haloes around bright stars originating from multiple internal reflections on the camera in a MATLAS image. *Middle*: Example of CCD gaps in a CFIS image. *Right*: Cirrus contamination in a MATLAS image.

Now that we are warned about the sources of contamination that make LSB studies more complicated, we now focus on the additional steps that are required to produce LSB-compliant images. The first is flat-fielding to correct the pixel response of the instrument. The goal is to correct the variation in the pixel-by-pixel sensitivity, to account for artefacts such as dust or scratches on the CCD camera, and to correct for vignetting (i.e. a loss in image brightness at the edges of the field of view): in the end, each pixel of each CCD will give an equal uniform response to a given source. To produce a flat-field frame, an image of a uniformly-illuminated screen is often taken. This technique is however not sufficient when the goal is to observe the faintest details in an image because of inhomogeneities in the illumination. In this case, the flat-field is created from the on-sky exposures (i.e. science images) themselves (e.g., Davies et al. 1989; Turner et al. 1993; Duc et al. 2015; Trujillo & Fliri 2016).

The second major point is background subtraction. To study galaxy formation and evolution, it is necessary to have reliable photometry, which is obtained through a precise estimation and subtraction of the sky background value. The idea is to detect the objects in an image, to mask them and to retain the remaining pixels where there should be no object as the background value. Yet, this leads to biased background, as some pixels in the outskirts of galaxies are not masked, hence contribute to increase the real background value. Usual pipelines are tuned for this type of point-source science, however it destroys the LSB signal. Indeed, for compact object science, the background is estimated on small patches of the images, which means that for extended objects like galaxies, their outskirts will be considered as being part of the background and will be oversubtracted (e.g., Aihara et al. 2018). An example of such oversubtraction is presented in Figure 1.9. It is therefore crucial to develop LSB-compliant pipelines that preserve the LSB signal. Several teams are working on this issue, for past or ongoing surveys but also for future missions such as Euclid or Rubin/LSST, such as Magnier & Cuillandre (2004); Duc et al. (2015); Ji et al. (2018); Borlaff et al. (2019); Euclid Collaboration et al. (2022); Kelvin et al. (2023), with various approaches. Only after this careful step of background subtraction can the images be analysed for LSB science.

To summarize this section, several challenges must be faced when dealing with deep images. Sources of contamination of the background arise from the instrument itself,

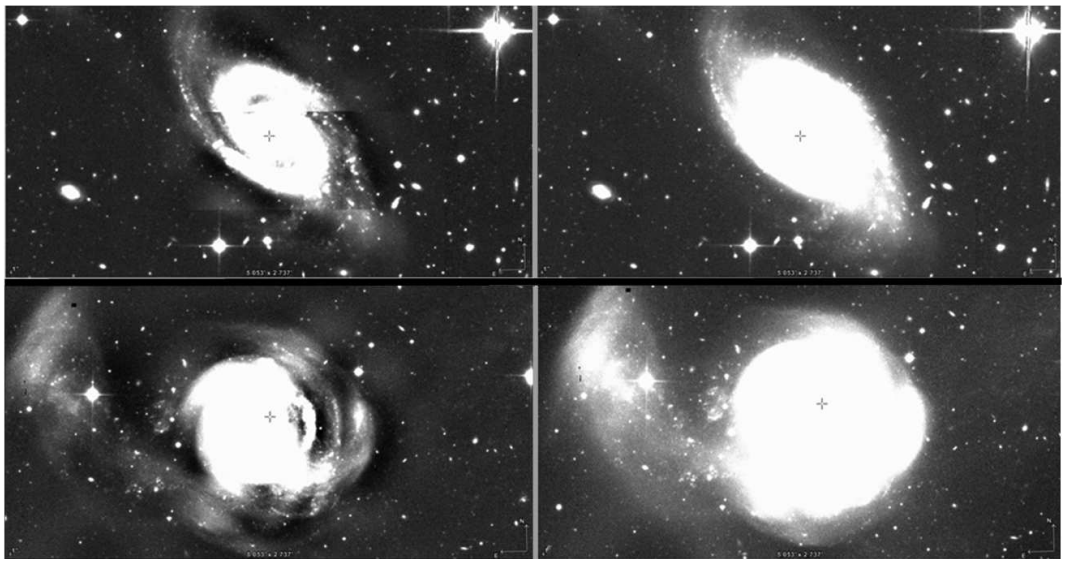

Figure 1.9: Illustration of an oversubtraction of the background value on two galaxies (on on the top panel and one on the bottom panel) in CFIS images. *Left*: Background value estimated locally, triggering an over-subtracted background in the outskirts of galaxies, erasing the LSB signal. *Right*: Background value estimated globally, that preserves the faint LSB structures.

the location at which observations are conducted as well as from astronomical sources, which must be corrected or at least limited. An important step of appropriate background subtraction that preserves the flux from LSB object is crucial for LSB studies, triggering the need for alternative LSB-compliant pipelines as traditional ones erase the LSB signal.

1.2.2 Detection of LSB structures

Once the deep images are available and treated with appropriate LSB-compliant pipelines, the next step is to start the study of LSB structures. The first point consists in detecting LSB features in deep images. There are different methods that can be divided between three categories: visual inspection, automated techniques and machine learning. The choice of the technique to apply depends on many parameters: the number of galaxies in the sample, the number of collaborators, the amount of time and computer resources available, or the specification of the task.

1.2.2.1 Visual inspection

The simplest approach consists in the visual identification of tidal features, performed either on real (pre-processed) images from telescopes but also from images that have undergone more data processing. We start by focusing on the former, with images from telescopes that are usually slightly modified (i.e., contrast, dynamics) to enhance LSB structures.

Many studies performed visual inspection of images in order to find LSB structures, with expert users examining from a few to thousands of galaxies. Starting with galaxy samples containing hundreds of objects, one can cite Jackson et al. (2021) who studied 118 low redshift, massive central galaxies in HSC images; Kluge et al. (2020) who focussed on 170 local brightest cluster galaxies from the Wendelstein Telescope Wide Field Images;

Bílek et al. (2020) who targeted 177 nearby massive ETGs in MATLAS fields; Sheen et al. (2012) who investigated post-merger signatures of red-sequence galaxies in four rich Abell clusters; or Kaviraj (2010) who focussed on ETGs with tidal features in SDSS Stripe 82 images from a sample of 902 bright nearby galaxies. Other teams identified tidal features in larger samples, such as Atkinson et al. (2013) who focussed on CFHTLS-Wide images of 1781 galaxies at $0.04 < z < 0.2$; Nair & Abraham (2010) who visually classified the morphologies of 14,000 galaxies in SDSS, including the presence of tidal features; or Bridge et al. (2010) who surveyed the 2 square degrees of the CFHTLS-deep images of 27,000 galaxies.

In addition to these works, there are other methods to detect tidal features in images on which additional steps of data processing were performed. A widely used technique is to inspect residual images: first, the light of the galaxy is modeled, and then it is subtracted from the original image. These models are symmetric and centred on the galaxy. If there are asymmetries, for instances signatures of past mergers, they will not be fitted correctly and will appear in the residual image. Several tools can be used to model galactic light profile, such as the *ELLIPSE* task from IRAF (Tody 1986) or the GALFIT (Peng et al. 2002) software. IRAF's *ELLIPSE* builds a model of the galaxy by fitting a series of elliptical and concentric isophotes to the image. Yet in the case of tidal features like shells, they might be included in the galaxy model; and there may be artefacts when the position angle of the galaxy varies significantly and brutally. GALFIT provides a different way to create a model by performing a multi-component and 2D-fitting procedure. For instance, Bell et al. (2006) and McIntosh et al. (2008) performed a 2D-fitting using GALFIT and visually inspected the residual images to respectively detect morphological signatures of mergers around 809 ETGs and signs of interactions in 221 galaxy pairs. Likewise, Tal et al. (2009) relied on the galaxy model created by IRAF to define a tidal parameter that quantifies the presence of tidal features, by comparing the values of pixels in the original and model images. Adams et al. (2012) followed this procedure to identify tidal features in 3551 ETGs from 54 galaxy clusters in CFHT images.

To this model subtraction can be added a previous step of unsharp masking. Introduced by Malin (1977) to be applied originally to photographic plates, unsharp masking consists in applying a filter on an image that enhances the high-frequencies, hence the details. To that end, a smoothed version of the original image is computed through Gaussian smoothing, and then the smoothed image is subtracted from the original one (with some level of contrast that can be added) in order to make more prominent the details, in our case the LSB structures. Several studies use a combination of unsharp masking followed by model subtraction to search for tidal features. One can cite for instance Colbert et al. (2001) who applied this technique to find collisional debris around 30 ETGs; Kim et al. (2012) around 65 ETGs from the S4G Spitzer survey or Giri et al. (2023) around 202 ETGs in SDSS Stripe 82 data.

Other methods exist and can be combined with the previous ones, such as structure maps (Pogge & Martini 2002) in which the original image is divided by a PSF-convolved copy of it in order to remove the large-scale smooth light distribution of the galaxy. In addition, Miskolczi et al. (2011) developed a search strategy for LSB structures around 474 galaxies from SDSS, taking into account the inclination, size and morphology of galaxies: after mosaicing and staking the images to improve the signal-to-noise ratio, they removed the point sources and then smoothed the image by a Gaussian kernel to enhance LSB features. This method was also used by Morales et al. (2018) to study 297 Local Volume galaxies in SDSS images. On a simpler approach, Hood et al. (2018) targeted 1048 galaxies

in the DECam Legacy Survey in images that were masked and smoothed before visual inspection.

For all these visually-based methods, a limitation arises when the size of the sample increases: it becomes impossible for a team of a few expert users to inspect all of them. One solution is to ask for the contribution of the general public through citizen science. A crowd of citizen scientists is able to classify hundreds of thousands of galaxies for simple tasks. A downsize of this approach is the lack of expertise of the users (it is not possible to ask for a precise distinction between the different types of tidal features if the users have not been trained to recognize them beforehand), but this is mitigated by the large number of annotations which makes the overall classification reliable. A famous example is the Galaxy Zoo platform (Lintott et al. 2008), which was extremely successful and gathered millions of galaxy morphology classifications. It was later complemented by the Galaxy Zoo 2 which collected 60 million classifications, and then it was expanded to other fields of science through the Zooniverse platform[2]. For instance, Casteels et al. (2013) used these Galaxy Zoo 2 classifications to study tidal features of 150,000 galaxies in SDSS images. An ongoing project, Galaxy Cruise[3] aims at bringing citizen scientists together with the task of identifying the presence of shells, streams, rings and disturbed haloes in Subaru-HSC images. However, once again, the precise distinction between sub-categories of tidal features (such as tidal tails and streams) requires previous knowledge of the field and of galaxy assembly. It is not something easily achievable through citizen science. Other methods are needed in these cases with too many objects to visually classify.

To conclude, there are several methods to find LSB structures in images based on visual inspection. They can be performed either on the images from telescopes or from images that have undergone more sophisticated data processing. Although visual inspection by expert users enables to determine the numbers and types of tidal features, an important limitation arises from the time required to do it. When there are too many galaxies to inspect, it is possible to rely on a crowd of citizen scientists but the tasks to perform are limited by the public's lack of expertise. Yet, there are cases where it is necessary to rely on more automated techniques, for complicated tasks or when the number of galaxies increases significantly.

1.2.2.2 Automated techniques

Automated methods become necessary when the sample of galaxies to classify becomes too important. There are several types of approaches. One idea is to study the level of asymmetry of the isophotes, as it can be the sign of on-going or past interactions. Pawlik et al. (2016) developed such a method and applied it to identify post-merger galaxies in a sample of 355 galaxies from SDSS. Other more general non-parametric methods were developed, as they have the advantage of not assuming any analytic function for the galaxy's light distribution, hence they can be applied to any type of galaxy. In particular, Abraham et al. (2003) and Lotz et al. (2004) introduced the Gini-M_{20} parameter to classify galaxy morphology. The Gini coefficient corresponds to the relative distribution of the galaxy pixel flux values (in most cases it is correlated with the concentration index), while the M_{20} is the second-order moment of the brightest 20% of the galaxy's flux (which is more sensitive to merger signatures). Likewise, the CAS system (e.g., Abraham et al. 1994, 1996; Schade et al. 1995; Conselice et al. 2003, 2008; Conselice 2009) is a

[2]Zooniverse, https://www.zooniverse.org/

[3]Galaxy Cruise, https://galaxycruise.mtk.nao.ac.jp/en/

non-parametric method that quantifies the shapes of galaxies, through the concentration index (C), asymmetry (A) to distinguish between early and late type galaxies, and the clumpiness (S) which is correlated with the galaxy's colour and star-formation rate. A limitation of these techniques is the fact that they are not very sensitive to LSB features as they are flux-weighted and dominated by the contribution of the most luminous parts of a galaxy. Wen et al. (2014) developed the AO-DO method that quantifies the outer structures of galaxies, and which is efficient to select galaxies with significant asymmetric features like tails. Furthermore, Mantha et al. (2019) developed a tool which is a first step to quantitatively identify and extract substructures from residual images, including tidal features from mergers. In addition, Kado-Fong et al. (2018) used an algorithm that separates the high and low spatial frequencies in an image, respectively corresponding to tidal structures and galaxy light, in order to pre-select galaxies that host potential tidal features, and then a visual inspection was performed.

1.2.2.3 Deep learning

Another promising approach when dealing with too large samples is to use machine learning (ML), and more precisely deep learning (DL). ML is a form of artificial intelligence that uses computers to automatically learn and improve their methods to find patterns in data. ML is often used to perform classification or regression tasks based on a set of parameters: for instance, it can be used to classify the morphology of galaxies from a set of parameters obtained after a previous step of feature extraction (e.g., Barchi et al. 2020). DL is a particular case of ML where the behaviour of the computers mimics the one of the neural system of living beings. DL is often applied directly on images, i.e. it starts from the pixels of the image to perform a classification and not from a set of parameters like classical ML methods. The architecture is based on artificial neural networks, where some neurons will be activated while other will regress as a function of the type of the input signal. The number of layers and the number of neurons per layer are two examples of parameters that describe the architecture: the more layers, the deeper the network. Convolutional neural networks (CNNs) are a type of architecture, based on convolutional layers, commonly used in DL as they are very efficient in processing images.

There are several types of ML, in particular supervised and non-supervised. In the supervised case, the network is given input data associated to their labels (i.e. what do the data represent) on which it will train to recognize the features of interest, and it learns to associate and predict labels of unseen data. Several metrics are used to characterize the performance of the network and to penalize wrong classifications. For instance, a supervised ML can be given as inputs many galaxy images, with labels corresponding to the morphological type, and the network will predict the morphological type of new galaxy images. Supervised ML hence requires data that were previously labelled, either manually or by other techniques. Another possibility is to use unsupervised ML. In that case, there are no labels associated to the input data: the network is more independent and learns on its own to classify data as a function of some characteristics it found in them. To follow up on the example of galaxy morphology, the network will learn to morphologically separate images of galaxies according to some pattern it found in the images.

Astronomy is a perfect field of application of ML and DL, because of the large number of images or data to process and of the huge number of possible applications. In particular, it has been widely used to classify the morphologies of galaxies, through supervised learning. Among the first successful tries, one can cite Dieleman et al. (2015), followed and extended by Huertas-Company et al. (2015), who developed a supervised CNN to

classify galaxy morphologies in SDSS, both in terms of overall morphology and presence of substructures like spiral arms. It was trained on labeled data from the Galaxy Zoo 2 dataset, and it achieved unprecedented classification accuracy of $> 90\%$ in most of the tasks. This work was important as it was the first to use CNN in astronomy, and it can be seen as the beginning of the wide use of DL in astronomy. Later, Vega-Ferrero et al. (2021) used CNNs trained on a small annotated DES dataset to distinguish between ET-G/LTG and face-on/edge-on orientations for 27 million galaxies in DES, providing the largest multiband catalogue of automated galaxy morphology at that time. Walmsley et al. (2022) provided more accurate classification (global morphology but also substructures, spiral arms, bar bulge, rarer features and mergers with major/minor disturbances) for 314,000 galaxies with a Bayesian CNN trained on Galaxy Zoo DECaLS. In the previous examples, the labels of the data were obtained manually, using the classifications by a crowd of citizen scientists. However, they can also be obtained by automated methods. For instance, Tohill et al. (2021) used a CNN to predict the concentration (C) and asymmetry (A) of galaxies, which was trained on CAS measurements obtained with the IRAF software for 150,000 galaxies from the CANDELS fields.

However, supervised learning requires a large amount of input labeled data, which can be time-consuming to obtain especially if it relies on human labelling. There are several options to tackle this issue while still performing supervised ML. The first one is to reduce the number of necessary input training data, with methods like few-shot learning (Zhang et al. 2022). The second is to train on simulations. Indeed, as the ground-truth is known in simulations, there is no need of manual classification as the data are already labelled. For instance, Pearson et al. (2019) trained a CNN on two datasets, one from SDSS and the other from the EAGLE simulation, to find galaxy mergers in images. The CNN performed well when trained and applied on real data (and likewise when trained and applied on simulations); but it was unsuccessful when trained on simulations and applied to real data, underlying the need for a better resolution in simulations and further refinement in the processing of mock images from simulations. Bickley et al. (2021) trained a CNN on images from the IllustrisTNG simulation in order to automatically detect post-merger galaxies in CFIS/UNIONS images. They insisted on the realistic aspect of simulated images without which the network would perform badly when applied to real data.

Another possibility to deal with the lack of labeled data is transfer learning. It consists in transferring the knowledge learnt by the network on a given field to another field: this reduces the number of input labeled data needed. For example, Domínguez Sánchez et al. (2018) showed that their DL network, whose goal was to classify galaxy morphology, could adapt from one instrument (SDSS) to another (DES) with a reasonable global accuracy. Likewise, Ghosh et al. (2020) developed a CNN to classify galaxy morphology based on their bulge-to-total light ratio which was first trained on simulated data and later successfully applied to SDSS and CANDELS images. Similarly, Ackermann et al. (2018) first trained a CNN to recognize everyday objects and then it was re-trained on SDSS images to recognize galaxy mergers (with labels from the Galaxy Zoo project). It appeared that transfer learning led to a small improvement of the results.

Finally, a different option to address this issue is unsupervised learning. Although there is less control over the output, the possibility to deal with unlabeled data is promising for surveys where manual classification is not an option. Martin et al. (2020) used an unsupervised network to classify the morphologies of galaxies from HSC-SSP images. The network found 160 morphological clusters that were later gathered in three broad classes (elliptical, S0/Sa and spiral) by visual inspection. Likewise, the network of Cheng

et al. (2021) found 27 morphological clusters for their galaxies, separated on Sérsic index, concentration, asymmetry and Gini coefficient. If gathered in two classes (ETG/LTG), the accuracy reaches 87%. This work underlined the fact that this unsupervised method may provide more physically meaningful classifications than visually-based ones. In a similar approach, Spindler et al. (2021) produced morphological classification from unlabeled data through an unsupervised clustering model that separates galaxies as a function of their orientation, surface brightness profiles, axis ratio and presence of companions.

All the studies mentioned above focus on the automated classification of galaxy morphologies, using various approaches. Yet, they do not focus much on the presence of tidal features, which are crucial to detect. Other works are designed for this precise task. In particular, it is worth noting the pioneer work by Walmsley et al. (2019) who trained a CNN for stream detection based on 1781 CFHTLS-Wide galaxies visually annotated for tidal features by Atkinson et al. (2013). The network achieved an overall accuracy of 76%, likely due to the small number of annotated dataset. This is still better than what automated methods from the literature can provide. In a similar approach, Domínguez Sánchez et al. (2023) based their CNN on a dataset of tidal features from 6000 simulated images annotated by a group of professional astronomers, presented in Martin et al. (2022). The CNN, trained and applied on simulated images, is able to reproduce human classification, with some features more easily recognizable than others. However, when trained on simulations and then applied to real data, the network is unsuccessful, which is likely due to a lack of realism in the simulated images. Both these works rely on supervised learning with manual labels. Storey-Fisher et al. (2020) tried an unsupervised method based on the detection of 'anomalous' galaxies, which after visual inspection appeared to present mergers, tidal features and extreme star-forming galaxies. This approach would be useful in a first step to automatically detect anomalous galaxies that are potentially hosting tidal features, and in a second step to identify them.

It is worth noting that all these studies on galaxy morphology and tidal feature classification are limited to their identification, i.e. saying whether the galaxy is elliptical, or whether there is a tidal feature. This is not sufficient if we want to characterize tidal features. It is necessary to have identification, census but mostly segmentation of such collisional debris. Segmentation is a different task: it is a technique where the algorithm not only detects a feature in an image, but automatically divides this image into pixel regions that are associated to a same object class. For instance, SExtractor (Bertin & Arnouts 1996) performs a simple segmentation task by determining whether a pixel belongs to a source (if its flux is higher than the background level) or to the background, but it does not give any indication about the type of the source segmented. To our knowledge, at the beginning of this PhD thesis there was no network dedicated to a task of tidal feature segmentation. There were only a few articles presenting segmentation algorithms for astronomical images, such as Hausen & Robertson (2019) with a supervised network tuned for simultaneous detection of morphological classification of astronomical objects in images. Other structures are important to segment. It is the case of pollutants such as galactic cirrus, as it makes the identification of tidal features more complicated. Indeed, as it has a LSB filamentary structure, cirrus can either resemble elongated collisional debris or hide it partially. Determining which regions of an image are contaminated is important to assess how reliable the tidal feature identification will be. This was the focus of a work by Richards et al. (2020) that took place at the same time as his PhD thesis, and which will be described in Section 7.2.2. More recently, Smirnov et al. (2023) also used supervised learning from manually-labeled cirrus regions to produce cirrus segmentation

maps.

To conclude, machine learning is a promising solution towards the automated detection of tidal features, and later their segmentation which is a crucial step towards their characterization in large samples of galaxies. Yet, for this specific task, a previous step of human labeling is necessary to provide labeled datasets for neural networks to be trained on.

1.2.3 Characterization of LSB structures

We remind here briefly the reasons why the study of LSB structures is important. As tidal features are the remnants of past mergers and interactions, studying them enables to probe the past assembly history of galaxies. Their characteristics, such as number, shape or morphology can be compared to the predictions of numerical simulations of galaxy formation and help constrain these models. Therefore, in order to have a comprehensive view of galaxy evolution, it is necessary to study a complete and statistically significant sample of galaxies, spanning a large range of mass and environment. Another key point is to analyse the results in a consistent manner.

In the previous section, we saw that there are many different methods to detect and study tidal features, either based on visual inspection or relying on more automated approaches. The choice of the method will depend on several criteria, such as the size of the galaxy sample, the number of collaborators, the amount of time available, and the type of study that will be performed. We can broadly group them in two categories: the ones aiming at producing a general census of tidal features around large samples of galaxies and the ones aiming at deriving more detailed information about them. For the former type, there can usually be a few expert collaborators or a crowd of citizen scientists, and the size of the samples can range from a few hundred to thousands, yet the classification task remains simple. For instance, the level of asymmetry of galactic morphology, the presence of tidal disturbance or even census of tidal features with a distinction between several types of features (yet restrained to expert users) are the type of information that can be collected (e.g., Tal et al. 2009; Miskolczi et al. 2011; Casteels et al. 2013; Morales et al. 2018; Bílek et al. 2020). Automated methods do not offer yet the subtle distinction between various types of tidal features, so this remains mostly done visually. For the latter category, a detailed analysis can be performed by a few expert users on a rather limited sample of galaxies, but with more information such as the surface brightness, morphology, or fraction of flux contained in tidal features and extended stellar halo (e.g., Trujillo & Fliri 2016; Huang & Fan 2022; Martínez-Delgado et al. 2023). However, to our knowledge there was no solution to bridge the gap between the two methods, i.e. the possibility to derive quantitative measurements of LSB tidal features around large and statistically significant samples of galaxies, which is needed to reconstruct the late assembly histories of individual galaxies.

One solution could be to combine the results from all previous works into a single study. However, this is not possible. Indeed, the definition of what a tidal feature is, the limiting surface brightness and optical bands as well as filters used for the surveys, the detection method or selection criteria vary hugely from one study to the other (e.g., Atkinson et al. 2013; Duc et al. 2015; Hood et al. 2018). Indeed, some works studied tidal features around galaxies that were already known to host LSB structures, without the goal of being a representative sample; while others focus on representative samples by selecting all galaxies matching a given criterion. This can lead to large discrepancies when comparing results from various surveys. For instance, the frequency of tidal features

around galaxies can range from a few percents (e.g., Malin & Carter 1983; Adams et al. 2012) to a very high fraction (e.g., van Dokkum 2005; Tal et al. 2009). The environment, distance or morphological types of galaxies also present huge variations. These reasons call for a statistical sample of galaxies that will be analysed in a consistent way. Several studies were carried out with this goal, such as the ones from e.g., Atkinson et al. (2013); Hood et al. (2018); Bridge et al. (2010); Bílek et al. (2020), but they restricted their analyses to the census of tidal features without a more quantitative approach.

With this in mind, the goal of my PhD work was to systematically quantitatively characterize tidal features around a complete sample of nearby massive galaxies in deep images, covering different environments and analysed in a consistent way. To that end, we developed an online annotation tool that enables expert collaborators to draw with precision the shapes of LSB structures superimposed on deep images. The goal is to annotate every LSB stellar structure (tidal features, extended stellar halo), but also sources of contamination such as ghost haloes from internal reflections or Galactic cirrus. The delineated features are stored in a database, which can be queried to retrieve quantitative measurements. In particular, the geometry, SB, colour or luminosity of tidal features can be computed. Such values were previously not well constrained on large samples of galaxies, although they are crucial to obtain as they can later be compared to the predictions of numerical simulations to put constraints on the late assembly history of galaxies.

1.3 Outline of this thesis

Through this introductory chapter, we have had an overview of the formation and evolution processes of galaxies. We have seen that collisions between them trigger tidal debris, whose properties are useful to probe the late assembly history of galaxies. In order to reveal this LSB collisional debris, dedicated instruments, observing strategies and data reduction pipelines are necessary. From these deep images, several methods aim at detecting and characterizing LSB tidal features. During this thesis, I have focussed on building a statistical sample of tidal debris around a large number of massive nearby galaxies in a quantitative approach. The goal was to probe galactic assembly for various masses and environments through a detailed analysis of their tidal features.

The first step of this work was to focus on a dedicated online annotation tool, as presented in Chapter 2. The annotation tool had already been developed, but I put it online, maintained and improved it during the PhD thesis. The analysis tools developed to derive quantitative measurements of LSB structures are presented.

Then, in Chapter 3 I present the deep images in which LSB features have been studied, as well as the sample of galaxies we targeted in Chapter 4. The characteristics of the instrument, of the data pre-processing and of some properties of the galaxies will be discussed.

With the online tool and the deep images, I classified LSB features around about 450 nearby massive galaxies, as did a few expert collaborators. Chapter 5 presents the results of the statistical analysis of these LSB structures. The implications of these results and their comparison to the literature are discussed in Chapter 6.

A last important point consists in the comparison of LSB structures in observations and in simulations, in order to constrain models of galactic evolution. In Chapter 7, I use the annotation tool to classify tidal features in simulated images. I present some preliminary results and compare them to observations, as well as discuss future prospects.

Finally, in Chapter 8 I present the general conclusions of my PhD work.

1.4 Résumé du chapitre

1.4.1 Evolution des galaxies

Galaxies - Les galaxies sont des systèmes composés d'étoiles gravitationnellement liées entre elles, ainsi que de poussières, de gaz, probablement situées au centre de halos de matière noire. Elles peuvent avoir des formes et des propriétés très variées. La classification de Hubble sépare les galaxies en fonction de leur morphologie: elliptiques, lenticulaires ou spirales, en incluant des galaxies irrégulières dont la morphologie n'est pas typique ou encore des galaxies naines. On sépare généralement les galaxies de type 'précoce' elliptiques (ETGs) des types 'tardifs' spirales (LTGs), même s'il est important de noter que ces dénominations n'impliquent pas une évolution des premières vers les secondes, au contraire.

Formation et évolution - En effet, selon les modèles hiérarchiques de formation des galaxies, les galaxies se sont formées par accrétion de gaz, fusions successives et accrétion de systèmes plus petits. Leur évolution dépend de leur type morphologique. Les ETGs résultent probablement d'une formation en deux phases, et résultent parfois de la collision entre deux spirales. Les LTGs semblent avoir une évolution plus douce, plutôt par accrétion de gaz intergalactique.

Un paramètre important à prendre en compte dans l'évolution des galaxies est leur environnement. Certaines galaxies sont isolées (i.e. dans le champ), d'autres sont regroupées dans des environnements un peu plus denses (des groupes), tandis que d'autres encore résident dans des environnements très denses (des amas de galaxies). Il a été confirmé de manière observationnelle et avec des simulations que les galaxies ont des propriétés différentes dans les amas et dans le champ. Par exemple, on trouve plus d'ETGs dans les amas, et plus de LTGs dans le champ. Plusieurs processus à l'intérieur des amas contribuent à des modifications drastiques de l'évolution des galaxies ainsi qu'à leur 'extinction'. En effet, plusieurs mécanismes (tels que la pression dynamique du milieu intra-amas ou des interactions rapides, mulitples et répétées avec les autres galaxies de l'amas) contribuent à arracher le gaz des galaxies. Or ce gaz est le carburant permettant de créer de nouvelles étoiles. Sans réservoirs de gaz et sans possibilité de le régénérer, les étoiles meurent peu à peu et ne sont pas renouvellées, ce qui laisse place à des galaxies dites éteintes.

Fusions et accrétions de galaxies - Les fusions successives et interactions entre galaxies modifient également leurs propriétés. En particulier, elles donnent lieu à la création de débris de collision: ce sont des étoiles arrachées à leur galaxie sous l'effet des forces de marée en jeu. Ces étoiles se retrouvent à former de grandes structures qui s'étendent autour des galaxies, on les appelle des structures de marée. La nature, morphologie et propriétés de ces débris dépend du type de collision qui a eu lieu. Ainsi, une fusion majeure entre deux galaxies de masses similaires génèrera des queues de marée, longues et relativement larges. Une fusion mineure avec une galaxie de bien plus petite masse créera des courants stellaires, plus petits et plus fins. Une fusion de masse intermédiaire aura tendance à créer des structures circulaires concentriques nommées coquilles. Ainsi, identifier le type des débris de marée permet de déterminer quel type de collision a eu lieu, et de remonter l'histoire passée de la galaxie. Ceci fait partie de l'archéologie galactique, dont le but est de comprendre comment les galaxies se sont formées et ont évolué à partir de ce que l'on peut observer aujourd'hui.

Ces débris de collision sont eux aussi affectés par l'environnement dans lesquelles les galaxies se trouvent. Dans les amas ils résultent de procédés opposés, certains tendant à

les générer (de par la proximité des galaxies) et d'autres tendant à les faire disparaître rapidement (interactions galactiques répétées à haute vitesse). Il est difficile de déterminer quel procédé domine.

Pour comprendre l'évolution des galaxies dans leur ensemble à travers l'étude des débris de collision, il est nécessaire d'avoir un grand échantillon de galaxies couvrant différents environnements. Il est également crucial de caractériser de manière quantitative ces débris, par exemple en termes de morphologie ou luminosité. En effet, ces données pourront être comparées aux résultats des simulations numériques de formation des galaxies: si observations et simulations diffèrent, les modèles pourront être ajustés et ainsi de suite jusqu'à ce qu'une concordance entre les deux soit possible. Avoir des mesures quantitatives sur de grands nombres de débris de collision permet donc de contraindre les modèles de formation et d'évolution des galaxies. L'un des problèmes majeurs est précisement d'obtenir ces mesures. Ceci est le but principal de ma thèse: caractériser quantitativement des débris de collision autour d'un grand nombre de galaxies, puis d'en tirer des interprétations physiques sur la manière dont les galaxies se sont formées.

1.4.2 L'Univers LSB

Faible brillance de surface (LSB) - L'étude des débris de collision est possible dans notre Galaxie et le Groupe Local en se basant sur des relevés du ciel où les étoiles sont résolues individuellement. En revanche, cette étude n'est plus possible pour des galaxies plus lointaines où les étoiles ne sont plus différentiables individuellement et il est nécessaire de se baser sur la lumière diffuse qu'ils émettent. Cependant, ceci est rendu compliqué par la très faible luminosité par unité d'aire des débris de collision: on parle de faible brillance de surface (LSB). De récents progrès dans les systèmes d'imagerie et des traitements de données particuliers ont permis d'ouvrir une fenêtre sur cet Univers LSB. Grâce à ces images dites profondes obtenues par des télescopes professionnels comme amateurs, les débris de collision et autres structures de faible brillance de surface sont révélés.

Identifications des structures LSB - Une fois que les images profondes sont disponibles, il s'agit de détecter et d'identifier les débris de collision. Il existe plusieurs méthodes, mais l'une des techniques les plus répandues est d'inspecter visuellement les images. Après un ajustement de leur dynamique et de leur contraste pour faire ressortir les structures LSB, des collaborateurs passent en revue les images et notent la présence ou non de débris. Quand le nombre de galaxies à étudier est raisonnable, plusieurs astronomes peuvent contribuer. En revanche, quand la taille des échantillons de galaxies augmente, il n'est plus possible pour une petite équipe de tout inspecter. Si la tâche demandée est simple, par exemple classer la morphologie des galaxies, il est possible de faire appel à des volontaires issus du grand public : des citoyens scientifiques. Cependant, pour des tâches plus complexes telles que la classification précise de débris de collision, bien souvent ambigüe, il est nécessaire de faire appel à des utilisateurs plus expérimentés qui ont déjà une connaissance du sujet. Une autre approche est d'utiliser des techniques automatiques. Diverses méthodes mesurant par exemple le degré d'asymmétrie des galaxies ou la déviation de leur profil de lumière par rapport à un modèle de galaxie non perturbée permettent d'identifier les galaxies avec débris. D'autres techniques d'apprentissage automatique (machine et deep learning) sont également prometteuses. Cependant, toutes ces méthodes ne permettent pas de distinguer de manière subtile les différents types de débris de marée, et ne donnent pas de résultats quantitatifs pour de larges échantillons.

Caractériser les structures LSB - Il est nécessaire d'obtenir des mesures quantitatives des débris de marée pour de larges échantillons de galaxies. Certaines équipes ont

fait des études détaillées mais de quelques galaxies uniquement, tandis que la majorité se contente de recenser les structures et leur type. Cependant, à notre connaissance il n'y a pas d'approche entre les deux qui permettrait de caractériser les débris de marée autour d'un grand nombre de galaxies, notamment à cause d'un manque d'outils dédiés.

Dans cette thèse, nous présentons un nouvel outil d'annotation qui permet à des utilisateurs de dessiner avec précision les contours des structures LSB superposées aux images profondes. Les annotations sont stockées dans une base de données, à partir de laquelle différentes mesures sont faites, telles que la morphologie, luminosité ou encore couleur. Nous avons classé les débris LSB autour de 450 galaxies dans des environnements différents. Nous nous sommes servis des résultats des annotations afin de déterminer l'impact de la masse et de l'environnement des galaxies sur les débris de collisions, ainsi que pour fournir des résultats quantitatifs sur les débris eux-mêmes.

1.4.3 Aperçu du contenu de cette thèse

Ce chapitre introductif nous a donné un aperçu des processus de formation et d'évolution des galaxies. Nous avons vu que les collisions entre les galaxies génèrent des débris de marée, dont les propriétés sont utiles pour sonder l'histoire de l'assemblage tardif des galaxies. Afin de révéler ces débris collisionnels LSB, des instruments dédiés, des stratégies d'observation et des pipelines de réduction de données sont nécessaires. A partir de ces images profondes, plusieurs méthodes visent à détecter et à caractériser les propriétés des structures de marée. Au cours de cette thèse, je me suis concentrée sur la construction d'un échantillon statistiquement significatif de débris de marée autour d'un grand nombre de galaxies proches et massives dans une approche quantitative. Le but était de sonder la croissance des galaxies, ayant des masses différentes et résidant dans des environnements différents également, à travers une analyse détaillée des caractéristiques de leurs débris de collision.

La première étape de ce travail a été de se concentrer sur un outil d'annotation en ligne dédié à l'étude des structures LSB, comme présenté dans le chapitre 2. L'outil d'annotation avait déjà été développé, mais je l'ai mis en ligne, maintenu et amélioré pendant le doctorat. Les outils d'analyse développés pour obtenir des mesures quantitatives des structures LSB sont également présentés.

Ensuite, dans le chapitre 3, je présente les images profondes dans lesquelles les caractéristiques LSB ont été étudiées, ainsi que l'échantillon de galaxies que nous avons ciblé dans le chapitre 4. Les caractéristiques de l'instrument, du prétraitement des données et de certaines propriétés des galaxies seront discutées.

Avec l'outil en ligne et les images profondes, j'ai classifié les structures LSB autour d'environ 450 galaxies massives et proches, ainsi que quelques collaborateurs experts. Le chapitre 5 présente les résultats de l'analyse statistique de ces structures. Les implications de ces résultats et leur comparaison avec la littérature sont discutées dans le chapitre 6.

Un dernier point important consiste en la comparaison des structures LSB dans les observations et dans les simulations, afin de contraindre les modèles d'évolution galactique. Dans le chapitre 7, je discute des perspectives possibles. J'utilise notamment l'outil d'annotation pour classer les débris de marée dans des images simulées, et je présente quelques résultats préliminaires et les compare aux observations.

Enfin, dans le chapitre 8 je présente la conclusion générale de mon travail de doctorat.

Chapter 2 An annotation tool for LSB structures

From Chapter 1, we learnt that LSB tidal features hold crucial clues about the history assembly of galaxies as their properties are a function of the collision that occurred. Their comparison to the predictions of numerical simulations will help put constraints on the models of galaxy formation and evolution. However, to carry out such an analysis, it is necessary to identify and precisely characterize them in deep images. Although automated methods exist, visual classification by expert users remains the preferred approach in particular for complicated tasks such as disentangling between various types of tidal features. Dedicated online tools ease these processes.

In this Chapter, we present the annotation tool with which quantitative measurements of LSB structures can be retrieved. We start by introducing the tool, then we present the annotation process itself and the analyses performed, and finally we focus on the application of machine learning using our annotation database.

2.1 Visual classification of tidal features

As mentioned in Chapter 1, there are several methods to classify tidal features, of which visual inspection is widely used. Even though more automated methods or machine learning (ML) can be applied, a step of human labelling is necessary to create large annotated datasets that can be used to train supervised ML algorithms. The visualization and annotations processes can be carried out in several ways, from a few collaborators to thousands, inspecting uneditable or modifiable images and using simple non-astronomical tools or dedicated softwares built for that purpose.

We will focus here on images from telescopes that have been pre-processed by appropriate data processing pipelines to preserve the LSB signal. We will not address images that have undergone additional data processing such as residual images from model subtraction or unsharp masking (as described in Section 1.2.2.1). By doing so, we are biased towards finding LSB structures in the outskirts of galaxies, and we may miss some internal structures (like shells) that could be disclosed in residual images. An important point to keep in mind when visually classifying LSB tidal features, whatever the tool used, is the intensity stretching of the images. Indeed, because of their low surface brightness, these structures do not appear in the original linear image scaling with full dynamics, so it is necessary to pre-process them by modifying the scaling and dynamics of the image intensity. For instance, applying an inverse hyperbolic sine enhances the faint outskirts of galaxies and potential tidal features.

There are two main approaches to inspect the images, each with their advantages and biases. The first is to use uneditable images, i.e. fixed images in which the dynamics, contrast or scaling cannot be modified. This has the advantage of having a single image for all the human classifiers, so they all visualize the same image, which enables a homogeneous treatment. This approach is biased towards finding features in the outskirts of galaxies (as the images were scaled to visualize the external parts). The second approach is to utilize modifiable images such as FITS files, i.e. where the intensity of the image can be edited. Being able to modify the image dynamics enables to search in a finer and more precise manner the faintest features. It also allows for better probing of the central regions, which would otherwise appear far too bright to distinguish structures inside. However, the different collaborators will search for debris in images that will not

be identical from one collaborator to the other, which may lead to discrepancies in the analysis.

We will mention here different types of studies that already have been carried out to characterize LSB tidal features, and their approach. Among the approaches relying on uneditable images, a rather simple process consists in opening the png/jpeg images with an image viewer, and having a corresponding table in which collaborators fill the number and type of tidal features they see around each galaxy. The output tables (in text format, ascii, csv, votable...) are then gathered and processed to obtain the result of the census. Among the many projects that rely on this organisation type, one can cite Martin et al. (2022) who proposed 45 expert users to identify the type and number of tidal features in simulated images through the inspection of png thumbnails (I contributed to this classification process). Likewise, Morales et al. (2018) visually detected stellar structures around nearby galaxies in uneditable but previously enhanced SDSS images. Valenzuela & Remus (2022) created a 3D-rendering view to display simulated galaxies, that were rotated to be viewed from all sides. The seven scientists who performed independent tidal feature identification were shown the same 3D-view (similar to an unmodifiable image) at the same time.

Among the studies working on modifiable images, Bridge et al. (2010) visually inspected galaxies for signs of interactions, tidal tails and bridges through the data visualisation software DS9[1] (Joye & Mandel 2003). Similarly, Atkinson et al. (2013) used DS9 to classify tidal features around 1789 galaxies, as it enables interactive manipulation of the image dynamics. One can also cite Jackson et al. (2021), where 5 expert users annotated features in 118 galaxies by changing the contrast of the images. Kluge et al. (2020) used a similar approach, by creating beforehand images with several scalings to maximize their changes to find features of various SB in 170 cluster environments. While these works rely on a simple organisation for the inspection (i.e. viewing and manipulating the image), others developed dedicated online tools to perform visual classification. For instance, the MATLAS web interface[2] was built for tidal features census (Duc et al. 2015; Bílek et al. 2020). It is composed of a navigation tool, based on the VisiOmatic web client (Bertin et al. 2015) with the possibility to adjust the scaling of the images, and of a questionnaire where collaborators were asked to produce a census and classification of tidal features around ETGs in MATLAS deep images. An illustration of the MATLAS interface is shown in Figure 2.1.

As single centralised systems, these online tools have several advantages. First, they provide an easy way for multiple collaborators to access the data to visualize. There can be a huge number of collaborators, who do not need to download the data on their personal computer. They can be connected and annotating at the time that suits them best, hence there can be many collaborators online at the same time. In addition to viewing the images, these systems generally include forms and/or tools to make the classification easier and faster, in order to centralise the entire process in a single app or online platform.

Several tools have been developed to enable human classification, segmentation and labelling, especially for the machine learning and computer vision communities. For instance, CVAT (Sekachev et al. 2020) is an interactive video and image annotation tool for computer vision; LabelMe (Russell et al. 2008) is a web-based tool enabling image annotation and sharing of the annotations, or VoTT[3] whose goal is to detect objects in

[1]SAOImageDS9, https://github.com/SAOImageDS9/SAOImageDS9
[2]MATLAS website, http://obas-matlas.u-strasbg.fr/
[3]VoTT, https://github.com/microsoft/VoTT

Figure 2.1: Illustration of the interface of MATLAS web server.

images and videos. Yet, these softwares are not suited for astronomical images because the projection from 2D images to the celestial sphere is not supported.

Other softwares were built to tackle this challenge in astronomical images. An example of such tool is the Zooniverse[4] platform. It was originally conceived as Galaxy Zoo (Lintott et al. 2008) with the aim of involving the general public, referred to as 'citizen scientists', to classify the morphologies of galaxies in SDSS images, a task impossible to carry out by a few collaborators given the huge number of galaxies to survey (roughly 900,000). This project was extremely successful: the numbers of citizen scientists, individual classifications, galaxies surveyed and required time have surpassed the expectations, and the results of the classifications have been used in many papers (e.g., see Table 1 of Fortson et al. 2012). Given the enthusiasm of the public and their reactivity, the project was expanded to Galaxy Zoo 2, which gathered over 60 million classifications. It was later broadened to other scientific fields through the Zooniverse platform (Fortson et al. 2012). In the context of LSB tidal features, the Zooniverse platform has not been used yet, even though some drawing tools are available. The future Galaxy Cruise[5] project aims at involving citizen scientists to identify tidal features and galaxy mergers in Subaru Hyper Suprime-Cam (HSC) images. Meanwhile, we already mentioned the MATLAS web interface whose goal was to provide a catalogue of tidal features and to survey their incidence. The website enables to navigate through the image, zoom in/out, modify the intensity, contrast and saturation, as well as a tool to measure the light profile along a line.

Therefore, several methods and/or dedicated online tools enable to collect information about the frequency and type of tidal features around galaxies. However, they do not enable to record quantitative measurements, such as geometry, position, surface brightness or colour, which is crucial to characterize LSB structures and the mass assembly of galaxies, as these properties can be compared to the predictions of numerical simulations to better constrain these models. To that end, a new online annotation tool has been developed with this dedicated purpose, and it will be presented in detail in the next Sections. This tool is in line with the first approach, as it is not possible to change the image

[4]Zooniverse, https://www.zooniverse.org/

[5]Galaxy Cruise, https://galaxycruise.mtk.nao.ac.jp/en

dynamics, but with a dedicated interface.

2.2 A new annotation tool

This Section describes the annotation tool used to record quantitative measurements of LSB tidal features. It presents the general purpose of the tool, its architecture, the steps to follow to annotate a galaxy and the database in which all the annotations are stored.

2.2.1 Why developing a new tool ?

The large goal beyond the development of annotation tools is to better understand how galaxies acquire their mass and evolve, through the characterization of LSB tidal features around large samples of galaxies. This classification process is complicated, and for now visual inspection remains one of the preferred solutions when the size of the sample remains relatively small, triggering the need for dedicated tools. These tools must provide an easy way to navigate through images and be able to record quantitative measurements.

To our knowledge there was no existing satisfying solution. To tackle this issue, Felix Richards developed the annotation tool and associated database during his PhD thesis, supervised by Dr. Adeline Paiement. My PhD work on the annotation server is based on his code[6], that I modified and made accessible by putting it online. F. Richards' work on the classification and segmentation of galactic structures in large multi-spectral images can be found in his PhD thesis (Richards 2023).

F. Richards' goals were multiple. As a former PhD student in machine learning, he was aiming at developing a neural network able to segment (i.e. delineate) and classify various galactic structures in deep images. Indeed, as future surveys will probe the sky over thousands of square degrees, it will no longer be possible to do human labelling and ML will become necessary. He focussed first on the detection of galactic cirrus clouds, which are important to segment (and later remove) as they pollute images and make the identification of faint structures more complicated. Then, he developed a network able to perform multi-class segmentation, i.e. segmentation of different types of objects at the same time, as well as cirrus detection. Yet, to train his supervised deep learning (DL) algorithms he needed an annotated dataset of structures in LSB images. As no such dataset was available at the time, he developed the annotation tool and database for this purpose. Once the LSB structures were annotated by astronomers, F. Richards could use the annotation database as his training set.

In the rest of this Chapter, we will apply the annotation tool to our problem of characterisation of LSB structures in deep images. In Section 7.2.2 we will discuss the prospects of using the annotation database to train DL algorithms.

2.2.2 Description of the annotation tool

2.2.2.1 Architecture

The goal of the annotation tool is to display uneditable deep images of galaxies, and on top of them to manually draw the shapes of LSB structures. This requires both a navigation and a drawing interface. The tool has to be online to enable collaboration and annotations by several astronomers.

[6]Dr Felix Richards' source code available at https://github.com/felixrichards/Project-Web-App

The navigation interface is based on Aladin Lite[7], which is a lightweight version of the Aladin Sky Atlas developed by the Centre de Données Astronomiques de Strasbourg (CDS). Aladin Lite can easily be embedded in a web app. Its aim is to visualize and explore pre-existing image surveys or to display custom images, as well as overlaying object information from astronomical databases such as SIMBAD[8]. Aladin Lite does not support FITS format for custom images, so it is necessary to convert the input images to the supported HiPS format. Standing for Hierarchical Progressive Surveys (Fernique et al. 2015) and developed by the CDS, HiPS is a hierarchical tiling mechanism to access and visualize astronomical data. The celestial sphere is divided in smaller and smaller tiles with a diamond shape, so each portion of the sphere is indexed by a number and an order. The higher the order, the smaller the tiles hence the better the resolution. Therefore, zooming in an image in HiPS format will reveal more and more details. The conversion between FITS and HiPS format can easily be done using the Hipsgen command[9]. The version used during this work is Aladin Lite v2. In this version, it is not possible to adjust the contrast and dynamics of the images on-the-fly. This means that for the deep images we had to adjust their scaling before converting them to HiPS format, i.e. if we wanted two scalings of the same image, we had to create two HiPS and then to add them to Aladin Lite. A more detailed description of the scaling used during this work is presented in Section 3.3. In the newer version of Aladin Lite v3 released early 2023, it is possible to adjust the image dynamics on-the-fly, yet there were still issues to fix and in particular it was not possible to include custom images by the time this manuscript was written.

On a more technical point of view, the annotation tool was coded in Python and Flask. Flask is a web application framework composed of libraries and modules designed to create web applications, services and APIs run in Python. This Python module respects the Web Server Gateway Interface (WSGI) norm that enables web servers (HTTP servers) to launch and run websites coded in Python. Therefore, Flask makes the link between the web interface presented to the user and the underlying Python code, and it can execute scripts such as reading/writing from a database, record the answer given in a form, perform computations on-the-fly, display interactive plots...

The annotation tool can be run off-line, i.e. locally on your own computer. Yet, it had to be accessible on the Internet. To that end, I had to put it online through a real HTTP server. This required three steps. The first was a dedicated web IP address, that could be queried to access the website: http://megacam.astro.unistra.fr. Then, we needed a web server to access the website: Apache web server was used. It deals with HTTP requests and HTML but cannot interact directly with the Python code. Finally, a WSGI server, in our case Gunicorn, was needed to make the link between Apache and the Flask application. The overall architecture of the server is presented in Figure 2.2.

Once the server was online and running, the next step was to ensure that the server had access to the images in HiPS format. The images could not be stored on the same server than the annotation tool, as the disc space required was too important. Indeed, each HiPS image approximately requires 400MB, and we had hundreds of images in different scalings. The images were therefore stored on a dedicated storage server, namely *space.astro.unistra.fr*, which is not accessible from outside the Strasbourg Observatory network. This storage server was 'mounted' on the *megacam* server so the HiPS would be accessible in the annotation server.

[7]Aladin Lite, https://aladin.u-strasbg.fr/AladinLite/
[8]SIMBAD, http://simbad.u-strasbg.fr/
[9]Create custom HiPS, https://aladin.u-strasbg.fr/hips/HipsIn10Steps.gml

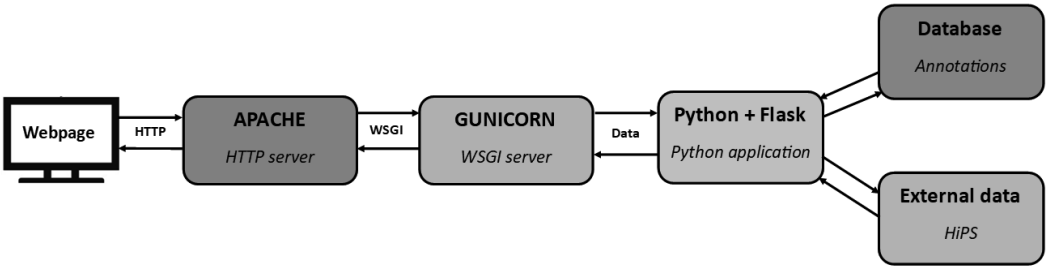

Figure 2.2: Illustration of the organisation of the annotation server.

Finally, the server is linked to a PostgreSQL database that is used to store the annotations. PostgreSQL is an open source object-relational database system which is compatible with SQL. The architecture of the database and the type of information stored in it will be described in Section 2.2.2.3.

2.2.2.2 Annotation process

Once the server was online and running, with the deep images in HiPS format created, the annotation process could take place. In this paragraph we detail the different steps to follow to annotate a galaxy. It starts with the user registering into the server, which sends an email to the server administrator who will create the account. Users need to be logged in to be identified and to record their annotations.

The first thing to do is to select on which project to work, and then to choose which galaxy to annotate. As shown in Figure 2.3, the user is presented with a list of the galaxies that have not been annotated yet, and another list detailing the annotations already performed and saved, but that can still be accessed to be modified. The list provides basic information about the galaxy, a small thumbnail image, its name, morphological type, survey and band from which the image was taken, a link to annotate the galaxy on this server and an external link redirecting to MATLAS webserver to classify this galaxy on that server. The goal of having both links is to have consistent classifications between the servers. Once the galaxy was annotated, another information appears, i.e. the date at which it was classified. The link now redirects to a page that enables the modification of the annotations. These lists are updated in real time: a script searches in the database for the already and not-already classified galaxies by this user, and puts them in one of the two lists.

Once the user selected the galaxy to classify, the link is redirecting to the main interface of the server: the navigation/drawing screen based on Aladin Lite. An image centred on the galaxy is shown, with the possibility to navigate, zoom in or out, in order to have a view of the local environment of the galaxy. It is possible to switch between different layers, either from already existing surveys or from custom HiPS with different intensity scaling. Additional information such as the coordinates of the mouse pointer or catalogues from the SIMBAD database can be shown.

The next step is to perform the annotation using the drawing tools, as displayed in Figure 2.4. The deep image with a given scaling is presented, and above it users are asked to delineate the contours of the features they see. In order to provide the most precise delineation, several drawing tool shapes (label 1) are available: *Rectangles*, *Circles*, *Ellipses*, *Curved lines*, and three different types of *Polygons*. The *Curved lines* are cubic Bezier parametric curves defined by four control points. In addition, shortcuts enable to delete an annotation, undo/redo an action, show a tutorial on how to draw the

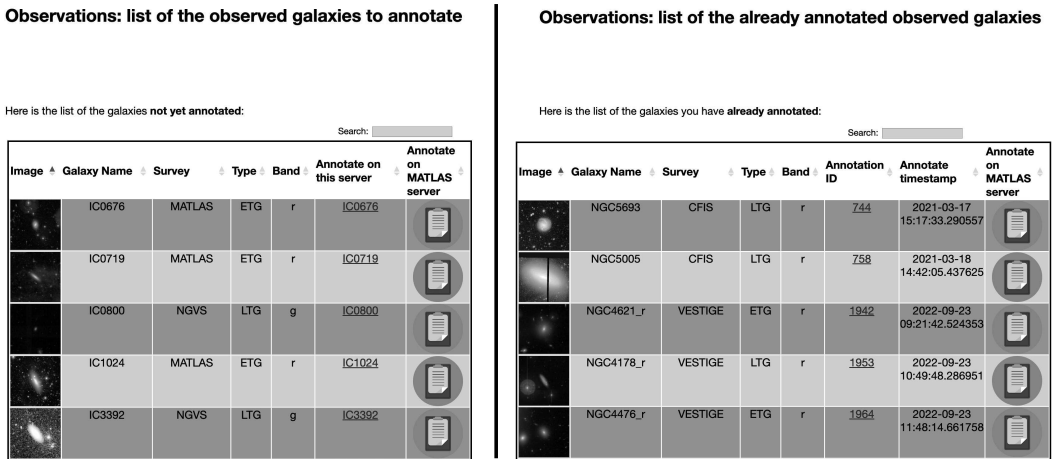

Figure 2.3: *Left*: Example of the list of galaxies that were not classified yet by the user. *Right*: Example of the list of galaxies already classified.

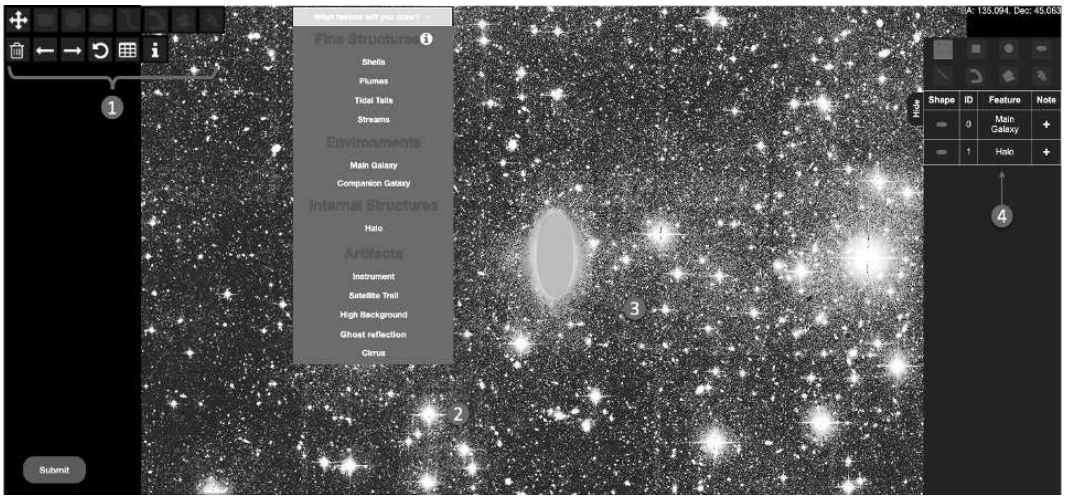

Figure 2.4: Adapted from Sola et al. (2022). Interface of the annotation tool, with its main facilities: drawing buttons (label 1), classification menu (label 2), examples of already drawn annotations (label 3) and summary table (label 4). The background is a deep image in asinh-scaled image.

annotations, or display a summary table (label 4) which contains all the drawn shapes. The semi-transparent red shapes (label 3) can be adjusted through their control points to be fitted as best as possible to the underlying feature. Then, a label defining the type of the annotated feature is chosen from a drop-drown menu (label 2). The shapes and labels can be selected and modified from the summary table.

About the features to delineate they are of several types, both stellar components and sources of pollution. The most important ones are the LSB tidal features and stellar halo, inner region, companion galaxies but also sources of contamination such as galactic cirrus clouds or artefacts. It is important to annotate pollutants as they make the identification of tidal structures more complicated and might bias the results. Hence, the number of annotated shapes on a single image can vary between a few and several. Sometimes, the field can be crowded with drawn shapes, which makes complicated the annotations of new features. A tick-box in the Summary table enables to hide some of the already drawn features, making easier the rest of the process. We define here more precisely what are

the types of features of interest. For tidal features, we follow the definitions established in previous papers of which this work is the continuation, namely Duc et al. (2015); Bílek et al. (2020).

- **Main Galaxy**: it corresponds to the inner region of the galaxy, as seen in shallow images (such as PanSTARRS DR1 colour image). It is the only feature which is not delineated on deep images. It encloses the most luminous features, such as spiral arms for LTGs. Its outer isophote is drawn with the *Ellipse* shape.

- **Halo**: it is the extended LSB stellar halo, only visible in deep images. Its outer isophote is generally elliptical and drawn with an *Ellipse*.

- **Tidal tails**: we define them as stellar material likely coming from the target galaxy, that originate from major mergers (e.g., Arp 1966; Toomre & Toomre 1972; Mihos 1995). Their shape should be relatively wide and elongated like antennae, but if they differ from it they are called *Plumes*. They are delineated with *Polygons*.

- **Streams**: they are stellar material which does not originate from the target galaxy; but either from a lower-mass galaxy progenitor or with no visible progenitor. Streams are generally thin and elongated, and likely trace minor mergers (e.g., Bullock & Johnston 2005; Belokurov et al. 2006; Martínez-Delgado et al. 2010). Like tails, they are drawn with *Polygons*.

- **Shells**: shells are arc-shaped features, often concentric and present in groups. They are likely formed during intermediate-mass radial mergers (e.g., Prieur 1990; Ebrova 2013; Duc et al. 2015; Pop et al. 2018). The shapes of shells were drawn with *Curved lines*, i.e. one pixel-wide lines.

- **Companion**: it is defined as a massive nearby galaxy, whose velocity is close to that of the target galaxy. We consider a threshold of around 200 $\mathrm{km\,s^{-1}}$ to be sure that the two galaxies would actually be involved in an interaction that could trigger the creation of tidal features. The outer isophote is delineated using *Ellipses*.

- **High background**: regions of the image where the background level is higher than the blank sky value. This is not an effect of defective flat fielding, but it is really due to higher emission regions, for instance due to dust illuminated by bright sources. They are delineated using *Polygons* as their shape is irregular.

- **Cirrus**: diffuse dust cloud from our own Galaxy, with filamentary structures, that scatter optical light. It is drawn with *Polygons*.

- **Ghost reflections**: artefacts due to internal reflections in the camera and detector of the light from bright sources. This creates artificial and extended round haloes around bright stars. They are delineated using *Circles*, i.e. pixel-wide circles (unlike *Ellipses* that are filled).

- **Instrument**: any artefact from the instrument itself, such as CCD gap. They can be delineated using *Rectangles* or *Polygons*.

- **Satellite trails**: the trails from satellites passing in the field of view. They are drawn with *Rectangles*.

It is important here to explain in more detail the distinction made between tidal tails and streams. The origin of both features is the same, but they differ with respect to the point of view. We always consider the point of view of the massive target galaxy. If the material seems to originate from the target galaxy, it is called a tail. Other indications were taken into account, such as the fact that tidal tails generally appear by pair: one tail directed towards the companion and a counter-tail at the opposite side of the galaxy. On the other hand, if the material seems to originate from either a lower-mass galaxy or from no visible companion, it is called a stream. In addition, contrary to tails, streams are not necessarily aligned with the massive galaxy's major axis and tend to wrap around that galaxy. The presence of an S-shape in the structure, around a condensation that may still be visible, also helps disentangle a stream from a tail.

Each annotation is linked to a single galaxy, which is the target galaxy that appears in the centre of the images. It means that if tidal features belonging to a remote background galaxy are annotated, they will still appear in the target galaxy's annotation. They can only be removed or not taken into account after computations during the data analysis. This is the reason why we asked users to only delineate tidal features associated to the target galaxy, with the exception of tails produced in on-going mergers between the target galaxy and its companion.

There is the possibility to write a few words, a note, about each feature if desired. Once users have delineated all features in the images, they click on a 'Submit' button and the annotations are stored in a database. It is later possible to review and modify an annotation by selecting it from the list of already classified galaxies.

2.2.2.3 Database

There are several pieces of information stored in the PostgreSQL database after submitting an annotation: the target galaxy around which the annotation was done, the name of the user, the type of the drawn shapes, the associated labels and notes and most of all the coordinates of the contours of the shapes. The coordinates are given both in pixel and world (right ascension (RA) and declination (DEC)) coordinates using the Aladin Lite function *pix2world*[10].

The contours of the drawn shapes will depend on their types. *Polygons* are the easiest as the control points are simply evenly-spaced points on the contours of the shape. For *Curved lines*, the four control points of the Bezier curve are stored. *Ellipses* are actually drawn using two Bezier curves plotted inside a rectangular bounding box: the coordinates of the bottom-left, bottom-centre, bottom-right, top-right, top-centre and top-left corners of the bounding box are saved, as well as its inclination angle. *Rectangles* are defined by their rectangular bounding box, defined in the same way than for *Ellipses*. *Circles* are defined by the bounding box with the circle inscribed in it, so four control points are saved (bottom-left, bottom-right, top-left, top-right).

Each drawn shape has a unique identifier *s_id*, and is stored in the *Shape* table, along with the *a_id* that corresponds to a unique annotation entry in the *Annotation* table. For each annotation, the galaxy and user identifiers are indicated, which enables to link it with the *Galaxy* and *User* tables. Figure 2.5 is a simple representation of the different tables in the database and the primary key values that link them together. In the *Galaxy* table are indicated the name, morphological type, survey or band of the target galaxy. The *User* table summarizes the username, email and access status (i.e. 'advanced', which

[10]Aladin Lite pix2wolrd, https://aladin.u-strasbg.fr/AladinLite/doc/API/

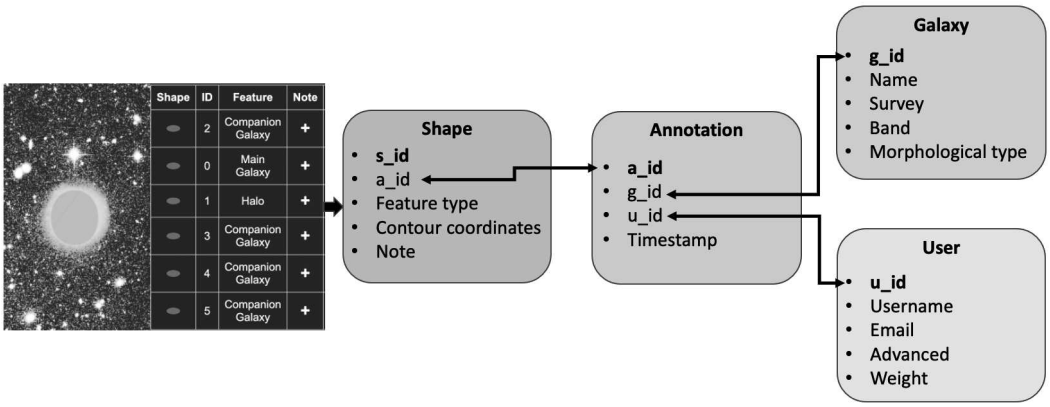

Figure 2.5: Illustration of the tables in the PostgreSQL database. Each drawn shape has its unique identifier s_id, stored in the table *Shape* along its coordinates, type and corresponding unique annotation identifier a_id from the *Annotation* table. From *Annotation*, there are links to the *Galaxy* and *User* tables with the values g_id and u_id.

grants different access rights to the server, from simple user to administrator). Each user can also be attributed a weight that corresponds to the level of expertise in the given classification task.

2.2.2.4 Multiple users

An important point to remind here is the fact that several users performed the annotations. This triggered several issues. The first one is the problem of combining the resulting annotations for each galaxy: this is addressed in Section 2.3.7. The second is the need to have expert users classifying our galaxies.

Indeed, we are not only asking collaborators whether a tidal feature is present, but rather to precisely draw its position and determine its type. It requires a previous knowledge of the field and already trained eyes. This has two main consequences. The first one is the fact that we need to train the users before letting them annotate. I started to train my eyes using the MATLAS website on which I indicated the presence and prominence of tidal features around ETGs. Only after that was I ready to classify galaxies with our annotation tool. I also contributed to the fine structure census in simulated images for the LSST collaboration whose results are presented in Martin et al. (2022). The second consequence is the fact that we cannot call for volunteer citizen scientists. We have to rely on either professional astronomers or on undergraduate students that would be trained. In addition, the annotation process is very time consuming: it took an expert user around 10-15 minutes for the simpler cases, and an even longer time was necessary for interns. It is therefore not possible to ask many astronomers to dedicate such a large part of their time to the delineation, and we are left with a few users only. However, the low number of contributors is compensated by their high level of expertise which makes the results more reliable.

In this work, I annotated all the galaxies. The annotations were also partly done by Dr Pierre-Alain Duc and three Master 1 students, Mathias Urbano, Julie Klehammer and Pierre Galois that I co-supervised during their internship. It quickly appeared that a great deal of discussion and exchange with interns was needed to explain what tidal features and contaminants looked like but more importantly to make sure that we annotated in a consistent manner.

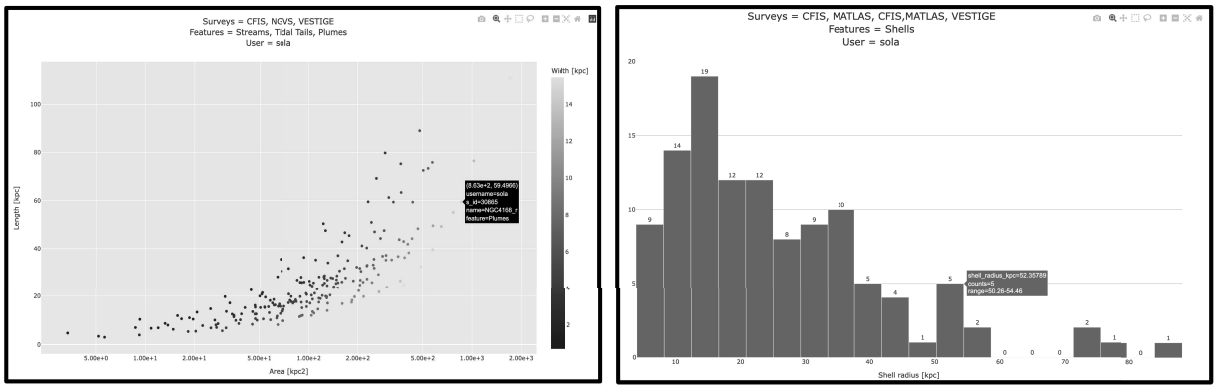

Figure 2.6: Examples of two interactive plots on the annotation server. Hovering the mouse over a point or histogram bin reveals additional information about it. *Left*: Length as a function of the area of tidal tails and streams, for a given user and surveys. *Right*: Histogram of the radii of shells for a given user and surveys.

2.2.2.5 Improvements

The annotation server was maintained and continuously improved during my PhD thesis. Some improvements were more aesthetic and other had more useful applications. I started by adjusting the CSS style to match MATLAS website in order to have a consistent graphic style. The automated send of emails to the administrators when an error occurred or when a new user wanted to register was added. I simplified the registration process. I created administration pages: one giving access to all the annotations from all users; another enabling the easy modification of existing users (access rights, reset passwords) or adding new ones; another page to create an export of the entire database to csv format and one to create and download the masks of the annotations (in FITS format or numpy array).

In addition, I wrote scripts to automatically perform on-the-fly computations on each drawn shape when the user clicks on the 'Submit' button, such as area, length or width of the structures. These computations were stored in another new table in the database. The next step was to display the results of this table: interactive plots (made with plotly[11]) were integrated in the server. It enables users to quickly see a few results based on their own annotations, and spot potential outliers that would need to be visually checked. Figure 2.6 shows two examples of interactive plots. Similarly, a 'Thumbnails' page was created with the goal of giving a fast way to visualise the contours of tidal features to have an overview of the shapes and to detect potential issues. This was motivated by the fact that some galaxies are crowded by annotations so it is complicated to clearly distinguish them. For a specific project, I added a page to display the histogram of the radii of shells.

For some of these tasks, I was helped by Olivier Marchal, engineer at the Observatory of Strasbourg, who worked with me on the server. In particular, he integrated the interactive plots, the administrator page to manage users and he simplified the registration process. One important point he developed was the possibility to hide some drawn shapes during the annotation process. It was not possible before and made the classification much more complicated when the field was already crowded with annotated shapes.

[11]Plotly Python Graphing Library, https://plotly.com/python/

2.3 Quantitative measurements

From the annotation database, we have access to the type and coordinates of the contours of LSB structures from which the goal is to obtain quantitative measurements. To that end, I computed several quantities, among which geometry, surface brightness, luminosity, colour or mass. This section details how these properties were derived. The results of these analyses will be presented in Chapter 5.

2.3.1 Morphology

2.3.1.1 Area

In order to characterize LSB structures, a first point is to investigate their morphology, in terms of area, length and overall shape. As mentioned in Section 2.2.2.2, there are several drawing shapes available to delineate fine structures: *Ellipse*, *Circle*, *Rectangle*, *Polygon*, and *Curved line*. Hence, each shape requires its own computations. They are all based on the RA,DEC coordinates of their contours. The distances between the control points were obtained using the on-sky separation[12].

We start by focusing on the area. For *Ellipses*, the length of the major and minor axis were computed from the coordinates of the rectangular bounding box, and then the area was derived. To compute areas of *Polygons*, we used the shoelace formula which is given by

$$A_{\text{polygon}} = \frac{1}{2}\left|\left(\sum_{i=1}^{n-1} x_i y_{i+1}\right) + x_n y_1 - \left(\sum_{i=1}^{n-1} x_{i+1} y_i\right) - x_1 y_n\right|$$

where A_{polygon} is the area, n is the number of sides of the polygon and $(x_i, y_i), i = 1, ..., n$ are the ordered planar coordinates of the vertices of the polygon. We had to use the sinusoidal projection to convert RA,DEC coordinates to Cartesian ones, obtained by multiplying the latitude (DEC) by the length of one degree of latitude, and the longitude (RA) by the length of a degree of latitude and the cosine of the latitude.

In order to save computation time, we did not compute the areas of *Rectangles* and *Circles*. The reason is that no stellar structure was drawn using these shapes, as they are used for contaminants (ghost reflections, satellite trail, high background regions...) for which we do not want to obtain their morphology.

2.3.1.2 Length and width

The next property to investigate is the projected longest length of LSB structures. For *Ellipses*, the longest length corresponds to the length of the major axis; for *Circles* it is the diameter; for *Rectangles* and *Curved lines* it is the length.

There is however no clear definition of the longest length of a *Polygon* whose shape can be irregular, with varying number of edges and vertices. We chose to use the medial axis as the longest length: it is the set of points that are equidistant to the contour of the polygon. The medial axis was obtained through the process of skeletonization using scikit-image (van der Walt et al. 2014): the image is eroded time after time until there is only one pixel-wide line left that represents the morphological skeleton of the polygon. The skeletons can have several branches but we are only interested in keeping the longest one. Therefore, we used the FilFinder package (Koch & Rosolowsky 2015) to explore the

[12] Astropy SkyCoord, https://docs.astropy.org/en/stable/api/astropy.coordinates.SkyCoord.html

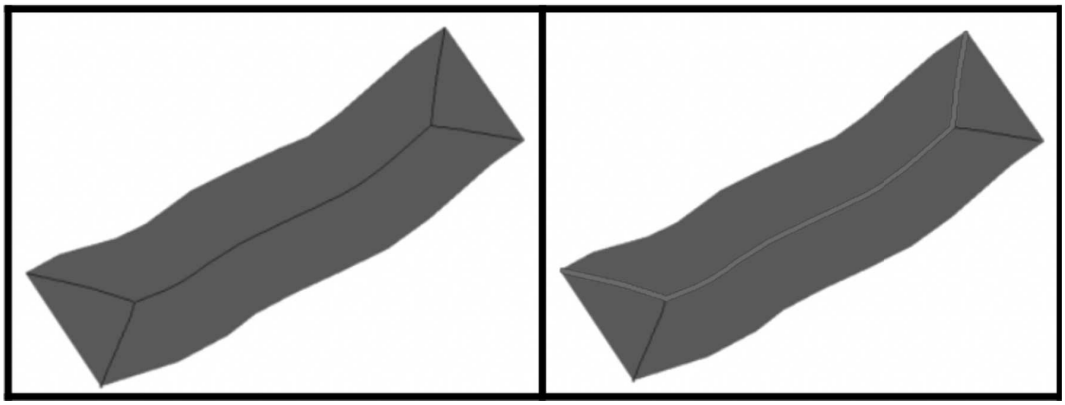

Figure 2.7: Adapted from Sola et al. (2022). Example of the longest length in a polygon. *Left*: the medial axis (black line) is obtained through skeletonization: it has several branches. *Right*: FilFinder is applied to retrieved the longest branch (red line).

branches and only record the longest one. An illustration of these steps is displayed in Figure 2.7.

This definition holds as long as the polygon is elongated. In the limit case where the polygon is round, the medial axis is reduced to a single point which is the centre of the circle. After inspecting the annotations, it appeared that a fraction had rather roundish or flattened shapes, which would lead to underestimate their longest length if skeletonization was used. We defined how round polygons are by their isoperimetric ratio q:

$$q = 4\pi \frac{\text{area}}{\text{perimeter}^2} \tag{2.1}$$

The rounder the polygon, the closer q is to 1. We defined a threshold of q above which the polygon is considered too round for our skeleton method to be incorrect. After some measurements, we chose a threshold of $q = 0.8$. For polygons with $q < 0.8$, the medial axis was computed, but for $q \geq 0.8$ we defined another method. Namely, we approximated the polygon by the inscribed ellipse with the maximum volume; and we took the longest length as the length of the major axis of the ellipse.

Then, from the area and length of the structures, we derived the width by dividing the area by the longest length. This relies on the assumption that polygons have ribbon or rectangular shapes. For round polygons, the width is taken as the minor axis of the ellipse approximating the polygons.

Finally, in order to have an idea of the general shape of polygonal annotations, whether they are straight or more curved, we computed their overall bending. This will be important when comparing tidal tails and streams as it will be discussed in Chapter 5. To that end, we estimated the curvature of polygons by fitting their topological skeleton by a linear function using a least-square regression. We then computed the R^2 correlation coefficient: the higher R^2, the closer the polygon is to a straight line. For round polygons, we could not apply this method as the topological skeleton is not representative of the longest length, but their isoperimetric ratio provides information on how close to a disc the polygon is.

2.3.1.3 Radii of shells

Shells are defined using one-pixel wide *Curved lines*, hence their area is not relevant. In complement to their length, an interesting quantity to compute is their radii, i.e.

their distance to the target galaxy. Indeed, there are several types of shells defined by their position and orientation with respect to the target galaxy (Wilkinson et al. 1987; Prieur 1990). Type I shells are concentric and centred on the galaxy; Type II shells are circular and randomly distributed around the galaxy while type III are irregular and not concentric, either because their shape is not typical or because there are too few shells. Here, we make the assumption that the *Curved lines* defining shells are circular arcs, and we take the radii as the projected distance between the centre of the galaxy and the point located in the middle of the *Curved line*.

2.3.2 Masks

From the coordinates of the annotations, it is possible to create masks of each individual shape. The RA,DEC coordinates of the contours are converted to pixel coordinates of the FITS image in which the annotation is located, using functions from Astropy World Coordinate System (WCS)[13]. Then, the interior of that boundary in pixel coordinates is applied to the input image, while all pixels outside are set to empty (NaN) values.

From it, we had large images with blank values almost everywhere except where the interesting annotation was. To reduce the sizes, we created cutouts of the masks: we cropped them in rectangular boxes around the masked annotation. Therefore, in the end we had as many masked cutouts as the number of annotated LSB features.

The advantages of these masks is that they can be used to retrieve surface brightness, colour or flux measurements very easily and in an automated way for thousands of annotations. The following Sections describe how these quantities are retrieved.

2.3.3 Automated aperture photometry

From the previously created masks, it is easy to retrieve flux and SB measurements. Automated aperture photometry is therefore possible on a large number of annotations with our masks, unlike more traditional methods were apertures have to be manually placed along each structure (preventing its use on large datasets). Yet, in these masks are present contamination sources such as bright sources or other overlapping LSB structures, which contribute to overestimate the flux and SB values. A first approach was to retrieve the median values to retrieve the most represented value and mitigate against extreme pixel values from bright stars for example. This approach has the advantage delivering very fast results from our masks. A more sophisticated but accurate approach is to use another way to retrieve these values, at the cost of more computational time. Briefly, we perform source-extraction in our cutout masks to remove bright sources and then re-estimate the value of the background. The different steps of this new method are detailed below for each annotation of each user, and an illustration of the process is shown in Figure 2.8.

Create cutouts - For each feature of interest, the process starts by the creation of two cutouts: one mask displaying only the feature (i.e. blank values outside it) and one of the feature and its surroundings (i.e. a cropped version of the original image). Then, we create the mask cutouts of other annotations made by the same user for that galaxy: halo, companion and dwarf galaxies and ghost reflections. These are the most important ones, as they may be overlapping with the feature of interest and their contribution is non-negligible. For the case where haloes are the feature of interest, we only want to conserve

[13]Astropy WCS, https://docs.astropy.org/en/stable/wcs/index.html

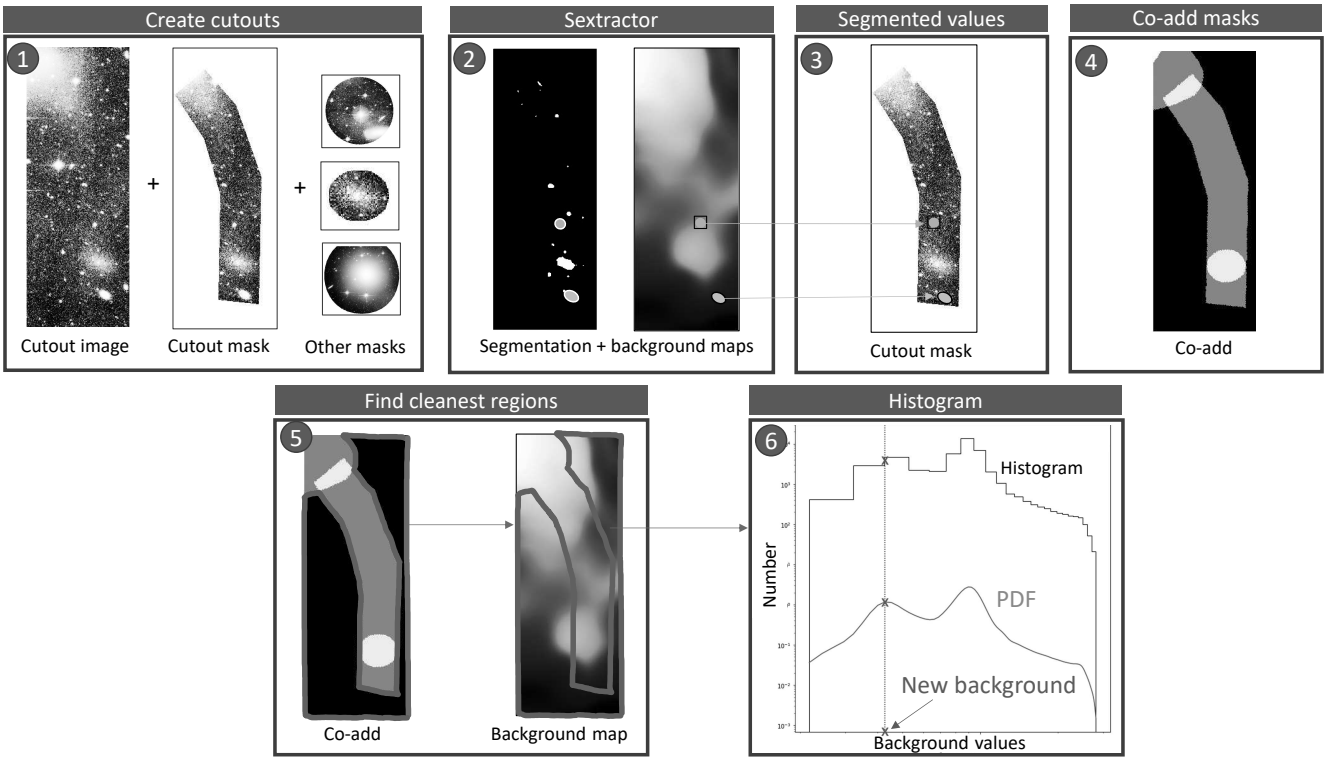

Figure 2.8: Adapted from Sola et al. (2023). Illustration of the refined luminosity estimation process, for one feature around one galaxy. *Step 1*: the cutouts of the mask and image of the feature of interest are created, as well as masks of other annotations around the galaxy. *Step 2*: source extraction and background measurement are performed using SExtractor. *Step 3*: segmented pixels are replaced by the corresponding background map values. Here are represented in green and orange two examples of such segmented regions. *Step 4*: the masks of the other annotations are co-added on the same field of view as the feature of interest. *Step 5*: the cleanest regions, with the least number of annotations overlapping, are retrieved from the co-add and applied to the background map. Here the cleanest regions are represented inside the red contour. *Step 6*: inside the cleanest regions, the histogram of the background map values is created (in black). Then, the PDF is estimated (in green), and the peaks of the PDF are computed. Only the first significant peak is kept (in blue), and its value is taken as the new background value.

the part corresponding to the LSB outskirt and not the centre of the galaxy. Hence, our halo masks contain non-blank values in the 'donut' region between the exterior boundary of the main galaxy and the external border of the halo. This first step is illustrated on the top left panel of Figure 2.8.

SExtractor - The second step (step 2, top) is to perform source extraction to remove bright sources. To that end, the SExtractor (Bertin & Arnouts 1996) software was run in double mode: the detection and segmentation was performed on the cutout mask of the feature while the measurement of the background was conducted on the cutout image. This yielded two outputs: the segmentation and the background maps.

To be more precise, SExtractor is a software that detects, deblends, measures and classifies sources from astronomical images. Its process for each image consists of six steps: estimation of the background, thresholding, deblending, filtering of the detections, photometry and star/galaxy separation (Bertin & Arnouts 1996). We are only interested

in two main points: the background estimation and detection. The background map corresponds to the background level of the image if no object was present in it. It is necessary to have a precise estimation of this background level in order to detect the faintest sources and to measure their flux accurately. It is measured in small meshes applied everywhere in the image and then combined through a bilinear interpolation. The size of the meshes is important: if it is too small, the background estimation will be perturbed by the presence of objects and random noise (and in particular the diffuse light from extended objects will be considered as being in the background), but if it is too large it will not be able to capture the local variations of the background. In our case, we tested different values and chose a mesh size of 32×32 with the parameters BACK_SIZE= 32 and BACK_PHOTOTYPE=Local (to recompute the background in a rectangular annulus around the object in order to have a more precise estimate). Then, all the pixels with a value higher than a given threshold are considered to be part of objects, while the others are part of the background. The threshold value is important: too high and it will only detect the brightest objects, too small and it will detect almost any source in the image. We had to adapt the threshold to our features of interest, as haloes are much brighter than tidal tails and streams. We ran some tests and it appeared that a threshold value of DETECT_THRESH=2 for tails and streams and of DETECT_THRESH=12 for haloes enabled to recover our features relatively well. The next step consists in separating all the detected pixels into individual objects (then photometry and other measurements (size, ellipticity...) are performed, and they are classified as star or galaxy by a neural network). The segmentation and background maps were the outputs useful for us, but we did not use their photometry, star/galaxy separation nor other measurements.

Segmented values - After these explanations on how SExtractor works, we now focus on the third step of our process. The pixels removed (masked) by the segmentation are replaced by the corresponding values from the background map and then applied to the mask cutout (step 3). Afterwards, there may still be erroneous pixels in the masks, coming from instrumental artefacts such as CCD gaps. These pixels do not have a physical signification and they should be removed so that they do not contribute to the flux of the feature of interest. After inspection of such pixels their values are largely negative, so all the pixels with a value smaller than a given threshold are replaced by the corresponding pixels of the background map. That threshold is arbitrary and was chosen to be -10 after the inspection of the histograms of some images with erroneous pixels.

At the end of this step, the cutout masks are considered cleaned. We could stop here and simply sum all the pixels to get the flux of the feature. Yet, a precise estimation of the background value is necessary when dealing with LSB structures. Our images are supposed to have a flat background equal to zero, after being processed by Elixir-LSB: this will be explained in Section 3.3. However, there may be local variations of that background value due to contamination sources like bright stars or cirrus: the background value is locally non-zero. Therefore, we added a phase of re-estimation of the background, which is divided in the following steps.

Co-add masks - We want to remove the influence of other structures that may be overlapping with the feature of interest. To that end, we use the other masks created during step 1 (halo, ghost reflection and companion galaxies) which may contribute the most to the contamination. They are co-added and reprojected (using the *reproject* Python package[14]) to the same field of view as the cutout mask of the feature of interest (step 4). The co-add only consists of the footprint of the annotations so the pixel values range from

[14] *reproject* Python package, https://pypi.org/project/reproject/

0 (no annotation) to n where n is the number of annotations overlapping in this pixel.

Find cleanest regions - The co-add is applied on SExtractor's background map. The goal is to find the cleanest regions in which the background will be estimated. We define as cleanest a region where there are the least number of annotations superimposed (i.e. 0 in the best case, otherwise the lowest value of the co-add).

Refined background estimation - The new background value is evaluated on these cleanest regions of the background map. The method to retrieve it was established after several inconclusive tests. A first attempt was to take the minimal (or 5% or 10%) pixel value but it is not robust as there can be a few low value pixels but a majority of high values. The second try was to take the mean, median, mode or even sigma-clipped mean of the pixel distribution. This does not work in the case where the distribution is multimodal which seemed to occur frequently. The third attempt was to work not on the histogram of the pixels but rather on its probability distribution function (PDF) obtained by kernel density estimation. The advantage of having the PDF is that representation is continuous and does not depend on the number of bins of the histogram. The next question was to determine which value of the PDF to consider. We tried to keep the maximum peak but it was not satisfying, so we rather took the value of the first prominent peak. A peak is considered significant enough if the computation of its prominence is higher than a given threshold. This condition enabled us to remove the effect of multiple small peaks and mostly to deal with multimodal distributions.

To summarize, we create the PDF of the histogram of the pixel values in SExtractor's background map in the cleanest regions. Then, we consider as the new background the value of the first prominent peak of the PDF (step 6).

Finally, the last step is to subtract this new background value to all the pixels of the cleaned cutout mask of the feature of interest: the final mask is obtained. Unless explicitly stated otherwise, the results presented in Chapter 5 are based on these background-corrected final masks.

From such background-corrected final mask, the total flux f is simply obtained by summing all the pixels of the mask, and we can compute the luminosity L by:

$$f = \sum_{i=1}^{N} p_i \tag{2.2}$$

$$L = 4\pi \times D^2 \times f \tag{2.3}$$

where D is the distance of the galaxy, f the flux, p_i the value of the i^{th} non-blank pixel and N the number of non-blank pixels in the mask. Likewise, the SB and colours can be computed from these final masks.

2.3.4 Surface brightness

The general formula to convert an image from linear scaling to surface brightness is the following:

$$\mu = -2.5 \log_{10}\left(\frac{p - bk}{\text{pixsize}^2}\right) + zp \tag{2.4}$$

where μ is the SB value in mag arcsec^{-2}; p is the pixel value in the linear-scaled image; bk is the background value, pixsize is the size of the pixel in arcsecond and zp is the value of the photometric zeropoint.

It is important to note that this formula depends on the size of the pixel. If the images are resized, as described in Section 3.3, this will trigger a modification of the zeropoint. For instance, if the image is downsized rebin_factor times, i.e. its pixels are binned together to a size of new_pixsize to form an image rebin_factor times smaller than the original one, the formula is modified as:

$$\mu = -2.5 \log_{10} \left(\frac{p - bk}{\text{new\_pixsize}^2} \right) + zp - 2.5 \log_{10}(\text{rebin\_factor}^2) \tag{2.5}$$

This warning also holds for the value of the flux: if each new pixel was obtained by averaging the rebin_factor × rebin_factor neighbouring pixels of the original images, the flux f must be computed from the new pixel values new_p as :

$$f = \left(\sum_{i=1}^{N} new\_p_i \right) \times \text{rebin\_factor}^2 \tag{2.6}$$

From our final masks, the background value bk is 0 as we had already subtracted it in the previous steps.

2.3.5 Colour measurements

If a galaxy was observed in at least two photometric bands, it is possible to derive its colours. For simplification purposes, we will take the examples of the r and $g-$bands as they are the ones used during the PhD work (see Chapter 3).

The first step is to create the background-corrected masks as described in Section 2.3.3, in both photometric bands. As the background values differ between the r and g bands, the segmentation and background maps, as well as the newly estimated background value, will be different. From them, the total fluxes flux_r and flux_g are retrieved. The integrated $g - r$ value (in mag) is obtained by:

$$g - r = 2.5 \log_{10} \left(\frac{\text{flux\_r}}{\text{flux\_g}} \right) \tag{2.7}$$

Another option is to create the $g - r$ colormap, and to compute the median $g - r$ from this colormap. The integrated $g - r$ is more precise than the median in the case where there is a colour gradient. To create the colormap, it is necessary to reproject one of the masks onto the other one in case the pixel sizes are different[15]. Then, the SB maps of the g and r-band reprojected files are created, and finally the colormap is obtained by subtracting the SB-r map from the SB-g map. The median value of the colormap is computed afterwards.

2.3.6 Stellar mass estimation

From the previous Section, we were able to derive the flux and colour of each individual feature of interest from the background-corrected masks of the annotations. With these two pieces of information, it is possible to compute the stellar mass in each annotation. Here, we take the example of the stellar mass computed from the r-band masks.

[15] During the reprojection the flux must be conserved, for instance using the *reproject_exact* function from the *reproject* package, https://reproject.readthedocs.io/en/stable/api/reproject.reproject_exact.html

Once the flux in the r-band f_r is retrieved, it can be converted to apparent magnitude m_r, extinction-corrected apparent magnitude $m_{0,\mathrm{r}}$ and absolute (extinction-corrected) M_r magnitudes using:

$$m_r = -2.5 \times log_{10}(f_r) + zp \tag{2.8}$$
$$m_{0,\mathrm{r}} = m_r - A_r \tag{2.9}$$
$$M_r = -5 \times log_{10}(D) + 5 - m_{0,\mathrm{r}} \tag{2.10}$$
$$M_r = -5 \times log_{10}(D) + 5 - A_r + m_r \tag{2.11}$$
$$\tag{2.12}$$

where zp is the value of the photometric zeropoint; D is the distance of the galaxy in pc and A_r is the extinction in the r-band, considering only the extinction from the Milky Way (and not the intrinsic extinction).

The computation of A_r is not straightforward, as we first need the total extinction in the V-band A_V (taken from the NED database[16]), and afterwards we use Equations 1, 3.a and 3.b from Cardelli et al. (1989) for the optical/NIR. For the r MegaCam filter centred at $\lambda_r = 0.64$ µm (see Chapter 3), the relation obtained from the equations from Cardelli et al. (1989) is:

$$A_r = 0.84295 \times A_V \tag{2.13}$$

Once we have the absolute magnitude, the luminosity L_r can be computed from it and from the Sun absolute magnitude in r-band $M_{\odot,\mathrm{r}} = 4.64$ (Willmer 2018).

$$\frac{L_r}{L_\odot} = 10^{0.4(M_{\odot,\mathrm{r}} - M_r)} \tag{2.14}$$

The next step is to compute the mass-to-light ratio γ from the $g-r$ colour following Table A.7 from Bell et al. (2003) (who constructed stellar mass-to-light ratio estimates using galaxy evolution model fits to SDSS $ugriz$ and 2MASS K-band fluxes, assuming a universally-applicable stellar initial mass function).

$$\gamma = 10^{-0.306 + (g-r) \times 1.097} \tag{2.15}$$

Finally, from the mass-to-light ratio and the luminosity, the stellar mass $M_\star$ can be computed:

$$M_\star = \gamma \times L_r \tag{2.16}$$

Therefore, the stellar mass of all LSB features available in r and g-bands can be computed directly from their flux in the background-corrected masks. It must be noted however that the uncertainty on the colour can be relatively high, which results in large uncertainties in the stellar mass estimation (the uncertainties on the mass can range from a few to several tens of percent, and even 100% for the least reliable measurements).

2.3.7 Annotations kept

2.3.7.1 Progenitor of tails

As described in Section 2.2.2.2, our annotation tool enables users to delineate LSB features in deep images centred around target galaxies. All the drawn shapes by one user in an

[16]NED, https://ned.ipac.caltech.edu/

image are attached to the corresponding target galaxy, even if other galaxies are present in the image. For haloes and main galaxy annotations they are by definition tied to the target galaxies. It is however more complicated for tidal features, and in particular tidal tails. In the case where there is a nearby massive companion interacting with the target galaxy, tidal tails can be formed around both galaxies and will be delineated and attached to the target galaxy. Yet, we want to distinguish tails as a function of their progenitor (target or companion galaxy). To that end, we developed a method to attribute a progenitor to each tail, depending on one of the four possible cases listed below.

In the first case, the tail is only overlapping with the halo of the target galaxy (but not with the companion), so the progenitor is the target galaxy. The second case is the opposite, when the tail only overlaps with the companion, which is then the progenitor. The third case occurs when the tail overlaps neither with the target or companion galaxy. Then, the distances between the centre of each galaxy and the position of the centre of mass of the tail are computed: the progenitor is the galaxy with the shortest distance. If both distances are very close and their difference smaller than an arbitrary threshold, the progenitor is set to *Unsure*. Finally, the fourth case occurs when the tail is overlapping with both the target and companion galaxies. Then the area of the intersection of the tail with each galaxy is computed, and the galaxy with the largest intersection is kept as the progenitor. If both intersection areas are very close and their difference is smaller than an arbitrary threshold, the progenitor is set to *Unsure*.

With this attribution process, we are able to conserve only the tails originating from the target galaxy. In order to count each tail once, in this work we only conserve the tails from the target galaxy. For streams, by definition they originate from lower-mass companions hence they are not linked to the target galaxy, so we kept all the annotations of streams.

2.3.7.2 Keeping the most representative annotations

The goal of our annotation tool is to collect delineations of LSB structures around galaxies made by several users. This means that for each galaxy, we have several annotations of the same feature type, at least for the halo and main galaxy. This is even more complicated for tidal features, as the label can differ between users (tails or streams); there may not only be one single stream or tail per galaxy and the delineations might differ depending on the user. Yet, our goal was to keep only the most representative annotation per galaxy, which means we had to develop a selection process gathering the annotations of all users. We based this selection on the basic assumption that most expert users tend to see fainter hence more extended features (Bílek et al. 2020). The aim is to have per galaxy one halo, one main galaxy and the most representative tidal features. The selection process depends on the type of the structure.

For haloes, we kept the annotation with the largest area. For the brightest parts (i.e. main galaxies), as they were delineated on shallow images, the extent is not relevant so we kept the annotation whose area was the closest to the median area. For tails and streams, the process is more delicate and is iterative: it is illustrated in Figure 2.9.

We start by considering tails and streams together (in order to remove the possible confusion in the labels for ambiguous cases) as seen in panel *a)*. Then we create pairs of features for the two first users (panel *b)*. The area of each individual feature and of their intersection region is computed, and an intersection score is defined as the area of the intersection divided by the area of the larger feature. If the percentage of intersection is higher than a given threshold (chosen as 25%), the two features are paired, hence they

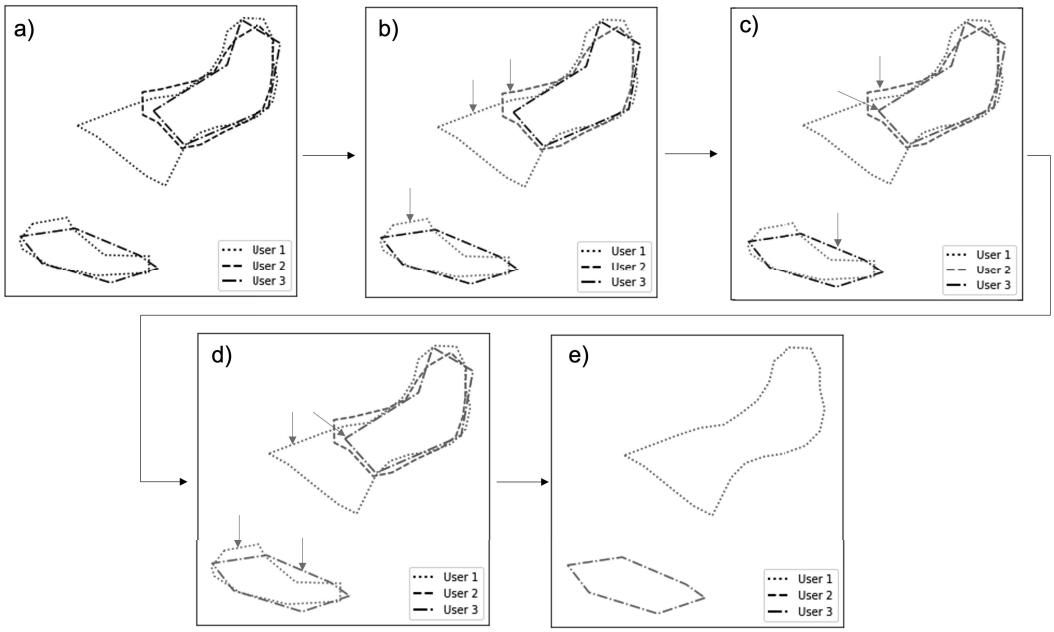

Figure 2.9: Adapted from Sola et al. (2022). Illustration of the selection process for tidal tails and streams for one galaxy. Black annotations indicate that the unique identifier has not been attributed yet, while each colour corresponds to one unique identifier. The linestyles correspond to the different users. The red arrows indicate which annotations are currently compared. *a)* Initialization: all the annotations of tidal tails and streams are considered. *b)* First iteration: the annotations of User 2 and User 1 are compared. *c)* Second iteration: comparison of the annotations of User 3 and User 2. *d)* Third iteration: comparison of the annotations of User 3 and User 1. *e)* Final step: for the annotations sharing the same unique identifier, only the one with the largest area is kept.

are considered as being the same feature. They are attributed a same unique identifier. Afterwards, we repeat the process with the annotations of the other users: the new features are paired with the previous ones, the percentage of intersection is computed and then the feature is either associated to an already existing unique identifier (if paired) or it is attributed a new one (if not paired with any other feature by any user). It is illustrated on the panels *a)* to *d)*. The final step consists for features sharing the same identifier to keep only the one with the largest area (this is similar as what we did for haloes). At the end of the process, each galaxy is left with tidal features having different unique identifiers, as visible in panel *e)*. In this work, we only take into account for the results the annotations that were kept after this selection process (unless explicitly mentioned otherwise).

2.3.8 Reliability of the annotations

2.3.8.1 Level of contamination

In addition to the tools mentioned in the previous Sections, it was necessary to have a way to quantify the amount of contamination in an image. Indeed, the detection of LSB features is made more complicated when sources of pollution overlap with them, such as bright stars, cirrus, high background regions or stellar components (either halo, companion galaxy or other tidal features). In such cases, annotations might be biased as users would be likely to underestimate the real size of the features. In order to quantify the contamination level, we created a reliability index based on the percentage of overlap

between the halo of a galaxy and pollutants. Some pollutants are more likely to perturb the annotations of LSB tidal features hence we ranked them: for instance, high background and ghost reflections have a higher impact than cirrus or satellite trails. The reliability index is defined as follows, ranging from 1 for the worst images to 9 for the cleanest:

- None: The annotation of the halo is impossible (for instance due to a high contamination by bright sources)
- 1: The entire halo is embedded in a ghost reflection or a high background region
- 2: The halo has an overlap with a ghost reflection or a high background region
- 3: The entire halo is embedded in a companion galaxy
- 4: The entire halo is embedded in a ghost reflection coming from the core of the galaxy
- 5: The halo has an overlap with a companion galaxy
- 6: The entire halo is embedded in a cirrus
- 7: The halo has an overlap with cirrus clouds
- 8: The halo has an overlap with a satellite trail
- 9: There is no pollutant overlapping with the halo

There can be several indices for the same galaxy. To keep a single value per halo, we computed a weighted average of the values. The indices [1,2,3,4,5,6,7,9] were associated to the weight [8,7,6,5,4,3,2,1] (satellite trails have such a low impact on the classification that they are not considered in the weighted index).

2.3.8.2 Similarity between annotations

The level of contamination is not the only way to determine whether an annotation is reliable. It is also interesting to determine whether several users annotated a feature in a similar way; and even test if the same user draws similar annotations of the same galaxy if it is presented in two images. The definition of the similarity index, ranging from 0 (no similarity) to 1 (total similarity), depends on the type of LSB structure.

For halo and main galaxy annotations: the Jaccard index (Hennig 2007) is used. It is defined as the area of the intersection of both structures divided by the area of their union. If an annotation is missing for one of the users the index is set to zero. For shells: since their number is relevant (and not their area as they are defined as one pixel-wide lines), the similarity index is defined as $\frac{\min(S_i,S_j)}{\max(S_i,S_j)}$, where S_i (respectively S_j) is the number of shells annotated by user_i (respectively user_j).

Then, for tails and streams: two metrics are used. The first one, like for shells, is an index computed over the number of features. However, the area covered is more relevant to characterize them (rather than the number of features, as they range from 0, 1, or 2 for tails and generally 0 or 1 for streams). Hence, a second metric was defined. It relies on the pairing of tidal tails and streams as defined in Section 2.3.7 with the unique identifier. The Jaccard index is computed on the structures annotated by user_i and user_j that share the same unique identifier.

Lastly, for contaminants (high background, cirrus): the union of all the contaminants of annotations of a given type is made, as sometimes a user will split a pollutant annotation into two. Then, the Jaccard index is computed between the unions of the contaminants.

2.4 Future developments

My applications of the annotation tool and database for the study of tidal features will be described in Chapter 5 and the future applications of the annotation server will be discussed in Chapter 7. In particular, in Section 7.2.2 we will focus on the applications of machine learning to automatically detect tidal features in deep images with algorithms trained on our annotation database. Here, we describe the future prospects of the annotation tool itself.

The annotation server relies on Aladin Lite version 2. Since the beginning of my PhD work, a new Aladin version has been released. Aladin Lite v3 is much faster than v2 and has new features such as improved display, new projections but mostly the possibility to modify on-the-fly the scaling and dynamics of the images. This is a major improvement in comparison to the current version: for now, if we want to display an image with different dynamics we have to create them beforehand, while with the version 3 this will not be needed. As discussed in Section 2.1, the two approaches are different and do not imply the same biases. This new Aladin Lite version 3 will soon be implemented with the help of Olivier Marchal, but it was not used during this work.

2.5 Conclusion

To summarize this Chapter, we focussed on the visual identification of LSB features in deep images. Considering the lack of tools able to retrieve quantitative measurements about these features, yet crucial to constrain models of galactic evolution, a new dedicated annotation tool was developed. It enables the precise delineation of LSB features, both stellar components and sources of contamination, by multiple users. The tool is web-based, user-friendly and the resulting annotations are stored in a database. During the PhD thesis, I put this server online, I maintained and improved it by adding new features. I developed analysis tools that performed computations on LSB structures from the database, amongst which geometry, surface brightness, flux, colour and stellar mass are important derived quantities. An important step of selection process to keep only the most representative annotations per galaxy from many users was defined. In addition, indices to quantify the reliability of the annotations were introduced. All these analysis tools will be applied to the data described in Chapters 3 and 4, and the results will be presented in Chapter 5. The future applications of the annotation server will be discussed in Chapter 7.

2.6 Résumé du chapitre

Dans ce chapitre, nous présentons l'outil d'annotation qui permet d'obtenir des mesures quantitatives sur les structures LSB. Nous commençons par présenter l'outil, puis le processus d'annotation lui-même et les analyses effectuées. Enfin, nous nous concentrons sur l'utilisation de notre base de données d'annotations en tant qu'échantillon d'entraînement pour des algorithmes d'apprentissage automatique (machine learning).

Nous commençons par constater le fait que l'étude des structures de marée repose aujourd'hui principalement sur une inspection visuelle des images profondes. Des méthodes plus automatiques existent, notamment le machine learning, et sont nécessaires quand l'échantillon de galaxies à classer devient trop important. Ça sera en particulier le cas des prochains grands relevés qui couvriront des milliers de degrés carrés dans le ciel avec une profondeur suffisante pour l'étude de l'Univers LSB. Cependant, les algorithmes de deep learning (DL) supervisés ont besoin de jeux de données d'entraînement précédemment créés manuellement. Il apparaît donc que la classification visuelle des structures LSB reste une solution privilégiée.

Or, afin de permettre cette étude, faite par de mutliples utilisateurs sur de grands nombres de galaxies, et ceci de manière simple et efficace, des outils dédiés sont nécessaires. En effet, même si des solutions plus simples existent (telles que l'ouverture des images avec des logiciels astronomiques dédiés et un tableur pour indiquer la classification), elles ne sont pas optimales. Des solutions centralisées avec une plateforme en ligne sont des options plus adaptées. D'une part, elles permettent aux utilisateurs de ne pas avoir à télécharger les images, et d'autre part elles permettent de collecter à un seul endroit les classifications d'utilisateurs qui peuvent travailler en simultané. Quelques outils de ce type existent, notamment pour recenser les structures LSB, cependant à notre connaissance aucun ne permet de lescaractériser de manière quantitative.

Il est apparu alors nécessaire de développer un outil spécifiquement dédié à cette tâche de classification. Codé par Felix Richards, ancien doctorant en informatique et science des données, le serveur permet de dessiner avec précision la forme des structures LSB superposées sur des images profondes. L'interface, basée sur Aladin Lite, permet de naviguer puis de dessiner en sélectionnant la forme de dessin la plus adaptée à la forme de la structure. Le but est de délimiter à la fois les structures stellaires mais également les sources de contamination afin de quantifier à quel point les images sont propres ou non. En effet, ces sources de contamination peuvent entraîner des biais dans la classification en masquant certaines structures très faibles. Les annotations (type de structure, position, forme) sont ensuite stockées dans une base de données. Durant cette thèse, j'ai mis cet outil en ligne sur Internet et j'y ai apporté plusieurs modifications. Entre autres choses, j'ai ajouté de nouvelles fonctionnalités facilitant le processus d'annotation, exécuté des codes permettant de faire des calculs en temps réel sur les annotations et créé des pages interactives pour visualiser ces résultats.

Ensuite, à partir de cette base de données j'ai développé des outils d'analyses et de calculs. En me basant sur les coordonnées des contours des annotations, j'ai obtenu des valeurs géométriques sur les structures (aire, longueur, largeur, courbure, rayon) et créé les masques correspondants. A partir des masques, j'ai pu faire de la photométrie d'ouverture automatique en récupérant le flux à l'intérieur, après une étape importante de ré-estimation de la valeur du fond du ciel. Avec des informations sur la couleur et le flux, j'ai ensuite déduit une estimation de la masse stellaire contenue dans les structures. Les outils développés dans ce Chapitre seront appliqués aux données puis les résultats seront présentés dans le Chapitre 5. Les futures applications du serveur d'annotation seront discutées au Chapitre 7.

Chapter 3 Deep imaging

Equipped with the annotation server and analysis tools presented in Chapter 2, the next step was to determine the set of deep images to utilize. We needed LSB-compliant images centred on nearby galaxies. Our galaxy sample will be described in Chapter 4. In this chapter, we detail the origin of the deep images we exploited, the telescope, surveys, as well as image processing steps needed for this work.

3.1 Observations with the CFHT MegaCam camera

As mentioned in Chapter 1, in order to detect LSB tidal features, it is necessary to have images which are deep enough to disclose these very faint structures. A combination of dedicated observing strategies, data reduction pipelines and telescopes (both amateur and professional) enable to obtain LSB-compliant images. Among the professional telescopes suitable for this task, one can cite the Canada France Hawaii Telescope (CFHT) from which I used images during my PhD thesis.

Placed on top of the Mauna Kea in Hawaii, at 4,200 metres above the sea, the CFHT is a 3.6-m telescope that is operational since 1979. There are currently five instruments that can be placed on the telescope to perform different types of science: imagers such as MegaCam and WIRCam (respectively an optical and infrared camera), spectropolarimeters with SPIRou (near-infrared) and ESPaDOnS, and an optical imaging Fourier

transform spectrometer (SITELLE). The CFHT runs both Principal Investigator (PI) programs and Large Programs.

During my PhD thesis I used images from four different surveys conducted with the MegaCam camera at CFHT. MegaCam (Boulade et al. 2003) is a wide-field optical imager, that had been put in place in 2003 and which was the largest imager ever built for astronomy between 2003 and at least 2013. This year, the 20$^{\text{th}}$ anniversary of MegaCam will be celebrated. It is composed of 40 back-illuminated CCDs (a total of 378 megapixels), with a field-of-view of $1^\circ \times 1^\circ$ and a pixel size of 0.187 arcsecond. It has 5 broad-band optical filters (u,g,r,i,z), as well as narrow-band filters, including Hα, OIII and CaHK. The broad band filters should actually be named u*, g', r', i', z' as they were designed to match the SDSS filters as close as possible, except for the u*. Due to the spatial extent of the filters, four CCDs were not completely illuminated: this vignetting issue reduced the number of usable CCDs to 36. Is is very important to note that after 2015, a new set of broad-band filters was put on MegaCam: the overall shapes of the filters and transmission curves differ between the old and new filters. The new filters were simply named u,g,r,i,z. These filters have a larger spatial covering than the previous ones, enabling the full illumination of all 40 CCDs. The impact of the change of usable CCDs and filters is discussed in more details in Section 3.4.1.

The original goal of MegaCam was oriented towards the science of relatively small objects, such as stars or galaxies (not nearby ones). The photometric calibration was later improved to reach 1% photometric precision (e.g., Regnault et al. 2009) which is necessary for precision cosmology. The standard pipeline of MegaCam is named Elixir (Magnier & Cuillandre 2004), and performs a local background subtraction. However, when dealing with extended objects, this pipeline cannot be directly applied to obtain LSB-compliant images: since the background is locally subtracted, over-subtracted areas near large galaxies are left. For instance, this issue is visible in the reference survey CFHT Legacy Survey (CFHTLS) images as they were processed by Elixir for sciences such as supernovae, gravitational lensing or quasars but not for extended objects. Figure 3.1 shows the difference between CFHTLS and a LSB-compliant survey (MATLAS): dark haloes or regions around CFHTLS appear around extended galaxies and erase the LSB signal. After the images were processed, they are stored on the Canadian Astronomy Data Centre (CADC[1]) where they can be accessed.

The following sections describe the surveys carried out at CFHT used during this work, namely NGVS, MATLAS, CFIS and VESTIGE, as well as the LSB-compliant pipeline that was implemented for these surveys. Table 3.1 is a summary table that presents the main properties of these surveys.

3.2 LSB-optimized surveys with the CFHT

In this section, we describe the four CFHT surveys that were used, in terms of observing strategies, available bands (the r-band being our reference band) and areas covered.

3.2.1 NGVS

The Next Generation Virgo cluster Survey (NGVS) (Ferrarese et al. 2012) is a CFHT Large Program that ran from 2008 to 2012 with the aim of imaging the entire 104 square degrees of the Virgo cluster. The Virgo cluster is the largest concentration of galaxies

[1]CADC website, https://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/

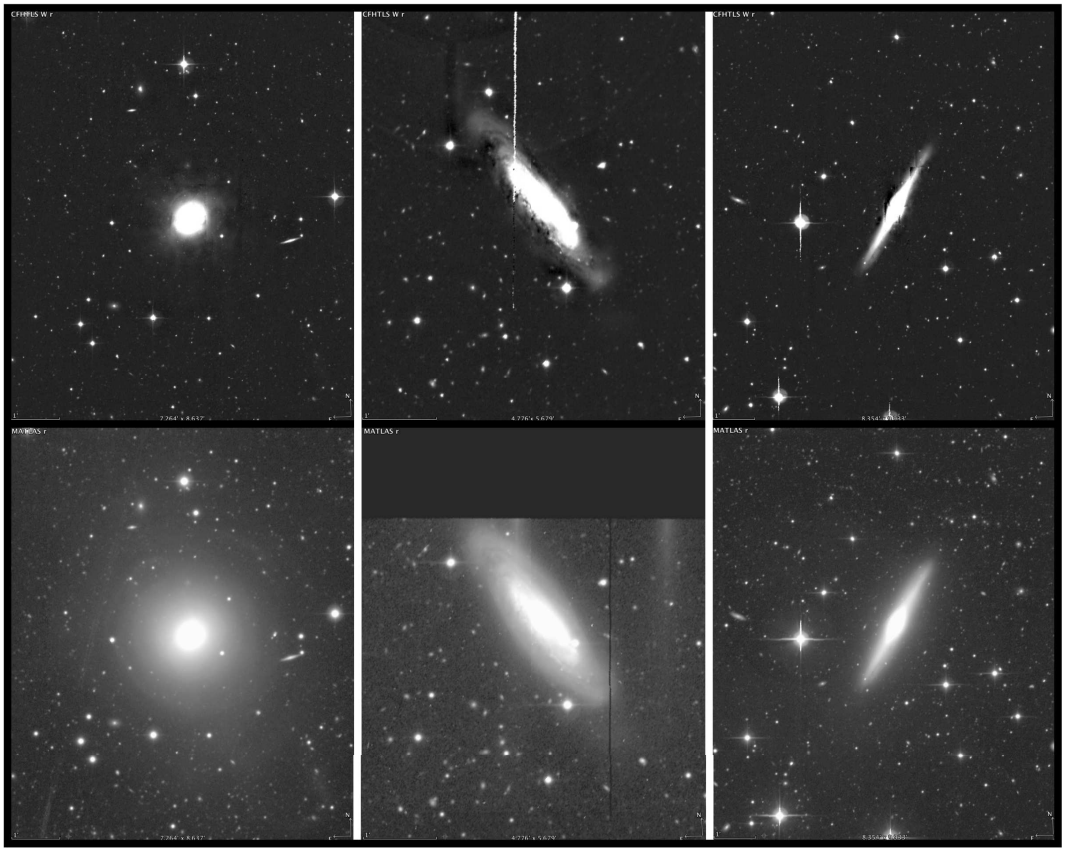

Figure 3.1: Illustration of the over-subtraction of the background after data processing by the Elixir pipeline for CFHTLS-Wide r-band images. *Top*: CFHTLS-Wide r-band images. *Bottom*: same galaxies observed by MATLAS in the r-band, treated with an LSB-compliant pipeline.

in the nearby (< 40 Mpc) Universe, and it had been studied extensively at X-ray, UV, IR, submillimetre and radio wavelengths. However, for the optical coverage, the data dated back to photographic plates in 1980's with the Virgo Cluster Catalog (VCC) from Binggeli et al. (1985). Therefore, there was a need for modern re-imaging of the cluster, from its core to its virial radius, which was done by NGVS in u,g,r,i,z filters. The original goal was to cover the cluster in all five bands, yet because of a lack of time the r-band had to be reduced to a few square degrees only. The surface brightness limit reached about 29 mag arcsec^{-2} in the g-band. The spatial coverage of NGVS is shown in blue in Figure 3.2. This spatial coverage was obtained through Multi-Order Coverage maps (MOC)(Fernique et al. 2014), which will be detailed in Section 4.2.

After the raw images were acquired, they were fed to the Elixir standard pipeline. The Elixir pipeline was developed for point-source like science, yielding a wrong background estimation for the extended objects targeted by the LSB science. Therefore, a new LSB-compliant pipeline had to be developed: Elixir-LSB, which is illustrated on Figure 10 from Ferrarese et al. (2012). The goal is to characterize the sky background on a large field-of-view, which can be achieved by median-combining exposures of different fields (provided that they are observed within a relatively short period so that the scattered light background does not vary significantly). At the core of this pipeline lies a peculiar dithering strategy. First, seven $1 \times 1^{\circ 2}$ images of a region are acquired in an uninterrupted sequence. Then, they are median-combined by Elixir-LSB, to produce an image of the

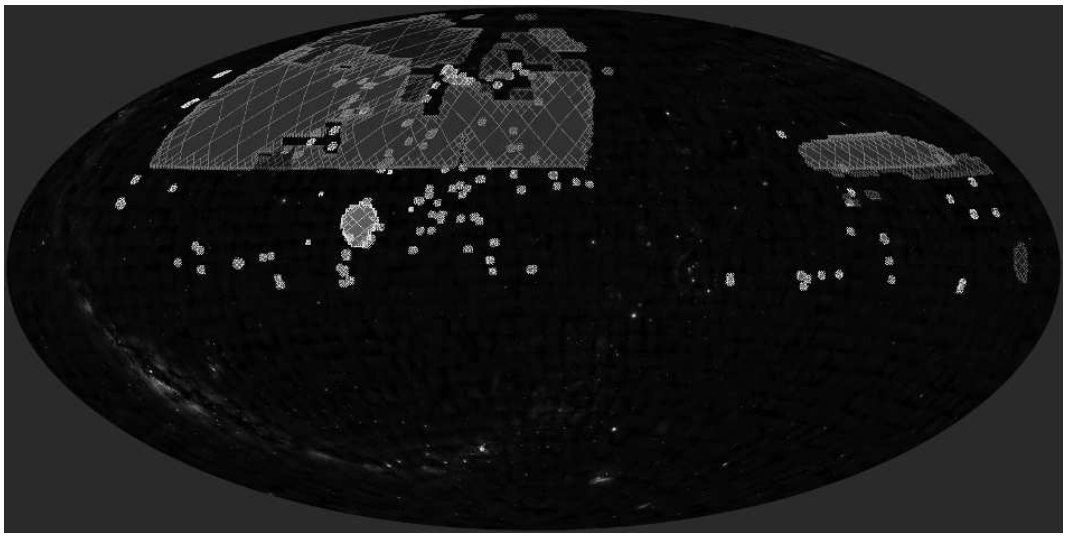

Figure 3.2: Spatial footprint (MOC) of NGVS and VESTIGE in blue, MATLAS in green, CFIS LSB-DR3 in magenta and CFIS LSB-DR4 in red. They are overlaid on a DSS2 image.

background scattered light (without any object in it). This global background is then subtracted from the seven initial images to obtain the Elixir-LSB frames. Later, the same sequence is repeated, but with a slightly different offset position. Several sequences can be acquired, all will be background-subtracted. These sequences of 7 images contribute to build the full Virgo map. For the final step, all the Elixir-LSB frames belonging to the same field are stacked together to produce the Elixir-LSB MegaPipe stacks. At the end of the Elixir-LSB step, the images are supposed to have a flat background. However, due to contamination sources the background may actually be locally not flat. The final images reach a depth of 29 mag arcsec^{-2} in the g-band Ferrarese et al. (2012).

With this optical coverage of the entire Virgo cluster, the scientific goals were multiple. To mention a few projects, NGVS enabled to study the structure of the Virgo cluster, the diffuse intracluster light and medium, the scaling relations for galaxies spanning a large range of mass, cosmic shear measurements, or even substructures in the Galactic halo that can be detected when observing in the direction of the Virgo cluster. For my PhD thesis, we defined the reference band to be the r one[2]. As NGVS does not provide r-band images for the entire cluster, another survey covering the Virgo cluster was needed. However, we utilized the NGVS g-band images in order to derive colour information.

3.2.2 VESTIGE

With NGVS, the Virgo cluster was imaged in the g-band, To perform a consistent study of tidal features, it was necessary to find another survey covering the Virgo cluster in the reference r-band. The Virgo Environmental Survey Tracing Ionized Gas Emission (VESTIGE) was utilized. VESTIGE (Boselli et al. 2018) is a CFHT Large Program that was allocated fifty nights of observing time, that ran between 2017 and 2020. VESTIGE is a blind narrow-band Hα+[NII] imaging survey of the Virgo cluster with MegaCam, designed as a follow-up of NGVS, with the goal of studying the effect of the environment

[2]The r-band was chosen as the reference one as for CFIS it is the only LSB-compliant band, as well as because the r-band was already available for MATLAS

on galaxy evolution. Additional broad-band r images were also acquired for calibration purposes in order to subtract the continuum emission. The images were processed in the same way than NGVS for consistency, with Elixir-LSB. The observing strategy and output stacked tiles are the same than for NGVS. We used these LSB r-band images for our study: although less deep than NGVS, they are deep enough to detect LSB tidal features.

3.2.3 MATLAS

Another survey from which deep images were derived is the Mass Assembly of early-type GaLAxies with their fine Structures (MATLAS) survey (Duc et al. 2015; Bílek et al. 2020). MATLAS is a CFHT Large Program that was allocated 300 hours of observing time on MegaCam, from 2013 to 2015. It is a part of the ATLAS^{3D} framework that will be described in Chapter 4. Its goal is to investigate the mass assembly of ETGs and the build-up of their scaling relations, through the study and census of LSB tidal features. To that end, MATLAS performed pointed observations of the target elliptical galaxies in r and g bands, as well as in i and u band for some galaxies. The images are deep and reach a surface brightness limit of 28.5-29 mag arcsec^{-2} in the g-band (Duc et al. 2015). Each 1 square degree image is centred on an ETG. The spatial coverage of MATLAS is represented with the green patches in Figure 3.2.

Once the images were acquired, the Elixir-LSB pipeline was run. It requires multiple input exposures, 7 for an optimal result, however the dithering strategy for MATLAS is different than the one for NGVS (an illustration of the dithering strategy is presented on Figure 1 from Duc et al. (2015)). The difference between the dithering strategies comes from the fact that NGVS needed to map the entire Virgo cluster, so building a sequence of 7 adjacent images served that purpose, while MATLAS performed pointed observations and the goal was not to lose too much observing time. For MATLAS, the target galaxy is observed in all 7 exposures, with a radial offset each time (the offsets in RA and DEC range between 2 and 14 arcmin): this enables both to remove the CCD gaps when median-combining and to create a global background map that is then subtracted. At the end of the stacking of the 7 Elixir-LSB frames, the image is cropped and centred around the target of interest. The full details are presented in Duc et al. (2015). Although the goal is to have one target ETG per image, in some cases two or more ETGs also appear to be very close on the same image, in which case they had not been observed again. Deep r-band MATLAS images were used, as well as g-band images in order to derive the $g-r$ colour information.

3.2.4 CFIS

The Canada-France Imaging Survey (CFIS) is a CFHT Large Program that has been allocated 271 nights. It started in 2017 and it is still on-going. It is a blind survey that covers the northern high Galactic latitude sky in the r-band with 5,000 square degrees ('CFIS-r') and in the u-band with 10,000 square degrees ('CFIS-u') (Ibata et al. 2017). CFIS is complementary to additional observations in other bands available from other observations made with other telescopes as part of the Ultraviolet Near Infrared Optical Northern Survey (UNIONS) project. The UNIONS project is a collaboration of wide field imaging surveys of the northern hemisphere. UNIONS consists of the Canada-France Imaging Survey (CFIS), conducted at the 3.6-metre CFHT on Maunakea, members of the Pan-STARRS team, and the Wide Imaging with Subaru Hyper-Suprime-Cam of the

Euclid Sky (WISHES) team. CFHT/CFIS is obtaining u and r bands; PanSTARRS is obtaining deep i and moderate-deep z band imaging, and Subaru/WISHES is obtaining deep z band imaging.

The main aim of CFIS is to address some questions in cosmology, such as the assembly of the Milky Way, the properties of dark matter and dark energy, the growth of structures in the Universe, or weak lensing studies, thanks to its large area covered and high quality images. CFIS data will be very useful as a complement of the future space mission Euclid that will tackle the dark matter and dark energy issues. Indeed, in addition to the infrared bands of Euclid, photometric redshifts for faint galaxies are needed, that is the reason why CFIS will provide them by adding the r-band from the ground. Other CFIS missions focus on the Solar System, the Milky Way dynamics, large scale structures or galaxy properties. However, the large spatial coverage and excellent r-band image quality make CFIS an ideal survey to study LSB structures around galaxies.

Once acquired, CFIS u and r-band images are processed by the standard Elixir pipeline. However, the r-band images were also processed by the Elixir-LSB pipeline, meaning that when finished, 5,000 square degrees of the sky will be available and LSB-compliant in the r-band. The surface brightness limit reached 28.3 mag arcsec^{-2} in the r-band (Cuillandre, private communication). These Elixir-LSB r-band images were the ones used during my PhD thesis. As of March 2021, 3,600 square degrees were covered and released in CFIS LSB-DR3. Later, the LSB-DR4 was released and covered 4670 square degrees, but the additional data was not used for my work, I only considered LSB-r images from DR3. The spatial coverage of the LSB-DR3 and current LSB-DR4 CFIS LSB-r images are plotted respectively in magenta and red in Figure 3.2.

CFIS images are processed by Elixir-LSB and then stacked on a set of tiles. Each tile is $10,000 \times 10,000$ pixels[3]. The tiles are spaced 0.5 degrees apart in declination and $0.5/\cos(dec)$ apart in right ascension. This leads to a small (3%) overlap between tiles. The adjustment of CFIS tiles is illustrated in Figure 3.3, using the UNIONS graphical search tool[4]. The border of each tile is materialized with the orange line, while the LSB-r images are displayed on the background, in addition to a galaxy of interest.

Survey	Filter used	Exposure time (s)	Targets	Area covered (deg^2)	Depth (mag arcsec^{-2})
NGVS	Old g	5×634	Virgo cluster	116	29
VESTIGE	New r	12×60	Virgo cluster	116	27-28
MATLAS	Old r + old g	7×345	Group ETGs	310	28.9
CFIS	New r	Varying[5]	Blind survey	5,000	28.3

Table 3.1: Summary of the main properties of the surveys we used. The reference band is the r one.

3.3 Data processing

The deep, Elixir-LSB processed images from NGVS, MATLAS, CFIS and VESTIGE were stored on the CADC website. The basics processing had already been done with Elixir-LSB, here I simply retrieved the images and only performed a few additional steps of

[3]The resolution of CFIS tiles is of 0.1857 arcsecond per pixel, which is close but slightly different from MegaCam's resolution of 0.1870 arcsecond per pixel

[4]UNIONS graphical search tool, https://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/community/unions

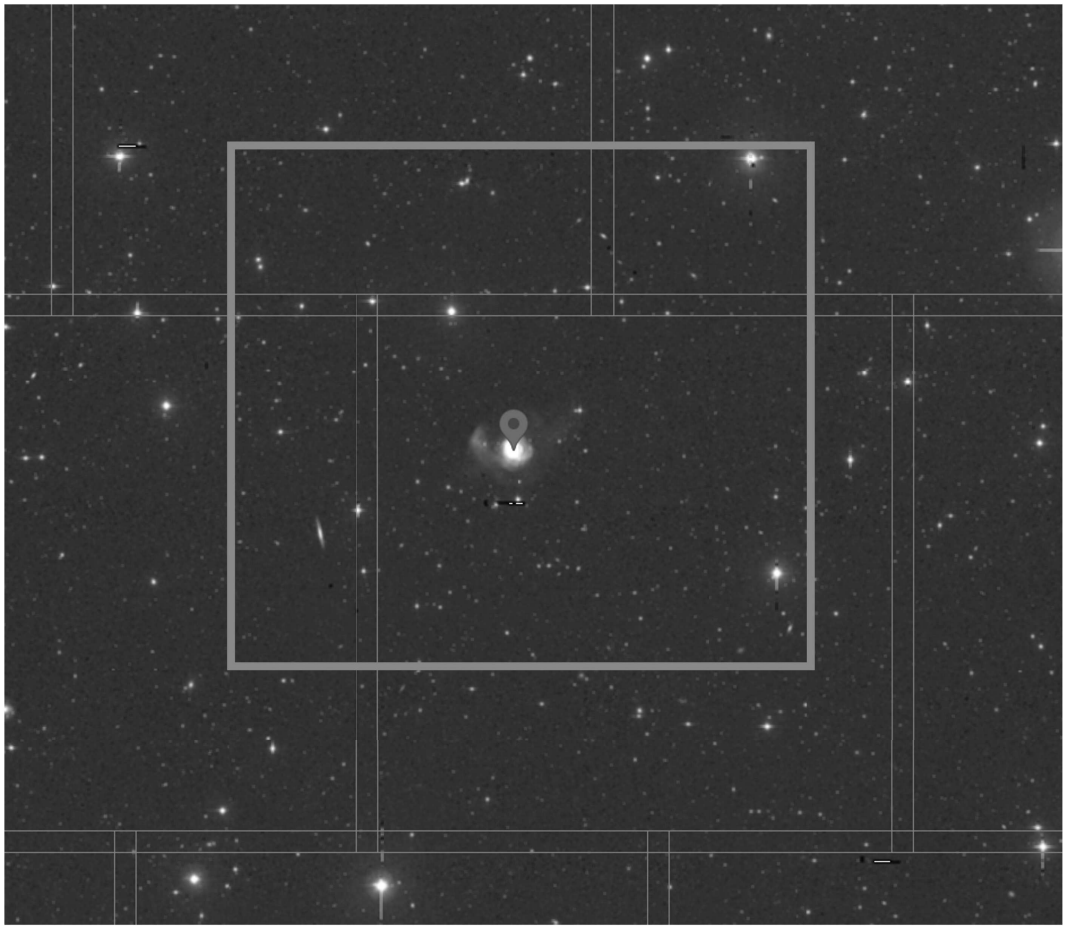

Figure 3.3: Illustration of CFIS tiling, using the UNIONS graphical search tool. An object of interest in displayed in the middle of the image, while the border of the tiles are represented in orange. The overlap between tiles is visible. The green square represents the 30×30 arcmin image centred on the galaxy of interest, which we used for the cropping.

processing in order for them to be added to the annotation server. The first step was to determine which images were interesting.

For MATLAS, the process was relatively simple. As the observations were pointed, each 1 square degree image is already centred on the target ETG. To reduce the amount of storage needed, we cropped the images to be 0.5×0.5 square degrees around the central ETG, which enabled to check the local environment and the presence of any companion that could create potential interactions. For the few ETGs located in another ETG's field, as well as for LTGs falling in these fields, we created 0.5×0.5 square degrees cutouts around them from the fullsize field. For CFIS, as the survey is blind, the process is different. On the CADC, the thousands of tiles are stored and labeled CFIS.xxx.yyy.r.fits, where xxx and yyy correspond to the RA,DEC coordinates of the centre of the tile following: $DEC = yyy/2 - 90$ and $RA = xxx/2/cos(dec)$. Each CADC tile is $30.95' \times 30.95'$, as illustrated by the orange tiles in Figure 3.3. With the coordinates of the target galaxy and a given size, I retrieved the names of the tiles containing and adjacent to the target galaxy. They were recombined together using SWarp (Bertin et al. 2002), and then cropped in order to centre the final image on the object of interest. For NGVS and VESTIGE, the naming convention of the tiles gives the offset in RA, DEC and to the nearest degree, from the (0,0) field which includes in its northwest corner, M87. From the header of the

images, I obtained the RA,DEC coordinates of the corners and only kept the images that contained the coordinates of the target galaxy. The images were recombined then using SWarp and cropped around the target galaxy. An illustration of this 30×30 arcmin cropped image is drawn in green in Figure 3.3.

Once I had the cropped images centred on the target galaxies, the next step was to rebin in order both to enhance the faintest structures and to save some storage space. The 'rebinned' images were obtained by performing a 3×3 binning, i.e. averaging the 9 surrounding pixels. This factor of 3×3 was chosen after some tests, as illustrated in Figure 3.4, as fine structures are more visible while keeping enough resolution. This factor must be taken into account when considering the flux or surface brightness of objects, as presented in Section 2.3.3.

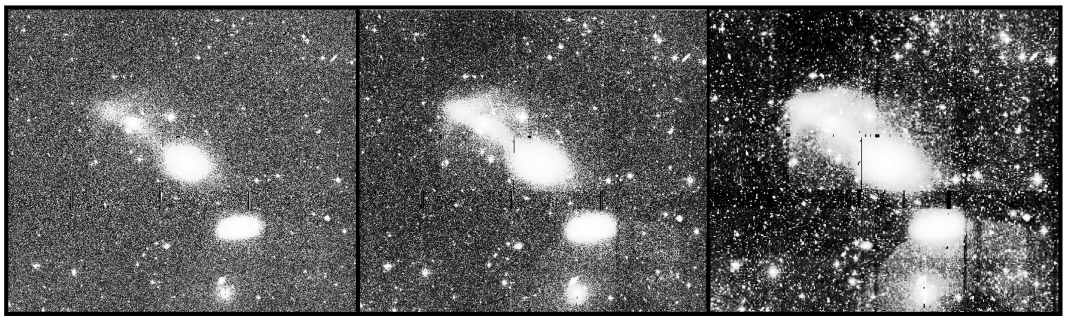

Figure 3.4: Illustration of an image rebinning. *Left*: Original image *Middle*: Image rebinned 3×3 times. *Right*: Image rebinned 24×24 times.

To further enhance the visual identification of fine structures, an adjustment of the image dynamics is usually performed such as the transformation from linear scale to an inverse hyperbolic sine. Here, we applied a slightly modified inverse hyperbolic sine, asinh. The goal is to crush the dynamics of the image to see the entire galaxy and the faint external regions, but by doing so, the risk is to over-interpret the importance of tidal debris. This is why a detailed analysis of the fraction of luminosity contained in these external regions must be conducted afterwards, as in Chapter 5, to properly re-evaluate their importance. An example of the comparison between linear and asinh scalings is presented in Figure 3.5. The transformation from linear to asinh scale follows the formula:

$$\mathrm{asinh}(ADU) = \log_{10}\left(\alpha \times (ADU - bk) + \sqrt{\alpha^2 \times (ADU - bk)^2 + 1}\right) \qquad (3.1)$$

where asinh is the value of the pixel in asinh scale; ADU is the pixel value in linear scale; bk is the background value and α is a parameter to tune, chosen here as 1. As explained in Section 2.3.3, the background level is supposed to be 0 after being processed by Elixir-LSB (that gives a flat background over a given field of view, after correcting for residual instrumental or large scale sky artefacts). However some local contamination of sources, such as star haloes or Galactic cirrus, remain after this processing. Therefore the real background may locally be non-zero. I performed a local determination of the background afterwards to get a precise photometry (see Section 2.3.3). For the asinh scaling, the only purpose is visualization so the background value bk in the formula is set to 0.

In addition to asinh scaled images, surface brightness (SB) maps scaled in mag arcsec^{-2} were produced and used to characterize the tidal features. The relation from linear to

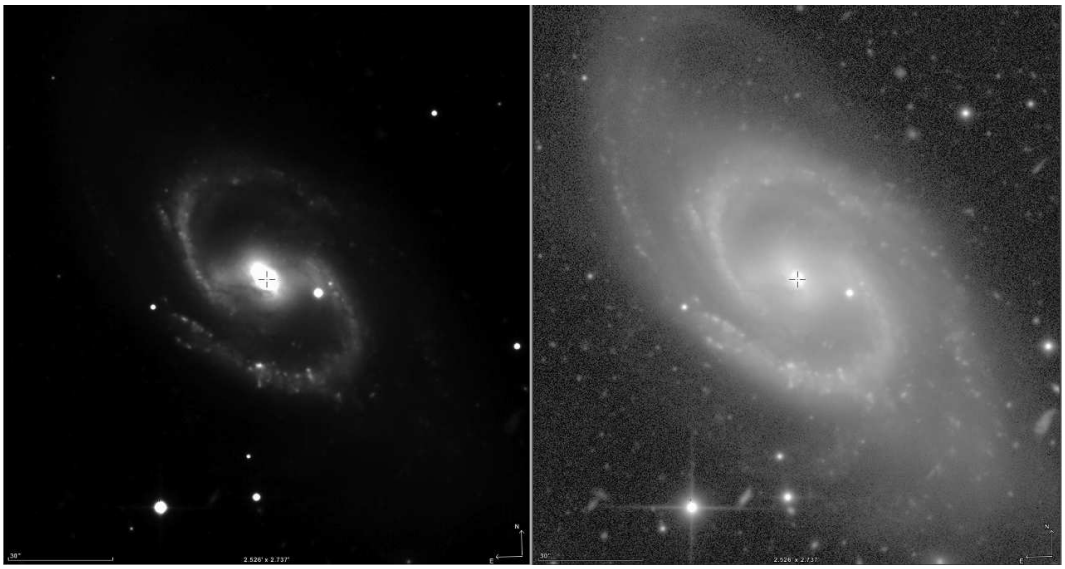

Figure 3.5: Illustration of the asinh scaling. *Left*: Image in linear scaling *Right*: Image in asinh scaling

surface brightness scale is the following:

$$\mu = -2.5 \log_{10} \left(\frac{ADU - bk}{\text{pixsize}^2} \right) + zp \tag{3.2}$$

where μ is the surface brightness value of the pixel in mag arcsec^{-2} in AB magnitude, ADU is the value of the pixel in the original linear image, bk is the value of the background, pixsize is the size of one pixel in arcsecond and zp is the value of the zero point (30 in our case for MegaCam) (see also Section 2.3.3).

The last step is to make the images compliant with the annotation server. Before this point, all the images are in FITS format, yet Aladin Lite on which the server is based only supports HiPS format (Fernique et al. 2015) as explained in Section 2.2.2. This was done with the hipsgen command[6], applied on the asinh (and SB) images.

As discussed in Section 2.2.2, the Aladin Lite version 2 on which the annotation server relies does not support on-the-fly adjustment of the contrast and dynamics of the images. We had to adjust the image scaling before converting them to HiPS format. After some tests, forcing the HiPS pixel values between 0 and 1 was helpful to enhance LSB structures. Having a single non-modifiable image enables a homogeneous classification between several users. Once created, the HiPS were stored on the server so they could be accessed through the annotation tool. To illustrate the storage space needed, one rebin FITS file is about 40 Mo (megaoctet) in linear scale, and 400 Mo for a HiPS. The sizes of HiPS can dramatically increase with the HiPS order (i.e. the maximum resolution of the HiPS) so a trade-off between resolution and file size was necessary.

Finally, other images were also used: the shallow PanSTARRS DR1 optical images but they were already integrated in Aladin Lite, hence directly accessible from the annotation server. We needed shallow images to annotate the *Main Galaxy*, and we chose PanSTARRS as the survey covers the entire sky north of -30 degrees in declination, and thus encompasses our galaxy sample as it will be described in Chapter 4.

[6]HiPS documentation, https://aladin.u-strasbg.fr/hips/HipsIn10Steps.gml

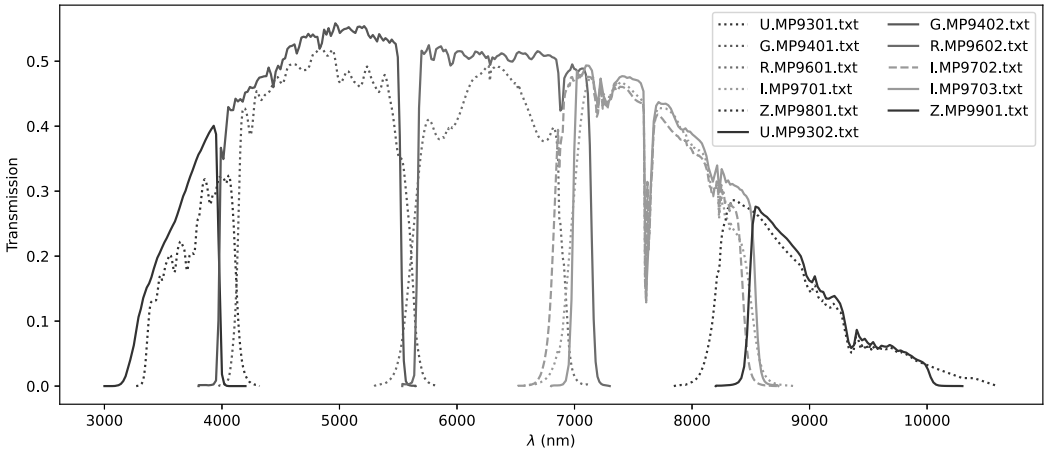

Figure 3.6: Comparison of the full transmission spectra of MegaCam old (dashed lines) and new (solid lines) broad-band filters. u, g, r, i and z are respectively plotted in blue, green, red, orange and magenta. The transmission curves include the filter transmission, the transmission of the different parts of the MegaCam optics, the telescope primary mirror and an average atmospheric transmission above the telescope. The data to recreate these curves can be accessed at https://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/megapipe/docs/filt.html

3.4 Dealing with multiple surveys

An issue we faced with images from four different surveys was to analyse them together. Several points had to be considered.

3.4.1 Change of MegaCam filters

Another point that has an important impact when analysing the four surveys is the change of set of MegaCam filters after 2014. The properties of these new broad-band *ugriz* filters differ from the old ones, with improved transmission properties (spatial and spectral) and different central wavelengths. Due to the larger spatial size of the filters, it was possible to utilize 4 CCDs that were previously not entirely illuminated, hence MegaCam now has a total of 40 usable CCDs (Prunet et al. 2015). Figure 3.6 presents the transmission curves of the old and new filters. This means that for MATLAS and NGVS, the old filters were in place whereas for CFIS and VESTIGE the new filters were used. This has an important impact on the photometry that we had to assess.

To that end, we selected 3 fields in r-band in common between CFIS and MATLAS, and one field in r-band in common between NGVS and VESTIGE. The goal was to determine whether we would find similar flux values between the old and new filters. To that end, for each field we applied SExtractor to automatically compute the segmentation maps, and we applied the same segmentation between the old and the new filters. We then performed a cross-match of the sources in order to select only the same regions in each pair of fields. From it, we only kept the segmented regions that were the cleanest, i.e. with a high enough signal-to-noise ratio, and which followed a relatively tight relation for the magnitude between the old and new filters. We called these regions kept 'best points' and we present their flux in the old and new filters in the left panel of Figure 3.7. From it, there is a clear offset between the new and the old magnitudes. We hypothesized that it could come from the extinction (we had four different fields so the extinction could differ

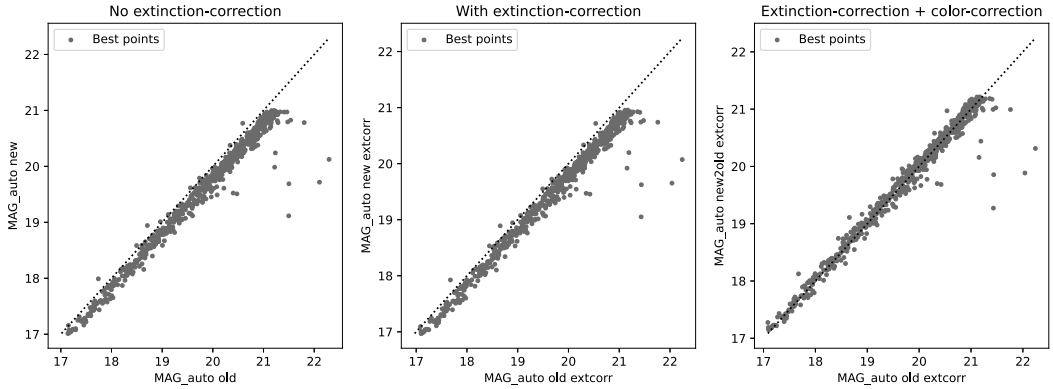

Figure 3.7: Estimation of the colour-correction between MegaCam old and new r-band filters, and effect of the correction of the extinction. Magnitudes are expressed in mag. *Left*: Scatter plot of the magnitude in the new filter (MAG_auto new) as a function of the old filter (MAG_auto old) without the correction of the extinction. *Middle*: Scatter plot of the magnitude in the new filter (MAG_auto new extcorr) as a function of the old filter (MAG_auto old extcorr) with the correction of the extinction. *Right*: Scatter plot of the magnitude in the new filter for which the colour-correction has been applied (MAG_auto new2old extcorr) as a function of the old filter (MAG_auto old extcorr) with the correction of the extinction and the colour-correction.

from one to the other), so we corrected these values for the extinction following Section 2.3.6: the result is shown in the middle panel of Figure 3.7. It appears that the extinction correction does not impact much our results. As we saw a linear relation between the points, we fitted a linear regression in order to convert the magnitudes from the new filter to the old filter, and we derived the formula:

$$mag_{\mathrm{r,new2old}} = 1.019 \times mag_{\mathrm{r,old}} - 0.147 \tag{3.3}$$

where $mag_{\mathrm{r,new2old}}$ is the extinction and colour-corrected magnitude[7] in the r-band new filter converted to be in the old filter, and $mag_{\mathrm{r,old}}$ is the magnitude in the old filter. When applying this correction, the results are shown in the right panel of Figure 3.7. Then, we applied this colour-correction to the flux measurements that will be presented in Chapter 5 in order to try to account for this difference of filters. However, it must be noted that this colour-correction is only a first attempt at correcting the differences and that it was estimated on four fields only. A more careful treatment of this effect would be needed to accurately compare the flux and colours between the new and old MegaCam filters[8].

Another point to take into account when working with LSB images are the sources of pollution. In particular, the scattered light from bright sources like foreground stars or the bright centre of galaxies can cause multiple internal reflections on the CCD, filters and telescope optics and generate large artificial haloes around the source in the image, that are called ghost reflections (Slater et al. 2009). Although ghost reflections

[7]The 'extinction and colour-corrected magnitude' corresponds to the magnitude of the object when corrected both for extinction and for the change of MegaCam filters which required colour-correction.

[8]Some calibrations have been computed between the MegaCam and SDSS systems for the old and new filters, and are presented in https://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/megapipe/docs/filt.html, but with our results we could not reproduce such trends. This likely comes from the differences in the sources used to perform the calibration from our method and theirs.

from bright stars (thereafter called external ghosts) can easily be identified and removed from the analysis, the additional amount of light from ghost reflections from the bright galactic centre (thereafter called internal ghosts) is more difficult to tackle (e.g., Karabal et al. 2017), as it is mixed with the intrinsic light from the stellar halo. The shape and prominence of the ghost reflections depends on the filter: as illustrated in Figure 3.8, the ghosts are much more prominent for the old MegaCam filters (MATLAS and NGVS) than for the new ones (VESTIGE and CFIS). Carrying out a precise estimation of the luminosity and sizes of tidal features and haloes requires to remove the contribution of the internal ghosts. To that end it is necessary to model the Point Spread Function (PSF) and then deconvolve images from the PSF. The PSF must be carefully modeled, with a focus on its outer wings where the light from ghost reflections comes into play. The PSF-deconvolution will impact the luminosity and size of galaxies. Although we did not have PSF-deconvolved images for this work, we will assess the impact of the PSF in Section 6.3.1.2 for a subset of a few MATLAS images PSF-deconvolved with a carefully modeled PSF by Karabal et al. (2017).

3.4.2 Depth of the surveys

Although the four CFHT surveys are LSB-optimized, they do not have the same observing strategies, exposure times nor surface brightness limit (see Table 3.1). In this section, we briefly discuss the impact of the difference of depth. A detailed assessment of this impact between CFIS and MATLAS (with a difference of depth of 0.6 mag) will be presented in Section 6.4.1. To summarize our results, we concluded that although there is a difference of depth, this does not trigger a major bias in our results (in terms of halo size, SB or number and geometry of tidal features).

Likewise, we compared the depths of NGVS and VESTIGE in Section 6.4.2. Although NGVS with a depth of 29 $\mathrm{mag\,arcsec^{-2}}$ in the g-band is at least 1 mag deeper than VESTIGE, a visual comparison between the images from both surveys, illustrated in Figure 3.8, revealed that VESTIGE seems almost as 'deep' as NGVS. Very faint structures are more easily detectable in NGVS images, such as cirrus clouds or tidal features (see the middle and right panels of Figure 3.8). However, the overall lower background and limited ghost haloes around bright stars in VESTIGE images tend to compensate this difference of depth. This is caused by the change of MegaCam filters (see Section 3.4.1). These reduced ghost haloes were also visible in CFIS compared to MATLAS.

Therefore, it appears that the exposure time is not the most important parameter when dealing with LSB images: the impact of filters used and the LSB-compliant pipeline seem to be more important. This underlines the difficulty of creating a homogeneous LSB sample when combining several surveys. In spite of this, we considered that we could combine the four surveys without inducing a major bias.

3.5 Conclusion

To summarize this chapter, we focussed on the deep images used during my PhD work. We utilized deep images from four CFHT surveys: NGVS (g-band) and VESTIGE (r-band) targeting the Virgo cluster; while MATLAS (r and g-band) focussed on ETG groups, and CFIS is a blind survey (r-band) in the Northern hemisphere. We described the observing strategy of the surveys and their overall characteristics, as well as Elixir-LSB, the LSB-compliant pipeline for MegaCam. The few steps of data processing needed to retrieve

Figure 3.8: Comparison between images of the same galaxies from NGVS g-band images (*top row*) and VESTIGE r-band images (*bottom row*).

and transform the Elixir-LSB images to integrate them to the annotation server were also detailed. Finally, we briefly discussed the difficulties arising from dealing with multiple surveys.

3.6 Résumé du chapitre

Dans ce chapitre, nous avons présenté les images profondes, préservant le signal LSB, qui nous ont permis d'étudier les structures de marée. Nous avons cherché quels relevés d'imagerie profonde pouvaient convenir. Plusieurs télescopes peuvent être utilisés pour la science LSB, en combinaison avec des stratégies d'observation et des pipelines de traitement de données dédiés, dont le Canada-France-Hawaii Télescope (CFHT), d'où proviennent les images utilisées pendant ma thèse. Le CFHT est équipé de plusieurs instruments dont l'imageur optique MegaCam. Composé à l'origine de 40 CCDs, MegaCam possède un large champ de vue de 1 degré carré avec une résolution de 0.1857 arcseconde par pixel. Il est équipé d'un ensemble de 5 filtres en bande large u,g,r,i,z ainsi que de filtres en bande étroite Hα, OIII et CaHK. Après 2015, le jeu de filtres en bande large a été changé et 4 CCDs (précédemment non entièrement illuminés et non utilisés) ont été rendu utilisables. Ce changement de filtres a un impact non négligeable sur la photométrie.

Nous avons utilisé les images du CFHT provenant de quatre Grands Programmes : NGVS, MATLAS, CFIS et VESTIGE. NGVS est un relevé en imagerie qui couvre l'entièreté de l'amas de galaxies de la Vierge, depuis le centre jusqu'à ses extrémités, en bande u,g,i,z, ainsi que r mais pour quelques régions uniquement. Pour obtenir des images profondes et permettre la science LSB, il a été nécessaire de développer une chaîne de traitement des données adaptée à la science LSB: Elixir-LSB. En effet, Elixir, le processus standard de traitement des données issues de MegaCam, est prévu pour une science qui

s'intéresse à des objects quasi-ponctuels dans laquelle une soustraction locale du fond du ciel est faite. Or cette soustraction détruit le signal LSB, ce qui rend inutilisables les images pour notre recherche. Elixir-LSB a donc été implémenté par les équipes du CFHT pour effectuer une soustraction globale du fond du ciel, permettant ainsi de préserver le signal LSB. Ce pipeline de traitement des données a été ensuite appliqué à tous les relevés d'imagerie profonde, dont MATLAS, CFIS et VESTIGE. Dans la continuité de NGVS, VESTIGE est un relevé en Hα+[NII] de l'amas de la Vierge, qui a aussi acquis des images r LSB. MATLAS est un relevé optique multibande (u,g,r,i) du ciel pointé, centré sur les ETGs proches et massives situées dans des groupes de galaxies. Enfin, CFIS est un relevé 'en aveugle', i.e. non-pointé, qui couvrira 5000 degrés carrés une fois terminé, en bandes u et r. Seule la bande r a été entrièrement traitée par Elixir-LSB.

Durant mon travail de thèse j'ai utilisé les images déjà traitées par Elixir-LSB en bande r de MATLAS, CFIS et VESTIGE, ainsi que la bande g de NGVS et de MATLAS. Avec ces filtres, il a été possible d'obtenir des informations sur la couleur $g - r$ des galaxies MATLAS et NGVS/VESTIGE.

Afin de pouvoir intégrer ces images au serveur d'annotation, j'ai dû procéder à un traitement supplémentaire des données. Il a fallu récupérer les images sur le site du CADC et les recombiner entre elles pour obtenir des images centrées sur les galaxies d'intérêt, puis les rogner. Elles ont ensuite été réduites en taille d'un facteur 3×3 à la fois pour faire ressortir les structures les plus faibles et pour réduire la quantité d'espace de stockage nécessaire. Afin de permettre l'inspection visuelle, un traitement non-linéaire asinh a été appliqué, tandis que les cartes de brillance de surface ont été crées. La dernière étape a été de transformer ces images FITS au format HiPS qui est celui supporté par le serveur d'annotations.

Chapter 4 Galaxy sample

Now that we know which deep images to utilize, the goal was to determine the sample of galaxies on which we would carry out a statistically significant analysis of tidal features. In this chapter, we present the reference sample of galaxies from which we derived our sample, resulting from the intersection between our images and the reference sample.

4.1 The ATLAS^{3D} project

The goal of this PhD thesis is to characterize LSB structures around a large number of galaxies, typically several hundreds, in order to probe their late assembly history. We wanted to extend the work already started by previous studies such as MATLAS (Duc et al. 2015; Bílek et al. 2020), focussed on group ETGs, to probe both ETGs and LTGs and different environments. In order for the study to be unbiased and statistically significant, we needed a complete sample of hundreds of galaxies.

To that end, we used a complete, volume-limited reference sample from the ATLAS^{3D} collaboration (Cappellari et al. 2011a). It consists of nearby (< 42 Mpc) and bright galaxies (K-band absolute magnitude M_K brighter than -21.5 mag); observable from telescopes in the Northern hemisphere, with a sky declination δ such as $|\delta - 29°| < 35°$, and excluding the dusty region near the Galaxy equatorial plane, with a galactic latitude b such as $|b| < 15°$. This reference sample contains 871 galaxies. From it, the ATLAS^{3D} collaboration had extracted a subset consisting of all the ETGs (i.e. the ATLAS^{3D} sample), the ellipticals/lenticulars being the subject of a complete study by this collaboration.

The ATLAS^{3D} project provides a wealth of information about the ETGs, with multi-wavelength coverage, spectroscopy, kinematics, stellar and gas content information, as well as a theoretical modelling effort. The galaxies from the reference sample are located in various environments, from the field, groups to the Virgo cluster. The brightness criterion of a K-band absolute magnitude brighter than -21.5 mag can be converted to

a stellar mass higher than $6 \times 10^9 M_\odot$. The total volume probed by the survey covers $1.16 \times 10^5 Mpc^3$. All these detailed information, in combination with deep images, enables to probe galactic assembly through the study of tidal features for galaxies spanning a large range of mass and environment, making the reference sample from the ATLAS^{3D} collaboration the ideal complete sample.

The selection of ETGs is based on the morphologies of galaxies. As discussed in Chapter 1, several classification schemes exist, based on morphologies, colours or kinematics. As not all galaxies from the ATLAS^{3D} reference sample had reliable colour information, and since morphology is more robust than colour to dust extinction and inclination effects, morphology was chosen as the distinction criterion between galaxies. However, thanks to available kinematic information, a more subtle distinction is made among ETGs between fast and slow rotators (Krajnović et al. 2011; Emsellem et al. 2011). The distinction between ETGs and LTGs is only based on the presence of spiral arms, following the revisited Hubble scheme (Sandage 1961). Cappellari et al. (2011a) visually classified galaxy morphologies using shallow, true-colour red-green-blue images from SDSS DR7 for 82% of their sample, and the B-band DSS2-blue images for the remaining objects. They compared their classification to the one provided by the T-type given by the HyperLeda database, and found an agreement in 97% of the cases. The parent sample consists of 8% E (defined as T-type≤ -3.5), 22% S0 (T-type> -3.5) and 70% spiral, hence a total of 260 ETGs and 611 LTGs.

A survey of LSB tidal features around the ATLAS^{3D} sample of ETGs had already been studied by MATLAS (Bílek et al. 2020), but our work extends this study to the entire parent sample. With this defined sample, the next step was to determine which deep images were available.

4.2 Sample selection

The previous section described the reference sample from the ATLAS^{3D} collaboration, consisting of 871 galaxies including 260 ETGs and 611 LTGs. In Chapter 3 we detailed the deep images that we used during this work. Here, the goal is to find the images on which the galaxies from the reference sample were located.

NGVS and VESTIGE surveyed the Virgo cluster and obtained deep images of 58 ETGs and 65 LTGs from the reference sample belonging to that cluster. MATLAS targeted a total of 179 ETGs. Its goal was to probe all the 260 ETGs from the reference sample, but it avoided galaxies in the Virgo cluster already probed by NGVS and galaxies too close to bright stars. MATLAS performed pointed observations centred on the target ETGs, but other interesting objects are located in these fields, such as 53 LTGs from the reference sample, 2000 dwarf galaxy satellites and many globular clusters. With MATLAS and NGVS, 237 over 260 ETGs from the reference sample were observed, but only 118 over 611 LTGs were targeted. To complement this sample, CFIS images were used to study a part of the remaining LTGs, as this survey covers a large area on the Northern sky in a blind and LSB-compliant way. With the combination of these four surveys, with the reference r-band and additional g-band for colour information when available, we probe the field (with CFIS), groups (with MATLAS) and the Virgo cluster (NGVS/VESTIGE). We retrieve 94% of all ETGs in the reference sample, but only 38% of all LTGs. The total number of galaxies of each morphological type and survey in which they were imaged are summarized in Table 4.1.

The next step was to determine on which images our reference sample galaxies were

located. We needed a way to represent the spatial footprint of these surveys, as the shape of their field of view on the sky is complex. To that end, Multi-Order Coverage maps (MOC) were used. Developed by the CDS, MOCs (Fernique et al. 2014) are a simple way to represent any region in the sky, even with a complicated shape, and to perform operations such as union or intersect of two regions very fast. It relies on a regular and hierarchical division of the celestial sphere into smaller and smaller diamonds: this technique is called tessellation, here the HEALPix tessellation. This partitioning is illustrated in Figure 4.1. At level 0, the sphere is divided in 12 diamonds. Then, each diamond is sub-divided into 4 smaller diamonds (level 1) leading to 48 diamonds, and so recursively. Each region of the sky is represented by the list of the indices of the diamonds it covers. The goal is to provide a very fast comparison mechanism between coverage maps.

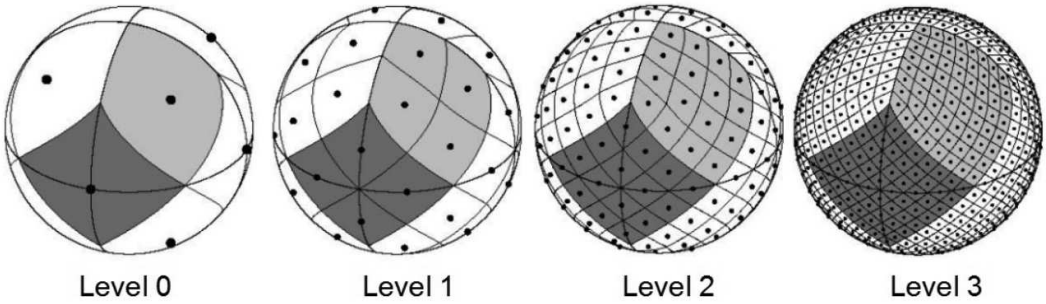

Figure 4.1: Illustration of the HEALPix tessellation: the sphere is hierarchically partitioned into diamonds.

With Aladin, it is easy to display or create a MOC for a given image collection, as well as performing operations (union, intersection) between catalogues and MOCs. Here, the goal is to perform the intersection between our images and a small square region of 30×30 arcsec around target galaxies. Using the mocgen command[1], I created the coverage map for all our images. This spatial footprint is represented in Figure 3.2, in blue for NGVS/VESTIGE, green for MATLAS, magenta for CFIS LSB-DR3 and in red for CFIS LSB-DR4. Once loaded in Aladin, the intersection between the MOC and galaxies in the reference sample was performed in order to retrieve all the galaxies lying in these footprints.

However, for two cases, a visual inspection of the images was necessary. Indeed, in some cases the galaxy of interest was in the MOC but the 30×30 arcmin field of view was not entirely in it: we actually want the intersection between the MOC and the 30×30 arcmin field of view centred on the target galaxy. In other cases where the galaxy is located near the edge of a survey, it is possible that not all the 7 tiles needed for Elixir-LSB to get a proper removal of the instrumental signature are available, hence the Elixir-LSB images are of poor quality and sometimes cannot be used. Some of these unusable images are illustrated in Figure 4.2. Due to these limitations, I had to visually inspect the images and reject the ones near the edges that were not usable. This means that some galaxies from the reference sample that lied in the MOC were actually not properly observed hence not kept in our sample.

During this work, we used images from the CFIS LSB-DR3 that were available before March 2021[2]. There are some overlapping regions between CFIS and MATLAS footprints.

[1]Aladin user manual, https://aladin.cds.unistra.fr/java/AladinManuel.pdf

[2]The CFIS LSB-DR4 release was available later but we did not use them for this work.

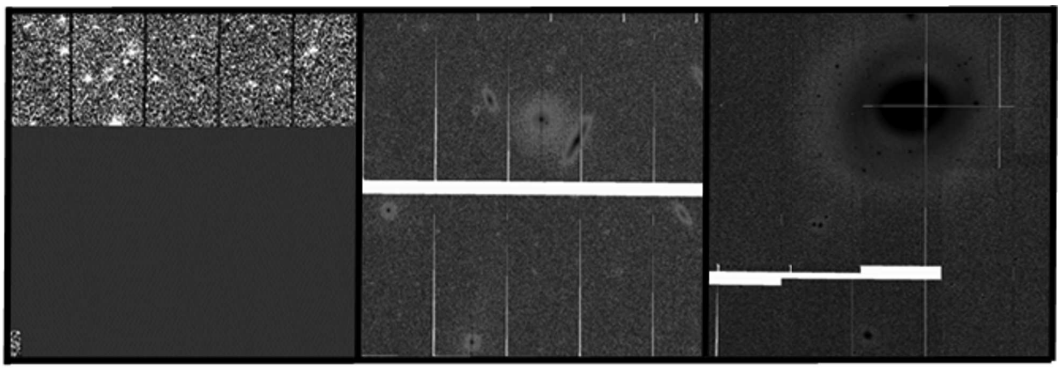

Figure 4.2: Illustration of different types of issues for three CFIS images near the edge of surveys. *Left*: One of the worse type of issues: the image is located right at an edge of the survey, which means that only 1/3 of the tile actually contains an image, the remaining fraction will never be available. *Middle*: Remaining CCD gaps, more or less pronounced, either due to a final exposure that had not yet been done or because the image is located near an edge where the 7 input tiles for Elixir-LSB are not available. *Right*: Same issues than in the *middle* panel.

In such cases, we kept MATLAS images as they are deeper than CFIS. The final CFIS sample consists of 113 LTGs and 7 ETGs avoided by MATLAS as they were too close to bright stars.[3] These stars are not an issue with CFIS as the filter set of MegaCam was changed between MATLAS and CFIS, and the new filters are less affected by the ghost reflections around bright objects (see Section 3.4.1). MATLAS galaxies are composed of 179 ETGs and 53 LTGs; while for NGVS and VESTIGE there are 65 LTGs and 58 ETGs.

Table 4.1 summarizes the number of galaxies studied as a function of their morphological type, survey and Virgo cluster membership (as defined in Cappellari et al. (2011a)). A few galaxies located in NGVS/VESTIGE images do not belong to the Virgo cluster, as there is a difference between being located in the direction of the Virgo cluster and being a member of it. For CFIS, we indicated in parenthesis and italics the total number of galaxies from the reference sample lying in the CFIS MOC (i.e. 42 ETGs and 118 LTGs), but we removed the ones that had already been observed by MATLAS so in the end we had only 7 ETGs and 113 LTGs for CFIS.

Table 4.1: Number of galaxies studied in this work, detailed by their survey, morphological type and Virgo membership. For CFIS, in parenthesis and italics is indicated the total number of galaxies from the reference sample lying in the CFIS MOC (i.e. 42 ETGs and 118 LTGs), but we removed the ones that had already been observed by MATLAS so we had 7 ETGs and 113 LTGs for CFIS.

Type	Virgo ETGs	Virgo LTGs	Field ETGs	Field LTGs	Total
MATLAS	5	4	174	49	232
CFIS	0	0	7 *(42)*	113 *(118)*	120 *(160)*
NGVS/VESTIGE	52	51	6	14	123
Total	57	55	187	176	475

Table 4.2 summarizes the percentage of galaxies we studied among the 871 galaxies of

[3]If we had updated our work with the latest CFIS LSB-DR4 release, we would have had only 8 new usable LTGs, but they were not included in the rest of our work.

the reference sample. Our final sample is composed of 475 galaxies, among which there are 244 ETGs and 231 LTGs, corresponding to 94% (resp. 38%) of all the ETGs (resp. LTGs) from the reference sample.

Table 4.2: Number of galaxies from the reference sample studied in this work, detailed by their survey and morphological type. In parenthesis is indicated the percentage of galaxies studied with respect to the total number of galaxies in each morphological type in the reference sample.

Type	ETGs	LTGs
MATLAS	179 (68.8%)	53 (8.7%)
CFIS	7 (2.7%)	113 (18.5%)
NGVS/VESTIGE	58 (22.3%)	65 (10.6%)
Total	244 (93.8%)	231 (37.8%)

4.3 Biases of our sample

In this section, we investigate the biases of our final sample. We already discussed the difficulty of combining several surveys (difference of depths, exposure times, MegaCam filters) in Section 3.4. Several other points must be taken into consideration. The first comes from the fact that we do not have all the galaxies from the reference sample: we have 94% of all ETGs, which can be considered as an almost complete sample, but we only have 38% of all LTGs. This comes from the fact that CFIS does not cover the entire Northern sky, unlike the reference sample where galaxies are located at a sky declination δ such as $|\delta - 29°| < 35°$ and a galactic latitude b such as $|b| < 15°$. Therefore, the fractions of CFIS ETGs and LTGs (respectively 2.7% and 18.5%) are not representative of the actual fraction of both morphological types in the field. Actually, if we consider all the field galaxies from the reference sample that are located in CFIS footprint (i.e. 118 LTGs and 42 ETGs, see Table 4.1), LTGs account for 73% of all field galaxies. However, an advantage arising from this situation is that we have a comparable number of ETGs and LTGs. For instance, if we consider CFIS and MATLAS together, we have 186 ETGs and 166 LTGs. Likewise, for NGVS/VESTIGE we have 58 ETGs and 65 LTGs. This is important for the statistical studies afterwards, when we will distinguish galaxies based on their morphological type. If we focus on galaxies of the reference sample belonging to the Virgo cluster, the NGVS/VESTIGE images cover all ETGs but one, and 80% of LTGs. The fact that NGVS/VESTIGE did not cover 100% of all Virgo LTGs is explained by the difference between observing in the direction of the Virgo cluster (what NGVS/VESTIGE did) and being an actual member galaxy of the cluster. This means that some Virgo galaxies are located slightly further away than the NGVS/VESTIGE footprint. In the end, our sample contains 87.5% of all Virgo galaxies from the reference sample.

The environment is also a source of bias. Indeed, when we consider the cluster environment, the only example we have is the Virgo cluster, i.e. a young cluster still in formation (e.g., Tully & Shaya 1984; Gavazzi et al. 1999; Karachentsev & Nasonova 2010; Sorce et al. 2016). Likewise, with MATLAS galaxies (both ETGs and LTGs) we are biased towards groups with ETGs. They are not representative of all groups, as for instance the Local Group contains three massive LTGs and no ETG. In addition, Hoyle et al. (2012)

studied the fraction of ETGs in groups and clusters, and found a higher fraction of LTGs than ETGs for galaxies with a stellar mass lower than $10^{11} M_{\odot}$, then the fraction of ETGs increases with galaxy stellar mass. Oh et al. (2018) found that less massive clusters have similar fractions of both morphological types, whereas massive clusters are dominated by ETGs.

In addition, for the galaxies in common between CFIS and MATLAS, especially for the ETGs, we decided to keep MATLAS images as they are in principle deeper (we refer to Section 6.4.1 regarding this comparison). We should therefore be cautious, as we are more likely to find fainter features around ETGs. Likewise, we should take into account the asinh scaling applied, as it crushed the dynamics in order to enhance the faintest structures. The importance of tidal features in terms of visibility does not represent their actual luminosity.

In addition, we did not subtract a galaxy model to our images, which means the innermost features such as internal shells will likely not be disclosed. An example of this bias is presented in Figure 4.3: it is an illustration of the unsharp masking technique that can be used to reveal the innermost features. In the original image (left panel), the most external shells are clearly visible, however the centre of the galaxy is bright and it is difficult to see any feature. After subtracting from this original image a smoothed version (middle panel), the residual image (right panel) discloses many internal shells. Therefore, without such a method we are biased towards finding features in the outer regions of the galaxy.

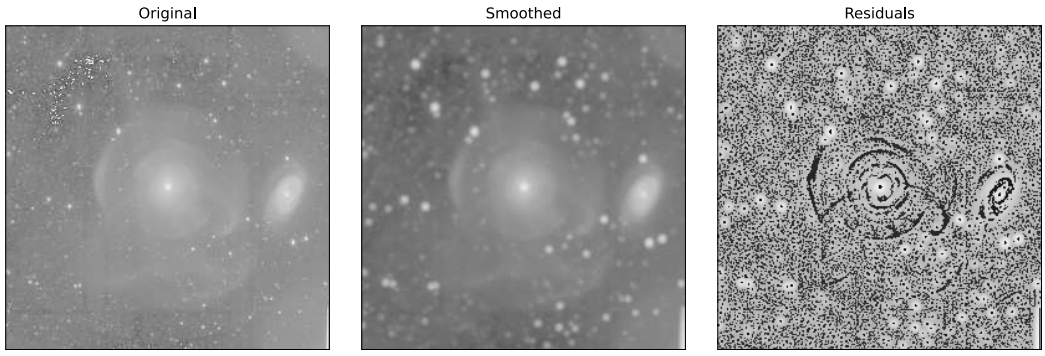

Figure 4.3: Illustration of the unsharp masking technique. *Left*: Original images, with external shells visible. *Middle*: Smoothed version of the original image. *Right*: Residual image after subtracting the smoothed image to the original one. The innermost shells are disclosed.

Finally, we cropped the images to have a fixed size of 30 × 30 arcmin around each target galaxy, but they are not located at the same distance. Our galaxies are located between 5 and 46 Mpc with an average of 25 Mpc, so our field of view (FoV) does not display the same fraction of local environment whether the galaxy is the closest or the furthest away. This FoV corresponds to an average physical size of 250×250 kpc and is equivalent to an average size of 50 effective radii around the target galaxies (with a minimum of 50×50 kpc, or 6.5 effective radii), which is considered enough to visualize the entire galaxy, its neighbourhood and potentially most of its extended tidal features. Likewise, because of the surface brightness dimming (e.g., Tolman 1930, 1934) we do not see the same structures around the closest and furthest galaxies.

4.4 Properties of our sample

Our sample consists of 475 galaxies from the ATLAS^{3D} catalogue, ETGs and LTGs, in the field, groups and Virgo cluster. This section presents some of their properties. As the goal is to study tidal features which are known to be affected by the mass and environment of their host galaxy (e.g., Mihos 2003; Tal et al. 2009; Hood et al. 2018; Bílek et al. 2020), we detail our computations of stellar mass and characterization of environment.

The environment is characterized by the Virgo membership presented in Cappellari et al. (2011a) and by the environmental density ρ_{10}, defined as the mean density of galaxies in a sphere centred on the target galaxy and containing the 10 closest neighbours, and defined in Cappellari et al. (2011b).

The stellar mass is computed differently for ETGs and LTGs. Following Cappellari et al. (2013b), ETGs are characterized by their dynamical mass $M_{\rm JAM}$, estimated by Jeans anisotropic modelling (JAM) (Cappellari 2008). This JAM method is a type of dynamical modelling technique, whose goal is to reproduce the surface brightness of galaxies in a detailed, non-parametric way. Yet dynamical models suffer from degeneracies in the model parameters that are not solved by observations, like for instance from deprojection effects (from observed SB to a 3-D stellar mass distribution); and they usually do not take into account the presence of bars (that are present in around 30% of ATLAS^{3D} sample, see Krajnović et al. 2011). Cappellari et al. (2013b) adopted a different modelling approach that makes empirically-motivated assumptions to reduce the range of model solutions and improve the accuracy of mass recovery. The dynamical modelling approach can be applied to any bound system of stars, provided that integral-field data is available in combination with a careful treatment of the observed SB and orbital distribution (e.g., Cappellari et al. 2006, 2013b). Unlike the estimation of stellar mass from luminosity, it does not require any assumption of a universal initial mass function, which was proven not to represent real galaxies (e.g., van Dokkum & Conroy 2010; Cappellari et al. 2012); even though an assumption on the fraction of dark matter is needed. The relation between the stellar mass of ETGs ($M_{\star,\rm ETG}$), the total galaxy luminosity (L), the mass-to-light ratio ($(M/L)_{\rm JAM}$) and dynamical mass is given by $M_{\rm JAM} \equiv L \times (M/L)_{\rm JAM}$ and $M_{\star,\rm ETG} \approx 0.87 M_{\rm JAM}$. The 0.87 factor linking stellar and dynamical masses comes from the estimated median fraction of dark matter mass of 13% enclosed in $M_{\rm JAM}$ (Cappellari et al. 2013b). For LTGs we did not have such $M_{\rm JAM}$ values, so we estimated their stellar mass $M_{\star,\rm LTG}$ from their K-band luminosity L_K and the stellar mass-to-light ratio $(M/L)_K$. Following Cappellari et al. (2013a), we assumed a fixed $(M/L)_K = 0.8 M_\odot/L_\odot$ for LTGs which ensures agreement between $M_{\rm JAM}$ and L_K at the lowest masses, so the stellar mass is computed as $M_{\star,\rm LTG} = (M/L)_K \times L_K$. Unless mentioned otherwise, we consider ETGs and LTGs together in the following analyses.

Figure 4.4 presents some properties of galaxies: the absolute K-band absolute magnitude M_K as a function of the distance, and the stellar mass as a function of the environment density. Galaxies in the Virgo cluster typically have $\log(\rho_{10}) > -0.4$, and are located at a distance of 16.5 Mpc (hence the concentration of points on the left panel plot). Galaxies from our sample are represented with filled circles, while the remaining ones from the parent sample of ATLAS^{3D} are indicated with empty circles. As mentioned in Section 4.3, our sample contains 94% of all ETGs and 38% of all LTGs from the ATLAS^{3D} reference sample. More precisely, in the Virgo cluster our ETGs represent 98% of all Virgo ETGs from the reference sample (which can be considered as complete) and 79% of all Virgo LTGs. In total, our sample includes 88% of all Virgo galaxies in the ref-

erence sample, which is visible on the right panel of Figure 4.4. For the field, our galaxies represent 49% (93% for field ETGs and 33% for field LTGs) of the field galaxies in the reference sample. We miss some of the LTGs located in the least-dense environment, and some of the most massive spirals in the field/group. We also miss a few of the closest LTGs and some of the most distant LTGs: in overall, LTGs in our sample are closer (median of 25.4 Mpc) than the ones from the reference sample (median of 29 Mpc). The statistical tests revealed that our ETGs and LTGs are located in similar environment densities in terms of means, medians and distributions. Our ETGs are statistically significantly more massive (median mass of $3.3 \times 10^{10} M_{\odot}$) than our LTGs (median mass of $2.2 \times 10^{10} M_{\odot}$), which was also the case in the reference sample. To summarize, our sample of ETGs is almost complete compared to the ATLAS^{3D} reference sample, but we miss many LTGs outside of the Virgo cluster. Our LTGs and ETGs are located in similar environments which makes their comparison simpler. It is also important to note that the ATLAS^{3D} reference sample only contains one cluster, the Virgo cluster which is still dynamically young and not representative of all the cluster types. It does not include isolated galaxies (i.e. with very low environment densities), so our sample is biased towards galaxies in groups and in the Virgo cluster.

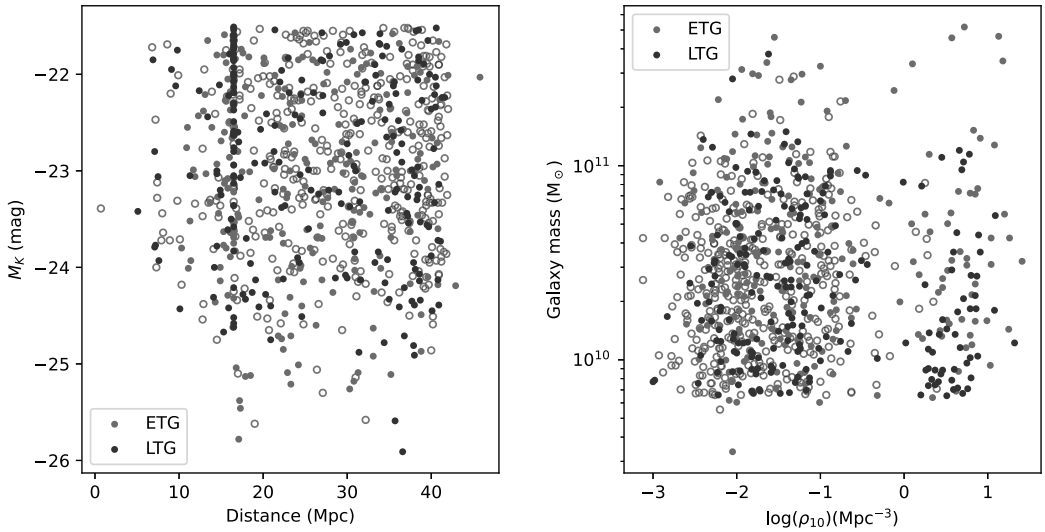

Figure 4.4: Presentation of some properties of our galaxy sample. ETGs are plotted in red and LTGs in blue. Filled circles correspond to our final sample of 475 galaxies, while open circles represent galaxies from the ATLAS^{3D} parent sample that were not selected for our work. *Left*: absolute K-band magnitude M_K in mag as a function of the distance in Mpc. *Right*: stellar mass (in $M_{\odot}$) as a function of the environment density ρ_{10} (in Mpc^{-3}). Galaxies in the Virgo cluster typically have $\log(\rho_{10}) > -0.4$.

4.5 Conclusion

To summarize, this chapter focussed on the galaxy sample used during my PhD work. As the goal was to study LSB tidal features around a large sample of galaxies, we needed a complete sample in order to perform a statistically significant and unbiased study of LSB features. We used the reference sample of nearby massive galaxies compiled by the ATLAS^{3D} collaboration. Then, we determined on which deep images (presented in Chapter 3) were located the galaxies of the reference sample. We performed the intersection

between the sample and the footprint of our images using MOCs. Our galaxy sample finally consists of 475 galaxies, both ETGs and LTGs, located in the field, groups and Virgo cluster. We discussed the biases in our sample and presented some of their properties.

4.6 Résumé du chapitre

Dans ce chapitre, nous nous sommes intéressés à l'échantillon de galaxies autour duquel l'étude statistique des débris de collisions sera effectuée. Afin de conduire une analyse statistiquement significative et non-biaisée, nous avons utilisé l'échantillon complet compilé par la collaboration ATLAS^{3D} de galaxies proches et massives. Il contient 260 ETGs et 611 LTGs, situées dans le champ, dans des groupes et dans l'amas de la Vierge. Cet échantillon dispose de nombreuses informations sur les galaxies (masse, distance, cinématique, gaz...) et avait déjà été utilisé par les travaux précédents dans le cadre desquels cette thèse s'inscrit.

Nous avons ensuite déterminé dans quelles images étaient présentes les galaxies de l'échantillon de référence. Les relevés NGVS/VESTIGE ont observé l'amas de la Vierge, donc la plupart des galaxies de l'échantillon de référence appartenant à cette amas ont été couvertes (87.5%). Le relevé MATLAS a quant à lui imagé 179 ETGs de l'échantillon de référence situées dans les groupes, ainsi qu'une cinquantaine de LTGs qui étaient présentes dans les mêmes champs. Afin de compléter notre échantillon de ETGs en ajoutant les LTGs, nous avons utilisé les images CFIS car elles couvrent une grande partie du ciel de l'hémisphère Nord.

A l'aide des MOCs, qui représentent la couverture spatiale des relevés, nous avons effectué l'intersection entre les galaxies de l'échantillon de référence et l'empreinte des relevés sur le ciel. Nous avons obtenu un échantillon final de 475 galaxies, contenant un nombre comparable de ETGs et de LTGs, situées à la fois dans le champ, dans des groupes et dans l'amas de la Vierge. Nous avons ensuite discuté des biais présent dans notre échantillon, puis nous avons présenté quelques caractéristiques de nos galaxies, en terme de distance, masse ou densité d'environnement.

Chapter 5 Probing the mass assembly of galaxies with LSB structures

Equipped with the annotation tool, the deep images and our sample of galaxies, we were able to annotate tidal features around massive nearby galaxies. In this chapter, we present statistical results of the annotations of LSB features around 475 galaxies. The sections below are based on the results presented in Sola et al. (2022) and Sola et al. (2023) in preparation. In particular, we derived geometrical properties, surface brightness and luminosity of stellar features, and we investigated the impact of the mass and environment of the host galaxy on LSB features. These values were obtained using the analysis tools described in Section 2.3 and from the background-corrected masks presented in Section 2.3.3. We opted for median values rather than mean ones in order to get the most representative values. For galaxies in the Virgo cluster, as the reference band is the *r*-band of VESTIGE, we performed the annotations on VESTIGE images (not NGVS) and report the values here. It is important to note here that in this PhD thesis we applied 'new' methods to retrieve colour and SB information compared to what was presented in Sola et al. (2022), on a larger galaxy sample, which explains the differences between the article and the results presented in this chapter.

In order to determine whether some trends were statistically significant, we applied statistical tests. We used the T-test[1] (Student 1908), that tests the null hypothesis that two independent samples have the same mean. We also applied the 2-sample Kolmogorov-Smirnov (KS)[2] (Hodges 1958) test testing the null hypothesis that the two samples come from the same distribution. In addition, we used Mood's test[3] (Mood 1950) to test the null hypothesis that two samples come from populations with the same median. To check whether the proportions from two populations are the same, the Z-test can be applied. Finally, to check whether two quantities are correlated, we used Pearson'r[4] (Pearson 1895) and Spearman's r[5] tests (Spearman 1904) to respectively test linear correlations and monotonic relationships. If the conditions of applications of the tests are verified, and if the p-value of the tests are lower than a given threshold, namely 0.05 for a confidence level of 5%, the null hypothesis can be rejected.

All galaxies have been annotated by at least two users, while a total of four users were involved in the annotation process. I have annotated 100% of them, Julie Klehammer 74%, Mathias Urbano 70% and Pierre-Alain Duc 22%. The database contains a total of 12119 annotations. Table 5.1 details the number of annotations per feature type, before and after applying our selection process (see Section 2.3.7). In this Table, all the tails are counted, even the ones associated to a companion.

[1]Scipy's stats.ttest_ind
[2]Scipy's 2-sample KS test
[3]Scipy's Mood's median test
[4]Scipy's Pearson's r test
[5]Scipy's Spearman's r test

Table 5.1: Total number of annotations stored in the database as a function of their type. In parentheses is indicated the number of annotations kept after our selection process on the main galaxy, halo, tidal tails and streams.

Annotation type	Number
Main Galaxy	1259 (475)
Halo	1193 (455)
Tidal Tails	616 (353)
Streams	202 (100)
Shells	311
Companion Galaxy	988
High Background	1307
Ghosted Halo	5343
Cirrus	317
Satellite Trail	114
Instrument	467
Total	12119 (10230)

5.1 Tidal features

5.1.1 Characterization of individual features

In this section, we present some global properties of the annotated LSB tidal features. The database contains a total of 353 tails, 100 streams and 311 shells after our selection process. Only 199 tidal tails have for progenitor the target galaxy (and not the companion), hence the analyses are based only on these tails.

5.1.1.1 Census of tidal features

We start by providing a census of the number of tidal features around each galaxy. To that end, we used the selection process described in Section 2.3.7 for tidal tails and streams, i.e. for almost similar features identified by several users around a galaxy, only the largest is kept at the end of the selection. A different process is needed for shells. Indeed, the selection is based on the percentage of intersection between features, yet shells are annotated using one pixel-wide lines so the intersection would be null. In addition, we are facing the issue of the precise delineation of shells which depends on the expertise of the users: if a shell is ambiguous and rather faint, it can be drawn as a single long arc by an expert user, while it can be cut in several parts by a less-expert user. Hence, for the census of shells, we consider the mean number of shells per user and per galaxy. The mean number of tidal features per galaxy, detailed by survey and morphological type, can be found in Table 5.2.

When combining the number of features per morphological type and including all the galaxies, even without tidal debris, tidal tails are more numerous around LTGs than ETGs (with a mean number of 0.55 and 0.3 per galaxy respectively). On the contrary, there are more shells around ETGs (with 0.4 as a mean number per galaxy, against 0.15 for LTGs). For streams, they seem to be more numerous around ETGs (0.25 per galaxy) than around LTGs (0.16 per galaxy).

To determine whether these differences are statistically significant, we applied several tests after checking their conditions of application. If the p-value is lower than the

Table 5.2: Mean number of tidal features per galaxy, taking into account all the galaxies, even the ones without tidal features. In parenthesis is indicated the mean number of debris per galaxy computed only for the galaxies that do present tidal debris.

Type	Tidal tails	Streams	Shells
MATLAS ETGs	0.28 (1.89)	0.28 (1.42)	0.44 (2.82)
CFIS ETGs	0.0 (–)	0.29 (1.0)	0.0 (–)
VESTIGE ETGs	0.38 (1.47)	0.16 (1.12)	0.30 (2.5)
Total ETGs	0.30 (1.74)	0.25 (1.35)	0.4 (2.76)
MATLAS LTGs	0.60 (2.0)	0.06 (1.0)	0.22 (1.48)
CFIS LTGs	0.50 (1.87)	0.27 (1.43)	0.14 (1.71)
VESTIGE LTGs	0.58 (1.65)	0.08 (1.67)	0.13 (1.42)
Total LTGs	0.55 (1.83)	0.16 (1.41)	0.15 (1.55)
Total	0.42 (1.79)	0.21 (1.37)	0.28 (2.28)

threshold of 0.05 we can reject the null hypothesis at a level of confidence of 5%. We could not apply the T-test to the number of tails and shells around ETGs versus LTGs, but for streams the T-test revealed equal means for ETGs and LTGs (p-value of 0.08). The 2-sample KS test revealed different distributions of the number of tails between ETGs and LTGs (p-value of 0.04); but the distributions of the number of streams and shells are similar for both morphological types (with a p-value of 0.54 and 0.46 respectively).

The higher fraction of tails around LTGs than ETGs can be explained by the fact that disc galaxies are more susceptible to produce large and visible antenna-like features, while elliptical galaxies will trigger more plume-like tails (if any feature is produced at all). However, some post-merger ETGs with features can also result from the merging of two spirals, hence the tidal tails will be thinner. The similar fraction of streams observed around ETGs and LTGs was expected. Indeed, as streams originate from lower-mass companions, the morphological type of the primary galaxy should not directly impact the companions. For shells, the important parameters for their formation are the impact parameters, and not so much the morphological type of the primary galaxy.

We computed the fraction of galaxies hosting tidal features. 36% of all galaxies show signs of interaction (i.e. one or more feature of any type); 23% host tidal tails, 16% host streams and 12% shells. A factor that is supposed to impact the census and fraction of tidal features is the galaxy mass and the environment. Their impact will be assessed in Section 5.1.2.

5.1.1.2 Morphology of tails and streams

Qualitative interpretation of global shapes

To have an initial overview of the morphology and location of tails and streams with respect to the target galaxy, we inspected their thumbnails as represented in Figure 5.1. It presents the thumbnails of galaxies displaying tidal features, ordered by increasing mass, from top left to bottom right. The contours of the main galaxy, halo, tails and streams are respectively plotted in purple, yellow, blue and red. Their shapes and sizes show a great variability. For a given galaxy, tidal tails often appear in pairs and are be rather located near the foci of the ellipse defining the halo, whereas streams seem more isolated. Globally, tidal tails seem rounder and broader than streams, while streams seem more elongated and thinner. Note that the thickest tails were referred as *Plumes* in our

study.

Figure 5.1: Thumbnails of all the galaxies having tidal tails or streams. Haloes are represented in yellow, the main galaxy in magenta, tails in blue and streams in red. Only galaxies showing tidal features are represented. The scalebar at the bottom of the plot represents 30 kpc. The galaxies are ordered by increasing mass, from top left to bottom right.

To better compare the individual shapes of each type of tidal structure, we present the footprint of tails and streams without their host galaxy, for LTGs in Figure 5.2 and ETGs in Figure 5.3. All thumbnails have the same physical size (namely 50×50 kpc) and they are sorted by increasing mass of the host galaxy. From these figures, more massive galaxies tend to host larger or more extended tidal tails. For ETGs, tidal tails seem slightly rounder and larger than for LTGs, but ETGs are on average more massive than LTGs. For streams, there is no clear trend neither as a function of the mass of the galaxy nor of its morphological type, which is expected as the material does not originate from the primary galaxy but from a companion. One important point to note is the fact that for a significant number of host galaxies tidal tails and streams look relatively similar.

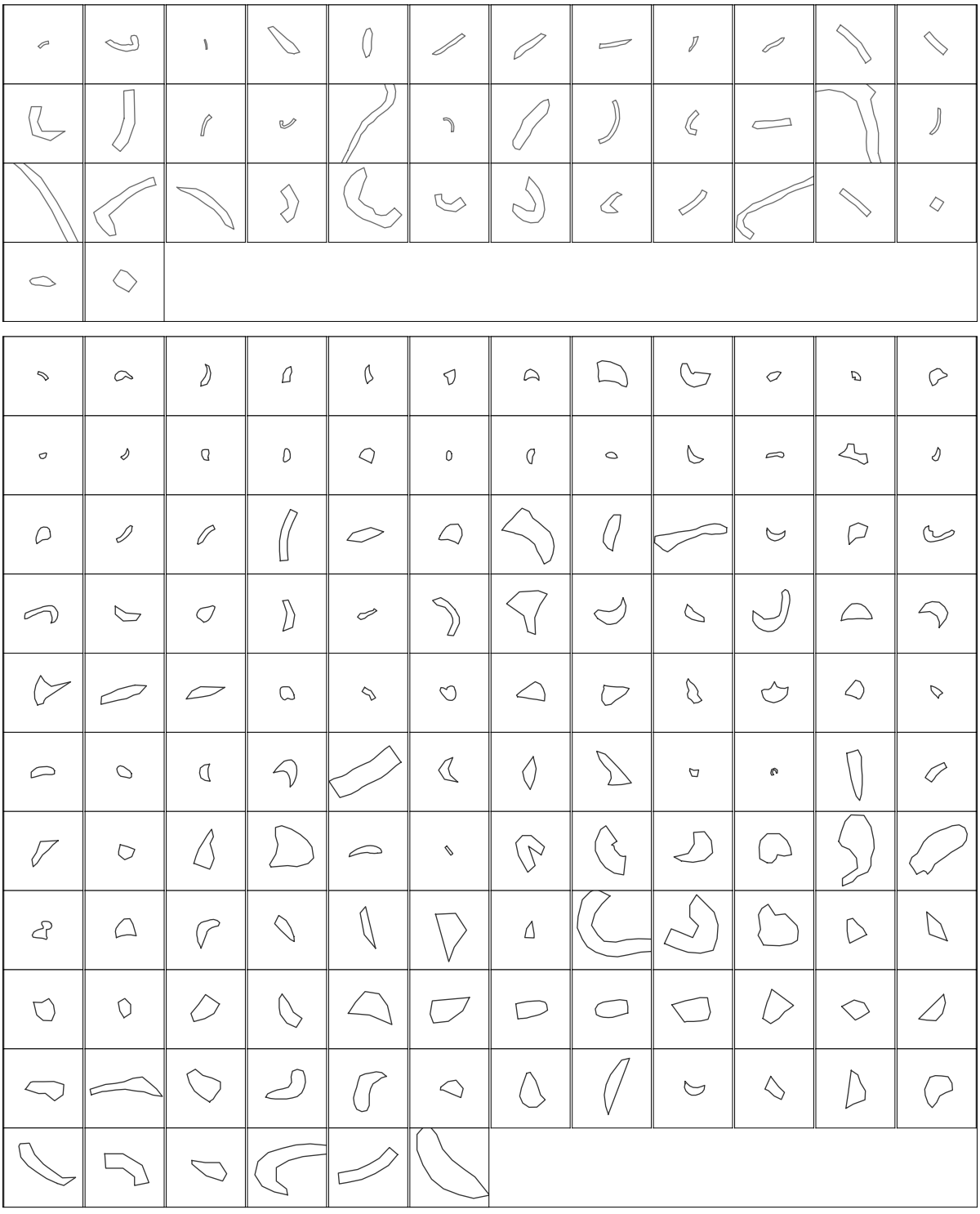

Figure 5.2: Thumbnails of the streams (*top, in red*) and tidal tails (*bottom, in blue*) identified around the LTGs, plotted in boxes of 50×50 kpc. They are sorted by increasing mass of the host galaxy, starting from the top-left for the lightest LTG to the bottom-right for the most massive LTG.

Although streams globally seem more elongated than tails, there is no obvious visual difference between them: some tails look like streams and conversely, for all mass ranges.

Though global trends are observed, clearly the large variability of structures (that

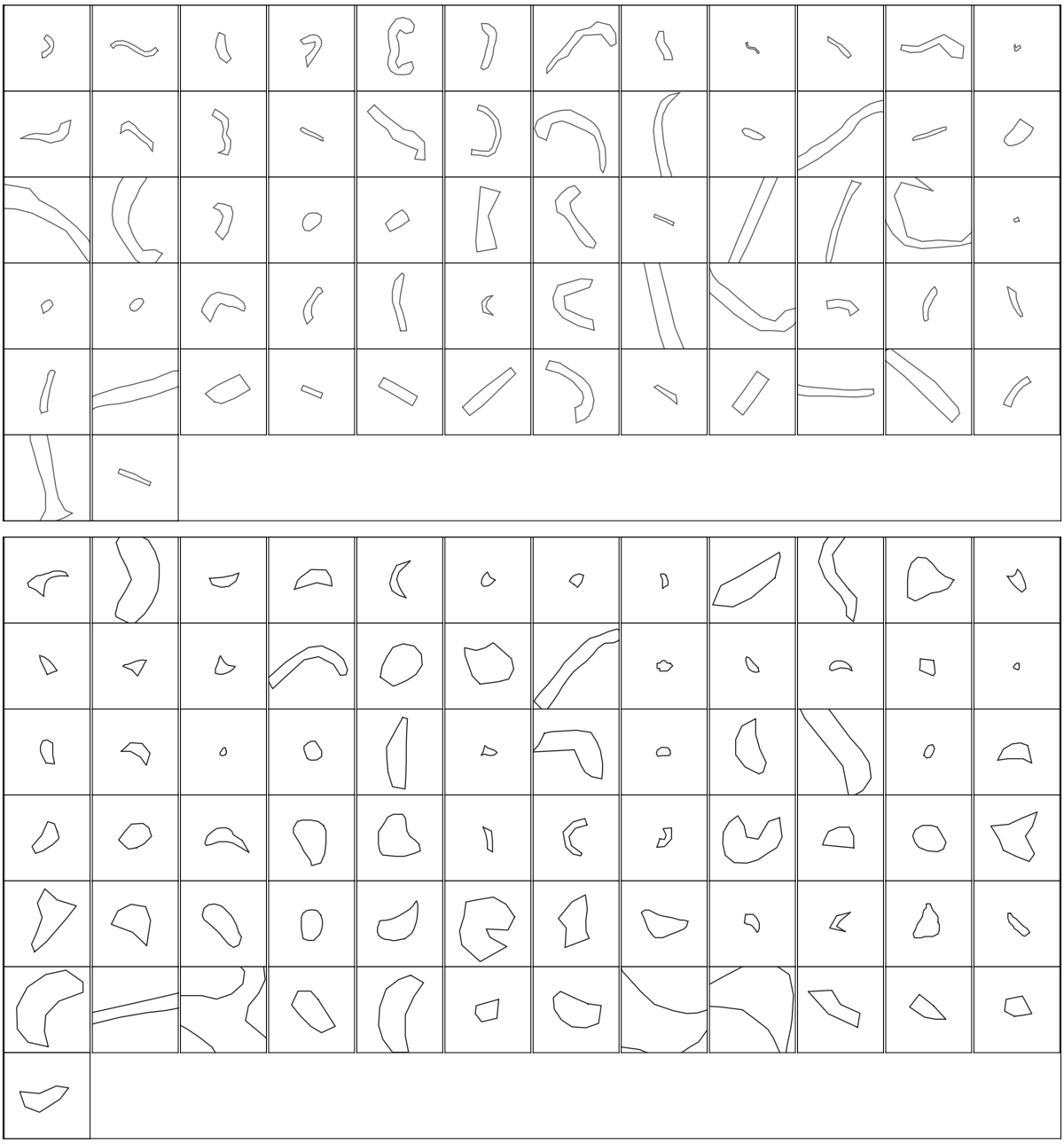

Figure 5.3: Thumbnails of the streams (*top, in red*) and tidal tails (*bottom, in blue*) identified around the ETGs, plotted in boxes of 50×50 kpc. They are sorted by increasing mass of the host galaxy, starting from the top-left for the lightest ETG to the bottom-right for the most massive ETG.

might be partly due to delineation or classification errors) does not allow us to make a sharp distinction between tails and streams simply based on visual inspection. In the following sections, we use a more quantitative approach to assess whether a statistically significant distinction between these two types of features can be found.

Quantitative interpretation: area

Table 5.3 summarizes the median area of tails and streams, making a distinction between ETGs and LTGs, Virgo and field. When considering all galaxies, the median area around

100-120 kpc^2 hardly differ for both tidal features. The large scatter visible is explained by the fact that features with the largest areas are associated to systems showing on-going interactions. Both the T-test and the 2-sample KS test showed that the areas of tails and streams have the same mean and similar distributions (with respective p-values of 0.4 and 0.6).

If we divide our sample as a function of the morphological type, structures surrounding ETGs seem more extended than structures surrounding LTGs. It is the case for tidal tails, as the distributions differ (with a p-value of 7×10^{-4} for the 2-sample KS-test while the T-test could not be applied). However, this difference is not statistically significant for streams (with p-values of 0.1 and 0.18 for the test on the mean and on the distribution). There is no difference between the Virgo cluster and the field for the area of streams (as revealed by the 2-sample KS test, Mood's test and T-test). For the area of tails, although the median values are different (with a p-value of 0.046 for Mood's test), the mean values and the distributions are similar. In order to interpret the results about the morphology of tidal features, we computed their length and width, as discussed in the next paragraph.

Table 5.3: Median area, length and width covered by tidal tails and streams. Associated standard deviations are indicated in parentheses to have an estimate of the spread of the distributions. Areas are expressed in square kiloparsecs, while length and width are in kiloparsecs.

	Tails			Streams		
Type	Area (kpc^2)	Length (kpc)	Width (kpc)	Area (kpc^2)	Length (kpc)	Width (kpc)
All galaxies	119 (323)	23 (20)	5.7 (3.9)	97 (213)	26 (27)	4.0 (2.3)
Virgo	97 (141)	20 (13)	5.2 (2.5)	85 (83)	23 (19)	3.8 (1.8)
Field	137 (357)	23 (21)	6.0 (4.2)	103 (228)	27 (29)	4.0 (2.4)
ETGs	166 (474)	24 (26)	6.9 (4.8)	128 (236)	29 (30)	4.3 (2.4)
LTGs	106 (157)	22 (14)	5.2 (3.0)	85 (163)	23 (23)	3.8 (2.0)

Quantitative interpretation: length and width

The median values of the measured length and width are summarized in Table 5.3, while their histograms are shown in Figure 5.4. Streams are slightly longer than tails (with respectively a median value of 26 kpc and 23 kpc when combining all galaxies), for all galaxy types. However, when applying the statistical tests, these differences are not statistically significant. The 2-sample KS test only revealed that the distributions of lengths of tails and streams differed (when considering all galaxies together, but the T-test could not be applied), but Mood's test confirmed that the median of the distributions were similar. This is visible in Figure 5.4 where the distribution of the length of streams is more extended and flatter than for tails. A few structures reach a length longer than 80 kpc.

If we focus on the width, tidal tails are on average wider than streams. This difference is statistically significant, as the three types of tests give p-values lower than the threshold of 0.05. This is also visible in Figure 5.4 where the shapes of the distributions differ and so do the medians. Almost all streams have a width less than 10 kpc with a peak around 4 kpc; while for tails the peak is around 6 kpc and the distribution is more extended. A few tails are very wide (higher than 15 kpc).

The fact that tidal tails are wider than streams was expected from models : indeed, the width of a tail or stream increases with the velocity dispersion of the stars that form

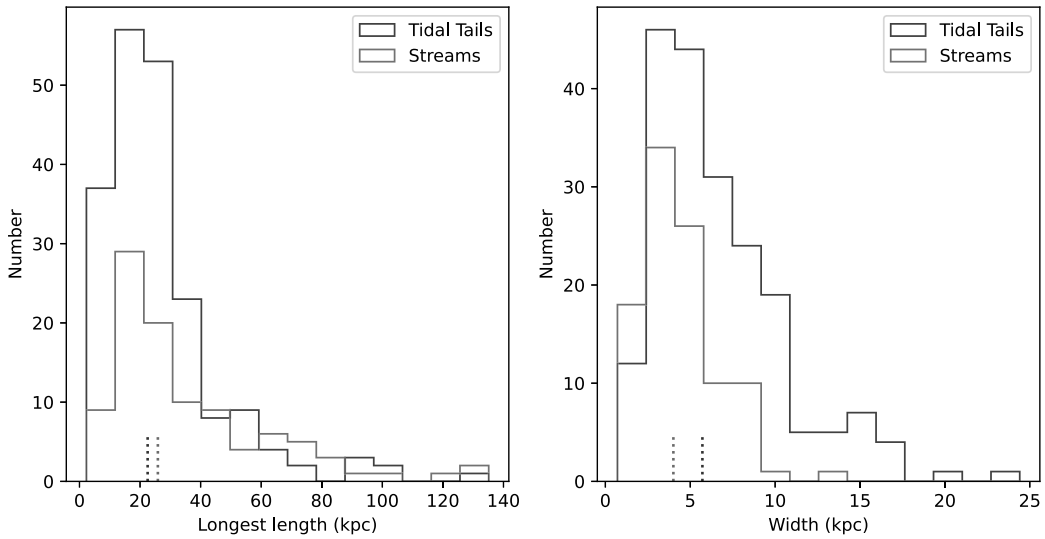

Figure 5.4: Histograms of the length (*left panel*) and width (*right panel*) in kiloparsecs of tidal tails (in blue) and streams (in red). The median of each distribution is represented by the dotted lines.

this structure (e.g., Johnston et al. 1996; Johnston 1998). Yet, the velocity dispersion of a galaxy depends both on its morphological type and on its mass (e.g., Bernardi et al. 2010; Bezanson et al. 2012). The more massive the galaxy, the higher the velocity dispersion and therefore the wider the tail. Hence, since streams originate from low-mass companions, their velocity dispersion is smaller. This contributes to their widths being smaller than for tails. This is consistent with Figure 6 from Hendel & Johnston (2015) who found in their simulation that if the infalling galaxy satellite had a higher mass, the width of the debris increases. Therefore, our results comfort our approach to make a distinction between tidal tails and streams.

If we now focus on the distinction between the field and the Virgo cluster, the statistical tests revealed similar distributions, means and medians for the lengths and widths of tails. Therefore, the environment does not impact the morphology of tails and streams.

To explore the results even further, the histograms of the width of tidal tails and streams as a function of the morphological type of their host galaxy are shown in Figure 5.5. The distributions of the width of tidal tails are very different for ETGs and LTGs: the distribution is flatter and more extended for ETGs, with a median value of 6.9 kpc and few tails having a width between 10 and 25 kpc, while for LTGs the distribution is peaked around 5.2 kpc, with fewer wider tails. This was also expected, as the velocity dispersion is higher for ETGs than for LTGs, producing wider tails. There is no real difference for streams between the distributions for ETGs and LTGs, which are relatively similar. This was also expected, as the morphological type of the primary galaxy is not related to the one of its small companion producing streams.

Overall bending

As seen in numerical simulations of galaxy mergers (e.g., Bullock & Johnston 2005; Cooper et al. 2010; Lux et al. 2013; Gibbons et al. 2014), streams follow approximately the orbit of their progenitors: wrapping around the primary galaxy, they appear as strongly curved. On the other hand, the shape of tails that emanate from the primary galaxy is mainly

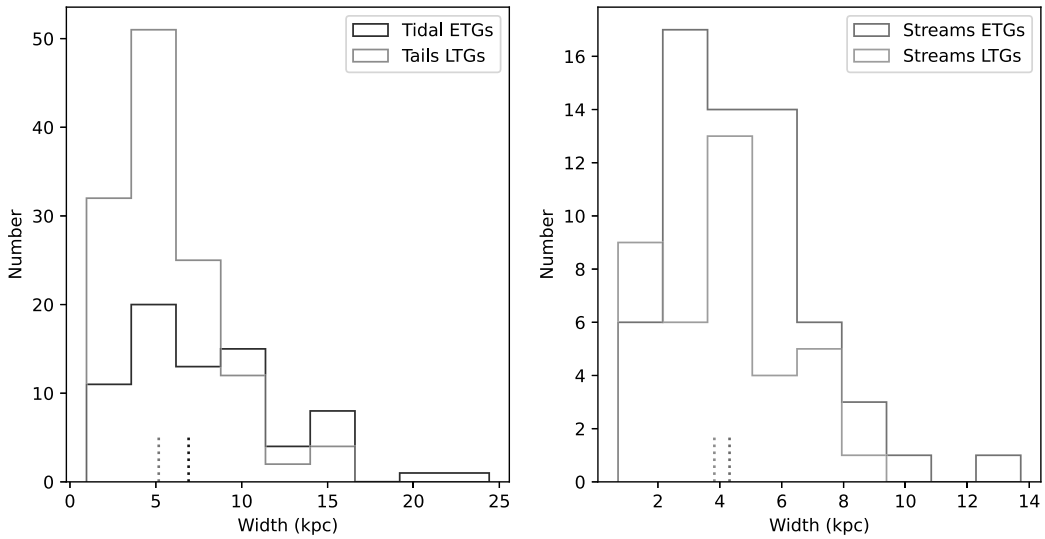

Figure 5.5: Histograms of the width in kiloparsecs of tidal tails (*left*) and streams (*right*) as a function of the morphological type of the galaxies: ETGs are represented by darker shades than LTGs. The median of each distribution is represented by the dotted lines.

driven by tidal forces. Depending on their orientation and until their material falls back on the primary, these structures may appear as relatively straight. Therefore curvature may be another criterion to disentangle streams and young tails.

To obtain a basic estimate of the curvature, we fitted the topological skeleton (as defined in Section 2.3.1.2, it is the medial axis and can be seen as a thinner version of the shape that is equidistant to its boundaries) of tidal features by a linear function, using a least-squares regression. Note that we performed this computation only for not-round polygons, as the skeleton for round polygons does not properly represent the shape. We compared the R^2 correlation coefficient determined for streams and tails: the higher R^2, the closer the feature is to a straight curve.

For streams, the mean and median R^2 are respectively 0.64 and 0.80 with an associated standard deviation of 0.36. For tails, the mean and median R^2 values are respectively 0.53 and 0.53, with a standard deviation of 0.33. There is a large variability for streams where the difference between the mean and the median value of R^2 is more important. In addition, the linear fit for streams is better than for tails. Hence, streams do not appear more curved than tails as we may have expected from simulations (e.g., Bullock & Johnston 2005).

One possible explanation for streams being less curved than tails, besides the projection effects, the uncertainties of the method and/or possible confusions between streams and tails, is that the depth of the survey may not be sufficient to follow the structures over large distances. They must be long enough to get a reliable estimate of the curvature, which is often not the case (see Figures 5.2 and 5.3). Besides, the bending of a stream might be invisible if it is partly hidden by the host galaxy.

Stream progenitors

If streams emanate from a disrupted satellite, remnants of the progenitor may still be visible. As a matter of fact, the presence of a condensation within a tidal structure was one of our criteria to label it as a possible stream, especially if a sign of an S-shape was

Table 5.4: Median inner SB measurements for tidal tails and streams for each type of galaxies, expressed in magnitudes per square arcsecond. Associated standard deviations are indicated in parentheses, to have an estimate of the spread of the distributions.

Galaxy type	Tidal tails SB	Streams SB
All galaxies	26.1 (1.0)	26.8 (0.7)
All ETGs	26.4 (0.8)	26.8 (0.7)
All LTGs	26.0 (1.0)	26.8 (0.6)
MATLAS ETGs	26.3 (0.9)	26.9 (0.7)
MATLAS LTGs	25.4 (1.1)	26.2 (0.2)
CFIS LTGs	26.0 (0.9)	26.8 (0.6)
VESTIGE ETGs	26.5 (0.4)	26.6 (0.5)
VESTIGE LTGs	26.2 (0.8)	26.9 (1.0)

present[6]. Hence, it is interesting to check the percentage of streams that have a progenitor from our annotation database. The progenitor of the stream is defined as follows. If there are no companion or dwarf galaxy in the annotations, the stream is considered orphan. If the stream has an overlap (partial or entire) with a companion or dwarf galaxy, then the progenitor is the companion galaxy. The absence of a progenitor in a genuine stream may indicate that the satellite has been totally destroyed, if the stream was formed long ago, or that it is hidden in the primary galaxy. Statistically, orphan streams should be older than those having a progenitor still visible.

In our results, when taking all CFIS and MATLAS galaxies with streams into account, about 70% of all streams are orphan. In comparison, for the Milky Way and M31, although there is no precise census of the percentage of orphan streams, most of the streams originating from companion galaxies do not present a progenitor, except for the Sagittarius stream (Ibata et al. 1994). As mentioned before, this percentage of orphan streams could be related to the age of these structures or to projection effects, but it might also be linked to misclassifications between tidal tails and streams. Indeed, the absence of a progenitor in the structure made the identification more complicated, which in some cases might have misled users during their classification. We do not see much of a difference between ETG (68% of orphan streams) and LTG (74% of orphan streams) hosts, but we did not expect the stream properties or its progenitor to depend on the morphological type of the primary galaxy.

5.1.1.3 SB of tails and streams

In Table 5.4, we give the overall median inner SB value for tidal tails and streams for each type of galaxy. The distributions of these values are visible in Figure 5.6. Having a median SB of 26.6 mag arcsec^{-2}, streams are fainter than tails by 0.7 mag. This difference is statistically significant, as Mood's test applied on the SB values of tidal tails and streams returns a p-value of 4.7×10^{-7}, and the 2-sample KS test also reveals different distributions (with a p-value of 1.3×10^{-10} (the T-test could however not be applied). As seen in Figure 5.6, the SB distribution of streams is narrower than that of tidal tails. None of them have SB fainter than 27.8 mag arcsec^{-2}.

Several reasons could explain why tidal tails are on average brighter than streams.

[6]Condensations and tidal dwarf galaxies may be present in tidal tails made in major mergers, but being formed in situ, they do not exhibit the S-shape typical of tidally disrupted dwarfs.

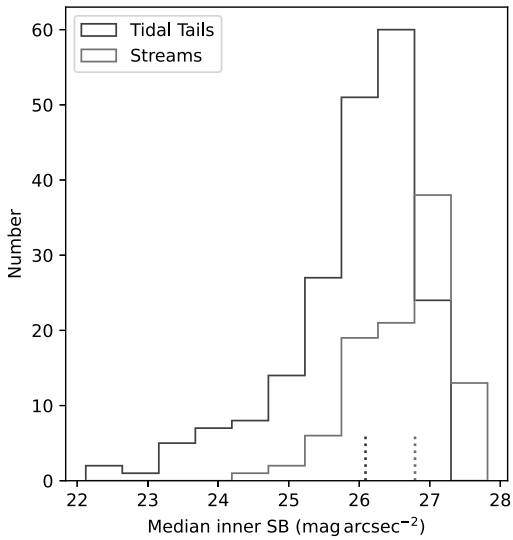

Figure 5.6: Histogram of the median surface brightness value in magnitudes per square arcsecond for tidal tails (in blue) and streams (in red). The median of each distribution is represented by the dotted lines

First, our sample includes on-going or recent mergers. Their tails are too young to face the evaporation process that fade collisional debris. In addition, by definition tails form from massive objects and should therefore be more massive and luminous than the streams which emanate from lower-mass satellites. A second explanation could be related to the different survival times of each type of structures (see e.g., Mihos 1995; Ji et al. 2014; Mancillas et al. 2019): streams tend to be visible for a longer time than tails. Tails could be more fragile than streams and therefore could disappear faster, i.e. when they are older they start to fall back on the galaxy, and combined to projection effects they can look curved like streams. We might be able to detect tails only when they are young enough to keep their typical morphology, hence bright enough.

In this Section, we do not present the SB of shells as we annotated shells as *Curved lines*, i.e. one pixel-wide circular Bezier curve. This line traces the sharpest edge of the shell, but the SB along this line may not be the best way to represent shells. We therefore decided not to include these values here, nor their $g-r$ colour.

5.1.1.4 $g-r$ colours of tails and streams

An additional information that can be retrieved from the background-corrected masks of the annotations and automated aperture photometry is the $g-r$ colour of tidal features. As explained in Section 2.3.5, the $g-r$ colour can be obtained in two different ways. The first consists in retrieving the total flux in g and r bands and computing the colour as $g-r=-2.5\log_{10}(\frac{flux\_r}{flux\_g})$: this is the integrated $g-r$ value. The second is to compute the $g-r$ colormaps from the surface brightness files in g and r bands, and then to get the median value of the colormap. The integrated and median colormap $g-r$ values should coincide in the case where there is no colour gradient. However, if a colour gradient is present, the median $g-r$ will only be equal to the most represented $g-r$ value and may differ from the integrated $g-r$. The advantage of colormaps is that they are directly obtainable from our annotation database, while integrated values are longer to derive but more accurate as it requires aperture photometry on background-corrected images.

It is important to note that in our case, the r-band filters have changed between the observations of MATLAS and VESTIGE, as explained in Section 3.4, therefore we applied a colour-correction to our VESTIGE images to match the old MegaCam filter set. Because of this change of filters, the PSF are different (with a much more extended PSF in MATLAS images), which makes the direct comparison of the colours between MATLAS and VESTIGE complicated. Therefore we chose to report separately the values for these two surveys, indicating the integrated $g-r$. For a comparison with the values computed in Sola et al. (2022), we included for MATLAS the median $g-r$ computed from the colormaps. We present the extinction and colour-corrected $g-r$ values of tidal tails and streams in Table 5.5, while their histograms are shown in Figure 5.7.

Table 5.5: Extinction and colour-corrected $g-r$ colours (in mag) for tidal tails and streams for each type of galaxies in MATLAS and VESTIGE surveys. Associated standard deviations are indicated in parentheses to have an estimate of the spread of the distributions. *Integrated* stands for the median $g-r$ integrated values; *Median* corresponds to the overall median of the median $g-r$ values from colormaps while *Number* indicates the number of structures annotated.

Galaxy type	Tidal Tails $g-r$			Streams $g-r$		
	Integrated	Median	Number	Integrated	Median	Number
MATLAS all galaxies	0.57 (0.20)	0.57 (0.18)	83	0.63 (0.48)	0.60 (0.24)	54
MATLAS ETGs	0.56 (0.21)	0.58 (0.17)	51	0.62 (0.48)	0.59 (0.23)	51
MATLAS LTGs	0.57 (0.19)	0.54 (0.19)	32	1.06 (0.32)	0.9 (0.18)	3
VESTIGE all galaxies	0.32 (0.35)	–	60	0.38 (0.62)	–	14
VESTIGE ETGs	0.34 (0.22)	–	22	0.42 (0.58)	–	9
VESTIGE LTGs	0.31 (0.41)	–	38	0.34 (0.69)	–	5

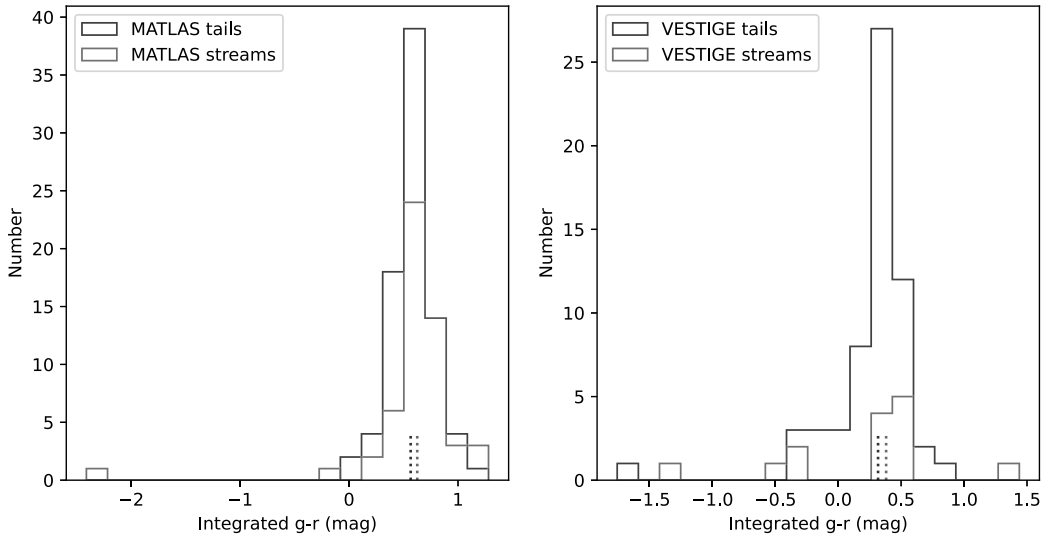

Figure 5.7: Histograms of the extinction and colour-corrected integrated $g-r$ values (in mag) of tails and streams. Tidal tails are plotted in blue and streams in red. The dotted vertical lines represent the median of the distributions. *Left*: Histograms of the integrated $g-r$ colours of MATLAS tails and streams. *Right*: Histograms of the integrated $g-r$ colours of VESTIGE tails and streams.

We now focus on MATLAS and VESTIGE separately (as a comparison between both is not straightforward). From the Table and the Figure, streams are redder than tidal tails by only 0.06 mag when considering the integrated $g-r$ of features around all galaxies.

The T-test, 2-sample KS test and Mood's tests did not reveal any statistically significant difference for the integrated $g-r$ between tails and streams. Although the trend is not statistically significant, we would expect to find tails bluer than streams due to the presence of young stars along the tails. In gas-rich mergers, gas is expelled along tidal tails, just like the stars, and may be compressed in the collisional debris and trigger star-formation (e.g., Olson & Kwan 1990; Jarrett et al. 2006). On the contrary, dwarf satellites are usually gas poor, and if they have been stripped the colour of their streams will reflect that of their old stellar populations. In addition, an age effect might come into play, as due to their lower visibility timescale, tails tend to be observed at a younger age than streams. This age effect might even be stronger when taking into account the age-metallicity degeneracy. Old stars of satellites are bluer due to lower metallicities, hence streams should appear bluer than tails. However, tails seems slightly (but not significantly) bluer than streams, hence they should be even younger than originally thought.

The statistical tests revealed no difference between the colours of tails around ETGs and around LTGs. It is not possible to properly compare streams around ETGs and LTGs as their number is too low (only 3 streams around MATLAS LTGs and 14 streams around VESTIGE galaxies). Our measure of the colour of streams surrounding MATLAS galaxies, with an integrated $g-r$ of 0.63 mag, is in agreement with Martínez-Delgado et al. (2023) who determined a $g-r$ value between 0.5 and 0.8 mag for 24 streams around local galaxies.

Finally, there is a difference between the $g-r$ colours of VESTIGE and MATLAS: features around VESTIGE galaxies are bluer than around MATLAS images. Although we applied a correction in the r-band to try to have comparable r-band, the effect of the PSF has not been accounted for and might explain the visible differences. We will investigate this in Section 6.3.1.2.

5.1.1.5 Luminosity and stellar mass

From the extinction-corrected and colour-corrected flux values, as well as from their integrated $g-r$ colour, we can compute the stellar mass contained in tidal features. The details of the computations are presented in Section 2.3.6. We present the results in Table 5.6 for tidal tails and streams. The median values of luminosity in the r-band and stellar mass are indicated, with their associated standard deviation.

Table 5.6: Median luminosities in the r-band (in $L_\odot$) and median stellar mass (in $M_\odot$) of tidal tails and streams. Associated standard deviations are indicated in parentheses.

	Tails		Streams	
Type	Luminosity (10^8 $L_\odot$)	Mass (10^8 $M_\odot$)	Luminosity (10^8 $L_\odot$)	Mass (10^8 $M_\odot$)
All galaxies	2.2 (6.6)	4.4 (18)	1.2 (10)	2.3 (23)
ETGs	2.0 (6.9)	4.5 (14)	1.2 (11)	2.3 (24)
LTGs	2.5 (6.3)	3.8 (21)	1.5 (1.4)	2.0 (11)

From this Table, tidal tails are almost twice as luminous and massive as streams, with an overall median mass of $2.2 \times 10^8 L_\odot$ against $1.2 \times 10^8 L_\odot$ for streams. We find again the same trends than in Sections 5.1.1.2 and 5.1.1.3 with tails being brighter than streams. It must be noted that there is a large scatter as a few structures are very massive, but they correspond to tails originating from on-going interactions or to features partly embedded in the halo (hence with a higher luminosity due to the contribution of the halo). There is no statistically significant difference between the luminosity of tails (and streams) around

ETGs and around LTGs. No difference is expected for streams, as they originate from a lower-mass companion.

In the following sections, we will rely on the luminosity of tidal features rather than their mass, as the estimation of the stellar mass relies on $g-r$ colour measurements which are difficult to properly compare between MATLAS and VESTIGE due to the change of r-band MegaCam filters.

5.1.1.6 Concentricity and radii of shells

A total of 311 shells have been annotated. Geometrical measurements of these features are presented here. Shells have been annotated using *Curved lines* and thus measuring their area is irrelevant. Such an annotation faces a major issue: the selection of the beginning and ending of a shell might be different for an expert or novice user. The former may be aware of the well-shaped circular shells in idealized numerical simulations and consider as a single structure a shell that might be divided into several arcs by the less expert users. Nevertheless, interesting metrics can still be computed, such as the concentricity or their radius. Indeed, there are several types of shells defined by their position and orientation with respect to the target galaxy (e.g., Wilkinson et al. 1987; Prieur 1990). Type I shells are concentric and centred on the galaxy; Type II shells are circular and randomly distributed around the galaxy while type III are irregular and not concentric, either because their shape is not typical or because there are too few shells.

Concentricity - From numerical simulations, shells are usually formed as concentric structures (e.g., Pop et al. 2018; Ebrová et al. 2021; Bílek et al. 2021), a prediction we can directly test with our observations. To compute the concentricity, we assume that the *Curved line* defining the shell is a circular arc, and we compute its centre[7]. Since the *Curved line* is a cubic Bézier curve, it may differ from a circular arc if the user did not draw the shell properly. In that case, we only consider the starting, middle and ending point of the curve and we compute the centre of the circle passing through these three points. We then compute the distance between the shell centre and the centre of the host galaxy. The histogram of these distances is displayed in the right panel of Figure 5.8. A distance equal to zero means the shell is centred on the galaxy (i.e. concentric), while larger distances indicate a higher deviation from concentricity.

There seems to be several shell populations. As the median difference between the galactic centre and the centre of the shell is 2.4 R_e, for the majority of shells their centre is still located inside the galaxy and we can consider them as relatively concentric. There are other cases where the relative distance between the centres is large ($> 5R_e$), that may correspond to bad approximations of the shells as circular structures or to real non-concentric shells such as those of Type III.

Radius - The shell radius is another metric which can easily be compared to predictions from simulations. It is computed as the on-sky separation between the point lying in the middle of the *Curved line* defining the shell and the centre of the galaxy host. Such a computation needs to take into account the fact that for a given galaxy, multiple users might have annotated the same shells. Simply computing the mean shell radius per galaxy averaged over all the users would introduce biases. Indeed, as mentioned earlier, the number of shells annotated depends on the expertise of the user and on the shape of the shell itself. Clearly defined shells will be annotated as one structure, while less-defined

[7]The centre of the shell is hence the centre of the circle that passes through the circular arc defining the shell.

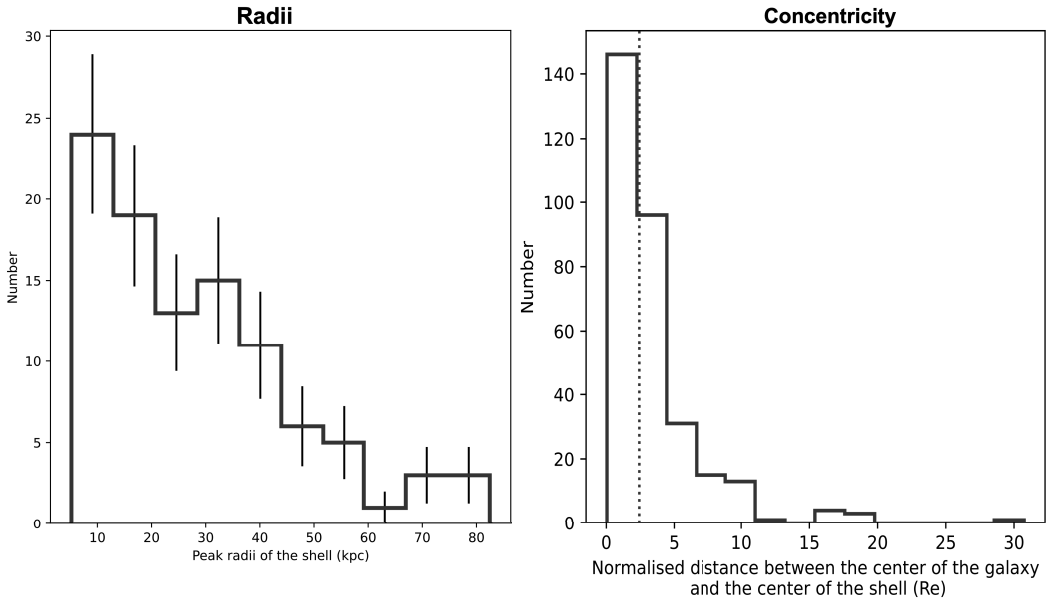

Figure 5.8: *Left*: Histogram of the peak radii of shells (in kiloparsecs) for all galaxies, with Poisson uncertainties plotted as the black error bars in each bin. *Right*: Concentricity test: histogram of the distance in effective radius[8] between the galactic centre and the centre of the shell. Higher distances correspond to larger deviations from concentricity.

ones will be annotated as several shells. This means that the less well-defined ones will have a higher impact and count for more features.

The goal here is to have an idea of the most representative radii, i.e. the typical radii of our shell populations. To do this, we start by plotting the radii of all shells from all users. We draw the corresponding density plot. We consider that the most representative radii correspond to the inflection points of the density plot. The inflection points, referred here as 'peak radii', are computed using a Scipy function to find peaks in a 1-D array using a Ricker wavelet transformation[9]. The histogram of the peak radii of shells for all galaxies are represented in the left panel of Figure 5.8.

From the left panel of Figure 5.8, most of the shells have a radius smaller than 60 kpc with a few extending to 80 kpc. For instance the shells of the prototypical galaxy NGC0474 range from 10 to 50 kpc (Bílek et al. 2021). We do not observe shells beyond 80 kpc, which seems at odds with some simulations. Whereas they extend to 120 kpc in Ebrová et al. (2019) and Karademir et al. (2019), some of the shells in Pop et al. (2018) have a radius reaching 150-200 kpc. Obviously the comparison is not straightforward as the shell orientation (not well constrained from our annotations of real systems) and differences in surface brightness need to be taken into account.

5.1.2 Effect of galaxy mass and environment on tidal features

5.1.2.1 Fraction of galaxies with features versus galaxy mass

Tidal features are the remnants of past or ongoing mergers and accretion events. In the hierarchical paradigm, more massive galaxies have undergone more of these events

[8]The values of the effective radii for all our galaxies are given in Tables 3 and 4 from Cappellari et al. (2011a).

[9]Scipy find_peaks_cwt, https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks_cwt.html

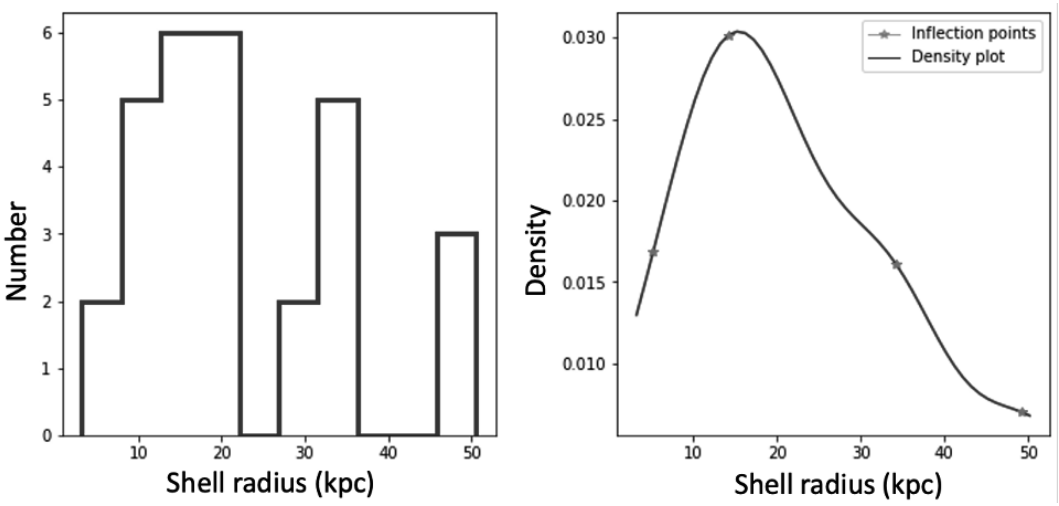

Figure 5.9: Figure 13 from Sola et al. (2022). Illustration of the peak radii of shells for a given galaxy (NGC0474). *Left*: Histogram of the shell radii for this galaxy. *Right*: Corresponding density plot with the inflection points of the curve indicated by red stars.

compared to less massive ones, so tidal features are expected to be more numerous around massive galaxies. To verify it, we investigated the frequency of galaxies with tidal debris as a function of the mass of the host galaxy. The results are represented in Figure 5.10.

A first point to note is that the total fraction of galaxies with any type of debris is 36%, it is of 23% for galaxies with tails, 16% for streams and 12% for shells. A direct comparison of this fraction with the literature is not straightforward: indeed, the definitions of the different types of tidal features and of their prominence, surface brightness limits of the surveys, distances of the galaxies, detection methods, image processing techniques, or sample selection differ from one study to the other (e.g., Atkinson et al. 2013; Hood et al. 2018). This results in large discrepancies in the fraction of tidal features in nearby galaxies, that can range from a few percent up to 70% (e.g., Malin & Carter 1983; Schweizer & Seitzer 1988; van Dokkum 2005; Tal et al. 2009; Bridge et al. 2010; Nair & Abraham 2010; Kaviraj 2010; Miskolczi et al. 2011; Kim et al. 2012; Sheen et al. 2012; Adams et al. 2012; Atkinson et al. 2013; Duc et al. 2015; Hood et al. 2018; Bílek et al. 2020; Jackson et al. 2021; Vázquez-Mata et al. 2022; Yoon et al. 2023). We can more easily compare our results with Duc et al. (2015) and Bílek et al. (2020) as they focussed on ETGs MATLAS: they found around 10-15% of their galaxies with tails, a similar fraction with streams and similarly for shells. If we restrict our results to MATLAS ETGs, we find similar results: 15% of galaxies host tidal tails, a similar fraction for shells, slightly more hosting streams (20%), and an overall 34% of galaxies hosting any type of tidal debris. On the simulation side, Pop et al. (2018) produced a census of shells in the Illustris simulation and found an incidence of 20-30% of shells which is higher than our finding, but they did not apply any SB limit so they would find more features. Martin et al. (2022) also produced a census of tidal features in simulated deep images considering different depths, and found fractions roughly similar to ours (between a few percents and less than 20%) for a SB limit of 29 mag arcsec^{-2}. However, here again we do not have the same criteria for the galaxy sample, images and detection and definitions of tidal features, without mentioning the fact that simulated images are not contaminated by sources of pollution such as artefacts, internal reflections in the optics of the telescope and camera or by Galactic cirrus.

The second point of interest is the fact that there is an increase of the fraction of

galaxies hosting debris with galaxy mass, for all types of tidal features. The fraction of galaxies with any type of debris is around 20-30% for galaxy masses lower than $3\times10^{10}M_{\odot}$, but it increases sharply up to 60% for the highest mass galaxies. This increase of the fraction of galaxies with debris with mass is statistically significant, as Spearman's test revealed a positive monotonic relation between both quantities (with a p-value of 0.02). In addition, one can see that this increase is driven by the three types of tidal features, as the fraction of galaxies with tails approximately goes from around 20% to 30%; from 5% to 22% for shells and from 12% to 25% for streams, with a sharper increase around the same mass threshold.

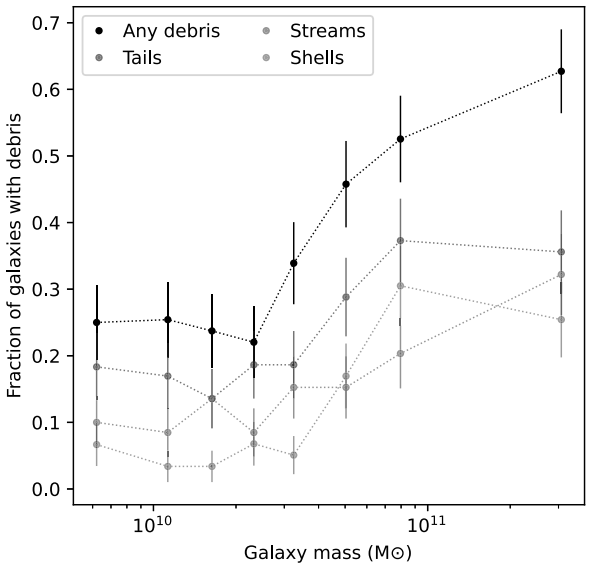

Figure 5.10: Fraction of galaxies that have tidal features as a function of the galaxy mass (in $M_{\odot}$), per bin of mass. Each bin approximately contains the same number of galaxies. The fraction of galaxies hosting any type of debris is plotted in black, it is shown in blue for tails, red for streams and green for shells. The error bars represent the standard deviation on proportions in each bin.

In the literature, the dependency between the appearance of tidal features and galaxy mass had already been observed. For instance, Atkinson et al. (2013) noted a mass-dependency, with an increased fraction of galaxies having linear features, shells and fans of stellar light for galaxy masses $> 10^{10.5} = 3.16\times10^{10}M_{\odot}$ which is very close to the mass threshold that we observe. Likewise, Duc et al. (2015) and Bílek et al. (2020) reported an increased fraction of shells, streams and disturbed isophotes with increasing galaxy mass, mostly for galaxies with a mass $> 10^{11}M_{\odot}$. Furthermore, other studies such as Yoon & Lim (2020) find a strong correlation between the fraction of ETGs with features and increasing galaxy mass: 30-40% of their massive ($M_{\rm dyn} > 10^{11.4}M_{\odot}$) galaxies have tidal features, while 2-5% of their less massive ($M_{\rm dyn} < 10^{10.4}M_{\odot}$) ETGs show features. Similarly, Vázquez-Mata et al. (2022) found the fraction of galaxies with any type of debris to increase with stellar mass, from 10% to 30% for their most-massive galaxies. On the simulation side, Martin et al. (2022) also noted a slight increase of the fraction of galaxies with shells, streams or tails with galaxy mass for a SB limit of 29 mag arcsec^{-2}. Although a precise comparison between studies is complicated, the trends and mass thresholds are globally in agreement with the one we found.

The fact that we see an increase of the fraction of galaxies with debris with galaxy

mass is consistent with hierarchical models of galactic evolution. As more massive galaxies have undergone more mergers, there should be more signs of interactions which is what we observe. The presence of a mass threshold above which the fraction of galaxies with debris increases in a steeper way will be investigated in the following sections.

5.1.2.2 Luminosity of features versus galaxy mass

To further explore the results, we investigated the evolution of the luminosity of tidal features as a function of galaxy mass. Luminosity, derived from the flux, is an important quantity to measure as it can be compared with that of simulated tidal features. In our case, the flux was automatically retrieved through aperture photometry with the masks of the annotations, as described in Section 2.3.3.

For each galaxy, we summed the flux in all the tidal features (tails and streams) and we divided it by the flux of the entire halo annotation. We remind here that the entire halo annotation includes the inner part of the galaxy as visible in shallow surveys, annotated as 'main galaxy'; and the external LSB part of the halo which is located between the external boundaries of the main galaxy and the external boundary of the galaxy seen in deep images, that has a 'donut' shape. The study of this luminosity fraction enables us to obtain clues about the assembly history of galaxies. Figure 5.11 presents the histogram of this fraction, named $f_{L,tidal}$, for the galaxies that do display tidal features. As a reminder, from Section 5.1.2.1 we showed that in overall 64% of our galaxies do not display tidal features so $f_{L,tidal} = 0$. It implies that the median value of $f_{L,tidal}$ is equal to zero if we consider all galaxies. This is why we represent in this Figure only the galaxies that do have tidal features and why we represent the mean values instead of the median. From Figure 5.11, it appears that the mean fraction of luminosity in tails and streams is 3.6% (and it is of 3.7% for tails and 2.1% for streams), when considering only the galaxies with tidal features. The distributions are dominated by very low $f_{L,tidal}$, as the median values for tails and streams is 1.8%, against 1.9% for tails and 0.7% for streams.

In Figure 5.12, we investigated the evolution of $f_{L,tidal}$ with galaxy mass for all our galaxies. As the majority of our galaxies do not host tidal features the distribution of $f_{L,tidal}$ is dominated by null values so we represented the mean value and not the median per bin of mass. The errorbars represent the 16% and 84% quantiles in each bin. The overall mean fraction of flux in tidal features around each galaxy is 1.2% when combining tails and streams together ($f_{L,tidal}$=0.9% for tails only and 0.3% for streams only). There seems to be a slight increase of $f_{L,tidal}$ with galaxy mass. This trend is statistically confirmed with Spearman's test, that revealed the presence of a positive monotonic relation (with a p-value of 3×10^{-5} and a correlation coefficient $r = 0.2$) when considering $f_{L,tidal}$ for tails and streams together. The same trend is verified when considering tails and streams separately. However, Pearson's test could not be applied, due to the non-normality of our data points. The increase of $f_{L,tidal}$ seems to be more important around 4-7$\times 10^{10} M_{\odot}$

A direct comparison with literature is complicated, as we do not have the same definitions nor the same methods. Jackson et al. (2021) found stellar material in the outskirts of their massive central galaxies, mostly in the form of tidal and merger features, to represent a few percent of the total stellar mass (8% outside 2Re, which likely includes a part of what we defined as halo). Huang & Fan (2022) also studied the fraction of the flux in tidal features around massive ETGs ($> 10^{11} M_{\odot}$) and found their $f_{L,tidal}$ increased by a factor of 2 from 0.5% for stellar masses around $10^{11} M_{\odot}$ up to 1% for masses $> 10^{12} M_{\odot}$. On the simulation side, Martin et al. (2022) studied tidal features from cosmological simulations and found the flux fraction in tidal features to range between 1% and 10% for their

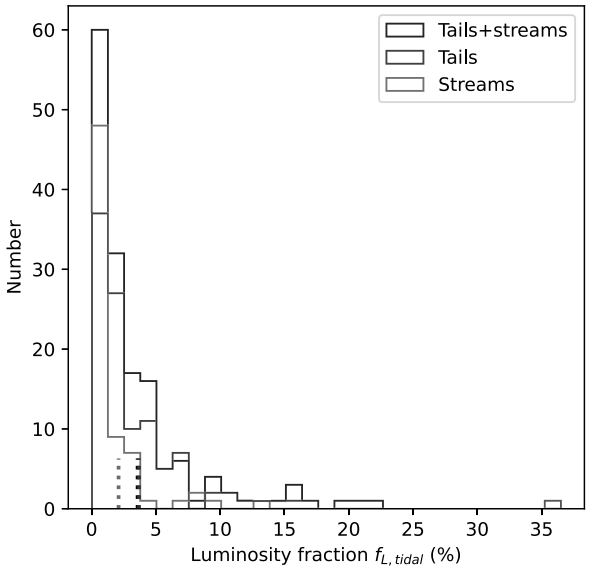

Figure 5.11: Histogram of the fraction of luminosity $f_{L,tidal}$ in tidal tails and streams, defined as the luminosity in these structures divided by the luminosity of the entire halo. $f_{L,tidal}$ is computed by summing the luminosity of all tidal features around each galaxy. Only galaxies with debris are represented. This fraction is represented in black when considering tails and streams together, in blue for tails and in red for streams. The vertical dotted lines represent the mean values of the distributions.

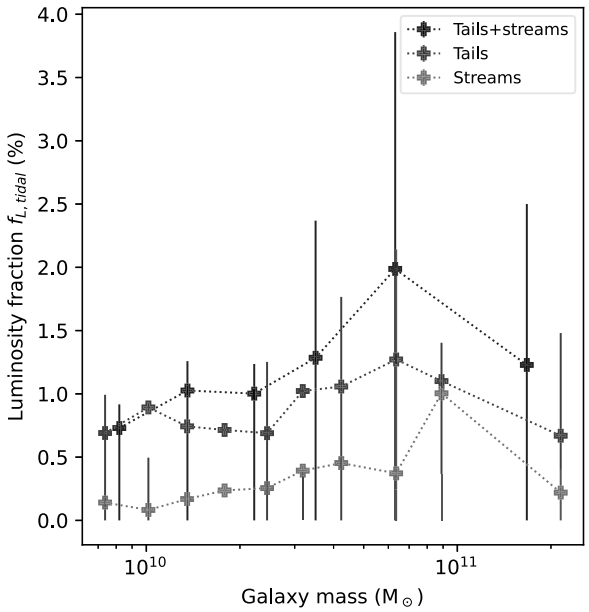

Figure 5.12: Scatter plot of the mean fraction of luminosity in tidal features, $f_{L,tidal}$, as a function of galaxy mass (in $M_\odot$). All galaxies are considered, even the ones that do not display tidal features. Each bin approximately contains the same number of galaxies. The mean value of $f_{L,tidal}$ in each bin of mass is shown in black for tails and streams together, in blue for tails and in red for streams. The errorbars represent the 16% and 84% quantiles in each bin.

most massive galaxies. They note a strong increase of $f_{L,tidal}$ with galaxy mass, around a threshold of $10^{10.1} M_{\odot}$ in stellar mass. Although we do not find such a sharp increase, we still observe a similar trend. This increase of $f_{L,tidal}$ with galaxy mass is consistent with hierarchical models, as more massive galaxies should have undergone more mergers, hence more features and a higher $f_{L,tidal}$. For instance, Stewart et al. (2008) showed with numerical simulations that more massive haloes have undergone more merger events, although the dependence they found is relatively weak. Likewise, Rodriguez-Gomez et al. (2015) found increased merger rates with galaxy mass in the Illustris simulations. From the results and the literature it seems that above the observed mass threshold, the growth of galaxies is dominated by mergers as accretion may no longer be possible (because of the too hot halo that prevents the infall of fresh inflows of gas), as the fraction of galaxies with debris and the fraction of flux contained in them increases sharply. Below this threshold, gas accretion dominates the evolution of galaxies, with lower $f_{L,tidal}$ and lower fraction of galaxies with debris.

5.1.2.3 Fraction of galaxies with features versus large-scale environment

We now focus on the impact of the environment on collisional debris. The environment can be parameterized in different ways. We start by considering the large-scale environment, characterized by the environmental density ρ_{10} (defined as the mean density of galaxies in a sphere centred on the target galaxy and containing the 10 closest neighbours, see Cappellari et al. (2011b) and Section 4.1): this definition is helpful to distinguish galaxies in the field, in groups and in the Virgo cluster. A second aspect that will be investigated later in Section 5.1.2.5 is the small-scale environment, i.e. the presence of close companions susceptible to generate tidal interactions.

Tidal features in clusters result from competing processes that either tend to generate them (e.g., galaxy harassment, interactions with the cluster's potential) or destroy them (e.g., dynamical heating, ram-pressure stripping). Although tidal features are expected to be phase-mixed more quickly in clusters, it is complicated to determine which effect dominates as they occur simultaneously. In the field, their lifetime will be longer as they will remain loosely bound to the galaxy, except if another merger or interaction occurs.

In a similar way than for the mass, we studied the evolution of the fraction of galaxies with tidal features as a function of the environment density: the results are presented in Figure 5.13, with galaxies in the Virgo cluster typically having $\log_{10}(\rho_{10}) > -0.4$. Several trends are visible. First, there is a net peak in the fraction of galaxies with any debris for galaxies characterized by $-2 < \log_{10}(\rho_{10}) < -1.5$: it rises from 25% to a maximum of 50%, then decreases to 30%. This trend is mostly driven by tidal tails (rising from 10% to 40% then 25%). A second smaller peak arises between $-1.5 < \log_{10}(\rho_{10}) < -1$. For the fraction of galaxies with shells or streams, there is also a peak (although less pronounced) for $-2 < \log_{10}(\rho_{10}) < -1.5$. Then, the fraction of galaxies with debris of any type decreases for $-1 < \log_{10}(\rho_{10}) < -0.5$, before rising sharply in the highest density bin: it goes from 30% to 50%. Again, tidal tails are driving this trend, while the fraction of galaxies with shells or streams keeps decreasing. If we divide galaxies located in Virgo and the ones in the field, the overall mean fraction of galaxies with debris of any type is about 35% in the field and 37% in Virgo. This is not statistically significant as shown by the Z-test on proportions (with a p-value of 0.75), and neither is it for each type of tidal features considered individually.

Therefore, there is a non-monotonic relation between the fraction of galaxies with debris and the environment density. The increase of this fraction outside the Virgo cluster,

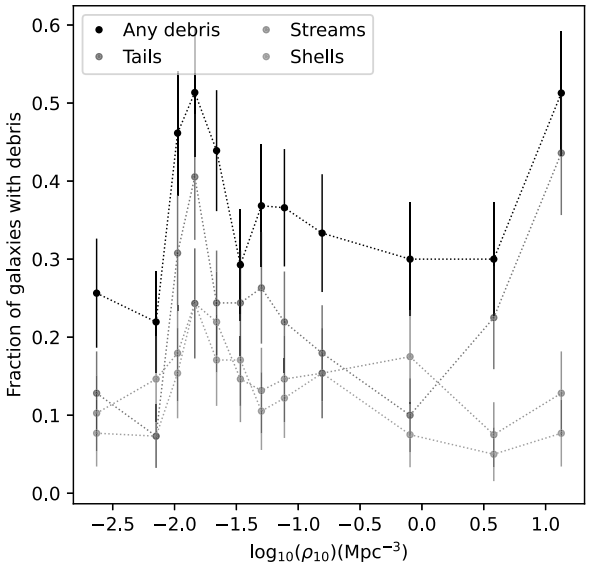

Figure 5.13: Scatter plot of the fraction of galaxies that have tidal features as a function of the environment density ρ_{10} (in Mpc^{-3}), per bin of ρ_{10}. Each bin approximately contains the same number of galaxies. The fraction of galaxies hosting any type of debris is plotted in black, it is shown in blue for tails, red for streams and green for shells. The error bars represent the standard deviation on proportions in each bin. Galaxies in the Virgo cluster typically have $\log_{10}(\rho_{10}) > -0.4$.

i.e. in the group environment for $-2 < \log_{10}(\rho_{10}) < -1.5$, was also seen by Bílek et al. (2020) for their shells and tails. We can make the hypothesis that this rise comes from interactions between galaxies inside galaxy groups. Likewise, as the net increase of that fraction in the highest density bin in Virgo is driven by tails, one can assume that interacting galaxies could be responsible for this trend. This assumption will actually be verified in Section 5.1.2.5, as the visible peaks originate from galaxies that are currently undergoing interactions. For shells, we see an overall decrease of the fraction of galaxies with shells with increasing environment density (except in the group environment where a peak is visible). This trend was also seen by Malin & Carter (1983) who found that the frequency of shells decreased with environment density. This could be explained by the fact that shells originate mostly from radial mergers with low velocity encounters (e.g., Quinn 1984), unlike the high velocity encounters in clusters. Additional tidal stripping from the cluster potential could also explain this decrease (e.g., Mihos 2003). Then, one may wonder about the decrease of the fraction of galaxies with debris in the intermediate region between the field/group and the Virgo cluster. This region corresponds to the outskirts of the Virgo cluster where the interactions between galaxies are less frequent (although the environment density is higher than in the field). In Figure 5.14 we see that there is indeed a decrease of the fraction of galaxies with debris with increasing distance to the centre of the Virgo cluster (taken as M87).

Therefore, we can say that the fraction of galaxies with debris is high for galaxies that have a nearby companion, independently of the large-scale environment. For Virgo, this can be related to galaxies in small groups that are currently infalling the cluster or to interactions near the centre of the cluster, while outside Virgo it is related to galaxy groups. The lower fraction of galaxies with debris at the outskirts of the Virgo cluster could either be due to a lack of mergers due to the lack of nearby companions, or to the

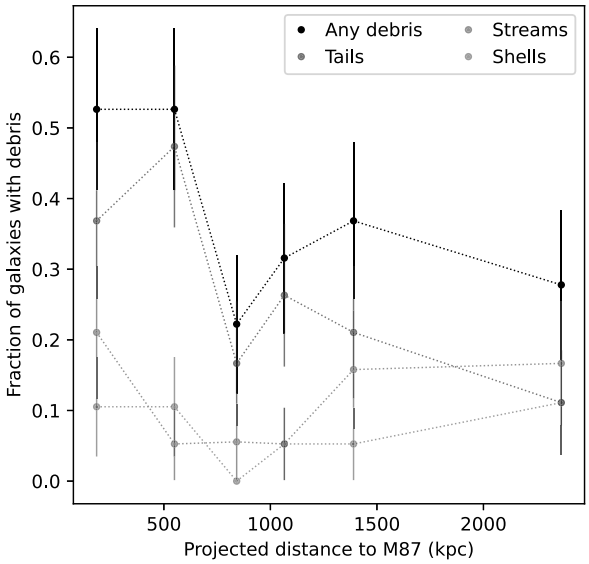

Figure 5.14: Fraction of galaxies inside the Virgo cluster that have tidal features as a function of the projected distance to M87 (taken to be the centre of the Virgo cluster). Each bin approximately contains the same number of galaxies. The fraction of galaxies hosting any type of debris is plotted in black, it is blue for tails, red for streams and green for shells. The error bars represent the standard deviation on proportions in each bin.

fact that tidal features tend to be phase-mixed more quickly than in the field, where they have longer lifetimes.

5.1.2.4 Luminosity of features versus large-scale environment

To push the analysis further, we studied the evolution of $f_{L,tidal}$ (i.e. the total luminosity of tidal features around each galaxy normalized by the luminosity of the whole galaxy) as a function of the environment density ρ_{10}, as shown in Figure 5.15. As mentioned in Section 5.1.2.2, the median value of $f_{L,tidal}$ is equal to zero in almost all bins of environment density, therefore we represent the mean value of $f_{L,tidal}$ per bin of ρ_{10}.

From this Figure, it appears that there is a peak of $f_{L,tidal}$ for bins of environment density such as $-2 < \log_{10}(\rho_{10}) < -1.5$ (with $f_{L,tidal}$ ranging from 1% to 2.5%), followed by a smaller peak around $\log_{10}(\rho_{10}) = -1$. This peak is similar to the one in Figure 5.13 which is related to on-going interactions in galaxy groups. The fraction of flux in tidal features then decreases for higher density bins (less than 1%). The overall trend is mostly driven by tidal tails. However, Spearman's test revealed a non-monotonic relationship between $f_{L,tidal}$ and ρ_{10} but Pearson's test could not be applied.

Unlike the trend visible in Figure 5.13, we do not see here an increase of $f_{L,tidal}$ in the Virgo cluster for tails although there are more galaxies with tails in the highest ρ_{10} bin. The fact that we see fewer luminous tidal features around Virgo galaxies compared to field ones might suggest that these structures are more evaporated due to interactions with the environment (e.g., Mihos 2003; Rudick et al. 2009; Janowiecki et al. 2010). However, there is no statistical difference in the area, length or width of tails and streams in Virgo compared to the field (see Table 5.3).

Therefore, tidal features have a higher luminosity fraction for galaxies in groups, due to on-going interactions, but otherwise no correlation is found between this fraction and the

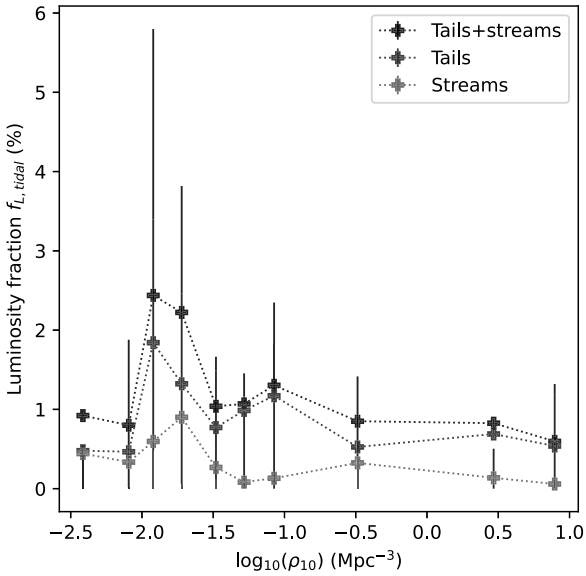

Figure 5.15: Scatter plot of the mean fraction of luminosity in tidal features, $f_{L,tidal}$, as a function of the environment density ρ_{10}. All galaxies are considered, even the ones that do not display tidal features. Each bin approximately contains the same number of galaxies. The mean value of $f_{L,tidal}$ in each bin of ρ_{10} is shown in black for tails and streams together, in blue for tails and in red for streams. The errorbars represent the 16% and 84% quantiles in each bin. Galaxies in the Virgo cluster typically have $\log_{10}(\rho_{10}) > -0.4$.

environment density. The lower fraction of $f_{L,tidal}$ in the Virgo cluster might be explained by environmental effects that tend to evaporate tidal features.

5.1.2.5 Tidal features versus small-scale environment

Another environment to consider is the small-scale one, such as the presence of nearby massive companions that can be responsible for on-going interactions or mergers. Indeed, it is important to remind here that tidal features originating from past mergers enable to probe the late assembly history of galaxies (they are the ones we are interested in), while features from on-going interactions rather give clues about the future of the galaxy.

In our annotation tool, we defined a companion galaxy as a massive[10] nearby galaxy which is considered close enough to be able to trigger detectable tidal interactions: we set a threshold of a projected distance of 100 kpc. Galaxies with no such companion are referred to as 'isolated', while galaxies having at least one companion are called 'in pairs'. For clarity purpose, even if the galaxy has several companions (hence if it is in a group), it is referred to as in pairs. With this definition, a galaxy in the Virgo cluster can be considered isolated, and a galaxy in the field/group can be in a pair. In our sample, 225 galaxies (47%) are in-pairs and 250 (53%) are isolated and genuine post-merger candidates.

Incidence of tidal features In Table 5.7, we present the percentage of galaxies with or without tidal debris for galaxies isolated and in pairs. The percentages are expressed with respect to the total number of galaxies in each category. Proportionally, there are more

[10]We were not able to check the actual mass of the galaxy (except if it was a galaxy from the reference sample of ATLAS^{3D}), we estimated the galaxy as massive if its size was close to the one of the target galaxy, i.e. not a dwarf or a much smaller companion.

galaxies with debris for galaxies in pairs (42%) than for isolated ones (31%). Likewise, in proportion, the fraction of galaxies with tidal tails is higher for galaxies in pairs (30% against 18% for isolated ones); but similar fractions of streams (around 15%) and shells (around 10-14%) are found for both categories. This was confirmed statistically with the Z-test on proportions: for isolated versus in-pair galaxies the proportion of galaxies with shells and streams is the same (with p-values of 0.9 and 0.2). These results are consistent with what was expected from our definitions. Indeed, when we keep isolated galaxies, there will not be tails originating from on-going major mergers, hence reducing both the percentage of galaxies with tails and with any type of debris.

Table 5.7: Percentage of galaxies that have any type of tidal debris, detailed for tails, streams and shells, as a function of whether the galaxy is isolated or in a pair. The first (resp. second) row is the percentage expressed with respect to the total number of galaxies isolated (resp. in pairs).

Galaxy type	Galaxies with any debris (%)	incl. galaxies with tails (%)	incl. galaxies with streams (%)	incl. galaxies with shells (%)
Isolated: total (%)	31.2	17.6	15.2	10.4
In pairs: total (%)	41.8	29.8	15.6	14.2

We extended our analysis by separating ETGs from LTGs. Indeed, as it will be discussed in Section 6.2.2, tails around ETGs and LTGs will not have the same morphology: tails around ETGs will be plume-like while LTGs will produce antennae-like features, likely brighter because younger. For streams, the morphological type of the primary galaxy should not impact them, as they originate from a lower-mass galaxy. Table 5.8 presents the percentage of galaxies with debris for ETGs (resp. LTGs) in the three top (resp. bottom) rows, normalised with respect to the total number of ETGs (resp. LTGs).

Table 5.8: Percentage of ETGs and LTGs that have any type of tidal debris, detailed for tails, streams and shells, as a function of whether the galaxy is isolated or in-pair. The first (resp. second) three rows present the results for ETGs (resp. LTGs) so percentages are expressed with respect to the total number of ETGs (resp. LTGs).

Galaxy type	Galaxies with any debris (%)	incl. galaxies with tails (%)	incl. galaxies with streams (%)	incl. galaxies with shells (%)	Total (%)
ETGs: isolated	14.3	6.6	8.6	6.1	48.8
ETGs: in pairs	20.1	10.7	10.2	8.2	51.2
ETGs: total (%)	34.4	17.2	18.9	14.3	100
LTGs: isolated	18.6	12.1	7.4	4.8	56.7
LTGs: in pairs	19.5	17.7	4.3	5.2	43.3
LTGs: total (%)	38.1	29.9	11.7	10.0	100

From this Table, the Z-test showed that the fraction of ETGs with debris (of any type and for each individual feature) is similar for galaxies in pairs and isolated ones. The test gives the same conclusion for LTGs. Hence, even though there seems to be more tidal tails around galaxies in-pairs (18% for LTGs, 11% for ETGs) than isolated ones (12% for LTGs, 7% for ETGs), the difference is not statistically significant.

However, there are statistically significantly more LTGs hosting tails than ETGs. This supports the fact that tidal tails are generated more easily and/or more visibly around LTGs (dominated by rotation) than ETGs (dominated by velocity dispersion). Among galaxies in pairs, there are more ETGs with streams than LTGs. This might be explained by the fact that ETGs are mostly located in groups where there can be more small satellites than in the field where most LTGs are. For shells, no difference is visible between ETGs and LTGs for galaxies isolated or in-pairs but we did not expect one as shells most likely originate from radial mergers.

Fraction of galaxies with debris versus mass: isolated versus in-pairs A second point to investigate is whether the increase of the fraction of galaxies with debris with galaxy mass we saw in Section 5.1.2.1 is still observed when considering separately galaxies in-pairs and isolated ones. To that end, we plot in Figure 5.16 the fraction of galaxies with debris as a function of the galaxy mass, for isolated (i.e. post-merger) and in-pair galaxies.

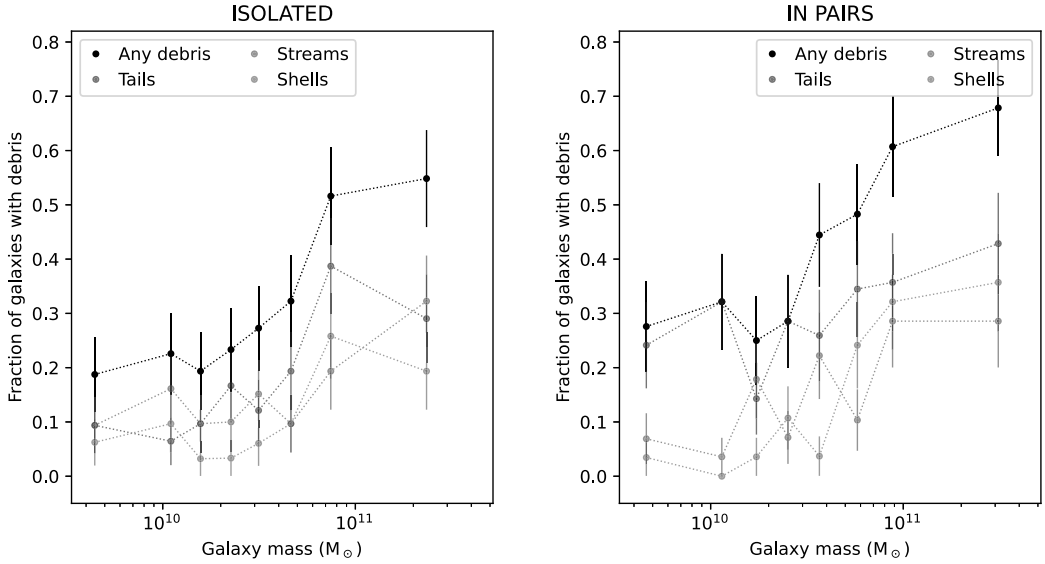

Figure 5.16: Fraction of galaxies that have tidal features as a function of the galaxy mass (in $M_\odot$), per bin of mass, for isolated galaxies (*left panel*) and galaxies in-pairs (*right panel*). Each bin approximately contains the same number of galaxies. The fraction of galaxies hosting any type of debris is plotted in black, it is shown in blue for tails, red for streams and green for shells. The error bars represent the standard deviation on proportions in each bin.

Several points arise from the comparison of the two panels of Figure 5.16. First, the fraction of galaxies with debris indeed increases with galaxy mass for both isolated and in-pair galaxies (as seen for all galaxies in Section 5.1.2.1). More precisely, for isolated galaxies it ranges between 19% and 55% with an overall fraction of 31%; while for galaxies in pairs it ranges between 25% and 67% with an overall value of 42%. The overall fractions are statistically different (from the Z-test, with a p-value of 0.016). This trend (the weighting) is actually driven by tidal tails. We find again a higher fraction of galaxies with tails among the in-pair ones; while no statistically significant difference is found for streams and shells between isolated and in-pair galaxies. An important point to underline is that the fraction of post-merger galaxies with debris indeed increases with galaxy mass.

This means that in the recent history (up to a few Gyr), more massive galaxies have undergone more merger events, both minor and major ones.

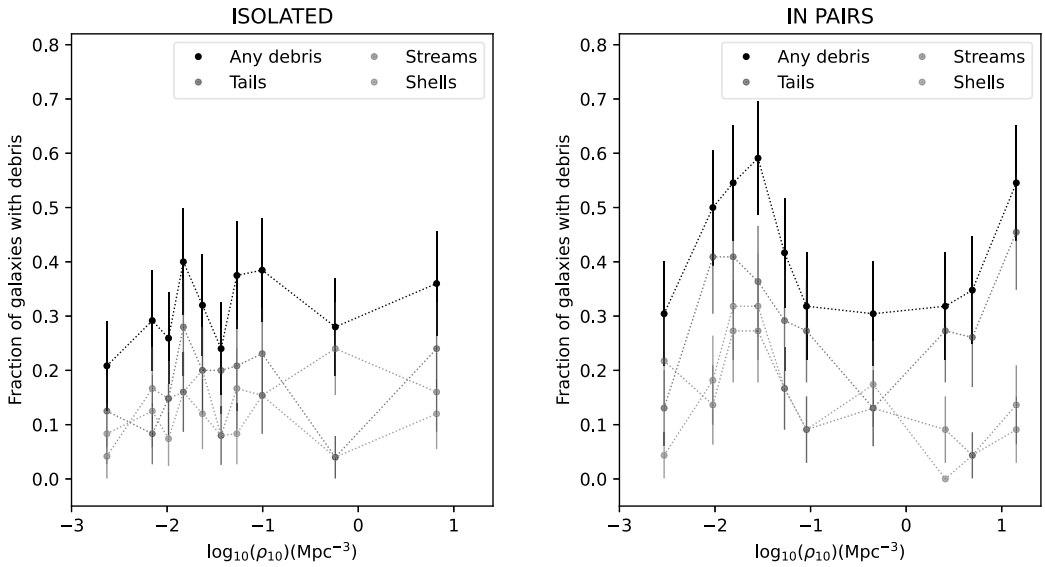

Figure 5.17: Fraction of galaxies that have tidal features as a function of the environment density ρ_{10} (in Mpc^{-3}), per bin of ρ_{10}, for isolated galaxies (*left panel*) and galaxies in pairs (*right panel*). Each bin approximately contains the same number of galaxies. The fraction of galaxies hosting any type of debris is plotted in black, it is blue for tails, red for streams and green for shells. The error bars represent the standard deviation on proportions in each bin. Galaxies in the Virgo cluster typically have $\log_{10}(\rho_{10}) > -0.4$.

Fraction of galaxies with debris versus environment density: isolated versus in-pair In Figure 5.17 we plot the fraction of galaxies with debris as a function of the environment density, for isolated and in-pair galaxies. We note that the peak in the fraction of galaxies with debris outside the Virgo cluster ($-2 < \log_{10}(\rho_{10})$) and the rise in the highest density bin originate from galaxies in pairs. This rise in the Virgo cluster is only driven by tails; while outside Virgo the rise is mostly driven by tails (from 10% to 50%) with a contribution from shells (up to 30%) and from streams (up to 20% before decreasing to less than 10%). The overall fraction of galaxies with debris is again higher for galaxies in pairs than isolated ones. For isolated galaxies, there is a small increase in the $-2 < \log_{10}(\rho_{10})$ bins (from 25% to 45% for all type of debris) and a small rise in the Virgo cluster (up to 40%). The fraction of galaxies with tails, streams or shells are relatively similar.

The conclusion that can be drawn from this Figure is that interacting galaxies are responsible for the increase of the fraction of galaxies with debris (due to tails) in Virgo as well as in the group, where there are also hints of multiple mergers occurring simultaneously (with increased fractions of galaxies hosting shells and streams). Therefore, the environment has a significant impact on tidal features only for on-going interactions and mergers. Indeed, post-merger galaxies do not seem much affected by the environment (no statistical difference between cluster and group).

Summary We separated galaxies in pairs and post-merger (isolated) ones. We found more tails around galaxies in-pairs and no difference for streams or shells: this was ex-

pected by definition, as tails originate from mergers with a massive companion but it is not the case for streams and shells. We also found more tails around LTGs than ETGs (both isolated and in-pairs) which is explained by the fact that tails are generated more easily in rotation-dominated systems, while for systems dominated by random motions, triggering large elongated features is more complicated.

Then, we found that the fraction of post-merger galaxies with debris increased with galaxy mass, which is the sign that more massive galaxies have experienced more merger events in the past few Gyr. Finally, the large-scale environment only affects tidal features around galaxies in-pairs (from on-going mergers and multiple interactions), as no statistical difference is found for post-merger galaxies in the group and cluster environment. **To conclude, the mass of the galaxies seems to be the most important factor to consider to understand the late assembly history of galaxies. The environment only seems to have an effect on tidal features generated during on-going interactions, i.e. the small-scale environment has a more important effect than the large-scale one.** As mass and environment are inter-dependent, we will estimate the joint impact of mass and environment on tidal features in Section 5.3.1.

5.1.3 Effect of internal kinematics on tidal features

In the previous sections we studied the effect of mass and environment on tidal features, and we separated ETGs and LTGs for more precise comparisons. However, in Chapter 1 we mentioned that ETGs do not form a uniform class of objects, as internal properties such as kinematics may vary. The majority of ETGs show regular rotation patterns (fast rotators, FRs) while others do not show any sign of rotation or complex kinematic features (slow rotators, SRs) (e.g., Emsellem et al. 2007, 2011). According to numerical simulations, the assembly history of these objects are expected to differ: ETGs should originally form as FRs and then transform into SRs through mergers (e.g., Naab et al. 2014; Penoyre et al. 2017; Lagos et al. 2018). It is therefore interesting to push the analysis further and study the potential correlations between internal kinematics of ETGs and the frequency of tidal features. Bílek et al. (2022), a work which I was involved in, combined kinematic information from the ATLAS^{3D} sample of ETGs and the census of tidal features in MATLAS deep images (Duc et al. 2015; Bílek et al. 2020) to study the mergers that transformed FRs into SRs, by probing merger-sensitive features (tidal features, kinematically distinct cores, stellar ages). This work showed that observational data and constraints from the literature would be best explained if these mergers were wet, minor and occurring at $z = 2$. Such mergers did not form the tidal features observed today.

Our annotation tool and database enable us to extend this work in a quantitative approach, for FRs versus SRs, and to increase the sample studied by adding ETGs in the Virgo cluster. This approach is complementary, as we probe the late mergers that were presumably not responsible for the transformation from FRs to SRs. We present here a preliminary result. Following Emsellem et al. (2011) and Bílek et al. (2022), we used the value of rotational support $\lambda^N_{\mathrm{R_e}} = \lambda/\sqrt{\epsilon}$ where λ quantifies the degree of ordered rotation of the galaxy and ϵ the apparent ellipticity[11]. Galaxies with $\lambda^N_{\mathrm{R_e}} < 0.31$ (resp. > 0.31) are considered SRs (resp. FR). The values were taken from Emsellem et al. (2011). With such a definition, we have 15% of SRs in our ETGs sample. SR are statistically

[11]This also includes 2-σ galaxies, which are characterized by two off-centre, but symmetric, peaks in the velocity dispersion, which lie on the major-axis of the galaxy (Krajnović et al. 2011).

significantly more massive than FRs. The fraction of SRs compared to FRs is 32% in the highest-density bin of Virgo, against around 16% in the field. However, the statistical tests did not reveal a significant difference between the environment of SRs and FRs, which may be due to the low number of SRs in Virgo.

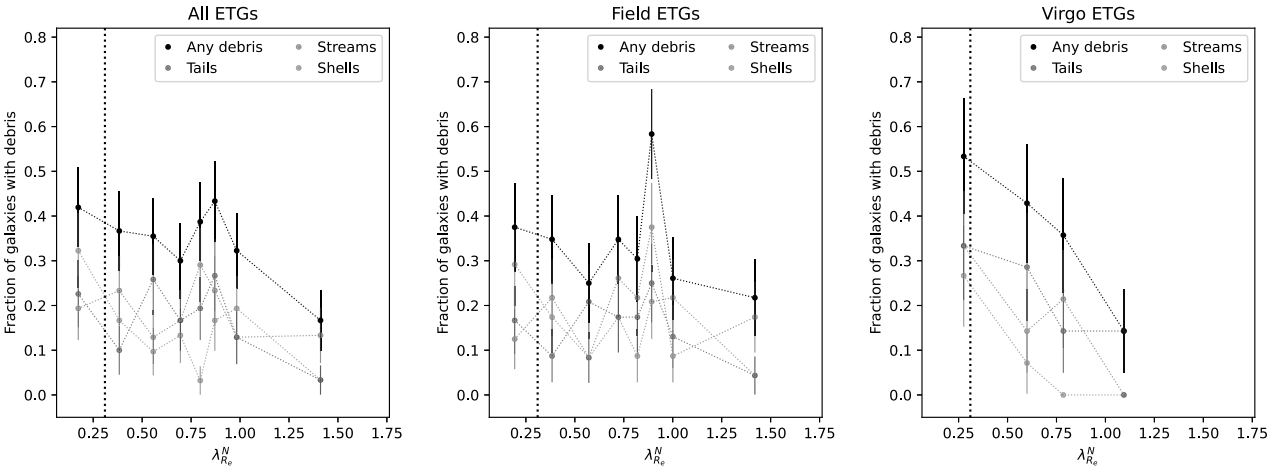

Figure 5.18: Scatter plot of the fraction of galaxies that have tidal features as a function of the rotational support $\lambda_{\mathrm{R_e}}^N$. Each bin approximately contains the same number of galaxies. The fraction of galaxies hosting any type of debris is plotted in black, it is shown in blue for tails, red for streams and green for shells. The error bars represent the standard deviation on proportions in each bin. The vertical dotted line separates SRs from FRs. Such fractions are presented for all ETGs (*left*), field ETGs (*middle*) and Virgo ETGs (*right*).

In Figure 5.18 we present the fraction of galaxies with debris as a function of the rotational support $\lambda_{\mathrm{R_e}}^N$ for all ETGs, and for field and Virgo ETGs separately. The fraction of galaxies hosting shells is higher for SRs than for FRs, and is higher than the fraction of galaxies with tails or streams (except in Virgo where the fraction of shells is smaller but there are too few SRs in Virgo to be significant). It could be explained by the fact that SRs are more massive than FRs, and more massive galaxies host more tidal features. About the fraction of galaxies with debris, we find two peaks: one for SRs with the lowest $\lambda_{\mathrm{R_e}}^N$, and one for field FRs. The former could be again related to mass effects, SRs being more massive. The latter could be triggered by multiple interactions in the group: indeed there are increased fractions of galaxies with all types of debris which are similar to the ones visible in Figures 5.13 and 5.17. For Virgo, there is a clear decrease of the fraction of galaxies with debris when $\lambda_{\mathrm{R_e}}^N$ increases: this could also be due to the higher mass of SR. In order to disentangle between the effect of the mass, environment and internal kinematics, it is necessary to study the correlations at fixed mass and environment density: Bílek et al. (2022) carried out such a careful comparison of SRs and FRs. They used tidal debris and other merger tracers (kinematic substructures, dust lanes, effective radii). They found a higher incidence of shells in SRs compared to FRs (like we did), and a decreasing probability of having tidal disturbances (tidal features) with increasing rotational support. They corrected for the mass effect and concluded that the trend was mainly due to mass. Their conclusion was that SRs were in place at times for which our method of detecting SRs is not sensitive.

5.2 Haloes

Although tidal features are the most prominent signs of interactions, their lifetime is limited and they will eventually phase-mix and contribute to the diffuse light of the extended stellar halo. Characterizing the luminosities and the sizes of haloes as a function of the galaxy mass and environment will enable to probe their assembly history over longer periods of time than accessible with tidal features. In the hierarchical paradigm, haloes are expected to grow in size and flux with mass, as they will have undergone more interactions and merger events.

We briefly remind here our definition of halo and inner galaxy. The inner galaxy corresponds to the central regions: for LTGs, it encompasses the spiral arms and the most luminous features. For ETGs, it encloses all the light visible is shallower surveys, so it is similar to an isophotal limit. Hence the inner galaxy actually corresponds to the entire galaxy visible in shallow surveys (such as PanSTARRS or SDSS). On the contrary, the halo corresponds to the light that extends further away than the inner galaxy, which is only visible in deep images. When we refer to the luminosity of the halo, it only corresponds to the light in the outskirts and does not include the inner galaxy (i.e. it has a 'donut' shape). This definition is arbitrary and motivated by what was doable with the annotation server. It is also important to note that for LTGs we cannot distinguish between a proper halo with a 3D structure and a thick disc no longer producing stars.

The results presented below are based on the 455 annotated haloes from the database. We have 475 galaxies but some haloes were impossible to annotate due to a contamination by a bright nearby star or overlapping companion.

5.2.1 Characterization of individual haloes

5.2.1.1 Radius and SB

We measured the surface brightness along the external contours of the annotated stellar haloes (i.e. outer SB) and the radius of haloes, as plotted in Figure 5.19 and shown in Table 5.9. The outer SB peaks at about 26.6 mag arcsec^{-2}, more precisely around 26.7 and 26.5 mag arcsec^{-2} for ETGs and LTGs respectively. LTGs have a statistically significantly lower SB than ETGs, as confirmed by the T-test, 2-sample KS test and Mood's test. We must remind here that we are not able to make a difference between a halo and a thick disc for LTGs. There can be star formation in the haloes of LTGs while it is not the case for ETGs that have an old population of stars. In addition, there might be an effect of the inclination of the galaxy: if the galaxy has a bright nucleus it can generate an artificial halo around the galaxy. For edge-on LTGs, which do not necessarily have a halo, the artificial halo will still appear and the SB will be overestimated; while for ETGs there is always a halo so the artificial one will be mixed with the stellar halo, with proportionally a less important effect on the SB. From additional analyses, galaxies involved in on-going interactions, more massive galaxies as well as galaxies with debris have brighter halo outer SB, which contributes to create a scatter towards brighter external contours. The presence of sources of pollution also contributes to this scatter, with cleanest haloes having higher (i.e. fainter) outer SB. In all cases, the outer SB of haloes is clearly much lower than the nominal SB limit of the survey, and that obtained when deriving integrated surface brightness profiles of galaxies. It just reflects the ability of the eye in delineating an external contour on our asinh images. However, what matters for this study are global trends and not individual measurements.

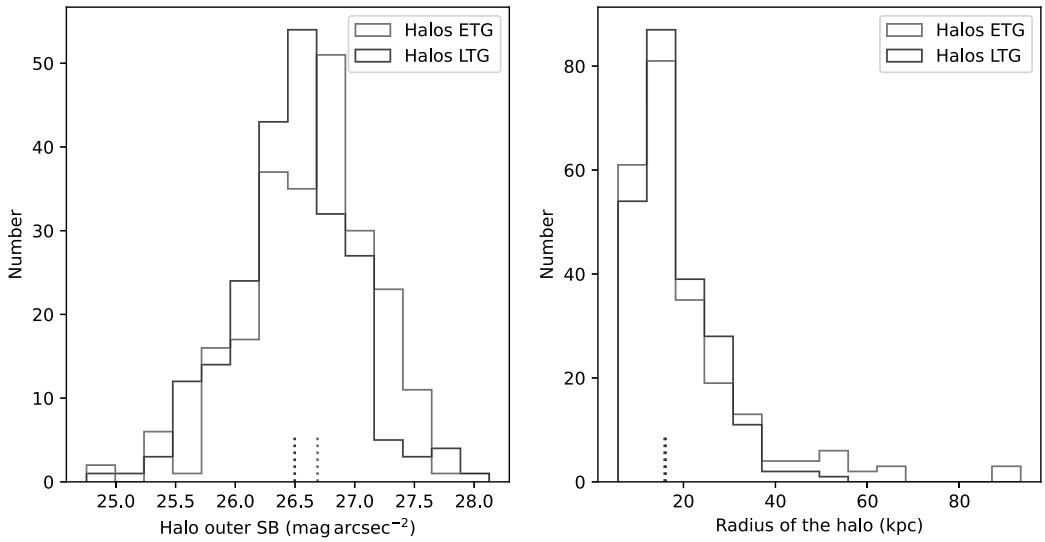

Figure 5.19: Histogram of the radius (in kiloparsecs) and outer SB (in magnitudes per square arcsecond) of the annotated haloes as a function of the morphological type, with ETGs in red and LTGs in blue. The median of each distribution is represented by the dotted line. *Left*: Histogram of the halo outer SB (in mag arcsec^{-2}). *Right*: Histogram of the radius of haloes (in kpc).

Table 5.9: Median halo outer SB, expressed in magnitudes per square arcsecond, for each type of galaxies. Associated standard deviations are indicated in parentheses.

Galaxy type	Halo outer SB
All galaxies	26.6 (0.5)
MATLAS ETGs	26.8 (0.6)
VESTIGE ETGs	26.5 (0.4)
All ETGs	26.7 (0.5)
MATLAS LTGs	26.4 (0.7)
CFIS LTGs	26.4 (0.4)
VESTIGE LTGs	26.6 (0.4)
All LTGs	26.5 (0.5)

The radii of annotated haloes (visible in the right panel of Figure 5.19) have a distribution peaked at 15.8 kpc and 16.1 kpc for ETGs and LTGs respectively. The majority of the haloes have a radius smaller than 40 kpc, with a few ETGs extending up to 90 kpc. We would expect ETGs to have larger radii than LTGs. Indeed, there are more satellite companions around ETGs than LTGs for a given mass, and more companions for more massive galaxies (Kawinwanichakij et al. 2014). In addition, the mass growth of present day galaxies seems to be driven by minor mergers (e.g., Oser et al. 2010), leading to the formation of streams that will eventually phase-mix in the outskirts of the haloes. The study of integrated SB profiles of ETGs and LTGs might reveal differences, but this is behind the scope of this work.

5.2.1.2 $g-r$ colour

In a similar way than for tidal features, we derived the $g-r$ colours from the masks of our annotations after background-correction of the g and r-band files. We also accounted for the extinction and for the difference of filter set between MATLAS (old r filter) and VESTIGE (new r filter). The colour-corrected and extinction-corrected $g-r$ colours for annotated haloes are presented in Table 5.10, while the histograms for ETGs and LTGs are shown in Figure 5.20.

Table 5.10: Extinction and colour-corrected $g-r$ (in mag) colours of annotated haloes for each type of galaxies in MATLAS and VESTIGE surveys. Associated standard deviations are indicated in parentheses. *Integrated* stands for the median $g-r$ integrated values; *Median* corresponds to the median of the individual median $g-r$ from colormaps values while *Number* indicates the number of structures annotated.

Galaxy type	Haloes $g-r$		
	Integrated	Median	Number
MATLAS all galaxies	0.58 (0.12)	0.58 (0.10)	217
MATLAS ETGs	0.60 (0.11)	0.60 (0.10)	168
MATLAS LTGs	0.53 (0.11)	0.53 (0.11)	49
VESTIGE all galaxies	0.48 (0.20)	–	119
VESTIGE ETGs	0.50 (0.14)	–	57
VESTIGE LTGs	0.45 (0.24)	–	62

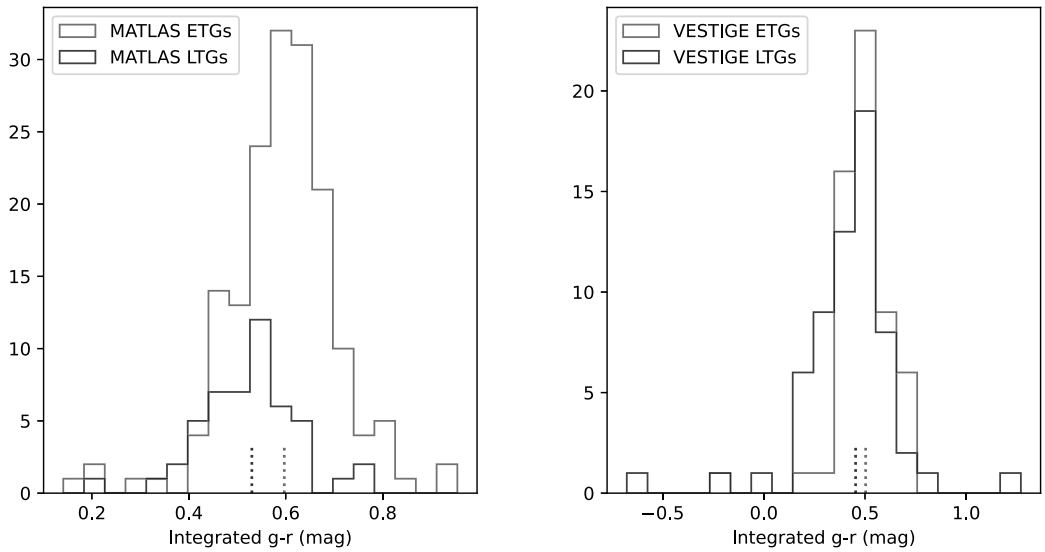

Figure 5.20: Histograms of the extinction and colour-corrected integrated $g-r$ values (in mag) of annotated haloes. The haloes of ETGs are plotted in red and in blue for LTGs. The dotted vertical lines represent the median of the distributions. *Left*: Histograms of the integrated $g-r$ colours of MATLAS haloes. *Right*: Histograms of the integrated $g-r$ colours of VESTIGE haloes.

Like for tidal features, we consider separately the $g-r$ colours of MATLAS and VESTIGE haloes (due to the change of filters). There is a difference of around 0.05 to 0.07 mag between the haloes of ETGs and LTGs. The haloes of LTGs are statistically

significantly bluer than ETGs (with p-values lower than 0.05 from the T-test, Mood's test and the 2-sample KS test for MATLAS, although Mood's test for VESTIGE revealed similar means). We expected ETGs to be redder, due to the low levels of star formation occurring in ETGs. The similar values of the integrated and median $g-r$ colour from colormap for MATLAS haloes indicate that no colour gradient is present. If we compare VESTIGE to MATLAS, VESTIGE haloes are bluer. This is likely linked to the PSF effect, with a much more extended PSF for MATLAS which causes a reddening of the halo. Accurate colour values would require PSF-deconvolution. We will discuss the impact of the PSF in Section 6.3.1.2.

5.2.1.3 Luminosity and stellar mass

From the flux values contained in haloes, corrected for extinction and colour-correction due to the change of MegaCam filters, and from the integrated $g-r$ colour, we derived the stellar mass in annotated haloes. The results are summarized in Table 5.11. We cannot directly compare the values between ETGs and LTGs as haloes were not annotated in a similar way: the luminosity of haloes of ETGs corresponds to the amount of light between two isophotes, while for LTGs it corresponds to the smooth part of the galaxy (compared to the spiral arms).

Table 5.11: Median luminosities in r-band (in $L_\odot$) and stellar masses (in $M_\odot$) of annotated haloes. Associated standard deviations are indicated in parentheses.

Halo type	Luminosity (10^9 $L_\odot$)	Mass (10^9 $M_\odot$)
All haloes	1.2 (9.9)	2.5 (25)
ETGs	1.4 (12)	2.9 (30)
LTGs	0.8 (3.2)	1.4 (6.8)

5.2.2 Effect of galaxy mass and environment on haloes

5.2.2.1 Luminosity of haloes versus galaxy mass

In a similar way than what was done for tidal features, a first point is to investigate the evolution of the fraction of luminosity contained in haloes as a function of the mass of the galaxy. To that end, with automated aperture photometry on our background-corrected images, we retrieved the flux in the halo (i.e. the 'donut' shape corresponding only to the outer halo revealed by LSB images) and we divided it by the flux of the whole galaxy (i.e. halo and inner galaxy). The histogram of this value, named $f_{L,halo}$, is represented on the left panel of Figure 5.21. Contrary to tidal features for which a majority of galaxies have $f_{L,tidal} = 0$, all galaxies have a halo hence it is possible to represent the median values of $f_{L,halo}$ instead of the mean. It appears that the median fraction of luminosity in all haloes is 10.6%, it is of 13.4% for ETGs only and 8.2% for LTGs only. Most haloes have a fraction of luminosity lower than 22%. There are a few systems with large values of $f_{L,halo}$ but these systems correspond either to very close and interacting pairs where it is impossible to distinguish the halo of the target galaxy and of the companion, or to very large ellipticals (for instance in the Virgo cluster). When comparing this fraction to the literature, we face the same issues than the comparison for tidal features, namely the differences in the definitions of haloes, in the methods used and in the survey depths. Trujillo & Fliri (2016) found from the surface brightness profile that the halo of their

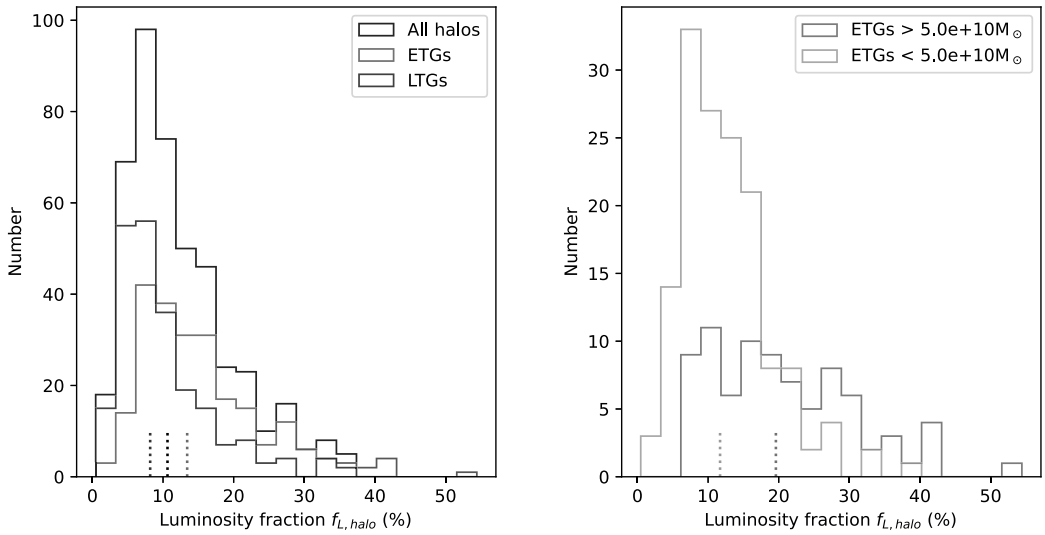

Figure 5.21: *Left*: Histogram of the fraction of luminosity $f_{L,halo}$ in haloes, defined as the luminosity in the 'donut' halo shape divided by the luminosity of the entire halo (including the inner galaxy). This fraction is represented in black when considering all galaxies, in red for ETGs and in blue for LTGs. The vertical dotted lines represent the median values of the distributions. *Right:* Histogram of $f_{L,halo}$ for the haloes of massive ETGs ($M_\star > 5 \times 10^{10} M_\odot$) in pink and for less massive ETGs in orange.

target galaxy represented 11% of the total flux in their original images (that had not been PSF-deconvolved). Likewise, from ellipse-fitting Duc et al. (2015) found the flux contained below the 25 mag arcsec^{-2} isophote for ETGs to be on average 4.5% with a maximum of 14%. Using integrated light, Gilhuly et al. (2022) estimated the halo fraction around 12 edge-on LTGs, with four different halo definitions: they ranged between less than one percent and less than 7%, but their haloes are more extended than ours and hence the comparison is not straightforward. Merritt et al. (2016) found a large variety in the halo fraction for the 8 spiral galaxies they studied, ranging from no halo detection to around 4.5% for the halo fraction. It is also important to keep in mind that in our case we do not account for the effect of the PSF, which some of the aforementioned studies do consider. The impact of the PSF will be investigated in Section 6.3.1.2. In overall, there seems to be an agreement between our values and the literature for the fraction of luminosity enclosed in the extended stellar halo.

In Figure 5.22, we study the evolution of $f_{L,halo}$ as a function of galaxy mass. The median percentage of $f_{L,halo}$ per bin of mass is represented with solid black markers for all haloes, it is plotted in red for ETGs and in blue for LTGs. From this Figure, there is a clear increase of $f_{L,halo}$ with galaxy mass when considering all the haloes together, from about 10% to 18% in the highest mass bin. For ETGs, the fraction rises from around 12% to 28%; while for LTGs the fraction remains roughly constant (about 8%). The increase visible for ETGs (and all haloes) is more important for masses higher than about $5 \times 10^{10} M_\odot$. Spearman's test revealed a statistically significant monotonic relation between $f_{L,halo}$ and galaxy mass when considering ETGs (with a p-value of 3×10^{-10}) and for all haloes; but no correlation was found for LTGs.

In order to better see the distribution of $f_{L,halo}$ for ETGs with a mass below $5 \times 10^{10} M_\odot$ and above, we plotted their distributions in the right panel of Figure 5.21, in orange for the less-massive and in pink for the most massive ETGs. We did not perform the same

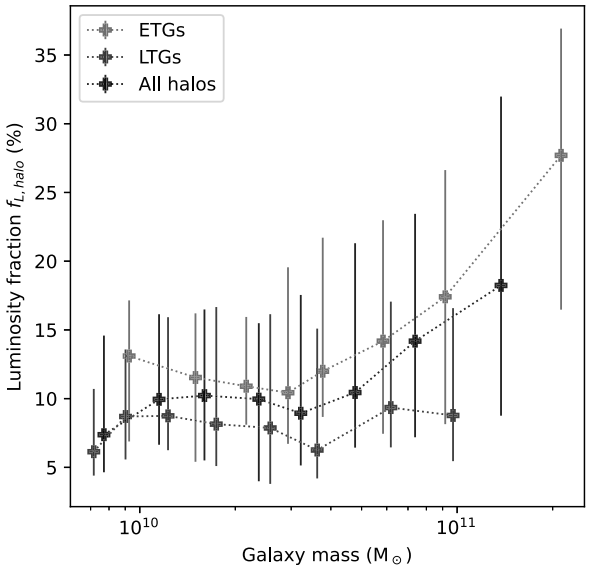

Figure 5.22: Scatter plots of the median fraction of luminosity of the haloes $f_{L,halo}$ as a function of the galaxy mass (in $M_\odot$). Each bin approximately contains the same number of galaxies. Large solid markers indicate the median values in each mass bin, in black for all haloes, in red for ETGs and blue for LTGs. The lower (resp. upper) errorbars represent the 16% (resp. 84%) quantiles in each bin.

plot for LTGs as there is no correlation between $f_{L,halo}$ and galaxy mass for LTGs. It appears that the distributions are clearly different: most massive ETGs have a median flux fraction of 20% and a wide distribution; against 11.5% for the less massive ones and a peaked distribution.

The overall increase of the halo luminosity with the mass of the galaxy is consistent with hierarchical assembly models. As the halo is built-up by the remnants of the items of debris from successive mergers, and as more massive galaxies are subject to more interactions (e.g., Murali et al. 2002), haloes are expected to be larger and contain more flux as the galaxy mass increases. This trend is dominated by ETGs, with a clear threshold mass above which the haloes of ETGs become almost twice as luminous as the less massive galaxies. This is consistent with the idea that the mass growth of the most massive galaxies is dominated by mergers, as gas accretion is not longer possible (due to too hot haloes that prevent the fresh infall of gas), hence more luminous extended haloes. The relatively constant luminosity for LTGs haloes could be explained by the fact that LTGs grow mostly from smooth accretion of gas rather than from mergers, according to the scenario proposed for instance by Sancisi et al. (2008) or Bílek et al. (2022). Indeed, if the LTGs had undergone major mergers, they would eventually have turned into ETGs, hence the LTGs visible today should have had a relatively quiet evolution in the past few gigayears.

To conclude on haloes, we found the fraction of the luminosity of haloes increases with galaxy mass, a trend which is driven by ETGs. This percentage of luminosity ranges from around 10% with a maximum of around 18% for the most massive galaxies. These trends and values are globally in agreement with the literature.

5.2.2.2 Luminosity of haloes versus large-scale environment

We studied the evolution of the fraction of flux contained in the halo as a function of the large-scale environment characterized by the environment density ρ_{10}: the results are presented in Figure 5.23.

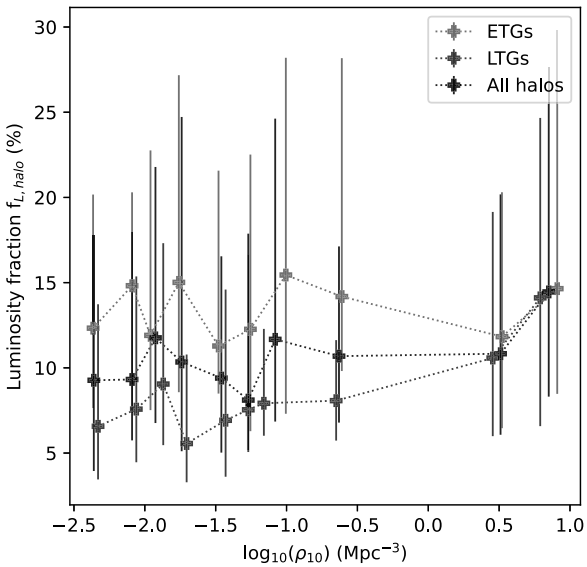

Figure 5.23: Scatter plots of the median fraction of luminosity of the haloes $f_{L,halo}$ as a function of the environment density ρ_{10} (in Mpc^{-3}). Each bin approximately contains the same number of galaxies. Large solid markers indicate the median values in each bin of ρ_{10}, in black for all haloes, in red for ETGs and blue for LTGs. The lower (resp. upper) errorbars represent the 16% (resp. 84%) quantiles in each bin. Galaxies in the Virgo cluster typically having $\log_{10}(\rho_{10}) > -0.4$.

From Figure 5.23, the overall median $f_{L,halo}$ is equal to 10% in the field and 12% in Virgo, and a small correlation is found between $f_{L,halo}$ and the environment density (with a p-value of 5×10^{-3}). If we focus on ETGs (resp. LTGs), the median $f_{L,halo}$ is 13.7% (resp. 7.6%) in the field and 13.1% (resp. 11.3%) in Virgo. There is no statistically significant relation for $f_{L,halo}$ versus ρ_{10} for ETGs, but it is significant (p-value = 0.0002) for LTGs. The increase of $f_{L,halo}$ with ρ_{10}, only visible for LTGs, suggests an effect of the environment. In clusters, the ram-pressure tends to quench spiral galaxies by removing a large fraction of their gas. Star formation is reduced (or even stopped), which will result in redder and more extended haloes. If ram-pressure affects our LTGs, we should observe brighter and more extended haloes. It does indeed seem to be the case for LTGs from Figure 5.23. We investigate in a more detailed approach the ram-pressure effect in the next section.

5.2.2.3 Ram-pressure stripping

In this section, we investigate another aspect of the effect of the cluster environment in comparison to the field or groups. Cluster galaxies are redder and less star-forming than their field counterparts, and there is a lack of gas-rich galaxies in clusters. Gunn & Gott (1972) first proposed ram-pressure stripping as an explanation for this observation. Since then, many studies demonstrated that ram-pressure stripping has an important impact

in shaping spiral galaxies in clusters by removing large fractions of their gas (see e.g., Boselli et al. 2022, for a recent review). As spirals are infalling into the cluster, they feel a wind from the intracluster medium that can strip their gas if the gravitational attraction between the galaxy and its gas is not strong enough. Since their gas reservoir is removed, star formation is quenched. The spiral arms will eventually disappear and the disc will redden. We remind here that for LTGs we cannot distinguish between a proper halo and a thick disc no longer producing stars.

To quantify the depletion of the gas reservoir in spirals, we can compare the observed HI mass of cluster galaxies and the expected HI mass of field galaxies of similar morphologies and sizes: this is called the HI-deficiency. HI-deficient galaxies will have smaller HI discs compared to their field counterparts and smaller star forming regions (hence larger stellar discs) than their field counterparts. Additionally they may present star forming clumps along filaments of stripped gas, such as the ones visible in jellyfish galaxies(e.g., Bekki 2009; Poggianti et al. 2016; Durret et al. 2021). Figure 5.24 illustrates with images from the CFHT (NGVS) the impact of ram-pressure stripping on the visual appearance of spiral galaxies. It starts with an unperturbed spiral (panel A), whose spiral arm and star forming regions cover the entire galaxy and halo. In panel B are shown examples of galaxies undergoing ram-pressure stripping, with blue starbursts along filaments of stripped gas extending outside of the halo. Finally, in panel C are presented some HI-deficient galaxies, with faint spiral arms, very small star forming regions and a larger halo compared compared to panel A. Here, we want to determine whether deep images associated to our annotation tool can provide information on this process.

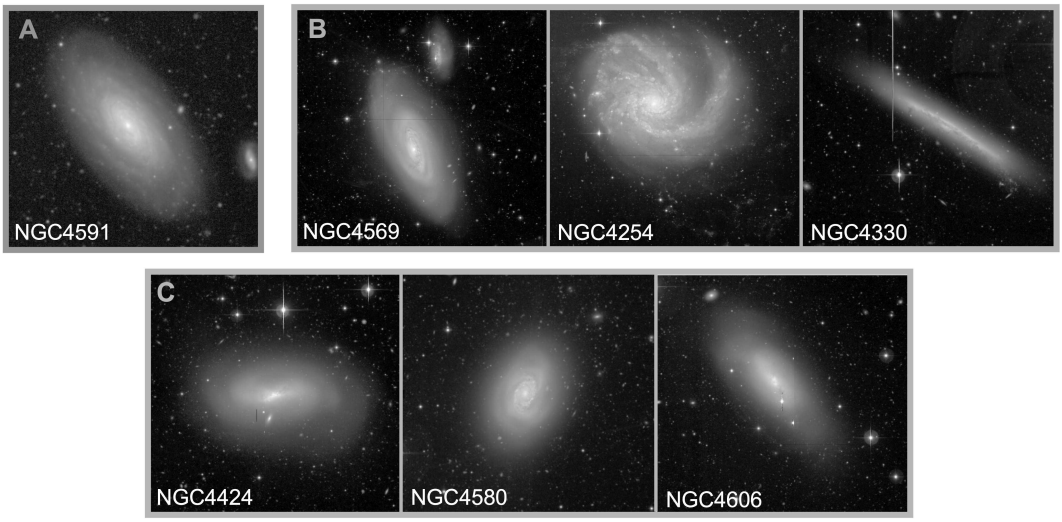

Figure 5.24: Figure adapted from Sola et al. (2023). Illustration of the effect of ram-pressure stripping on spiral galaxies, using NGVS composite *ugi* colour images in logarithmic scale. Panel A represents the phase prior to ram-pressure stripping, with an unperturbed spiral whose spiral arms extend up to the halo. In Panel B are represented galaxies that are currently undergoing ram-pressure stripping: blue filaments of star forming regions along the stripped gas are extending outside the disc and halo of the galaxy. Panel C represents the last stage of ram-pressure stripping, with HI-deficient galaxies having faint or no spiral arms, much smaller star forming regions, located near the center and proportionally a much larger halo than in panel A.

To that end, we studied the size of the haloes of LTGs in the Virgo cluster as a function of their HI-deficiency. If LTGs are experiencing ram-pressure stripping, they will be more

HI-deficient and we expect to observe more extended haloes. We used the HI-deficiency values from Table 3 of Chung et al. (2009), but we only had such values for 27 LTGs in Virgo. We investigated the evolution of the normalised area of the haloes as a function of the HI-deficiency. We normalized by the area of star-forming regions, i.e. regions enclosing spiral arms and bluer regions. Such regions are useful as they characterize the spatial extent of where star formation occurs. They are are generally smaller than the inner galaxy, especially when the galaxy is seen edge-on. These star-forming regions were annotated only for the LTGs with available HI-deficiency data.

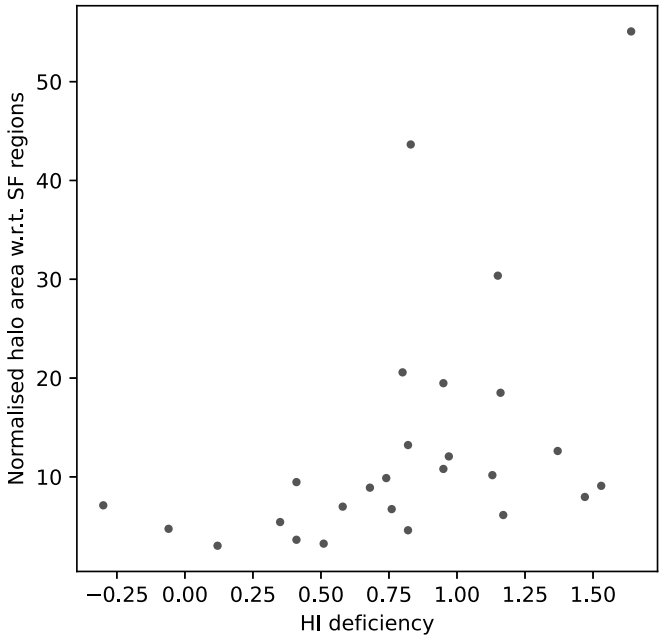

Figure 5.25: Figure adapted from Sola et al. (2023). Scatter plot of the halo area, normalised with respect to star forming (SF) regions, as a function of the HI-deficiency for LTGs in the Virgo cluster with available HI-deficiency data.

The results are presented in Figure 5.25. There is a clear increase of the halo normalized with respect to star-forming regions with increasing HI-deficiency. Hence, HI-deficient galaxies tend to have haloes much larger than the size of the star-forming regions than galaxies which are less HI-deficient. This trend can be interpreted as the effect of ram-pressure stripping: as the gas is removed from the outskirts of the galaxy, the gas is left only in the central regions where star forming can still occur, but the size of these regions will be smaller compared to the halo than galaxies that did not suffer from ram-pressure stripping. A complementary visual inspection of colour images of the most HI-deficient galaxies revealed clear signs of ram-pressure stripping, such as blue clumps oriented along small filaments perpendicular to the morphological major axis of the galaxy, that resembled smaller versions of jellyfish galaxies see panel B from Figure 5.24.

This comparison showed that our tool combined with deep images can be useful to quantify the effect of ram-pressure stripping. Indeed, even if no information on the gas is available, by comparing the extent of the stellar halo to the size of the central star-forming regions, it is possible to have information on the ram-pressure effect.

5.2.3 Effect of internal kinematics on haloes

A complementary approach is to study the correlations between haloes and internal kinematics for ETGs. We present here a preliminary result. Following Section 5.1.3, we separate fast and slow rotators based on their degree of ordered rotation. In Figure 5.26 we present the fraction of luminosity in annotated haloes, $f_{L,halo}$, as a function of the rotational support $\lambda^N_{R_e}$. From it, there is a clear decrease of the luminosity fraction in haloes with increasing $\lambda^N_{R_e}$. The annotated haloes of SRs (with a median $f_{L,halo} = 23\%$) are almost twice as luminous as the ones of FRs ($f_{L,halo} = 12\%$), which is statistically significant. This trend can be compared to the evolution of $f_{L,halo}$ as a function of galaxy mass (see Figure 5.22, where the most massive ETGs roughly have twice as luminous haloes as the less massive ones). We can make the hypothesis that the trend seen for SRs can be explained by galaxy mass, as more massive galaxies have undergone more merger events, hence have more extended and brighter haloes. A disentanglement between the effect of mass, environment and internal kinematics would be needed to properly study the results, such as the one carried out by Bílek et al. (2022).

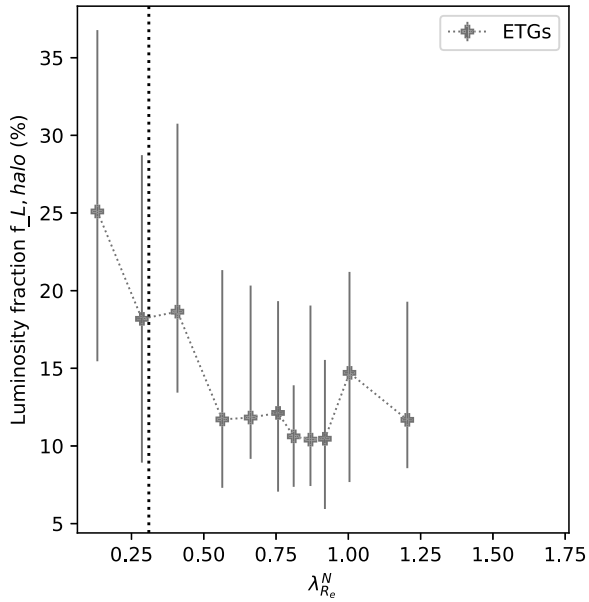

Figure 5.26: Scatter plots of the median fraction of luminosity in the annotated haloes of ETGs $f_{L,halo}$ (in %) as a function of the rotational parameter $\lambda^N_{R_e}$. Each bin approximately contains the same number of galaxies. The lower (resp. upper) errorbars represent the 16% (resp. 84%) quantiles in each bin.

5.3 Discussion

5.3.1 Disentangling the effects of galaxy mass and environment

In the previous sections, we studied separately the effect of galaxy mass and environment on the flux and size of tidal features and haloes. However, it may be argued that mass and environment are correlated, with more massive galaxies residing in denser environments. Here, we want to check this by studying the 2D-histograms of some properties (fraction of galaxies with debris, $f_{L,tidal}$, $f_{L,halo}$) as a function of the mass and environment simultaneously. To that end, we computed the edges of 2D-bins so they contain approximately

the same number of galaxies per 2D-bin. We plotted the value of the properties in each of the 2D-bins, as presented in Figure 5.27.

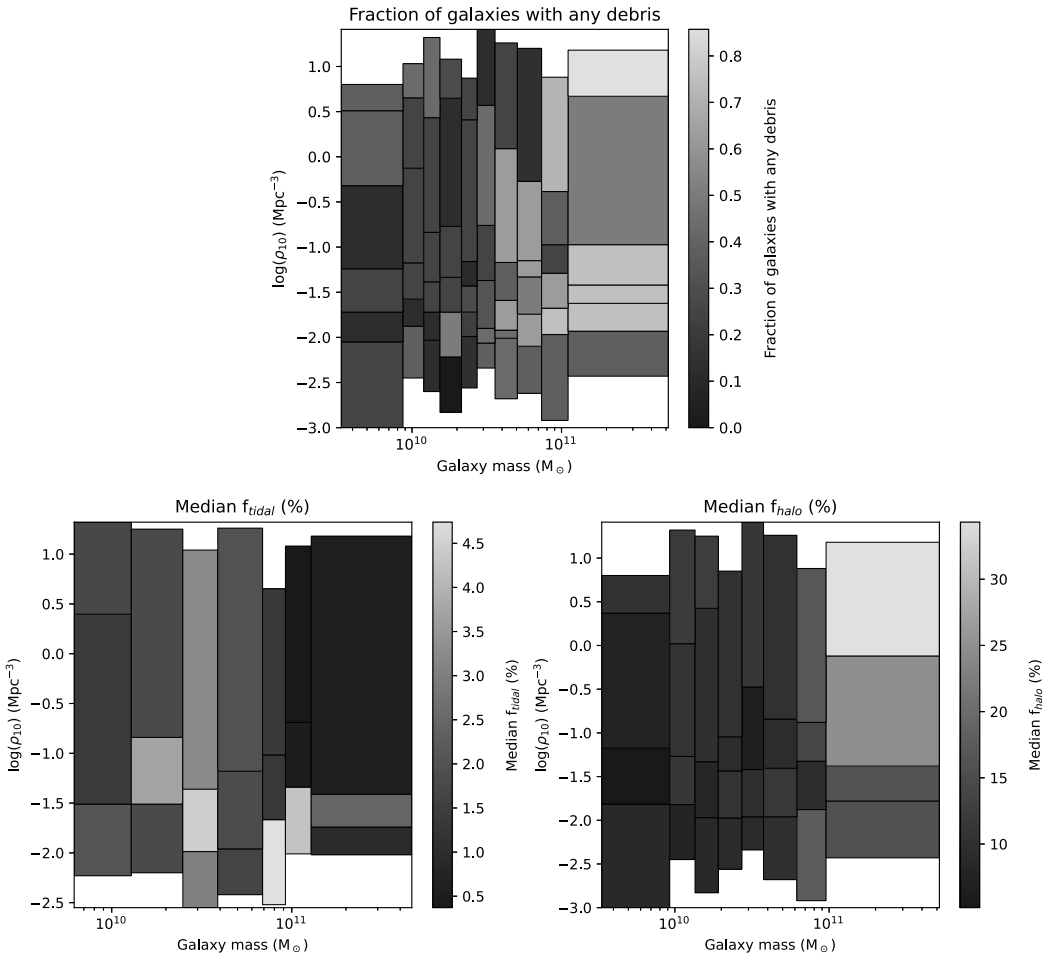

Figure 5.27: 2D-histograms of some properties of tidal features and haloes, as a function of the galaxy mass (in $M_\odot$) and environment density ρ_{10} (in Mpc^{-3}). Each 2D-bin contains approximately the same number of galaxies. *Top*: 2D-histogram of the fraction of galaxies with any type of debris. *Bottom left*: 2D-histogram of the median percentage of flux contained in tidal features $f_{L,tidal}$. *Bottom right*: 2D-histogram of the median percentage of flux in the halo $f_{L,halo}$.

The top panel represents the evolution of the fraction of galaxies with any type of debris, and it appears that there is a clear trend that this fraction increases more with the mass than with the environment density. For the fraction of the flux in features $f_{L,tidal}$, shown in the bottom left panel, there is no clear correlation visible, neither with the environment density nor with the mass. For the fraction of the flux in the halo $f_{L,halo}$, it appears from the bottom right panel that the correlation is mainly with the mass, although the highest $f_{L,halo}$ is reached in the highest environment density and highest mass 2D-bin. From this Figure, it appears that the mass of the galaxy is the dominant effect, so it seems possible to disentangle the effect of mass and environment as we did in the previous sections.

5.3.2 Evolution of LSB structures with galaxy mass, environment and internal kinematics

Here, we summarize and discuss the joint impact of galaxy mass and environment on tidal features and haloes. We started our analysis by studying each effect individually. Although one may argue that both quantities are linked, as we expect to find more massive galaxies in denser environments, Figure 5.27 shows that for our sample the correlation seems strong with the mass, but not so much with environment density. It is then possible to disentangle the effects of mass and environment.

From the previous sections, the galaxy mass appears to be the predominant factor impacting the evolution of the properties of tidal features and haloes, rather than the large-scale environment. From Figure 5.10, the fraction of galaxies hosting any type of tidal debris increases with galaxy mass, from around 25% for the less-massive up to 60% for the most massive galaxies. Similarly, the fraction of luminosity contained in tidal features, $f_{L,tidal}$, increases with galaxy mass, as visible in Figure 5.12 from about 0.7% to 2%. These two trends indicate that more massive galaxies have undergone more merger events than less massive ones, which is consistent with hierarchical models of galactic evolution. Although we do not have the same definitions of tidal features, depth or detection methods, the dependence of the fraction of galaxies with tidal features and their luminosity with stellar mass had also been reported in the literature, both on the observation (e.g., Atkinson et al. 2013; Duc et al. 2015; Bílek et al. 2020; Yoon & Lim 2020; Huang & Fan 2022; Vázquez-Mata et al. 2022) and simulation sides (e.g., Rodriguez-Gomez et al. 2015; Martin et al. 2022).

These findings about tidal features are consistent with what was found for haloes, and in particular the increase of the fraction of luminosity in haloes, $f_{L,halo}$, with galaxy mass (see Figure 5.22). The median flux fraction ranges from less than 10% for the less massive galaxies up to 18% for the most massive ones. This trend is actually driven by ETGs (with $f_{L,halo}$ ranging from 12% to 28%), while for LTGs this fraction remains constant with galaxy mass. This also supports the hierarchical model, where ETGs are more 'evolved' systems which have undergone more merger events than LTGs. These mergers generated tidal features that eventually phase-mixed and contributed to form the diffuse and extended stellar halo. On the contrary, LTGs have a smoother assembly history, mostly through gas accretion (e.g., Sancisi et al. 2008; Bílek et al. 2022). If LTGs had undergone major merger events, they would have been morphologically impacted and would have looked like ETGs, which means the LTGs observed at $z = 0$ did not encounter any major interaction in the past few Gyr.

Another point to note is the presence of a mass threshold above which the increase of the previous quantities with galaxy mass sharpens. This mass threshold is around $3 \times 10^{10} M_{\odot}$ for the fraction of galaxies with debris, it is of around $4 - 7 \times 10^{10} M_{\odot}$ for $f_{L,tidal}$ and about $5 \times 10^{10} M_{\odot}$ for $f_{L,halo}$. The presence of such a mass threshold had also been reported in the literature, with a similar order of magnitude (between a few 10^{10} to $10^{11} M_{\odot}$) (e.g., Atkinson et al. 2013; Duc et al. 2015; Bílek et al. 2020; Yoon & Lim 2020; Martin et al. 2022; Vázquez-Mata et al. 2022; Huang & Fan 2022). This threshold has also been seen by Lim & collaborators (2023) with changes in the globular cluster (GC) distributions, sizes and colours above a threshold of $10^{10.81} M_{\odot}$. They interpret this as high-mass galaxies having accreted their GC systems while lower-mass galaxies have formed them in situ.

This mass threshold also coincides with the galaxy mass where merger events become the dominant factor of galactic growth from numerical simulations (e.g., Stewart et al. 2008; Rodriguez-Gomez et al. 2015). Likewise, Kauffmann et al. (2003) noted a change in

galaxy properties around $3 \times 10^{10} M_{\odot}$, with lower-mass galaxies having young stellar populations, disc-like structures, and low surface mass densities. These findings are consistent with the idea that the mass growth of the most massive galaxies is dominated by mergers (e.g., Newman et al. 2012; Hilz et al. 2013; Robotham et al. 2014; Rodriguez-Gomez et al. 2016; Vulcani et al. 2016; Hill et al. 2017a,b; Nevin et al. 2023), as gas accretion may no longer be possible (due to too hot haloes that prevent the fresh infall of gas), hence more luminous extended haloes as well as increased fraction and luminosity of tidal features. Kereš et al. (2005) and Dekel & Birnboim (2006) noted that around a similar mass threshold, lower-mass galaxies are dominated by cold accretion (i.e. gas accreted from cold dense intergalactic filaments) and higher-mass galaxies are dominated by hot accretion (i.e. gas shock-heated to high temperature that later cools). In comparison to massive ETGs, LTGs seem to have a smoother accretion history, mostly through gas accretion (e.g., Sancisi et al. 2008).

In a second step we investigated the impact of the environment on tidal features and haloes. We considered two types of environments: the large-scale one defined by the environment density ρ_{10} which makes a distinction between the field, groups and the Virgo cluster; and the small-scale one defined by the presence of a very nearby companion susceptible to generate tidal interactions. If we focus on the fraction of galaxies with debris as a function of ρ_{10}, from Figure 5.13 two peaks are visible, one in the group environment and one in the Virgo cluster. As tidal tails dominate this trend, a hypothesis was that the trend was driven by on-going interactions. We saw a decrease of the fraction of galaxies with shells in higher density environments, like it was seen by Malin & Carter (1983). This suggests that environmental factors are at play and tend to destroy tidal features more easily. Figure 5.15 presents the evolution of $f_{L,tidal}$ with ρ_{10}: a peak is visible in the group environment, related to the one from Figure 5.13, but we do not see an increase of $f_{L,tidal}$ in the Virgo cluster. This may suggest that tidal features in the Virgo cluster are more evaporated due to interactions with the environment (e.g., Mihos 2003; Rudick et al. 2009; Janowiecki et al. 2010).

The next crucial step was to investigate the small-scale environment to make a distinction between the origin of tidal features. Indeed, tidal features around 'isolated' (post-merger) galaxies are the ones that give clues about the past assembly history of galaxies; while features originating from on-going interactions give information about the future history. We separated isolated galaxies from the ones in pairs (with a visually-based distinction): galaxies in pairs host more tidal tails, preferentially around LTGs, but no difference is visible for shells and streams. By our definitions, we expected to see more tidal tails around interacting LTGs as they originate from on-going interactions and tails appear preferentially around rotation-supported systems. When focusing on $f_{L,tidal}$ with galaxy mass (see Figure 5.16), we confirmed that the fraction of post-merger galaxies with debris increases with galaxy mass, which is consistent with hierarchical models. When focusing on the evolution of $f_{L,tidal}$ with environment density, the peaks visible are indeed related to on-going interactions, while no statistically significant trend is visible with ρ_{10} for isolated galaxies (see Figure 5.17). From it, it appears that the large-scale environment does not impact much tidal features, unlike the small-scale environment (the presence of a companion) that triggers peaks in the fraction of galaxies with debris due to on-going interactions.

We now focus on the effect of the environment on haloes. From Figure 5.23, there is no trend of the fraction of luminosity of haloes $f_{L,halo}$ versus the environment density for ETGs. However, $f_{L,halo}$ increases with ρ_{10} for LTGs. This can be interpreted as the effect

of ram-pressure stripping (e.g., Gunn & Gott 1972; Boselli et al. 2022) that depletes LTGs of their gas reservoir, resulting in brighter and more extended haloes in comparison to the star forming regions that become smaller and smaller; which was also seen in Section 5.2.2.3.

Finally, we investigated the correlations between internal kinematics of ETGs and the frequency of tidal features and the luminosity of haloes (we presented preliminary results of this analysis). Indeed, ETGs can be divided between slow and fast rotators, depending on their degree of ordered rotation (e.g., Emsellem et al. 2007, 2011), and their history assembly is expected to differ. We found that SRs have a higher fraction of galaxies with debris (in particular shells) and have more luminous haloes than FRs. This trend could be due to a mass effect, as SRs are more massive and reside in denser environments than FRs (e.g., Emsellem et al. 2011; van de Sande et al. 2017; Veale et al. 2017; Brough et al. 2017; Wang et al. 2020) and as more massive galaxies have undergone more merger events. A trend with the small-scale environment (with multiple mergers) might be visible for tidal features for some FRs. However, a proper study of the correlations with internal kinematics require a disentanglement of mass and environment. For instance, at fixed stellar mass, the trends with the environment are not clear anymore (e.g., Brough et al. 2017; Greene et al. 2017; Graham et al. 2019). A detailed analysis at fixed mass and environment density was carried out by Bílek et al. (2022) and we refer to that paper for the conclusions.

To summarize, it appears from our results that the mass seems to be the dominant factor that affects tidal features and stellar haloes. More massive galaxies host more tidal features and more luminous ones, which is consistent with the hierarchical paradigm. The haloes of ETGs are brighter for more massive galaxies, which is also consistent with the hierarchical growth, while the growth of LTGs would be dominated by gas accretion. As in the literature, a mass threshold is visible above which the trends aforementioned are sharper. The large-scale environment (field, group, cluster) has an impact in clusters with signs of ram-pressure stripping for LTGs as well as hints of a more pronounced evaporation of tidal features. The small-scale environment has a more pronounced effect as the presence of a nearby companion tends to generate tidal features, both in the group and in the cluster environment.

5.4 Conclusion

In this chapter, we have provided a characterization of tidal features and extended stellar haloes for a sample of 475 nearby massive galaxies. From our annotation database, we derived quantitative information such as geometrical properties, surface brightness, colours or census of tidal tails, streams, shells and haloes. Thanks to the masks of our annotations, we derived the fraction of luminosity and then stellar mass in these LSB tidal features. Such values were so far not very well constrained on large samples of galaxies. They may be compared to simulations to better understand the type of mergers that took place, and more generally to constrain models of galactic evolution. We started by providing these properties for the individual features, and then we studied the impact of galaxy mass and environment on them in order to obtain clues about the past assembly history of galaxies.

From the individual properties, the annotated structures show an apparent large variety of shapes and sizes when put and compared together. There is a large overlap between tails (tidal structures emanating from the primary galaxy and made during major merg-

ers) and streams (tidal debris from disrupted low-mass companions), which may be biased due to the definition and the difficulty to distinguish one type from the other. However, a statistical analysis revealed real differences:

- Tidal tails are wider than streams, with a typical width of 5.7 kpc against 4 kpc for streams. This was expected from models: since streams originate from low-mass companions with lower velocity dispersion, their width should be smaller.
- Tidal tails are brighter than streams, with a difference between their median surface brightness of almost 0.6 mag arcsec^{-2}. This may be due to some age effects, with streams having a longer survival (i.e. visibility) time than tails, which are more easily identified as such when they are young.

These statistical differences comfort the approach we used to specifically distinguish tidal tails from streams. We also found that tidal tails and streams account for less than 2% of the total galaxy luminosity. The overall fraction of galaxies hosting any type of tidal features is 36%. Then, we focussed on the properties of the extended stellar haloes and shells.

- The annotated stellar haloes have radii extending mostly between 5 and 40 kpc with a similar distribution for ETGs and LTGs.
- Identified shells have typically radii lower than 40 kpc, with few extending to 80 kpc: we do not observe the very external extended shells found in some simulations, even though the comparison is not straightforward.
- The stellar haloes account for about 10% of the total galaxy luminosity, with the haloes of ETGs being brighter than the ones of LTGs.

Other tendencies interesting for the analysis for future surveys are worth highlighting:

- The faintest structures identified as tidal tails and streams have a median surface brightness about 1 mag brighter than the nominal limiting surface brightness of the survey. This should also be kept in mind when comparing observations to simulations.

In a second part, we studied the impact of galaxy mass and environment on tidal features.

- The fraction of galaxies with debris increases with galaxy mass, as well as the fraction of luminosities contained in tidal tails and streams.
- More massive ETGs have the more luminous haloes, while no correlation is found for LTGs.
- There seems to be a mass threshold above which the increase of the fraction of galaxies with debris and the fraction of luminosities in tidal features and haloes increases more sharply with galaxy mass.

These points are consistent with the hierarchical paradigm: more massive ETGs have undergone more merger events, triggering more tidal features that eventually contribute to form the extended stellar haloes. This is also consistent with the idea that for the most massive galaxies, mass growth mostly occurs through mergers in comparison to lower-mass galaxies and LTGs where the growth is dominated by smooth accretion of gas.

- The small-scale environment (presence of a nearby companion) has an important impact on tidal features, as they are susceptible to generate tidal tails during ongoing interactions. This occurs both in groups and in the Virgo cluster.

- The large-scale environment (cluster versus field/group) impacts LTGs: the haloes of LTGs are larger and brighter in Virgo, and they correlate with HI-deficiency, which can be interpreted as the effect of ram-pressure stripping.

- The large-scale environment (Virgo cluster) affects tidal features: the number of shells and streams drop in Virgo, and tidal features are less luminous. This may be due to faster evaporation because of environmental processes (numerous high-speed flybys, interaction with the cluster potential).

When we investigate the joint effect of mass and large-scale environment (as well as internal kinematics for ETGs), it seems that galaxy mass has the more important impact on tidal features and haloes. In the next Chapter, we will discuss the limitations of our study.

5.5 Résumé du chapitre

Dans ce chapitre, nous avons fourni une caractérisation des débris de marée et des haloes stellaires étendus pour notre échantillon de 475 galaxies proches et massives. Nous avons obtenu à partir de notre base de données d'annotations des informations quantitatives telles que des propriétés géométriques, la brillance de surface, les couleurs ou le recensement des queues de marée, des courants stellaires, des coquilles et des halos. Grâce aux masques de nos annotations, nous avons calculé la fraction de luminosité puis de masse stellaire dans ces structures LSB. Jusqu'à présent, ces valeurs n'étaient pas très bien contraintes sur de grands échantillons de galaxies. Elles pourront être comparées aux prédictions des simulations pour mieux comprendre le type de fusions qui ont eu lieu, et plus généralement pour contraindre les modèles d'évolution galactique. Nous avons commencé par fournir ces propriétés pour les structures de marée individuelles, puis nous avons étudié l'impact de la masse et de l'environnement des galaxies sur celles-ci afin d'obtenir des indices sur l'histoire passée des galaxies.

Les principales conclusions que nous avons tirées de notre étude sont les suivantes. D'après les propriétés des structures individuelles, les débris de marée annotés présentent une grande variété apparente de formes et de tailles lorsqu'ils sont mis et comparés ensemble. Il existe un large chevauchement entre les queues de marée (structures émanant de la galaxie primaire et créées lors de fusions majeures) et les courants stellaires (débris provenant de compagnons perturbés de faible masse). Cependant, une analyse statistique a révélé de réelles différences :

- Les queues de marée sont plus larges que les courants stellaires, avec une largeur typique de 5,7 kpc contre 4 kpc pour les courants stellaires. Ceci était prévu par les modèles : puisque les courants stellaires proviennent de compagnons de faible masse avec une dispersion de vitesse plus faible, leur largeur devrait être plus petite.

- Les queues de marée sont plus brillantes que les courants stellaires, avec une différence entre leur brillance de surface moyenne de 0.6 $\mathrm{mag\,arcsec^{-2}}$. Ceci peut être dû à un effet d'âge, les courants stellaires ayant une durée de vie (c'est-à-dire une

visibilité) plus importante que les queues de marées, qui sont plus facilement identifiables en tant que telles lorsqu'elles sont jeunes.

Ces différences statistiques confortent l'approche que nous avons utilisée pour distinguer spécifiquement les queues de marée des courants stellaires. Nous avons également constaté que ces structures représentent moins de 2% de la luminosité totale des galaxies. La fraction globale des galaxies présentant des débris de marée est de 36%. Nous nous sommes ensuite intéressés aux propriétés des haloes stellaires étendus et des coquilles.

- Les coquilles identifiées ont typiquement des rayons inférieurs à 40 kpc, avec quelques coquilles s'étendant jusqu'à 80 kpc : nous n'observons pas les coquilles étendues très externes trouvées dans certaines simulations, même si la comparaison n'est pas directe.

- Les halos stellaires annotés ont des rayons qui s'étendent principalement entre 5 et 40 kpc avec une distribution similaire pour les ETGs et les LTGs.

- Les halos stellaires représentent environ 10% de la luminosité totale des galaxies, les halos des ETG étant plus lumineux que ceux des LTG.

D'autres tendances intéressantes pour l'analyse des futurs relevés méritent d'être soulignées:

- Les structures les moins brillantes identifiées comme des queues de marée et des courants stellaires ont une brillance de surface médiane d'environ 1 mag plus brillante que la brillance de surface limite nominale du relevé. Ceci doit être gardé à l'esprit lorsque l'on compare les observations aux simulations.

Dans une deuxième partie, nous avons étudié l'impact de la masse et de l'environnement des galaxies sur les débris de marées et les halos.

- La fraction de galaxies avec des débris augmente avec la masse de la galaxie, ainsi que la fraction de luminosité contenue dans les queues de marée et les courants stellaires.

- Les ETGs les plus massives ont les halos les plus lumineux, alors qu'aucune corrélation n'est trouvée pour les LTGs.

- Il semble y avoir un seuil de masse au-delà duquel l'augmentation de la fraction des galaxies avec des débris et la fraction des luminosités dans les débris de marée et les halos augmentent plus fortement avec la masse de la galaxie.

Ces points sont cohérents avec le modèle hiérarchique : les ETGs plus massives ont subi plus de fusions, créeant plus de débris de marée qui contribuent à former les halos stellaires étendus. Ceci est également cohérent avec l'idée que pour les galaxies les plus massives, la croissance en masse se produit principalement par des fusions par rapport aux galaxies de plus faible masse et aux LTGs où la croissance est dominée par l'accrétion régulière de gaz.

- L'environnement à petite échelle (présence d'un compagnon proche) a un impact important sur les débris de marée, car ils sont susceptibles de générer des queues de marée durant des interactions en cours. Cela se produit à la fois dans les groupes et dans l'amas de la Vierge.

- L'environnement à grande échelle (amas ou champ/groupe) a un impact sur les LTGs : les halos des LTGs sont plus grands et plus brillants dans l'amas de la Vierge, et ils sont corrélés avec une déficience en HI, ce qui peut être interprété comme l'effet de la pression dynamique dans l'amas.

- L'environnement à grande échelle (amas de la Vierge) affecte les débris de marées: le nombre de coquilles et de courants stellaire diminue dans l'amas, et les débris des marée sont moins lumineux. Cela peut être dû à une évaporation plus rapide en raison des processus environnementaux (nombreux survols à grande vitesse, interaction avec le potentiel de l'amas).

Lorsque nous étudions l'effet conjoint de la masse et de l'environnement à grande échelle (ainsi que de la cinématique interne des ETGs), il semble que la masse des galaxies ait un impact plus important sur les débris de marées et les halos. Dans le chapitre suivant, nous discuterons des limites de notre étude.

Chapter 6 Limitations of the study

In this chapter, we will discuss the limitations of our methods and of their impact on some results presented in Chapter 5. The originality of our approach relies on the use of an annotation database of low surface brightness features compiled by a group of users who have precisely delineated a large number of individual structures (representing a total of 12119 annotations) directly on displayed images and classified them. The method raises a number of issues partly posted in the previous chapters: (1) the difficulty of matching individual annotations made by different users (2) the reliability of the annotations when considering users with different levels of expertise (3) the presence of overlapping structures, in particular the contaminants which have a large covering factor, which prevent us from determining with precision the boundaries of some structures (4) the ambiguity in the classification of the various types of tidal features (5) the fact that we have used

images coming from various CFHT surveys with varying depth, surface brightness limits and filters and (6) the estimation of the luminosity of LSB features. We address all these issues in this discussion.

6.1 Biases of the annotation process

6.1.1 Generalisation to users with different levels of expertise

In the analyses conducted during this PhD thesis, the annotations were made by four users with a similar level of expertise. However, the annotation tool can be used by anyone and this study can be extended to less expert users. Taking the level of expertise of the user into account in the classification of tidal structures is important, as it reflects the degree of confidence that we can have in the annotation of this user (Bílek et al. 2020). Therefore, it is possible with our annotation tool to attribute weights to the users, and these weights would be inherited by the annotations. In that case, the results should take the weights into account, by computing for instance the weighted median and weighted standard deviation instead of simply the median and standard deviation.

Applying weights to the users could also modify our selection method described in Section 2.3.7. For instance, instead of keeping the largest annotation for haloes, tails and streams, one could think of weighting the annotation masks and combining the shapes into a weighted combination of the different annotations.

Since several users have annotated the same galaxies, it is necessary to assess their reliability. Following Section 2.3.7, if the users have different levels of expertise, this assessment could be useful to adjust the user weight. Indeed, it is important to characterize whether they annotate like the majority of the group or if their annotations are too different from what is expected (especially for non-expert users). In that case, the weight associated to that user will be lowered in order to avoid outliers in the results due to non-reliable classifications. This step could be part of a methodological process for future studies. For instance, we did not use here some annotations made by interns as they were made in an inhomogeneous way, but with a low user weight we could have taken them into account. We remind that in this work, all users have a similar level of expertise and no weight was assigned. To that end, a similarity index was defined.

6.1.2 Similarity between annotations

A similarity index ranging from 0 (no similarity) to 1 (total similarity) was computed to assess the similarity between the annotations of two users. It depends on the annotation type and is defined as follows:

- For halo and main galaxy annotations: the Jaccard index (Hennig 2007) is used. It is defined as the area of the intersection of both structures divided by the area of their union. If an annotation is missing for one of the users the index is set to zero.
- For shells: since their number is relevant, the similarity index is defined as $\frac{min(S_i,S_j)}{max(S_i,S_j)}$, where S_i (respectively S_j) is the number of shells annotated by user$_i$ (respectively user$_j$).
- For tails and streams: two metrics are used. The first one, like for shells, is an index computed over the number of features. However, the area covered is more relevant to characterize them (rather than the number of features, as they range from 0, 1,

or 2 for tails and generally 0 or 1 for streams). Hence, a second metric was defined. It relies on the pairing of tidal tails and streams as defined in Section 2.3.7 with the unique identifier. The Jaccard index is computed on the structures annotated by user_i and user_j that share the same unique identifier.

- For contaminants (high background, cirrus): the union of all the contaminants of the given type is made, as sometimes a user will split a pollutant annotation into two. Then, the Jaccard index is computed between the unions of the contaminants.

We computed the mean similarity index between each pair of our four users. The mean index for haloes ranges around 0.78 (from 0.76 to 0.81), while the mean index for tidal tails and streams is 0.49 (ranging from 0.44 to 0.52). The relatively high similarity between halo annotations is an indicator that users annotated in a comparable way, which is important for our study. The lower similarity index for tidal features was expected, since annotating tidal features is not as clear and easy as annotating haloes.

It is also interesting to determine whether users annotate in a comparable way in the presence of pollutants. When we consider only the cleanest annotations, that is to say annotations with a weighted reliability index as defined in Section 2.3.8.1 equal to 9, we obtain a similarity index of 0.82 (between 0.79 and 0.86) for haloes, and 0.58 (between 0.48 and 0.67) for tidal tails and streams. However, the number of clean annotations clearly decreases with the total number of annotations, so the statistics are computed on smaller samples (between 14 and 79 haloes and 16 to 77 tidal features). When keeping only the most polluted annotations (with a weighted reliability index smaller or equal to 2), the similarity index for haloes is 0.79 (between 0.77 and 0.8), and of 0.44 (between 0.38 and 0.52) for tidal features. One can see that for haloes, the cleaner the image, the higher the similarity index, which is likely related to the fact that users annotate in a more similar manner when the image is less polluted. The trend is similar for tails and streams. Therefore, it seems that pollutants make the annotation of all structures, especially tidal features, more complicated and less similar between users. This was expected, as they are the feature the most likely 'hidden' behind the additional amount of light from the sources of contamination. However, the relatively low similarity index between tidal features even in the cleanest image underlines the existing variability in the delineation of tidal debris.

6.1.3 Limits of our approach

Like any other visual classification process, our annotations would be more reliable if tens or hundreds of people would have participated. In this paper, we rely on the annotations of 4 users only. However, we realized this process is complicated and the users need to be trained both to recognize LSB structures and to draw their shape with precision, making any citizen science project like Galaxy Zoo difficult to implement. This task is made more complicated by the presence of many pollutants that overlap with the interesting features. The time necessary to annotate is quite important as well (around 10-15 minutes per galaxy). Therefore, automated techniques will be needed to analyse larger samples of galaxies with few annotations. The annotations we made could be used to train machine learning algorithms. In fact, it has already been the case with Richards et al. (2020), who used our cirrus clouds annotations to train a new machine learning algorithm to detect cirrus clouds on deep images as it will be presented in Section 7.2.2. Yet, the small number of annotations is a problem as large datasets are needed to train such algorithms. Appropriate data augmentation or annotations of cosmological simulations may be necessary to have large enough training datasets.

Another limitation of our method can arise from the annotation task itself. Indeed, in order to delineate a structure, the user must be certain that it exists, so only features with a high level of confidence will be delineated. It also requires an effort to draw the shape with precision, so it may be possible that in some ambiguous cases, the user preferred not to annotate a feature (by lack of time and/or confidence). This is less likely the case for surveys where users are asked to say if the feature is present or not, and/or when they can indicate their level of confidence about the presence of a structure. Therefore, simple censuses may in overall find more features than we do with the annotation tool.

6.2 Biases in the identification of tidal features

6.2.1 Survey sensitivity to identify tidal features

One important point to note is the fact that no annotated tidal tails or streams are fainter than 27.3 and 27.8 mag arcsec^{-2} respectively, even though the nominal depth of the images is at least of 28.3 mag arcsec^{-2}. This discrepancy can be explained by several factors. First, the nominal survey depth was estimated from measures done on boxes of 10×10 square arcsecond, while the structures of interest are more extended. Fluctuations of the SB brightness along the most diffuse structures (possibly above or below the SB limit) make it break into several sub-structures on our images, and its identification and classification as a single genuine stellar feature is very difficult. The presence of artefacts of similar SB as the structures has the same consequence of apparently breaking them into pieces.

This discrepancy of about 0.5 to 1 mag between the faintest tidal structures that may be identified and classified and the nominal SB limit of the survey must be kept in mind when making comparisons with simulations or estimating their visibility with other facilities. Similarly, the outer SB of haloes is smaller than the depth of the survey, as the eye is not able to detect the faintest structures compared to what can be obtained with aperture photometry.

6.2.2 Disentangling tidal tails from streams

As mentioned in Section 2.2.2.2, we defined tidal tails as structures with stellar material apparently coming from the primary galaxy, while streams originate from a less-massive smaller companion which may still be visible, hidden or have been destroyed. Most probably the users have adopted this definition in different ways according to their expertise and used a variety of observables (location, shape, amount of overlap with the closest galaxy, etc.) to assess the classification. Getting quantitative measurements from our annotation database, we are able to determine whether these two classes of tidal objects show different properties, may thus be really distinguished from our images and retrospectively check whether the basic initial criterion for disentangling them was relevant.

As presented in detail in section 5.1.1.2, we found that streams are narrower than tidal tails, a difference expected from models, as the width of tidal debris largely depends on the velocity dispersion of the progenitor, itself linked with its total mass. Just considering tidal tails, those found around ETGs appear wider than those associated to LTGs, another result at first order consistent with the expectations, since the velocity dispersion of galaxies decreases with their morphological type. In fact this explanation holds only for ETG-ETG collisions which naturally produce plume-like tidal tails. A merger involving

one or two LTGs will produce more narrow tails. Conversely, tails coming from late-type galaxies (that have kept their stellar disc) are necessarily relatively thin. Taking into account all configurations, statistically, there should be more wide tidal tails around ETGs, as observed.

Note that the area and length measured (from which we estimated the width) are only based on the projected shape of the structures since it is the only thing that can be annotated. It does not take the inclination and orientation along the line-of-sight of the galaxy into account, so the real intrinsic size of each structure is not known. Some tidal features are likely to be overlooked, especially when the galaxy is seen edge-on or when a tidal tail is hidden behind the galaxy. We also found a statistically significant difference between the median surface brightness and colour of tails and streams. This is also consistent with having progenitors of different masses and ages. The fact that we see a statistical distinction in the physical properties of tidal tails and streams is a validation of our definitions of these features (see Section 2.2.2.2).

However, from visual inspection of the thumbnails with all the individual tidal features shown in Figures 5.2 and 5.3, it can be seen that their shapes can vary significantly from one to the other. Some trends are emerging (tidal tails seem broader than streams while streams seem more elongated), but the great variability observed makes difficult a clear visual separation between the two structures. This might suggest that an automated classification of streams and tails solely based on their individual shape could be very complicated. The human expert classifier might have used a number of criteria to classify streams and tails, in addition to calling his physical intuition on their origin. Thus results might be better when providing the machine all relevant information, including the properties of the primary galaxy, but still requires reliable labels. Deep learning techniques are good candidates to achieve this inclusion of properties thanks to its inherent accounting for visual context.

6.3 Quantifying the impact of the sources of pollution

6.3.1 Impact of internal ghost reflections

As mentioned in the previous Sections, sources of pollution are present in our images. In particular, the scattered light from bright sources like foreground stars or the bright centre of galaxies can cause multiple internal reflections on the CCD and generate large artificial haloes around the source in the image, that are called ghost reflections (Slater et al. 2009). Although ghost reflections from bright stars (thereafter called external ghosts) can easily be delineated and removed from our analysis, the additional amount of light from ghost reflections from the bright galactic centre (thereafter called internal ghosts) is more difficult to tackle (e.g., Karabal et al. 2017), as it is mixed with the intrinsic light from the stellar halo. In this section, we evaluate the impact of these internal ghosts on our stellar haloes annotations.

6.3.1.1 Shapes of haloes

A first approach towards identifying haloes that are potentially affected by internal ghost reflections from the galactic centre is to study their morphology. Haloes that tend to be round and whose diameter is close to the one of external ghost reflections from bright stars are more susceptible to be contaminated.

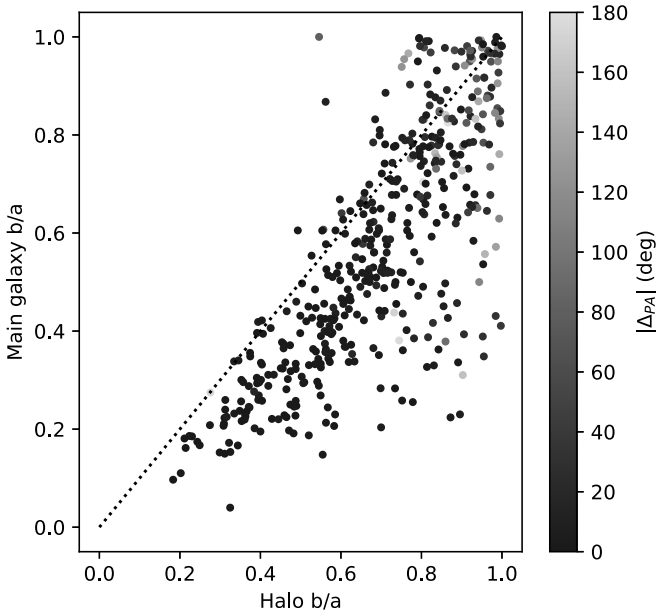

Figure 6.1: Visualization of the morphologies of haloes. The ratio of the minor axis b over the major axis a of the main galaxy is plotted against this ratio for the halo. Points are colour-coded by the difference of position angle (in degrees) between the ellipse representing the main galaxy and the ellipse of the halo.

Our annotation tool is useful when exploring data and finding potentially problematic or interesting haloes. It enables us to quantify the morphological shape of the extended halo with respect to the main galaxy. We can represent the ratios of the minor axis b over the major axis a for the halo and the main galaxy, and visually identify some categories of galaxies. We consider that haloes are potentially affected by internal ghosts when they are round enough (defined by a given b/a threshold) and when their size is close to the size of an external ghost from a bright star (with a typical radius of 3.5 arcmin for MATLAS Karabal et al. 2017). In Figure 6.1, we plot the axis ratio of the main galaxy versus that of the halo. We colour-coded the points by the difference in position angle between the ellipse representing the main galaxy and the ellipse of the halo. A high difference can indicate a peculiar shape, worth visual inspection. This Figure shows our tool's usefulness when exploring data and finding potentially interesting or problematic haloes. However, in order to estimate the effect of internal ghosts on these potentially problematic haloes, instead of simply comparing their size, flux and surface brightness to the rest of the sample, we chose to work on a subsample of galaxies for which the PSF-deconvolved images were available.

6.3.1.2 PSF deconvolution

To remove the contribution of the internal ghosts, it is necessary to model the Point Spread Function (PSF) and then deconvolve images from the PSF. The PSF must be carefully modeled, with a focus on its outer wings where the light from ghost reflections comes into play. The PSF-deconvolution will impact the luminosity and size of galaxies. Indeed, the deconvolution will remove the additional light from internal ghosts, hence deconvolved galaxies will have lower luminosities and are expected to be smaller. These effects should be more important for haloes whose size is close to the size of the PSF, as

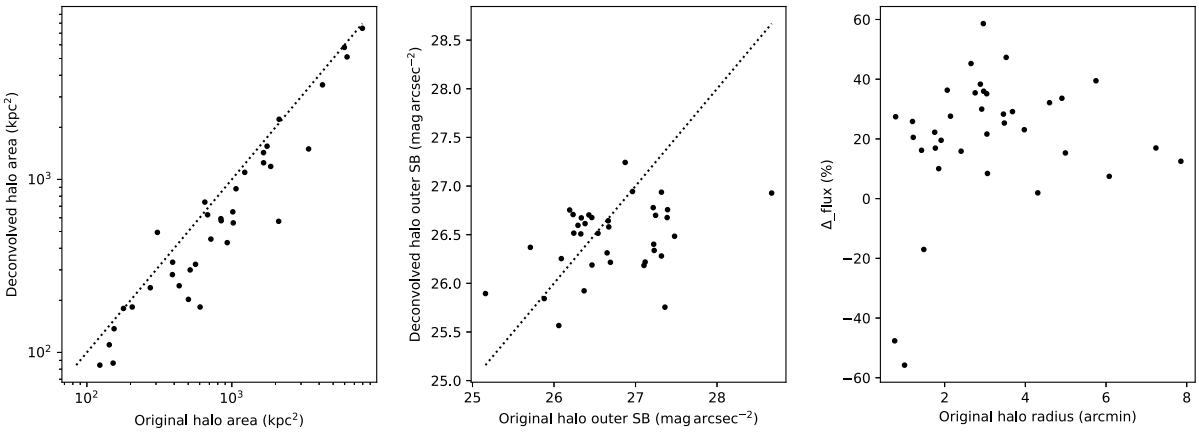

Figure 6.2: Figure adapted from Sola et al. (2023). Impact of PSF deconvolution on the area and fluxes in haloes. *Left*: Scatter plot of the areas covered by PSF-deconvolved versus original haloes, in kpc^2. *Middle*: Scatter plot of the halo outer SB (in magnitudes per square arcsecond) for PSF-deconvolved versus original haloes. *Right*: Percentage of difference between the flux of original and deconvolved haloes, as a function of the radius of the original halo in arcminutes.

the outer wings are mixed with the halo outskirts.

Fluxes The PSF for MATLAS images were carefully modeled by Karabal et al. (2017): they created a database of PSF models for MegaCam, and then deconvolved their images from these PSFs. Here, we focus on a subsample of 36 MATLAS galaxies that have been PSF-deconvolved. We annotated these 36 PSF-deconvolved MATLAS images after applying the same scaling and transformation than for the other images. Each galaxy was annotated both in the original and PSF-deconvolved images. In order to have comparable annotations, we only kept the annotations of one user (namely E. Sola) instead of performing our selection process. The flux (hence luminosity) inside the halo masks (i.e. without the contribution of the inner galaxy) and their size were compared for PSF-deconvolved and original images.

The results are presented in Figure 6.2, where two main points of interest arise. The first one is about the size of haloes. The left panel represents the areas covered by original and PSF-deconvolved haloes, and it appears that PSF-deconvolved haloes are smaller than the original ones. This was expected for two reasons. The first one is that if the observed halo was actually created by the PSF, the deconvolution will remove it, so the halo size will drastically decrease. The second one occurs when the observed halo is the actual stellar halo, then the deconvolution should not alter its size. However, the deconvolution process adds noise in the image which makes the identification of the external boundaries more complicated: as the outer isophotes are lost in the noise, it may lead to the annotation of slightly smaller haloes. In addition, haloes whose size is close to the PSF size are more likely to be impacted in their annotation (i.e. the halo would look as round and large as nearby external ghosts from bright stars), hence we would expect a higher area difference for a radius close to the typical radius of ghosts (3.5 arcmin for MATLAS). Both larger and smaller haloes should not be affected: for larger haloes the outer wings of the PSF do not extend up to the outskirts of the halo, while for smaller haloes the PSF wings extends much further than the outskirts of the halo, so in both cases there is no risk of confusion.

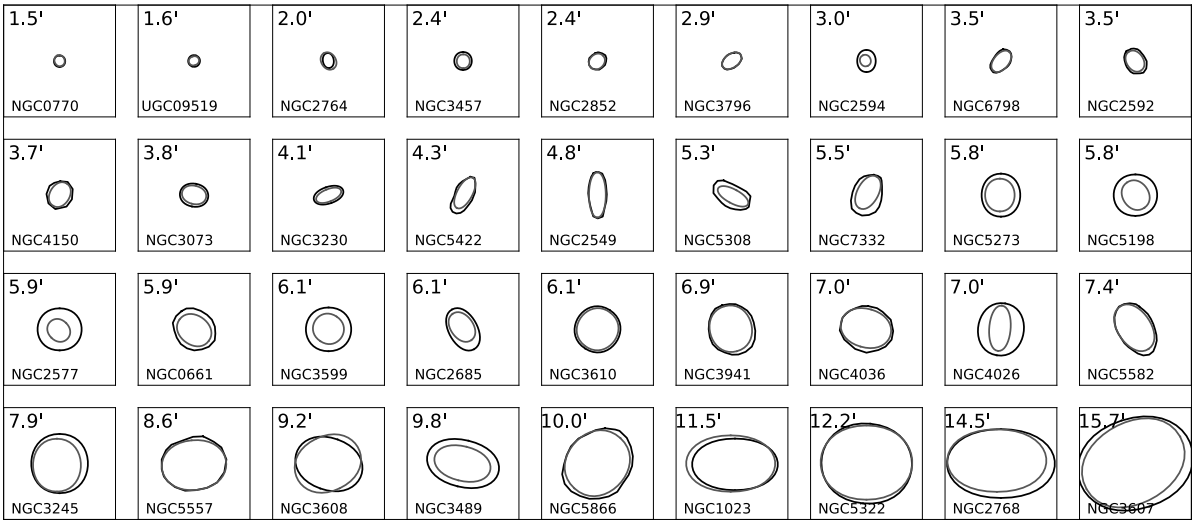

Figure 6.3: Comparison of the contours of original (in black) and PSF-deconvolved (in green) haloes. Each box has a fixed size of 15 arcmin, and haloes are ordered by increasing original radius. The name of the galaxy is indicated at the bottom of each box and its original diameter (in arcminute) is indicated at the top.

The middle panel of Figure 6.2 represents the percentage of difference of areas before and after deconvolution, as a function of the size of the original haloes. It appears that there is a large scatter visible, but the scatter is even higher for original haloes whose size is around 3 arcmin, which is what we expected. It must be noted that part of the scatter comes from the variations in the annotations (even between a single user) although this effect is relatively small and it is an intrinsic limitation of our method.

Morphologies Another way of representing this is to investigate the appearance of original and deconvolved haloes as their size gets closer to the typical radius of an external ghost. Figure 6.3 presents the contours of original (in black) and deconvolved (in green) haloes, ordered by increasing original radius size, in boxes of 15 arcmin. From it, it appears that there are no large differences between the contours as a function of the halo size, except for NGC 4026 with a diameter of 7 arcmin (i.e. the typical size of a ghost). The deconvolved halo is much more elliptical than the original one, and after inspection the original halo has the same characteristics than nearby external ghosts, so it had clearly been contaminated. For haloes with a radius smaller than 3.5 arcmin, there is no difference, with two exceptions: NGC 2594 and NGC 2577. NGC 2594's halo is overlapping with a very bright star which was removed after deconvolution (so the smaller deconvolved halo is not due to an internal ghost). NGC 2577, with a radius of 3 arcmin, is on the contrary clearly embedded in a ghost halo hence the deconvolved halo is smaller. For radii larger than 3.5 arcmin, there is no big difference between the sizes of original and deconvolved haloes. Therefore, these plots confirm our expectation that original haloes whose size is close to the one of an external ghost are more likely to be affected by the PSF.

A second point to study is the percentage of difference between the luminosity of the original and deconvolved haloes. The percentage of flux difference as a function of the radius of the halo is shown in the right panel of Figure 6.2. The median luminosity difference before and after PSF-deconvolution is around -15%. In addition, we expect a

larger luminosity difference for haloes close to the PSF size (3.5 arcmin for MATLAS), as the deconvolution will remove the flux contribution from the ghost. Although a large scatter is visible, it would seem that a higher luminosity difference is visible for a halo radius around 3 arcmin, which is in agreement with our expectations.

Overall, the PSF deconvolution accounts for a difference of luminosity of 15% in the halo, which must be kept in mind when interpreting the results. Since the luminosity in the halo itself represents about 10.6% of the luminosity of the entire galaxy, as discussed in Section 5.2.2, the impact of the PSF on the total galaxy luminosity is less than 2%. This is consistent with the 1.6% found by Karabal et al. (2017) for the external parts of their simulated galaxy. Trujillo & Fliri (2016) also find similar results for their haloes, with the original halo accounting for 11% of the total galaxy light, but only 3% for the PSF-corrected halo.

Colours A last point to investigate, related to the previous ones, is the impact of the PSF on colours. We saw in Section 5.2.1.2 that MATLAS haloes are redder than VESTIGE by about 0.1 mag and we hypothesized that the PSF was responsible for this. Indeed, as MATLAS is in the old r-band MegaCam filter, the outer wings of the PSF trigger large and bright ghost haloes due to bright sources (including the centre of galaxies) which causes a reddening of the halo. As an illustration of the more prominent ghost haloes in the old filters compared to the new ones, Figure 6.4 shows a comparison of three images from MATLAS and CFIS[1]. Ghost haloes are more prominent and there is more scattered light in the old filter set. Here, we check whether indeed this can be

Figure 6.4: Comparison between MATLAS images (*top panel*) with the old r-band MegaCam filter and the corresponding CFIS images (*bottom panel*) with the new r-band MegaCam filter.

[1]As a reminder, all our CFHT images (MATLAS, CFIS, NGVS, VESTIGE) images have been processed by the Elixir-LSB pipeline

the explanation. To that end, we retrieved the contour of the halo annotations (made on PSF-deconvolved images) and we applied the same contours to the corresponding original image (i.e. we created masks). This way, we ensure that the same regions are covered in both images. Then, we derived the flux inside the masks. We also retrieved from our previous computations the g-band flux value of the halo annotated in the g-band. We finally computed the integrated $g - r$ values and the results are presented in Figure 6.5.

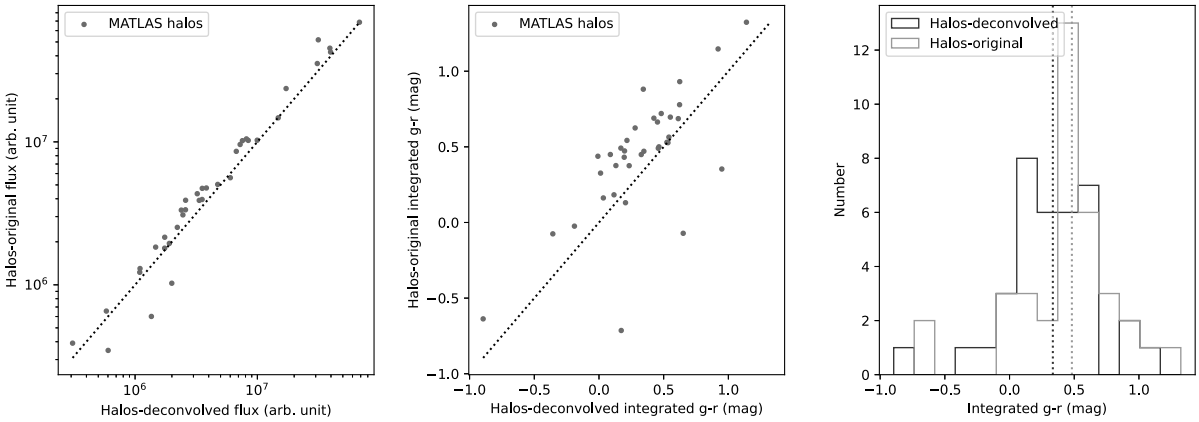

Figure 6.5: Impact of the PSF-deconvolution of MATLAS haloes on fluxes and colours. *Left*: Scatter plot of the flux (in arbitrary units) of MATLAS haloes in the original versus PSF-deconvolved images. *Middle*: Scatter plot of the integrated $g - r$ colour (in mag) of MATLAS haloes in original versus PSF-deconvolved image. *Right*: Histogram of the integrated $g - r$ colour (in mag) of MATLAS haloes in original (in orange) versus PSF-deconvolved (in purple) image.

As already mentioned, the PSF-deconvolved images contain less flux than the original ones because the additional amount of light from the PSF was removed. The middle and right panel of Figure 6.5 show the difference in the integrated $g - r$ values. It appears that PSF-deconvolved haloes are bluer than the original ones, by around 0.15 mag. This is consistent with what we hypothesized in Section 5.2.1.2: MATLAS haloes are redder than VESTIGE ones due to the artificial reddening coming from the PSF. If we remove this contribution (a difference of 0.15 mag), we would find similar $g - r$ values for the haloes of MATLAS and VESTIGE (which also needs to be deconvolved).

Finally, a last point to mention is the fact that the colour-correction we computed in Section 3.4.1 in order to account for the change of r-band filters between MATLAS and NGVS (old) and CFIS/VESTIGE (new) is only a linear relation. We computed it using four fields and selecting what seemed to be the best points, while an accurate colour-correction would require a more careful treatment.

To conclude, PSF-deconvolved haloes are smaller and about 15% less luminous than original ones, this must be kept in mind when characterizing the amount of flux in haloes. A higher difference is visible for haloes whose size is close to that of external ghosts from bright stars, confirming our expectation that these haloes are the ones that are most impacted by the PSF. From this, it appears that a simple size estimate can be a first indicator of whether the galaxy is contaminated by ghosts. In addition, not accounting for the PSF leads to overestimated $g-r$ colours due to the artificial reddening of the halo. Therefore, a precise analysis requires PSF-deconvolution but this is beyond the scope of this PhD work. This must be kept in mind when interpreting the results presented in Chapter 5.

6.3.2 Level of contamination

As mentioned in Section 2.3.8.1, we defined a reliability index to take the pollutant sources in the vicinity of the halo of the galaxy into account, as they might lead to biased annotations. The higher the reliability index, the cleaner the image around the galaxy. The histogram of the weighted reliability index can be seen in Figure 6.6. One can see that only 28% of the images are completely free of pollutants around the halo, while 58% are polluted (with an index smaller or equal to 3). This indicates that a majority of our annotations are embedded in a polluted region which might have biased our delineations, but more importantly this could be a major issue for automated classification methods.

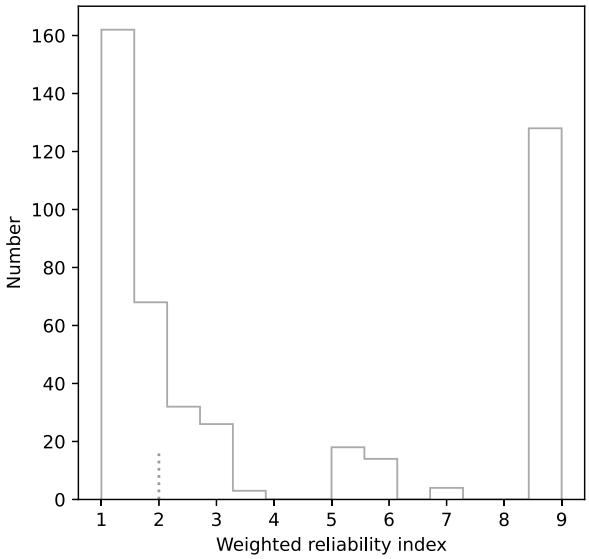

Figure 6.6: Histogram of the weighted reliability index for the galaxies. A higher index corresponds to a cleaner image around the halo. The median is indicated by the dotted line.

6.3.3 Covering factor

A complementary approach to the level of contamination is to determine the relative importance of pollutants. This can be done using the 'covering factor', defined as the percentage of the pixels belonging to one type of structure with respect to a given field of view. For the latter, we considered boxes of side length $20{R_e}^2$ centred on the primary galaxy. The covering factor can be computed both for sources of contamination but also for stellar features: this enables to obtain a proxy of the importance of tidal features based on their spatial extent without having to compute their stellar mass (as this computation from the annotation database is not straightforward).

For this measurement, we selected the annotations of a given galaxy and a given user. The median values of the covering factor for different types of structures are given in Table 6.1. A covering factor of 0 (respectively 1) means that the feature is not present on (respectively entirely covers) the given field of view. It is only computed in the images that do display the features. Results are shown in Figure 6.7. Within the selected boxes,

[2]The values of the effective radii for all our galaxies are given in Tables 3 and 4 from Cappellari et al. (2011a).

the stellar halo has a median covering factor of 0.18: it is computed for the entire outer ellipse annotation of the halo, without subtracting the main galaxy. This value is to be compared to 0.0035 for the central regions ('main galaxy'), 0.004 for tidal tails and 0.002 for streams, considering only galaxies that do have tidal tails or streams.

Table 6.1: Median of the covering factor in selected boxes of side length 20 R_e around the primary galaxy. The covering factor is only computed in the images that do display such features. In parenthesis is indicated the percentage of images around which the structures are found.

Feature type	Median covering factor
Main Galaxy	0.0035 (100%)
Halo	0.180 (94%)
Tidal Tails	0.004 (15%)
Streams	0.002 (12%)
Ghost reflections	0.140 (57%)

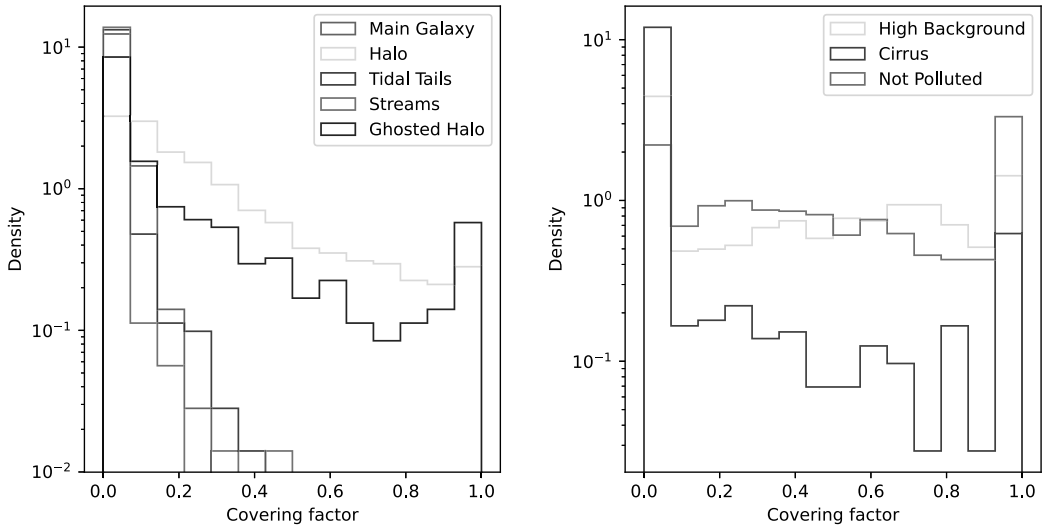

Figure 6.7: *Left*: histogram of the covering factor for stellar structures and ghost reflections in a selected box of side length 20 R_e centred on the primary galaxy. *Right*: histogram of the covering factor for high background, cirrus and clean ('Not Polluted') regions in a field of view of 30×30 arcmin around the primary galaxy.

For pollutants such as high background and cirrus regions, it is more relevant to determine their covering factor for larger fields of view, to get predictions on the contamination levels for other surveys. We considered a field of view of 30x30 arcmin, which corresponds to the size of our cropped CFHT images. For regions with high background, the covering factor has two peaks respectively near 0 and 1, with a relatively uniform distribution between these two values. It is higher than 0.9 (fully polluted images) for 12% of our galaxies and lower than 0.1 (clean images) for 33% of our galaxies. Restricting the analysis to the cirrus clouds, the histogram reveals images that are either fully covered or completely absent with very few intermediate cases. Overall, regions free of any contamination sources (except the ghosted haloes), i.e. with a covering factor higher than 0.9 for the clean pixels, correspond to about one fourth of our images, while 18% of our images are almost fully contaminated, i.e. clean regions cover less than 10

6.4 Combining multiple surveys

6.4.1 Impact of the depth of the survey

In this work, we used images from four different surveys with different depths, MATLAS and NGVS being deeper than CFIS and VESTIGE in terms of integration time (see Table 3.1). We already saw visually in Section 3.4 with Figure 3.8 or Figure 6.4 that although the faintest structures are more easily detectable in NGVS and MATLAS, the overall lower background level and less prominent ghost haloes in VESTIGE and CFIS tend to compensate for the lower integration time. This is due to the change of MegaCam filters, with the new r-band filter being less impacted by internal reflections of the light.

Here, we assess in a quantitative manner the impact of depth. In order to remove any effect from the environment, we kept galaxies outside clusters, i.e. CFIS and MATLAS. In Figure 6.8 we compare the environment density and mass of galaxies from CFIS and MATLAS. The large-scale environments (characterized by ρ_{10}) are comparable (same distributions, means and medians from the statistical tests) between both surveys. A warning must however be raised for the mass: some MATLAS galaxies are more massive than CFIS galaxies (the distributions of masses differ although the medians are similar). This may lead to more tidal features and more massive haloes for MATLAS galaxies. In addition, since most of our ETGs are drawn from MATLAS images and most LTGs from CFIS, we are in principle biased toward finding fainter features in ETGs. Here, we study the impact of the depth of the survey on tidal features and haloes properties.

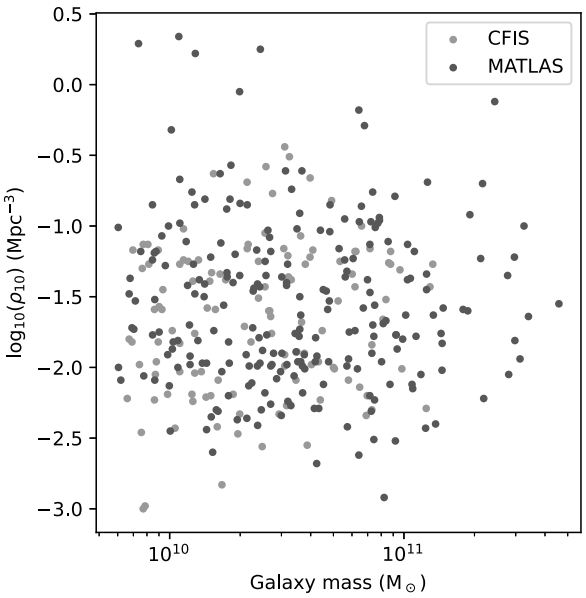

Figure 6.8: Scatter plot of the environment density ρ_{10} (in Mpc^{-3}) as a function of galaxy mass (in $\mathrm{M}_{\odot}$) for MATLAS (in green) and CFIS (in orange) galaxies.

Area and length vs SB of tails and streams To that end, we plot the 2D-histogram of the SB values of tails and streams as a function of their area (Figure 6.9), for ETGs and LTGs. A similar 2D-histogram of the length as a function of the SB is represented in Figure 6.10. One must note that we cannot compare these 2D-histograms between tidal tails and streams as other processes than the depth are involved (see Section 6.2.2), and

we must keep in mind the fact that we do not have the same number of tidal features around ETGs and LTGs.

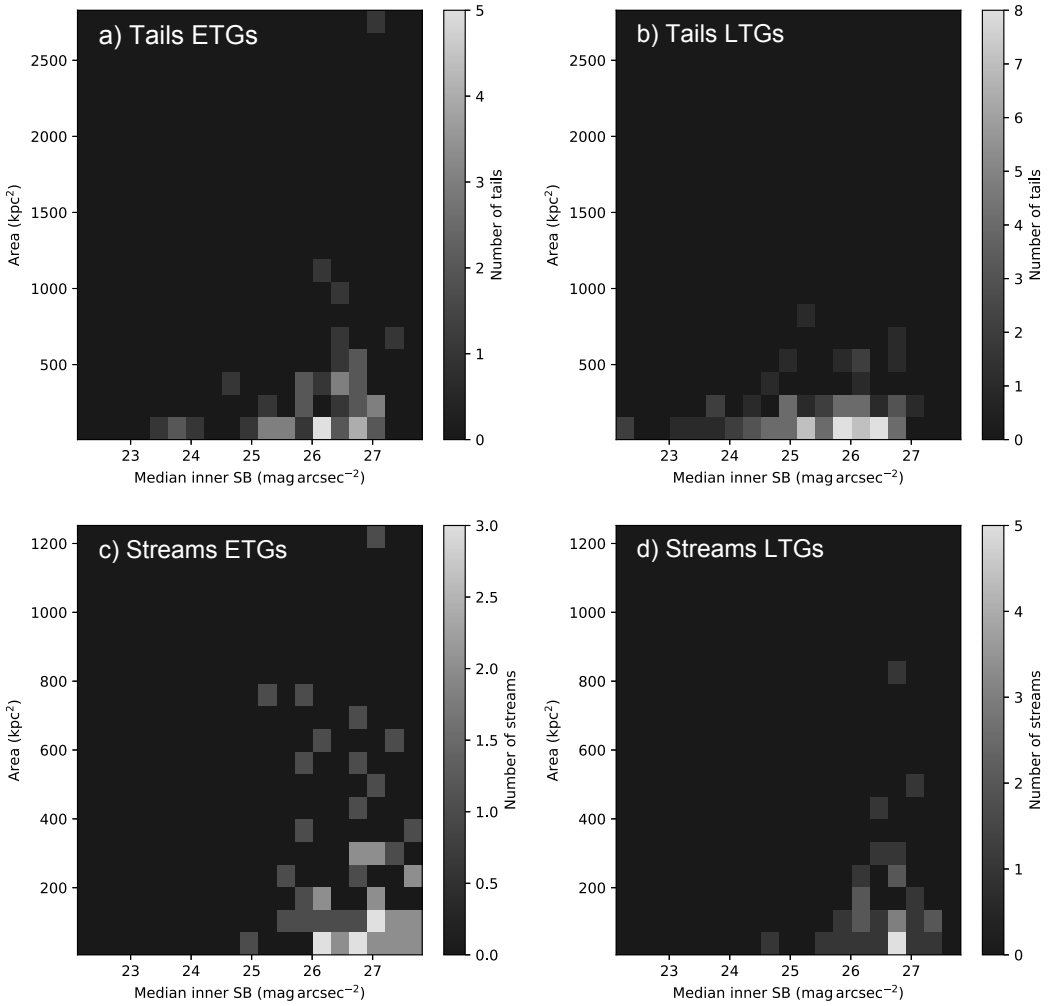

Figure 6.9: 2D-histograms of the area (in square kiloparsecs) of tidal tails and streams as a function of the median surface brightness (in mag arcsec^{-2}). This 2D-histogram for tidal tails around ETGs is presented in panel *a)* and for LTGs in panel *b)*. For streams, the histogram for ETGs is visible in panel *c)* and in panel *d)* for LTGs.

For the area (Figure 6.9) of tidal tails (panels a and b), there seems to be a slight trend that fainter structures are larger. Tails around ETGs are larger than around LTGs, while a few tails around LTGs are brighter, but for both morphological types the majority of the distributions are between 25 and 27 mag arcsec^{-2}. The bright tails around LTGs can be explained by the ones created during on-going mergers. The difference in the SB of tails around LTGs and ETGs is statistically significant. For the area of streams (panels c and d), streams around ETGs seem slightly fainter and a few are more extended than for LTGs. There is actually no statistically significant difference for the distributions of the SB of streams between ETGs and LTGs as revealed by the T-test, Mood's test and the 2-sample KS-test.

For the length (Figure 6.10) of tidal tails (panels a and b), it seems that fainter structures are longer (the trend is clearer for ETGs than for LTGs). The maximum SB of tails hardly differs: around ETGs it reaches about 27.3 mag arcsec^{-2} for one structure, compared to and 27.1 mag arcsec^{-2} for three tails around LTGs. For the length of

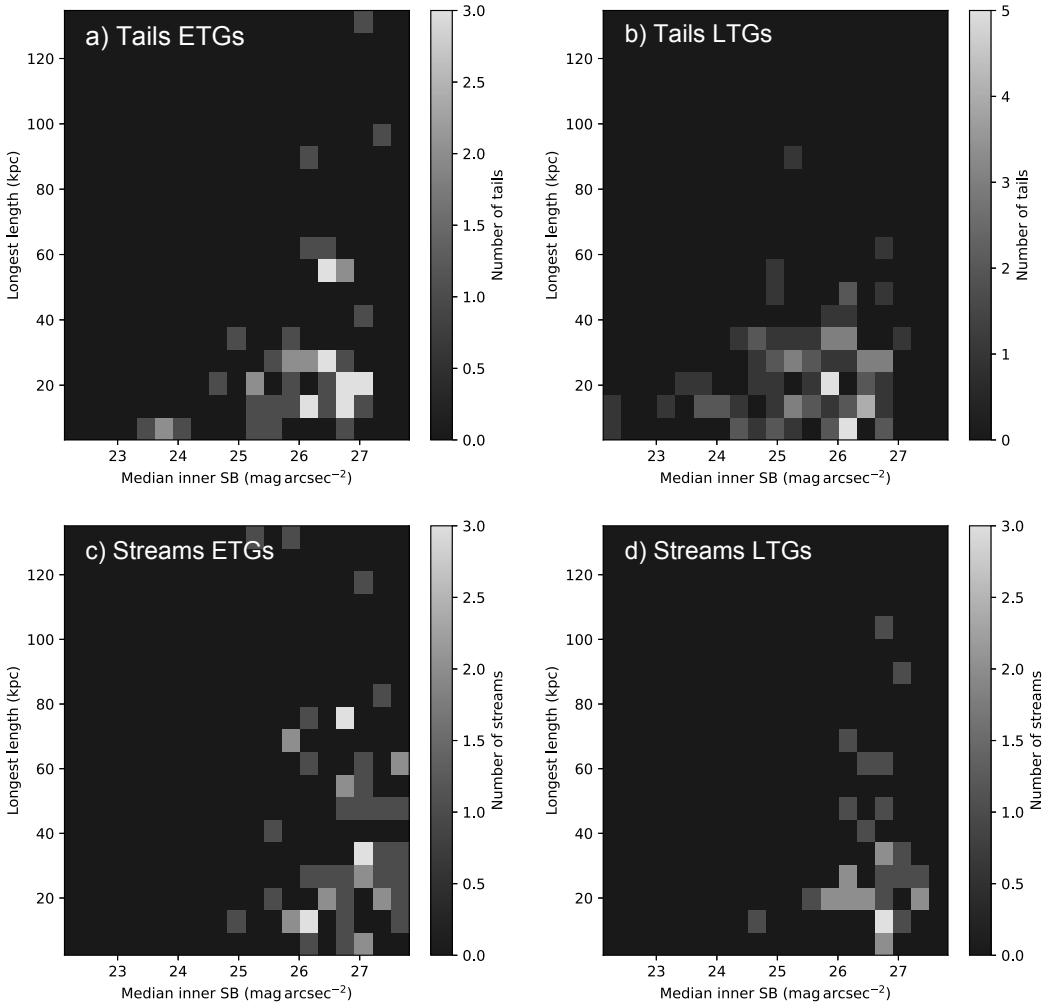

Figure 6.10: 2D-histograms of the longest length (in kiloparsecs) of tidal tails and streams as a function of the median surface brightness (in mag arcsec^{-2}). This 2D-histogram for tidal tails around ETGs is presented in panel *a)* and for LTGs in panel *b)*. For streams, the histogram for ETGs is visible in panel *c)* and in panel *d)* for LTGs.

streams (panels c and d), for ETGs several streams reach 27.8 mag arcsec^{-2}, which is 0.4 mag fainter than the faintest stream around LTGs. The longest streams (with a length around 80-125 kpc) are not necessarily the faintest ones, and the majority of streams have comparable length and SB for ETGs and LTGs.

From these 2D-plots, the difference of depth between CFIS and MATLAS does not seem to bias much our results. The fainter tails around ETGs than LTGs are explained by physical processes (i.e. on-going interactions or none) rather than by MATLAS being deeper than CFIS. There seems to be a trend that more elongated structures are fainter, especially around ETGs, but the trends are not so different between ETGs and LTGs.

Haloes We could also extend this analysis by combining other indicators, such as the outer SB of haloes, their size and the fraction of galaxies with debris. In order to remove any effect that could be due to the morphological type of the galaxy, we only kept LTGs for CFIS and MATLAS. We could not statistically compare ETGs as there are too few CFIS ETGs. We only visually compared three of the common images between CFIS and MATLAS in Figure 6.4. Using the statistical tests, we checked whether the mass and

environment of CFIS LTGs and MATLAS LTGs were similar: this is indeed the case (similar distributions and median values), so no bias on mass or environment should be introduced.

A first indicator to study are haloes. As MATLAS is deeper than CFIS, more extended haloes with fainter outer SB (i.e the SB of the contour of the halo annotation, similar to the outer isophote) are expected if the depth has an important effect. The left panel of Figure 6.11 shows the histogram of this value for LTGs in CFIS and MATLAS. With an overall median outer SB of around 26.4 mag arcsec^{-2}, there is no difference between both distributions, which is confirmed by the 2-sample KS test (p-value of 0.2), but the T-test could not be applied. One should note that even the faintest haloes do not reach a SB higher than 28 mag arcsec^{-2}, which is 0.3 to 0.9 mag fainter than the depth of the surveys.

In addition, from the right panel of Figure 6.11 and from the T-test and 2-sample KS test, it appears that there is no statistical difference in the area covered by haloes of LTGs from CFIS and MATLAS (with p-values of 0.97 and 0.2 respectively). Therefore, the difference of depth between CFIS and MATLAS does not significantly impact haloes.

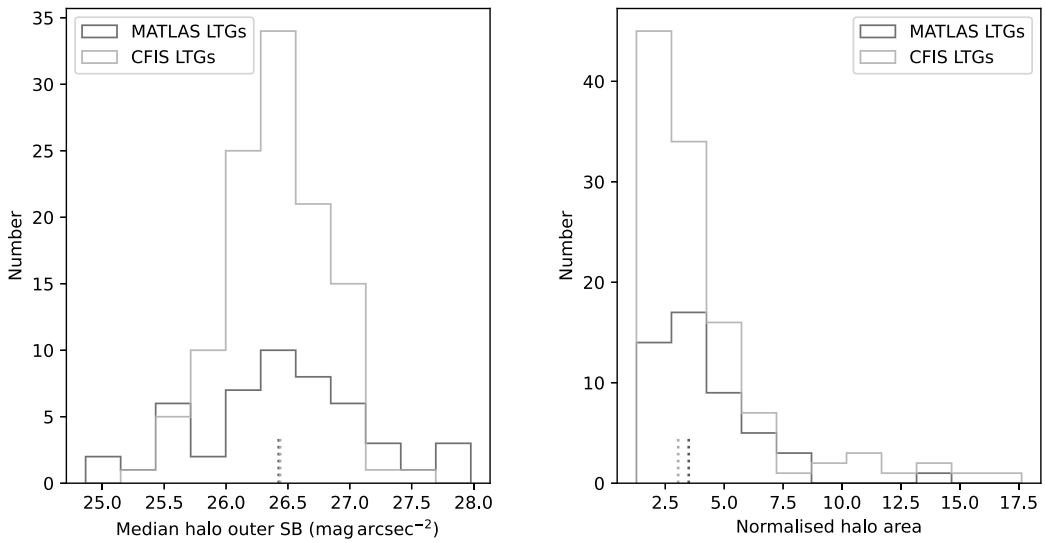

Figure 6.11: Histograms of the halo outer SB and size for MATLAS LTGs in green and CFIS LTGs in orange. The median values are represented by the dotted lines. *Left*: Histogram of the median haloes outer SB (in mag arcsec^{-2}). *Right*: Histogram of the normalised area of haloes (with respect to the area of the centre of the galaxy).

Galaxies with debris We expect to see more and fainter tidal features in deeper surveys. Yet, the fraction of tidal features shown in Table 6.2 does not present any substantial difference, which is also supported by the statistical tests. Indeed, for all feature types, the 2-sample KS-test shows similar distributions between CFIS and MATLAS, while the T-test revealed similar means for the fraction of galaxies with any type of debris, tails and shells (but it could not be applied to streams as the condition of equal variances was not verified). Our structures might be so bright (compared to the depths of the surveys) than any difference between the depth of CFIS and MATLAS would not affect much the visibility of the tidal features. This could also be explained by our difference of depths

Table 6.2: Fraction of galaxies that have the indicated tidal feature, for CFIS LTGs and MATLAS LTGs.

Type	Fraction any debris	Fraction with tails	Fraction with streams	Fraction with shells
MATLAS LTGs	0.36	0.3	0.06	0.15
CFIS LTGs	0.38	0.27	0.19	0.08

that is too small (0.5 mag only) to significantly affect our results.

To summarize this section, globally we found no statistically significant difference about the halo size and outer SB, or the number of tidal features between CFIS LTGs and MATLAS LTGs. All the clues tend to support the trend that the difference of depth between surveys does not bias much our results. In fact, background subtraction and systematic effects (such as the presence of prominent ghost haloes) have a much more important impact for LSB studies as they make the identification of LSB structures more complicated. This has a major impact for future surveys: a particular attention should be paid to the design of the instrument and of the filters (with the use of specific telescope design and/or special coating) as well as to the background subtraction pipelines, rather than integration time.

6.4.2 Effect of the change of the set of filters

In Section 6.3.1.2 we investigated the impact of the PSF on the measured fluxes, sizes and colours. Another point related to this is the comparison between the new and old filters in terms of cleanness of the images. We already saw that the old filters are subject to prominent ghost haloes from bright sources (see also Section 3.4) and here we want to quantify this effect. To that end, we compared the weighted reliability index (i.e. how

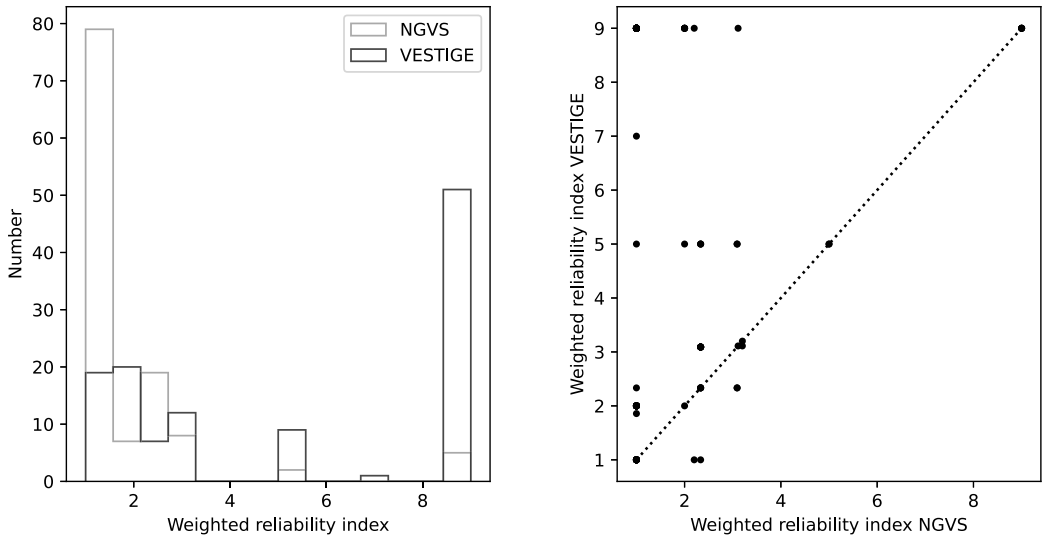

Figure 6.12: *Left:* Histogram of the weighted reliability index for NGVS (in orange) and VESTIGE (in purple). The higher the index, the cleaner the image. *Right:* Scatter plot of the weighted reliability index of VESTIGE as a function of the one for NGVS for the same fields.

clean images are) between NGVS (g-band, old filter) and VESTIGE (r-band, new filter). We only kept my annotations (as no other user had annotated every NGVS and VESTIGE

images) to have a consistency in the annotations. In Figure 6.12 we compare the weighted reliability index between both surveys. It clearly appears that NGVS images are much more impacted by contamination sources such as ghosted haloes and bright stars: 72% of images have an index lower than 2. On the contrary, 43% of VESTIGE images are clear of any contamination source, and 33% have an index lower than 2. The r-band new filter images are much cleaner than NGVS, although g-band images are less affected by PSF contamination (i.e. less bright ghost reflections) than the r-band. The identification of tidal features is therefore on average easier in VESTIGE images, even though VESTIGE is less deep than NGVS. This shows that instrumental effects including internal reflections due to filters (new versus old) have a more important impact than the depth of the survey in the detection and identification of tidal features.

6.5 Effect of masking and background-correction

We investigate here the impact of the step of masking the bright sources in our images and then applying the background correction to re-estimate locally the background value, as detailed in Section 2.3.3. To that end, we compared the fraction of flux in each individual tidal feature between the initial cutout images (i.e. with bright sources and no background correction), the clean images (i.e. with bright sources masked) and the final background-corrected images (i.e. with bright sources masked and background correction, which are the values presented in this PhD thesis). Such values are presented in Figure 6.13.

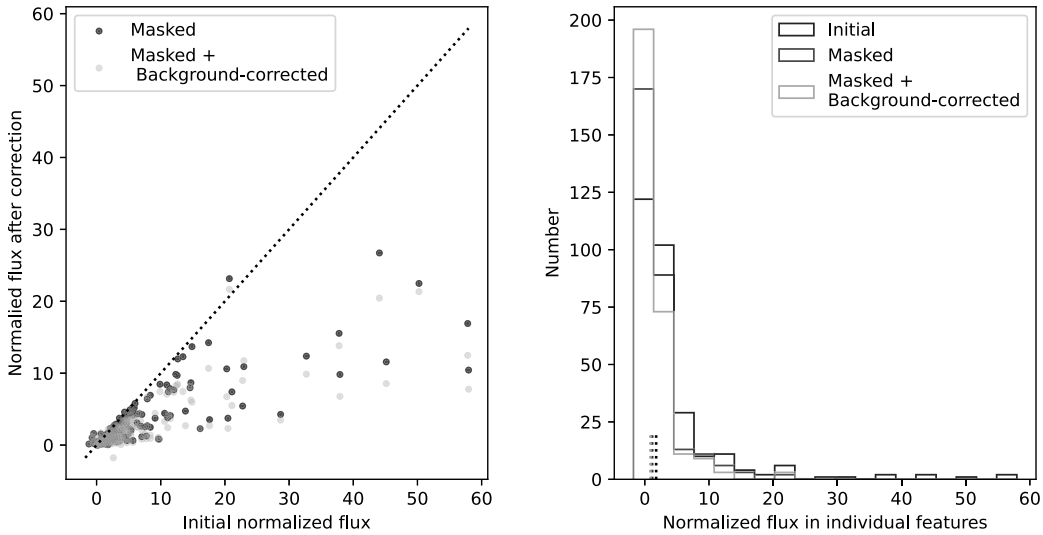

Figure 6.13: Comparison of the fraction of flux in individual tidal features (tails and streams), normalized with respect to the flux of the entire annotated halo. This fraction is computed on the initial cutout images, on the clean (i.e. bright sources masked) and on the final background-corrected (i.e. bright sources masked and background correction) images. *Left*: Scatter plot of the fraction of initial versus clean (in purple) and background-corrected (in orange) fluxes for each individual feature. *Right*: Histogram of the fraction of flux in tidal features, computed on the initial (in black), clean (in purple) and final background-corrected (in black) images.

From this Figure, bright sources account for 67% of the flux of tidal features, while an additional 33% comes from not performing the background re-estimation. Indeed, the

median (resp. mean) flux fraction in individual tidal features is 1.8% (resp. 4.5%) for the initial images; against 1.2% (resp. 2.4%) for the clean images and 0.9% (resp. 1.8%) for the background-corrected images.

6.6 Conclusion

In this chapter, we discussed the limitations of our study and their impact on the results presented in Chapter 5. We estimated the reliability of the annotations of the users, and we saw that there is less agreement for tidal features than for haloes due to the complicated task of their delineation. This task is even more complicated when the image is highly polluted. We discussed the possibility of attributing weights to users to take into account their level of expertise. We also discussed the intrinsic limit of our approach, namely the fact that manual annotation is very time consuming and will not be possible to carry out in future large astronomical surveys. This is why deep learning methods are needed, but our annotation database can be used as a training set for such algorithms.

Then, we focussed on the biases related to the identification of tidal features. We insisted on the discrepancy between the nominal depth of the surveys and the faintest annotated stellar features, linked to our visual detection itself, that must be kept in mind when comparing observations and simulations. Then, we discussed the distinction we made between tails and streams which was validated by the statistical differences we observed, even though there is a large variability in their individual shapes.

Another important aspect was to assess the impact of pollutants on our annotations. A first type of contamination sources are ghost haloes from internal reflections of bright sources in the telescope and camera. It is necessary to deconvolve images from the PSF in order to mitigate this effect. Indeed, haloes are more extended and 15% more luminous because of the artificial amount of light generated by the PSF. Likewise, the PSF leads to a reddening of the haloes which prevents the direct comparison between different surveys. A precise estimation of the flux and sizes of LSB features requires PSF-deconvolution, which is beyond the scope of this PhD work. We also assessed the impact of other contamination sources (cirrus, high background) and it appears that many of our images are polluted by various sources of contamination.

We discussed about combining different surveys together, with different filters and depth. It appears that the difference of depth we have (around 0.5 mag) does not impact much our results, neither in terms of geometrical properties, SB nor fraction of galaxies with debris. The change of MegaCam filters has however a much more significant impact (as already seen with the sections on the PSF). It appears that the quality of the filters, related to the shape of the PSF, is more important than the depth in the detection and identification of tidal features.

Finally, we investigated the impact of masking bright sources and re-estimating the background during automated aperture photometry. Without these two steps, the fraction of flux in tidal features is overestimated by a factor of two. The step of masking is the most important one in comparison to the re-estimation of the background.

6.7 Résumé du chapitre

Dans ce chapitre, nous avons examiné les limites de notre étude et leur impact sur les résultats présentés dans le Chapitre 5. Nous avons estimé la fiabilité des annotations des utilisateurs, et nous avons constaté qu'il y a moins d'accord pour les débris de marée que

pour les halos en raison de la complexité de leur délimitation. Cette tâche est encore plus compliquée lorsque l'image est très polluée. Nous avons discuté de la possibilité d'attribuer des poids aux utilisateurs pour tenir compte de leur niveau d'expertise. Nous avons également discuté de la limite intrinsèque de notre approche, à savoir le fait que l'annotation manuelle prend beaucoup de temps et ne pourra pas être réalisée dans les futurs grands relevés astronomiques. C'est pourquoi les méthodes d'apprentissage profond sont nécessaires, mais notre base de données d'annotation peut être utilisée comme un ensemble d'entraînement pour de tels algorithmes.

Ensuite, nous nous sommes concentrés sur les biais liés à l'identification des débris de marée. Nous avons insisté sur l'écart entre la profondeur nominale des relevés et la brillance de surface des structures les plus faiblement lumineuses, lié à notre détection visuelle elle-même, et qui doit être gardé à l'esprit lors de la comparaison des observations et des simulations. Ensuite, nous avons discuté de la distinction que nous avons faite entre les queues de marée et les courants stellaires, qui a été validée par les différences statistiques que nous avons observées, même s'il y a une grande variabilité dans leurs formes individuelles.

Un autre aspect important était d'évaluer l'impact des polluants sur nos annotations. Un premier type de sources de contamination est constitué par les halos fantômes provenant des réflexions internes des sources lumineuses dans le télescope et la caméra. Il est nécessaire de déconvoluer les images à partir de la PSF afin d'atténuer cet effet. En effet, les halos sont plus étendus et 15% plus lumineux à cause de la quantité artificielle de lumière générée par la PSF. De même, la PSF conduit à un rougissement des halos qui empêche la comparaison directe entre différents relevés. Une estimation précise du flux et de la taille des caractéristiques des LSB nécessite une déconvolution de la PSF, ce qui dépasse le cadre de ce travail de thèse. Nous avons également évalué l'impact d'autres sources de contamination (cirrus, arrière-plan lumineux) et il apparaît que beaucoup de nos images sont polluées par diverses sources de contamination.

Ensuite, nous avons discuté de la possibilité de combiner plusieurs relevés, avec différents filtres et différentes profondeurs. Il apparaît que la différence de profondeur que nous avons (environ 0,5 mag) n'a pas beaucoup d'impact sur les résultats, ni en termes de propriétés géométriques, de SB, ni de fraction de galaxies avec des débris. Le changement des filtres de MegaCam a cependant un impact beaucoup plus important (comme nous l'avons déjà vu avec les sections sur la PSF). Il semble que la qualité des filtres, liée à la forme de la PSF, soit plus importante que la profondeur des images pour la détection et l'identification des débris de marée.

Chapter 7 Prospects

The annotation tool we developed, the associated database and results presented in Chapters 2 and 5 enabled us to probe the late assembly of galaxies through the study of LSB tidal features. In this chapter, I present some future prospects which will extend the current work and benefit from our annotation tool. This includes annotated numerical simulations, machine learning and other potential applications.

7.1 Simulations

7.1.1 Numerical simulations and galactic evolution

The goal of my PhD thesis was to build a statistically significant sample of LSB tidal features around hundreds of massive nearby galaxies, to retrieve quantitative measurements and to carry out a detailed statistical analysis. We interpreted the results according to hierarchical models of galactic evolution, in which galaxies grow through mergers and accretion of gas. Numerical simulations are a way to test these models: by comparing the predictions of simulations (where the ground truth[1] is known) about tidal features and

[1]The ground truth in simulations means the entire histories, merger trees and properties of galaxies are known.

extended stellar haloes to the observed ones, it is possible to constrain models of galactic evolution.

Numerical simulations, carried out over several decades, have shown that baryonic matter (gas and stars) can be used to trace back merger events through the formation of tidal features (e.g., Toomre & Toomre 1972; Dubinski et al. 1996; Cooper et al. 2010; Ebrova 2013; Hendel & Johnston 2015). For instance, Howard et al. (1993) created a 'dictionary' of numerical simulations with various encounter parameters so that observers could compared the morphology of observed tidal features to the predictions of simulations in order to find the parameters that best match observations. Conversely, this dictionary can be used with observational samples for statistical studies of system parameters. Johnston et al. (2008) used numerical simulations to investigate which parts of the assembly history of galaxies could be inferred from the observed properties of stellar haloes. They found that the morphology and number of substructures in haloes (related to a recent merger) give hints about the mass, orbit and accretion time of progenitors. Hendel & Johnston (2015) also showed that the morphology of tidal debris could be used to obtain information about the orbit distributions of the progenitors. In particular, streams are highly studied: as they approximately follow the orbit of their progenitor (e.g., Dehnen et al. 2004; Eyre & Binney 2011), they can be used to derive information about the underlying gravitational potential and dark matter distribution (e.g., Hernquist & Quinn 1987; Johnston 1998; Lux et al. 2013; Ebrova 2013; Thomas et al. 2017; Nibauer et al. 2023). Numerical simulations can also be used to infer the time at which the merger occurred (e.g., Lotz et al. 2008; Ji et al. 2014; Mancillas et al. 2019). Another advantage of simulations is the possibility to investigate projection effects (e.g., Martin et al. 2022): indeed, tidal features are not detected equally along different inclination angles (under which the galaxy is observed) (Mancillas et al. 2019) and their shape may be mistaken for another type of feature, like streams resembling shells in some cases (Hendel & Johnston 2015; Greco et al. 2018).

However, major limitations arise when comparing observations to simulations. Like when comparing different observational surveys, different teams do not necessarily have the same definitions of what a tidal feature or halo is, neither the same depths, bands, or detection methods. In addition, images from simulations are clean of any contamination sources, so realistic mock images are needed to be comparable to observations (e.g., Bickley et al. 2021). Our annotation tool enables to annotate and later analyse simulations and observations in the same way. This enables a more straightforward comparison between both types of images, and later a direct comparison of LSB structures.

During my PhD thesis, I started to work on the comparison between observations and simulations. This work was conducted as part of a Master 1 internship with Pierre Galois, whom I co-supervised, and of external collaborations with Marteen Baes, Crescenzo Tortora and their team, as well as with Ivana Ebrova. We also had the contribution of Ilona Morel for some annotations of simulated galaxies, during a one-month internship. The aim of P. Galois' internship was to determine whether the available simulated images could address the issue of comparing simulations and observations and whether they would be adapted to it. We used a total of two sets of simulated images, and I contributed to annotate them. The characteristics of the simulated images and the preliminary results derived from the analysis of the annotations will be presented in the following sections.

7.1.2 Mock images from simulations

Illustris TNG50 The goal of this project was to determine whether the simulated images that we had access to could be used to compare observations to the predictions of numerical simulations.

This project was conducted with P. Galois during his internship, in collaboration with Maarten Baes, Crescenzo Tortora and their team, who provided us with mock images and galaxies' properties. They created mock images from the Illustris TNG50 simulations. The IllustrisTNG project (Nelson et al. 2017; Springel et al. 2017; Naiman et al. 2017; Marinacci et al. 2017; Pillepich et al. 2017, 2019; Nelson et al. 2019) is a suite of state-of-the-art cosmological galaxy formation simulations. The project consists of three volumes (cubic volumes of 50 (TNG50), 100 (TNG100) and 300 (TNG300) Mpc of side length), with a total of 18 simulations. The physical size, mass resolution, and complexity of physics included vary between the simulations. The simulations were run up to $z = 0$. M.Baes and C.Tortora's team provided us with images of 136 galaxies from Illustris TNG50 at $z = 0$, seen from 6 different viewpoints. They made their images to be 'Euclid-like', both in terms of photometric band (Euclid VIS) and resolution (0.1 arcsecond per pixel). They also created 'LSST-like' images (photometric bands g, r, i, u). The field of view of the images is 80×80 kpc, with galaxies artificially placed at 82.5 Mpc. We applied a SB limit of 29.5 mag arcsec^{-2} to match Euclid's estimated depth. An illustration of the TNG50 images is presented in Figure 7.1.

In a first step, we studied the TNG50 simulated images and the galaxies in them, to determine whether they had characteristics comparable to our observational sample. The use of a radiative transfer code enabled to obtain more realistic spiral galaxies with dust taken into account, leading for instance to the formation of dust lanes. The number of simulated galaxies (136) was relatively large for a first simulated dataset. The characteristics (photometric band, resolution, $z = 0$) of the mock images were approximately comparable to our observations. However, we also discovered during the exploration of the data that the TNG50 images also suffered from various issues. The first is the fact that they only contained the main galaxies and not companions. Indeed, the objects were selected from the simulation using a friend-of-friend algorithm and images have been reconstructed after it. Any particles belonging to gravitationally bound subhaloes were left out, which means there are no companions in the images. This has major consequences, as we cannot see the tidal tails originating from a massive companion nor streams from a lower-mass galaxy, and we cannot see the local environment of the target galaxy. A second issue comes from the transformation from particles in the simulation to 2D images. The team of M.Baes and C.Tortora used a projection algorithm within the radiative transfer code SKIRT (Camps & Baes 2014) that first interpolates everything on a grid and subsequently integrates along the line of sight. This resulted in numerical artefacts in some images, especially at the faintest SB where there are few particles. Following on the targets of the simulation, the team mainly based the sample selection on the wish to properly sample the relation between stellar mass and star formation rate. It resulted in a bias towards massive large spiral galaxies and a lack of early-type objects. A third problem came from the field of view. The most massive galaxies were larger than the field of view, which made their annotation impossible as we wanted to probe the outskirts of galaxies. This resulted in some galaxies being rejected from our sample.

Newer versions of these images with less issues were created by M.Baes and C.Tortora after our initial study but we did not have the time to exploit them yet. The annotations and the results were done on the first version of the images. From the original sample of

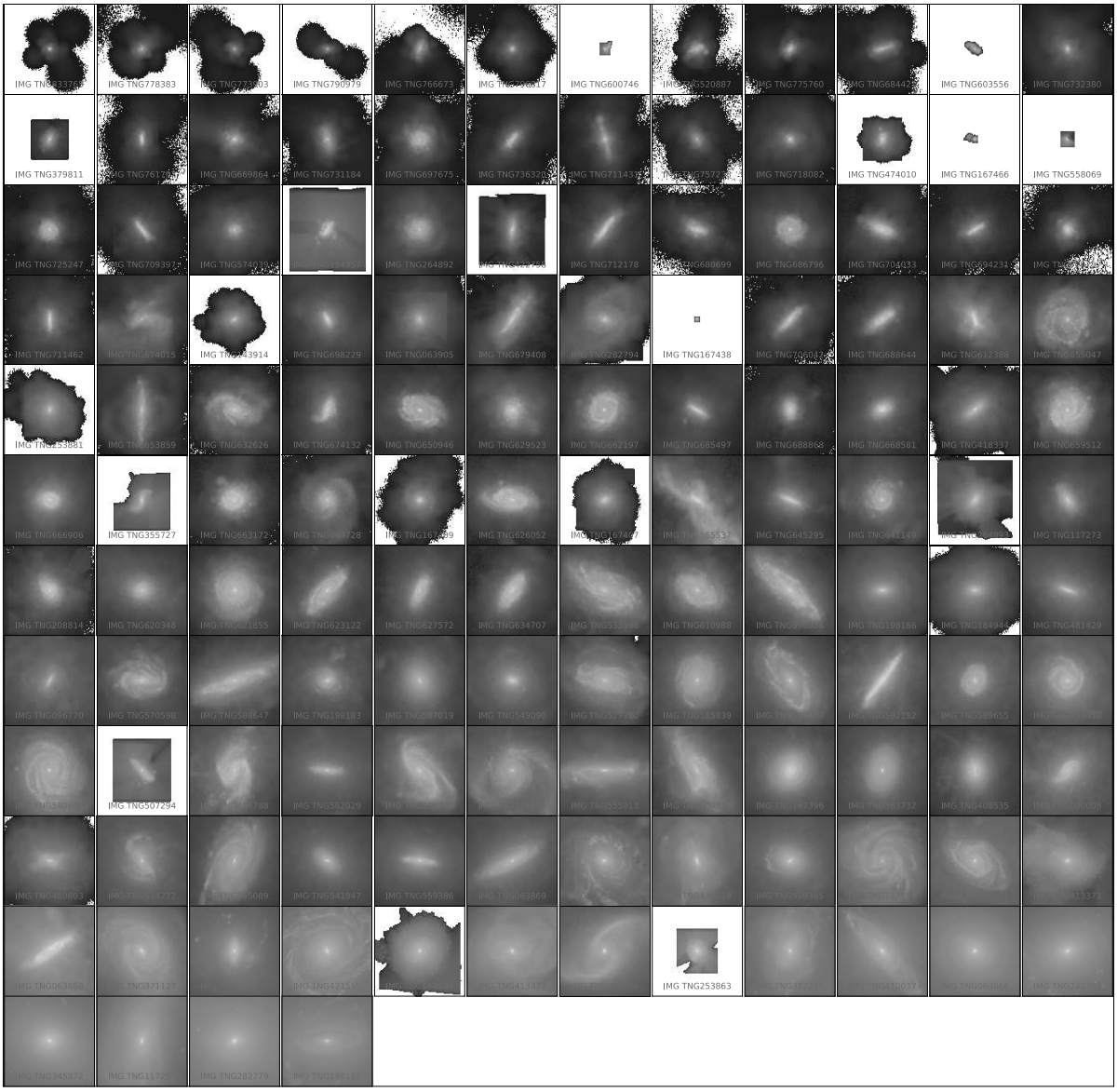

Figure 7.1: Thumbnails of the TNG50 simulated images, seen along the A05 viewpoint and ordered by increasing mass. The name of each simulated galaxy is indicated at the bottom.

136 galaxies, we kept only 70 that were small enough compared to the field of view, that did not contain large artefacts and that were in a mass range comparable to that of the ATLAS^{3D} observational reference sample. As mentioned, almost all the simulated galaxies are LTGs. For each galaxy we kept three viewpoints, which makes a total of 210 simulated images. In order to integrate them to the annotation server, we had to artificially place the galaxies at RA,DEC=(180°,0°), and we converted the images to asinh scaling and then to HiPS format. We applied a cut in the HiPS pixel values, between 0 and 2. The simulated galaxies were added to the database. We also created 'shallower' images to reproduce the equivalent of PanSTARRS to annotate the main galaxies, by combining u, g, i LSST-like images. The annotations of these simulated images were done by Pierre Galois during his Master 1 internship and myself.

Illustris TNG100 This project has been conducted in collaboration with Ivana Ebrova (the results will be published in Ebrova & collaborators 2023), with the goal of comparing the radius of shells in simulations, as they seem to be too large compared to observations. In observations, shells can extend from a few to 100 kpc. As a reminder, we found in Section 5.1.1.6 that shells had radii smaller than 60 kpc with a few extending to 80 kpc. On the simulation side, shells extend further than this, with for instance radii up to 120 kpc in Karademir et al. (2019) and some reaching 150-200 kpc in Pop et al. (2018).

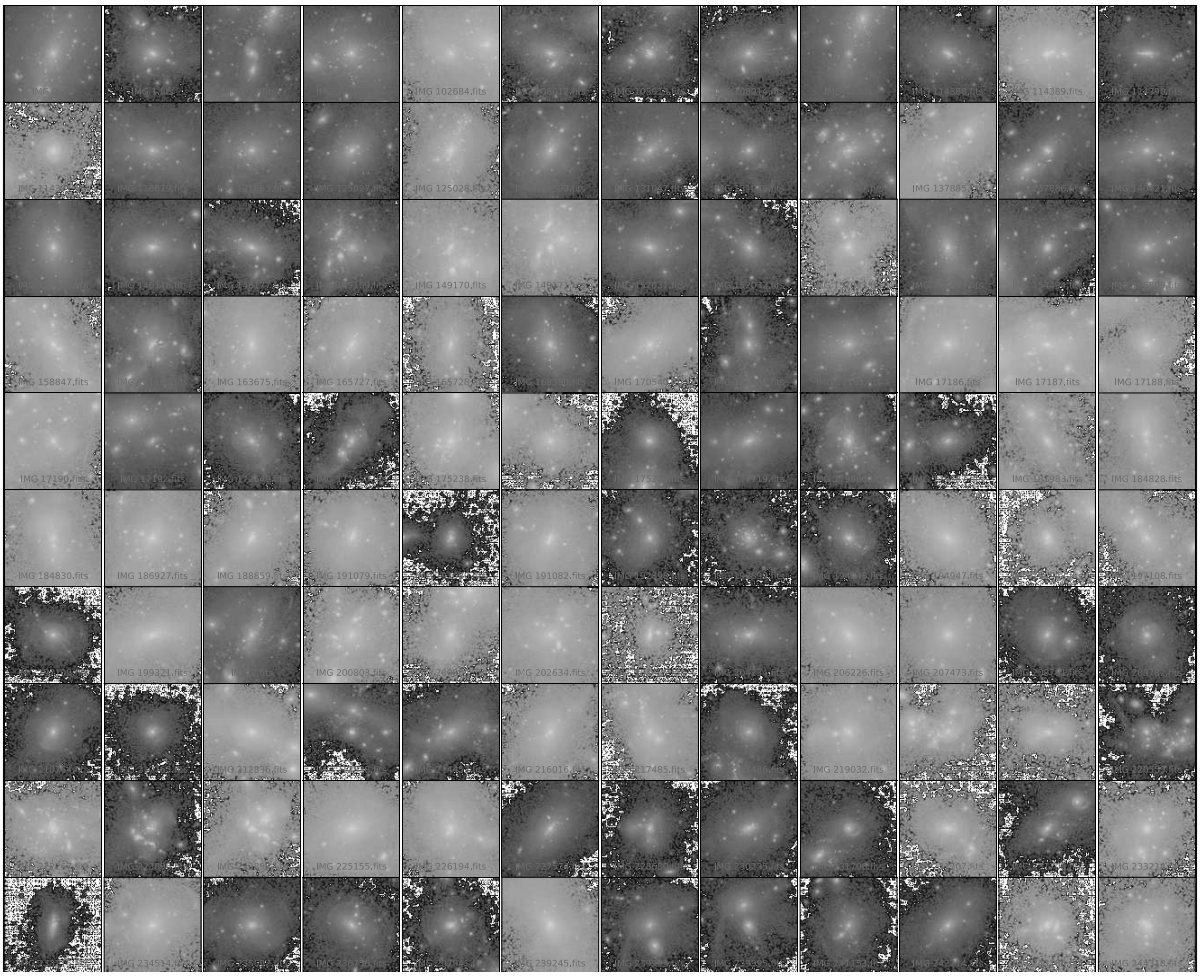

Figure 7.2: Thumbnails of some TNG100 simulated images. The name of each simulated galaxy is indicated at the bottom.

I.Ebrova provided us with mock images from the Illustris TNG100 simulations for 300 galaxies at $z = 0$. The procedure to generate the mock images was different than for the TNG50 images, therefore companion galaxies (and their potential tidal tails) can be present. The simulated galaxies are placed at a distance of 40 Mpc. The field of view of each image is 800×800 kpc, with a resolution of 4 arcseconds per pixel. An illustration of some TNG100 simulated galaxies are shown in Figure 7.2. In order for the images to be comparable to our observations, we applied a SB limit of 28.5 mag arcsec^{-2} (between the depth of CFIS and MATLAS). The images are in the i-band. A small fraction of galaxies are slightly larger than the field of view. In a similar manner than for the TNG50 images, we artificially placed the galaxies at RA,DEC=(180°,0°). After some tests to make the images look like observational ones, we created the asinh images with $\alpha = 1$ and the HiPS files from them (without applying a further cut). We also created shallower colour images

to reproduce the equivalent of PanSTARRS to annotate the main galaxies. To that end, we combined g, r, i images of the galaxies and we applied a SB cut at 25 mag arcsec^{-2}. We also created the HiPS of these images, and we uploaded them to the annotation server.

Our galaxy sample is composed of 300 galaxies. At the time this manuscript was written, we did not have additional information about them (like morphological type, mass or half-light radius) so our analysis was restricted to the tidal features themselves and not to correlations with the host galaxy properties. I.Ebrova annotated all the galaxies in the i-band, focusing only on main galaxies and shells (without any other label). We also had the contribution of Ilona Morel for a few galaxies, during a one-month internship where I briefly co-supervised her, but we only kept for the results presented here the annotations by I.Ebrova as she is an expert user. The annotations by I.Morel were useful as she annotated some simulated and observational images in a similar way, but the sample is not large enough to be statistically significant.

7.1.3 Preliminary results

Here, we present some preliminary results based on the comparison between observed and simulated tidal features and haloes. A statistical study would require more annotations of simulated images, in the same band that the observed images and with the overall same properties (distances of the galaxies, resolution, SB limit, presence of artefacts..). The available simulated images still enable us to perform a first analysis. We compared simulations to the results of our annotations of MATLAS, CFIS and VESTIGE images, which were presented in Chapter 5.

Thanks to our annotation tool, several quantities can be compared similarly between observations and simulations. Since the galaxies are located at different distances (40 Mpc for TNG100, 82.5 Mpc for TNG50 and between 5 and 45 Mpc for the observations), the angular and physical sizes are not comparable. We needed to find a normalization that enables us to compare features in the same way. As we did not have information about the effective (or half-light) radius for TNG100 galaxies[2], we decided to use the properties of the annotated main galaxy (area, length) as the reference. Indeed, the main galaxies were annotated on shallower images, which should be comparable between observations and simulations. We also tried to use the haloes as the reference, but as they were sometimes too large for the field of view, they were less reliable than main galaxies. Although this normalization is not perfect (due to delineation errors or variability), it is still a first approximation.

Shells We started by working on shells. Their number, radius and distributions can be compared to numerical models to constrain the mass distribution of the galaxy and to time the merger event (e.g., Quinn 1984; Dupraz & Combes 1986; Hernquist & Quinn 1987; Canalizo et al. 2007; Pop et al. 2018). We compared the radii of shells between observations and Illustris TNG100. We used the same definition than the one introduced in Section 5.1.1.6, i.e. the 'peak' radii: they are the most representative shell radii for each galaxy when taking into account all the shells annotated by all users. The results are visible in Figure 7.3. In the left panel of this Figure, we plotted the peak radii of shells in kiloparsecs. There is a clear difference between the distribution of the observed shells, with a median peak radius of 25 kpc and no shell extending further than 100 kpc,

[2]Due to the limited amount of time allocated to this project and the preliminary analyses, we did not compute the effective radii of the simulated galaxies.

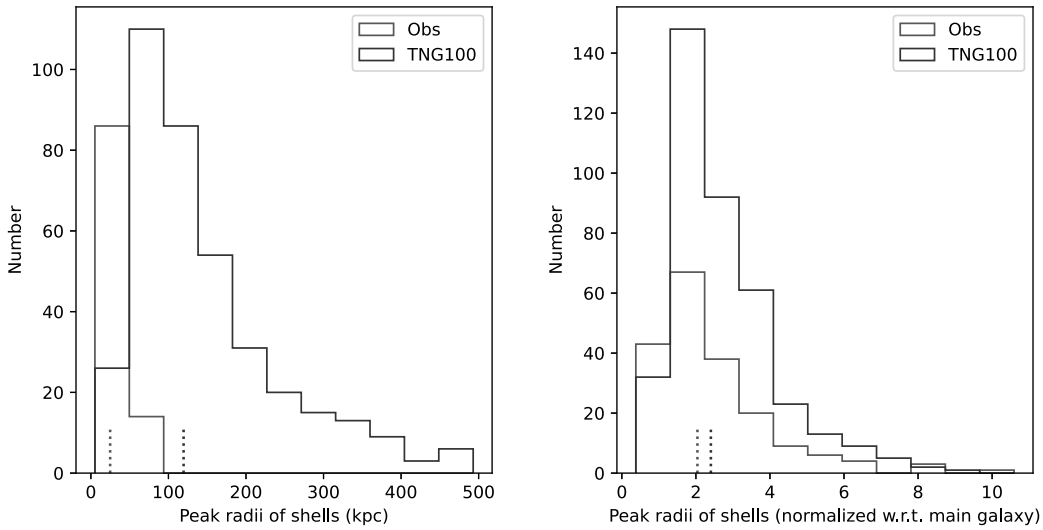

Figure 7.3: Histograms of the peak radii of shells. Observed shells (from CFIS, MATLAS and VESTIGE) are plotted in green, while simulated shells from Illustris TNG100 are plotted in purple. The median of the distributions are represented by the dotted lines. *Left*: Histogram of the peak radii of shells in kiloparsecs. *Right:* Histogram of the peak radii of shells normalized with respect to the radius of the main galaxy.

and the simulated ones extending up to 500 kpc with a median of 120 kpc. If we only look at this result, we find strong disagreement between observations and simulations, with simulated shells being much more extended. However, galaxies do not have the same physical size hence some normalization is needed. As the haloes were not annotated in TNG100, we normalized with respect to the radius of main galaxies, as presented in the right panel of Figure 7.3. It appears that once normalized, the distributions of the peak radii of shells look more similar than before, between observations (with a median of 2 times the radius of the main galaxy) and TNG100 (with a median of 2.4 times the radius of the main galaxy). The statistical tests (2-sample KS test, T-test and Mood's test) however revealed statistically significant differences in the distributions, median and mean values: simulated shells are more extended (i.e. further away from the galaxy centre) than observed ones when normalized with respect to the main galaxy. However, the comparison is not that straightforward as the properties of our samples of galaxies differ.

A proper analysis would be to compare these values or the incidence of shells for galaxies of similar mass or similar size but we did not have these pieces of information for these preliminary results. Likewise, we could not compare the fraction of galaxies hosting shells. Indeed, the selection criterium for TNG100 was to include both galaxies with and without shells, without the goal of being complete because the aim was to study the radii of shells (not their census). In addition, I.Ebrova is an expert in shells so she may detect more structures than other annotators, and she did not look at observations (conversely, I did not look at TNG100 images). This can add an important variability in the number of shells detected.

Mass in tidal features and haloes A second aspect that we investigated is the comparison of the percentage of stellar mass contained in tidal features and haloes between observations and simulations. The mass of observed features was derived following 2.3.6.

For simulated ones, we had access to the stellar mass maps, so we only had to apply the masks of the annotations to retrieve the total mass enclosed in features. We only had such information for TNG50 galaxies (and tidal tails and streams were not annotated in TNG100 images). In addition, in TNG50 we had several viewpoints for each galaxy so we only conserved one (named A05) in the preliminary results below. This means that we have a total of 70 galaxies, with a relatively low number of tidal features identified. As TNG50 galaxies are almost all LTGs, we kept LTGs from observations (CFIS, MATLAS and VESTIGE). Likewise, as there cannot be companion galaxies or smaller satellites in TNG50 images, we only kept tidal tails from the observations. We present the mass fractions for tails and haloes in Figure 7.4.

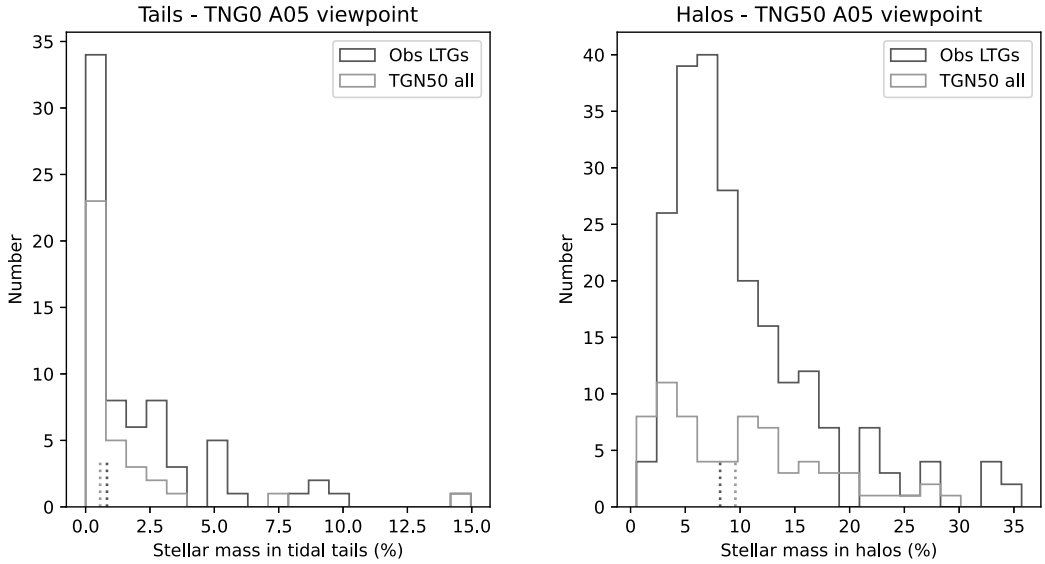

Figure 7.4: Histograms of the percentage of mass contained in haloes and tails. Observational features (CFIS, MATLAS, VESTIGE) for LTGs are plotted in green, while simulated ones from the Illustris TNG50 (mostly spirals), for the viewpoint A05, are plotted in orange. The median of the distributions are represented by the dotted lines. *Left*: Histogram of the percentage of stellar mass contained in tidal tails around LTGs. *Right*: Histogram of the percentage of stellar mass contained in haloes of LTGs.

Concerning tidal tails one can see from the left panel of Figure 7.4 that they only account for a very small percentage of the total stellar mass, both in observations (with a mean value of 2%) and simulations (with a mean value of 1.4%). There are fewer very massive tails in TNG50, which is due to the fact that we cannot see ongoing interactions, hence bright tails, with a nearby companion because of the way simulated images were created. If we concentrate now on haloes on the right panel of Figure 7.4, the observed LTGs and TNG50 galaxies account for similar fractions of galaxy mass (around 8-9%). The statistical tests revealed similar means, medians and distributions. Here again, there are a few more massive observed LTGs but they are the ones involved in major mergers, whose halo can be mixed with the companion's. The distribution of TNG50 haloes is less peaked but we have only 70 galaxies. Although a statistical study would require more annotated haloes, we find a good agreement between observations and the predictions of numerical simulations for the mass in tidal features and haloes. This comforts our approach and the usefulness of the annotation server to compare them.

Effect of the viewpoint One of the advantages from simulations is the possibility to study features as a function of the orientation under which the galaxy is seen (viewpoint). In this preliminary analysis we kept three different viewpoints for each of the 70 TNG50 galaxies. However, we do not have all images annotated, which makes the comparison of features between different viewpoints complicated. An example of what can be studied is the percentage of mass in all tidal features and haloes around a given galaxy as a function of the viewpoints. To that end, we only kept the TNG50 galaxies that had been annotated along 2 or 3 viewpoints.

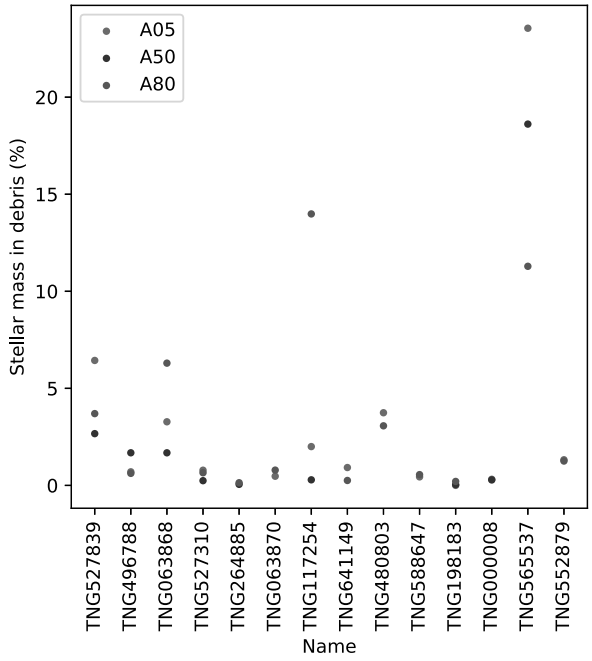

Figure 7.5: Percentage of mass contained in tidal tails and streams for each galaxy as a function of the viewpoint of the simulated image. The viewpoint A05 is plotted in red, A50 in blue and A80 in green. The name of the galaxy is indicated on the x-axis.

In Figure 7.5 (resp. Figure 7.6) we present the percentage of stellar mass contained in tails and streams (resp. haloes) as a function of the viewpoint. The histograms of the difference of mass between viewpoints are presented in Figure 7.7. From Figure 7.5, there is around 1-2% of difference between the mass of tidal features recovered between two viewpoints, but in a few cases it can increase up to 13%. From Figure 7.6, for some galaxies there can be difference of up to 10% of the stellar mass of haloes recovered between two viewpoints. The histograms in Figure 7.6 show that the median (absolute) difference is around 2.5% and the mean around 4.5% between each pair of viewpoints for haloes. For tails and streams, the median difference is 0.5-1.3% and the mean difference is around 1.7-3.5%. Although we do not have a statistically significant sample, especially for galaxies with tidal features, the scatter visible indicates a dependency of the viewpoint, although the majority show differences of 1-3% which is a reasonable agreement. We could not produce a census of tidal features as a function of the viewpoint but it would be interesting with more annotations to study it and compare the results to the literature about the detectability of features depending on the projection (e.g., Mancillas et al. 2019).

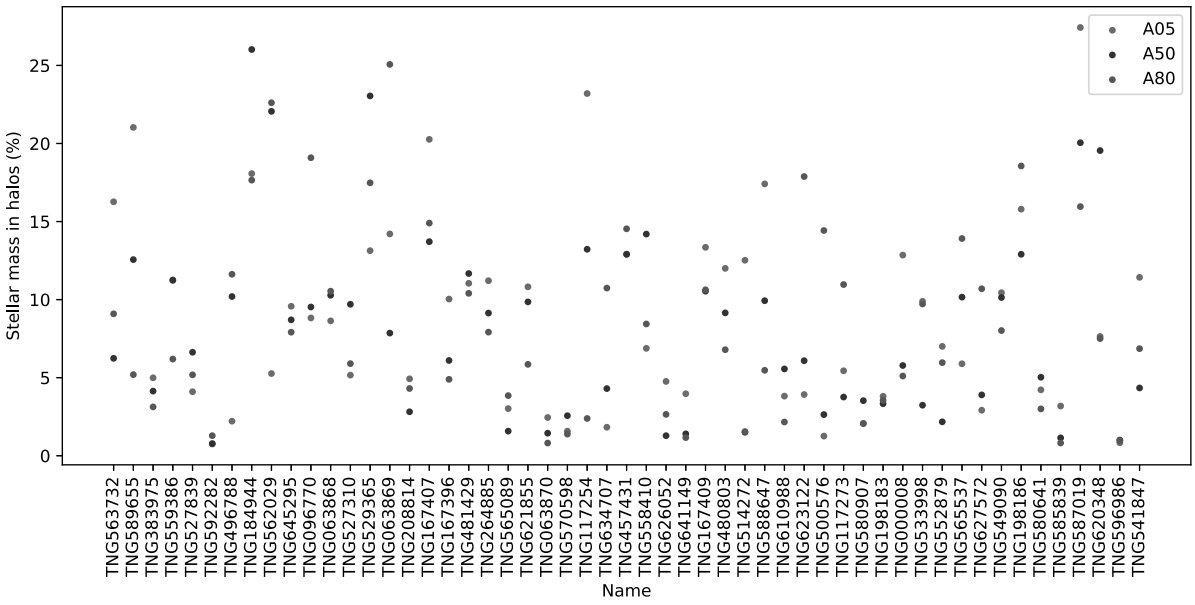

Figure 7.6: Percentage of mass contained in haloes for each galaxy as a function of the viewpoint of the simulated image. The viewpoint A05 is plotted in red, A50 in blue and A80 in green. The name of the galaxy is indicated on the x-axis.

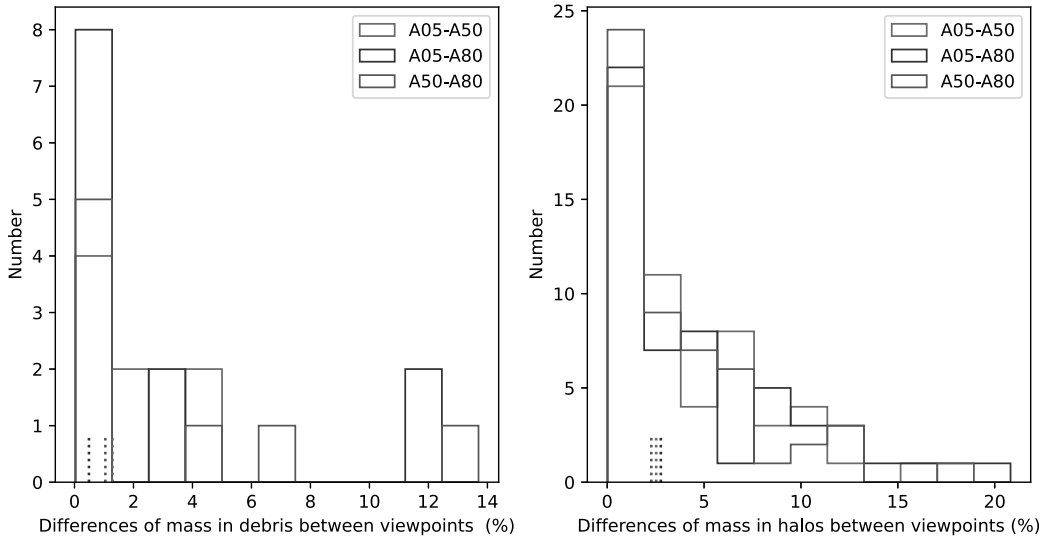

Figure 7.7: Histograms of the difference of the percentage of mass contained in tidal tails and streams (*left*) and in haloes (*right*) for each galaxy as a function of the viewpoint of the simulated image. The difference between the viewpoints A05 and A50 is plotted in red, between A05 and A80 in blue and between A50 and A80 in green.

7.1.4 Future prospects on numerical simulations

The next step in the comparison between observations and simulations using the annotation server is to have multiple users to annotate all the simulated images and searching for all the types of LSB features. The preliminary study was performed on a small number of annotated galaxies by one or two users only so the results are not statistically significant and we do not account for the variability between users. Once this would be done, we could perform again the analysis presented above and in particular focus on the

detectability of tidal features as a function of the viewpoint of the galaxy. The goal would be to study the projection effects and to correct for this bias in observations.

Another avenue to explore is to study LSB features as a function of the redshift. From simulations where the ground truth is known, we could annotate simulated images at various redshifts to probe galactic assembly through cosmological times. Martin et al. (2022), a collaboration in which I took part, produced a census of tidal features at four different redshifts from mock images from the NewHorizon simulation. Applying the annotation tool to such images would enable to extend that study in a quantitative manner.

A more relevant and complementary step is to use mock images from other simulations, created in a way that enables to keep other subhaloes (i.e. where nearby companions and potential tidal features can appear next to the target galaxy). A potential collaboration with Aman Khalid and Sarah Brough would be very interesting, as they are working on the detection of LSB tidal features in simulated images from four different simulations. Their results will be presented in Khalid & collaborators (2023). Likewise, a collaboration with Garreth Martin (with whom I have already worked) would be very beneficial as he created mock LSST-like images from the NewHorizon simulation (Martin et al. 2022).

7.2 Machine learning and upcoming surveys

7.2.1 Upcoming surveys

The frontier of the LSB Universe will be pushed back in the coming years thanks to the advent of future large surveys such as Rubin/LSST or Euclid. The Rubin Observatory is a 8.4 metre telescope being built in Chile. It has 6 filters (u, g, r, i, z, y), a wide field of view (9.6 deg^2) and an expected pixel size of 0.2 arcsecond per pixel. The Rubin Observatory will conduct a 10-year survey, the Legacy Survey of Space and Time (LSST). LSST's main Wide-Fast-Deep survey will cover 18,000 deg^2 and will reach after ten years a surface brightness limit of 32 mag arcsec^{-2} (LSST Science Collaboration et al. 2009; Ivezić et al. 2019). The combination of these characteristics (large field of view, depth, small PSF) make Rubin/LSST an ideal LSB survey (e.g., Brough et al. 2020). Nevertheless, Rubin/LSST will face the same issue than any other deep imaging survey (see Section 1.2.1.3) such as the diffuse light from bright sources, cirrus contamination and the need for LSB-compliant pipelines. Euclid (e.g., Euclid Collaboration et al. 2022) is an ESA mission aiming at investigating the nature of dark matter and dark energy, that has successfully been launched in July 2023. It will cover 15,000 deg^2 for the Wide Survey and 40 deg^2 for the deep survey using the VIS (optical broadband filter with a single bandpass of 560-900 nm) and NISP (near-infrared, Y, J and H bands) instruments. VIS has a high spatial sampling of 0.1 arcsecond per pixel. Like for the LSST, the combination of large area covered, high spatial resolution and depth of the survey makes Euclid ideal for the study of the LSB Universe. However, important efforts[3] are needed to develop an LSB-compliant pipeline which is not yet an official pipeline of the mission.

Both Euclid and Rubin/LSST will cover thousands of square degrees on the sky with an exquisite depth, revealing the LSB Universe. It will be possible to perform statistical studies of LSB structures around unprecedented samples of galaxies, within a representative volume and in various environments (field, groups, clusters). This implies that the number of galaxies targeted will dramatically increase, going from hundreds to tens

[3]The team at the Observatory of Strasbourg (M.Urbano, P.-A.Duc, O.Marchal, A.Lançon) is currently leading one of the developments of an LSB-compliant pipeline.

of thousands of objects of interest in the nearby Universe. Visually detecting and classifying tidal features will no longer be an option. It is even less realistic to manually annotate more than a few hundred galaxies in a reasonable time. Automated techniques, and in particular machine learning (ML) and deep learning (DL), will become necessary to automatically detect and then delineate tidal features in deep images.

7.2.2 Annotation database and machine learning

As mentioned in Chapter 1, DL has been widely used in astronomy in particular to classify the morphologies of galaxies but less studies are focusing directly on tidal features. Among the works that aim at detecting LSB tidal features, the pioneer work by Walmsley et al. (2019) proved to be a promising approach: they used a CNN to detect streams in CFHTLS-Wide images. To train their network, they relied on the visual classification of these galaxies by Atkinson et al. (2013). Likewise, Domínguez Sánchez et al. (2023) trained their CNN on a larger annotated dataset of tidal features from 6,000 simulated images annotated by many professional astronomers (Martin et al. 2022). Yet, both techniques require a dataset of previously manually annotated tidal features, and only perform detection but no segmentation. Current unsupervised approaches (e.g., Storey-Fisher et al. 2020) do not necessitate labeled data, but are not yet able to identify tidal features (only 'anomalous' galaxies).

To our knowledge, no satisfactory solution existed to detect and segment tidal features in deep images. The PhD work carried out by F. Richards, to which I contributed, aimed at filling this gap. It will be presented in the following paragraph, see also the article Richards et al. (2020) in which I am a co-author

7.2.2.1 Automatic segmentation of tidal features

F. Richards' goal was not only to assess the presence of a tidal feature in an image, but also to segment it, i.e. find and delineate its position. To that end, he used supervised learning and trained a dedicated CNN to this precise task. This is the reason why he created the annotation tool and associated database. Once the LSB structures around our sample of galaxies were annotated he could use our annotated dataset of LSB structures. It is currently the only database containing this type of very detailed information of the features: not only their type is available, but their location and external boundaries.

A first challenge with this project is the presence of contaminants in deep images, in particular galactic cirrus. It is an issue when searching for LSB tidal features, both because it adds an additional amount of light that superimposes with the light of faint features and hide them and because its filamentary structure can be mistaken for tidal features. It is crucial to automatically identify regions contaminated by cirrus. F. Richards addressed this issue by developing a ML pipeline to segment cirrus clouds. His CNN was improved to handle orientation-dependent features (i.e. to capture the overall orientation of cirrus) through Learnable Gabor Convolutional Networks (LGCNs). LGCN is a complex-valued CNN architecture which is highly sensitive to rotation transformation. After being trained on our annotated cirrus dataset, his network performed better in segmenting cirrus than more traditional and simple CNN approaches without complex values or Gabor modules. The performances of the network could be increased with a larger annotated dataset: indeed, the training dataset was very small (only 48 images containing cirrus were used) compared to the size of the ones classically used for DL approaches (typically tens or

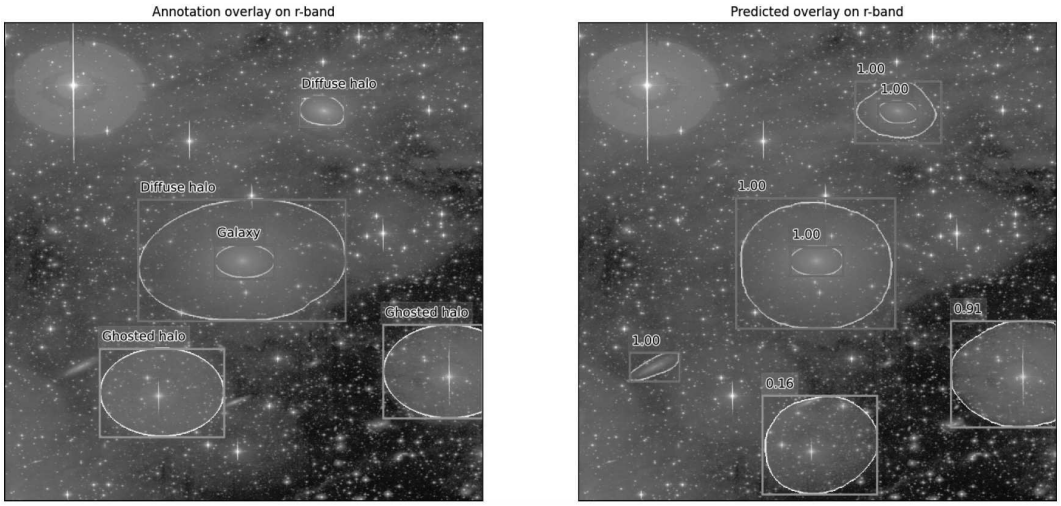

Figure 7.8: Figure 7.6 from Richards (2023). Example of multi-class segmentation with Dr Richards' DL network. *Left*: annotations of haloes (green), galaxy (orange) and ghost reflection (blue) around a given galaxy. *Right*: Predictions of the neural network. The number around each prediction indicates how confident the network is about its classification. On the bottom left part of the image there is an example of a correctly predicted but unannotated galaxy.

hundreds of thousands of training samples), or maybe it could be improved by modifying even more the network architecture.

In a second step of his PhD work, F. Richards investigated whether his network could simultaneously detect and segment stellar structures and galactic cirrus in deep images. These represent two types of segmentation: the former is instance segmentation (detecting distinct objects, countable quantities like galactic structures) while the latter is semantic segmentation (detecting uncountable amorphous regions, like cirrus). Performing both at the same time is called panoptic segmentation. To that end, he modified his previous Mask R-CNN model to detect five classes of objects: galaxy (including main galaxy and companions), halo, elongated tidal structure (including tails and streams), cirrus and ghost reflection. The training dataset is again based on our annotation database. An example of the input annotations and of the network's output is presented in Figure 7.8.

The first step was to focus on instance segmentation, with only galaxies, haloes and elongated tidal structures. From F. Richards' tests, it appears that elongated tidal structures are the hardest to recover, as the network was not able to correctly predict them. On the contrary, ghost reflections are the most successfully segmented type, which could be explained by the fact that they are bright with a sharp boundary, and their shape and size is globally the same. Galaxies and haloes are predicted with roughly the same performance, which is likely due to the fact that both categories are intertwined (as haloes are the LSB outskirts of galaxies). Yet, the network is able to predict reasonable contours for haloes and galaxies, even when the image is highly contaminated by cirrus as in Figure 7.8. This is a significant improvement compared to traditional methods. A better detection of tidal features could be helped by adding other annotations to the training dataset.

Another interesting point to note is that the network also predicted correct but unannotated objects (one example is visible in Figure 7.8). It is normal to find many objects in the images that were not manually annotated by users: indeed, it already took a non-

negligible amount of time to delineate the features of interest, so if we had asked users to delineate every single object it would not have been possible to do for many images. For instance, background galaxies were not annotated, and only the ghost reflections close to the target galaxy were delineated. But here, the interesting point is that the network is able to recover them. This means that the network's correct but unannotated predictions can be used to increase the size of the training dataset, and later to re-train the network. To check that the predictions were correct, they were visually checked and kept or rejected before being integrated to the annotation dataset. This type of human intervention and feedback to adjust the network is referred to as human-in-the-loop training. It appeared that 67% of the new output predictions were correct and added to the dataset, with around 89% of correct predicted galaxies. With this additional training, the network performed significantly better for the ghost reflection segmentation. Therefore, this shows the usefulness of a human-in-the-loop step to increase both the size of the training dataset and the network performance; and which must be kept in mind for future studies. The final step of F. Richards' work was to perform panoptic (cirrus and LSB features) segmentation. With this unification of tasks, it appeared that the network performed better in both segmentation tasks.

A conclusion that can be drawn from F. Richards' work is that automated segmentation of galactic structures in deep images using dedicated deep learning algorithms is possible, even in highly contaminated images. Galaxies, haloes and ghost reflections are detected reliably, with reasonable boundaries even in regions highly contaminated by cirrus. Even though the network performs badly on elongated tidal structures, a larger number of annotations in the training dataset would likely increase its performances.

7.2.2.2 On-going project: dwarf galaxy identification

The type of networks used by F. Richards is very promising in the automated search for LSB features in deep images. His network was trained on our annotation database that focussed on tidal features, extended stellar haloes and contaminants. Yet, there are other LSB objects worth studying, such as dwarf galaxies. Dwarfs are essential to study to build a comprehensive view of galaxy formation and evolution across large mass ranges. In addition, their properties (such as number, morphology, spatial distribution, presence of a nucleus, HI content...) provide fundamental tests of the hierarchical model of galaxy formation (e.g., Klypin et al. 1999; Moore et al. 1999; Bode et al. 2001; McConnachie et al. 2009; Kravtsov 2010; Belokurov 2013) as well as constraints on the nature of dark matter (e.g., Koposov et al. 2009; Peñarrubia et al. 2010; Nadler et al. 2019).

Therefore, the search and detection of dwarfs is a crucial question. It has been pushed forwards thanks to large photometric surveys such as SDSS (Abazajian et al. 2003) or DES (Abbott et al. 2018); dedicated dwarf detection algorithms (e.g., Koposov et al. 2008; Martin et al. 2013; Doliva-Dolinsky et al. 2022); source extraction and/or manual cataloguing (Chiboucas et al. 2009; Venhola et al. 2017; Habas et al. 2020; Müller & Jerjen 2020). The number of dwarf galaxies to be studied will dramatically increase with future surveys such as Euclid or LSST. As for tidal features, it will no longer be possible to manually identify dwarfs or run computer-expensive algorithms, so machine learning will be necessary. It has been shown that ML could be used to retrieve these LSB galaxies (e.g., Tanoglidis et al. 2020; Müller & Schnider 2021).

To extend this automated search, F. Richards' network is well-suited, as it is already able to predict correctly and relatively reliably haloes that had not been annotated. As dwarfs and haloes have similar shapes (but different sizes and luminosities), it is interesting

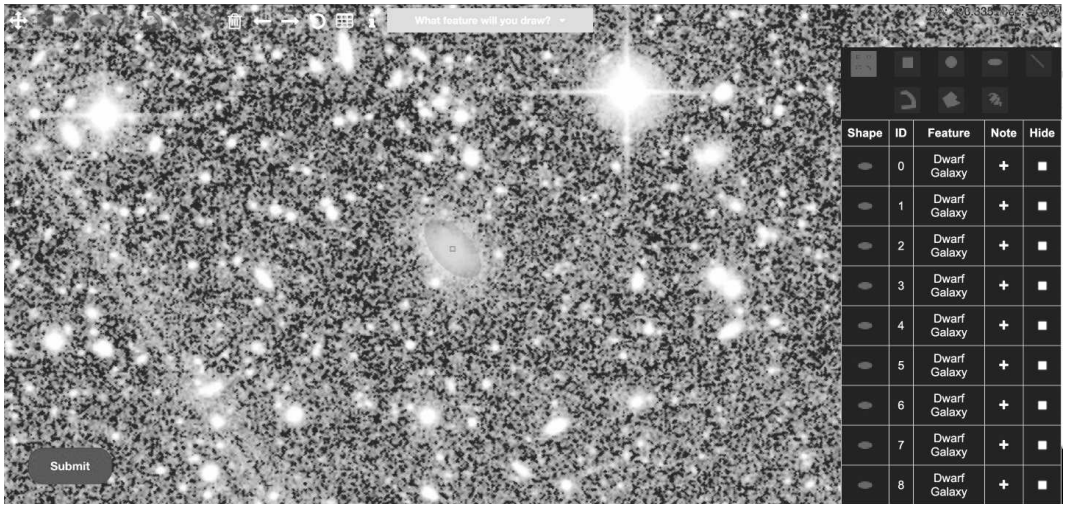

Figure 7.9: Example of a mock dwarf annotation, artificially integrated in the database from the GALFIT measurements of the dwarf (Poulain et al. 2021).

to test if the network is able to recover dwarf galaxies. This is the subject of an on-going collaboration between Dr. S. Sreejith (who is taking over F. Richards' work), A. Paiement, F. Marleau, P.-A. Duc and myself.

About the annotations on which the network is trained, we asked users to delineate dwarf galaxies only if they were close enough to the target galaxy or if they seemed to be the progenitor of streams. Therefore, we do not have enough annotated dwarfs to train the network. To tackle this issue, we could have either asked user to re-do the classification for this type of objects, or we could rely on already existing dwarfs catalogues (e.g., Sharina et al. 2008; Geha et al. 2012; Ann et al. 2015; Müller et al. 2017; Carlsten et al. 2020). In particular, Habas et al. (2020) and Poulain et al. (2021) gathered a catalogue of 2210 dwarf galaxies in MATLAS images. Using the GALFIT sotfware (Peng et al. 2010), they modelled the dwarf galaxies' light profiles. In particular, the effective radius R_e, position angle, ratio of semi-minor over semi-major axis (b/a) and Sérsic index were obtained for 1589 dwarf galaxies (for the others the software was not able to converge).

From these dwarfs measurements, the idea was to create mock dwarf annotations drawn with *Ellipses*, in order to be able to train the neural network on them afterwards, following the same process than what was done for LSB structures. To artificially integrate them to the database, I had to slightly adjust the input values. Indeed, the database expects for *Ellipses* to have the value of semi-major axis (in addition to the position angle) which differs from the effective radius. After some tests, I found an empirical relationship (that may not work in all cases): if the Sérsic index is higher than one, the semimajor axis a is set to $a = R_e \times 1.5 \times Sersic\_index$ (or $a = R_e \times 2 \times Sersic\_index$ if the dwarf is nucleated). Otherwise, $a = R_e \times 2.5 \times Sersic\_index$.[4] The mock dwarf annotations are attached to a mock user, and then the annotations can be displayed on the server. An example of such a dwarf is presented in Figure 7.9.

The next step is to train the network developed by F. Richards to find these dwarfs. It is still on-going work, as it needs to be fine-tuned for this specific case. In addition, in order to increase the number of annotated dwarfs, a step of human-in-the-loop training is considered. The goal is to see what are the predictions of the network and if it is able to

[4]Another approach would be to derive this value from the profile itself but this information was not directly available in the catalogue we used.

find correct but unannotated dwarfs. Expert users will be asked to accept or reject these new dwarfs, and the accepted ones will be included in the database on which the network will train again.

If the network performs well in this task, it will be an important tool for the automated dwarf detection in future large surveys.

7.2.2.3 Future prospects on machine learning

The projects presented in the previous sections are still ongoing, as detecting tails, streams or dwarfs remain a very complicated task even if the DL network performs well on segmenting stellar haloes, so I intend to pursue the collaboration with the ML community around A. Paiement in this direction.

Still in collaboration with A.Paiement's team, I intend to train a DL network both on simulations and observations to fine-tune it, in order to detect features but also learn the noise in real images (from artefacts or cirrus). Training on simulations will enable to derive the tidal origin and lifetime of tidal features. Indeed, the network could be given additional information about tidal features (like age, progenitor, morphological type) from simulations to infer these values in observational images. Data augmentation will be applied, as well as an iterative process to integrate DL-generated annotations, as our database only contains a few hundreds of annotations but a larger training dataset is needed.

7.3 Further analyses

From the annotation database and the deep images, it is possible to further probe galactic assembly by studying the haloes. For instance, Deason et al. (2013) showed that the presence of a break in the shape of the density profile can be related to the accretion history of galaxies. Likewise, other studies used numerical simulations to probe the relation between the shape of the radial density or SB profiles and the accretion history of galaxies (e.g., Pillepich et al. 2014; Rodriguez-Gomez et al. 2016; Merritt et al. 2020; Sánchez-Alarcón et al. 2023). I could perform a similar analysis on our galaxies, and combine these profiles to the information we already have in order to build a more detailed view of the late assembly history of galaxies, through the study of a large sample of galaxies.

In addition, Nibauer et al. (2023) developed a fast method to infer the shapes of dark matter haloes from the curvature of 2D projected stream tracks without simulations. This method would be very useful for extragalactic streams, as unlike resolved stellar population studies we usually only have access to 2D images without kinematic information for streams outside the Local Group. It would be interesting to apply this method to our sample of well-defined thin streams to determine the dark matter halo shape parameters. This would enable to test cosmological models, by comparing the theoretical predictions of simulations to the inferred distributions of dark matter haloes from observed streams.

7.4 Conclusion

To summarize this chapter, there are many interesting ways to extend the work carried out during this PhD thesis. The first is the comparison between numerical simulations and observations, which is crucial to constrain models of galactic evolution, and which is the direct continuation of my work. The annotation tool can be used to annotate in a similar

manner observational and simulated images, which enables a straightforward comparison between them. I presented some preliminary results of LSB features from the annotations of two sets of simulated images and the comparison to the results from observations (from Chapter 5). This project should be continued, with more annotators and mock images more suitable for our study. A second major prospect is machine learning. Indeed, as it will no longer be possible to manually annotate tidal features in future large surveys such as Euclid or Rubin/LSST, automated techniques will be needed to detect LSB features in images, including dwarf galaxies. Our annotation database can be used as a training set for such deep learning algorithms. DL algorithms can also be trained on simulated images to overcome some issues due to the complicated task of automated detection of LSB features. Finally, it would be interesting to further probe the accretion history of galaxies through the study of the radial density profiles of the haloes; as well as deriving parameters of the dark matter halo shape using the morphology of streams in our images.

7.5 Résumé du chapitre

Dans ce chapitre, nous nous sommes intéressés aux perspectives futures liées au travail de thèse réalisé. Plusieurs pistes sont intéressantes et prometteuses. La première est la comparaison entre les simulations numériques et les observations, qui est cruciale pour contraindre les modèles d'évolution galactique, et qui consitue le but direct de mon travail de thèse. L'outil d'annotation peut être utilisé pour annoter de manière similaire les images d'observation et les images simulées, ce qui permet une comparaison directe et cohérente entre elles. Les études précédentes ne permettaient en effet pas de comparer de manière similaires observations et simulations, à cause de différences dans les profondeurs des images, les bandes, les définitions des structures LSB et les méthodes de détection. J'ai présenté quelques résultats préliminaires sur les structures LSB à partir des annotations de deux ensembles d'images simulées et de la comparaison avec les observations (voir le Chapitre 5). Ce projet prometteur devrait être poursuivi, avec davantage d'annotateurs et des images simulées plus adaptées à notre étude. Une deuxième perspective majeure est l'apprentissage automatique. En effet, comme il ne sera plus possible d'annoter manuellement les débris de marée dans les futurs grands relevés à venir comme Euclid ou Rubin/LSST, des techniques automatisées seront nécessaires pour détecter les structures LSB dans les images ainsi que les galaxies naines. Notre base de données d'annotations peut être utilisée comme un ensemble d'entraînement pour de tels algorithmes d'apprentissage profond. Les algorithmes d'apprentissage profond peuvent également être entraînés sur des images simulées afin de surmonter certains problèmes liés à la complexité de la détection automatique des structures LSB. Enfin, il serait intéressant de sonder davantage l'histoire de l'accrétion des galaxies par l'étude des profils de densité radiale des halos stellaires; ainsi que d'inférer les paramètres de la forme du halo de matière noire en utilisant la morphologie des courants stellaires dans nos images.

Chapter 8 Conclusions

8.1 English version

The work carried out during my PhD thesis aims at characterizing low surface brightness tidal features around nearby massive galaxies, in order to probe the late assembly history of galaxies. In this chapter, we summarize what was presented in this thesis manuscript.

In Chapter 1, we introduced the general context about the formation and evolution of galaxies. Hierarchical models of galactic evolution predict that galaxies grow through successive merger events and accretion of gas (Section 1.1). The environment in which galaxies reside also has a strong impact on their evolution. For instance, the cluster environment tends to quench spiral galaxies through multiple processes that deplete their gas reservoir. The gravitational interactions and mergers between galaxies then strip them of their stars, which form diffuse structures: they are called tidal features. The properties of this collisional debris depend on the type of merger that occurred, which means that studying tidal features gives hints about the past history assembly of galaxies. A complication in the study of tidal features comes from their low surface brightness (Section 1.2). Dedicated observing strategies and data reduction pipelines are needed to reveal LSB features in deep images.

Once the deep images are acquired, the question of the detection and characterization of LSB tidal features arises (Section 1.2.2). Different approaches can be applied, ranging from visual inspection to automated methods, including machine learning techniques. However, automated techniques do not offer yet the subtle distinction between various types of tidal features, which remains mostly done visually. In addition, reconstructing the assembly history of galaxies requires detailed statistical studies on a large number of tidal features, which to our knowledge did not exist before the beginning of this PhD work. Therefore, the goal of my PhD thesis was to systematically quantitatively characterize tidal features around a complete sample of nearby massive galaxies in deep images, covering different environments and analysed in a consistent way.

To that end, a new annotation tool has been developed in order to retrieve quantitative measurement about LSB structures (Chapter 2). It enables multiple collaborators to precisely delineate LSB features, both stellar components and sources of contamination, superimposed on deep images. I contributed to make this tool accessible by putting it online, maintaining and improving it throughout the PhD thesis. I developed analysis tools to perform computations on LSB features from the database, amongst which geometry, surface brightness, flux, colour and stellar mass are important derived quantities (Section 2.3).

Equipped with the annotation server and analysis tools, the next step was to determine which deep images to utilize (Chapter 3). We used images from the Canada-France-Hawaii Telescope coming from four different surveys: MATLAS, CFIS, VESTIGE (all in the r-

band, chosen as the reference one) and NGVS (g-band). We described their characteristics and LSB-compliant reduction pipeline, as well as the few steps of data processing needed to retrieve the images and integrate them to the annotation server. Then, we determined the sample of galaxies on which we would carry out the statistical analysis of LSB features (Chapter 4). From the reference sample compiled by the ATLAS^{3D} collaboration of massive nearby galaxies, we performed the intersection between it and the deep images we had. Our final sample consists of 475 galaxies, both early-type and late-type galaxies, located in the Virgo cluster, in galaxy groups and in the field. Our final sample can be considered almost complete with respect to the ATLAS^{3D} reference sample for the ETGs, but we miss many LTGs outside the Virgo cluster. We are biased towards galaxies in groups and we only have one cluster environment.

Once we had the galaxy sample, the images and the annotation tool, we annotated LSB features around 475 galaxies. The results are presented in Chapter 5 and published in Sola et al. (2022) and Sola et al. (2023) (in preparation). From our annotation database, we derived quantitative information such as geometrical properties, surface brightness, colours or census of tidal tails, streams, shells and haloes. Thanks to the masks of our annotations, we derived the fraction of luminosity and then stellar mass in these LSB tidal features. Such values were so far not very well constrained on large samples of galaxies. We started by providing these properties for the individual features, and then we studied the impact of galaxy mass and environment on them in order to obtain clues about the past assembly history of galaxies.

We saw in Section 5.1.1 that there is a large variability in the shapes of tidal tails and streams, but statistically significant differences are found in their morphology. Indeed, tails are wider and brighter than streams, which is expected from models and which comforts the approach we used to distinguish them. We found that although these structures can extend over several kiloparsecs, tails and streams only account for less than 2% of the luminosity of the galaxy. We also computed the radii of shells and we did not recover the most extended shells visible in simulations. We found that 36% of our galaxies host any type of tidal debris. About extended stellar haloes (Section 5.2.1), they account for about 10% of the total galaxy luminosity, and as expected ETGs are redder than LTGs.

In a second part, we investigated the impact of mass and environment on tidal features (Sections 5.1.2 and 5.2.2). We found that the fraction of galaxies hosting tidal debris, as well as the fraction of luminosity in debris, increases with galaxy mass. Likewise, more massive ETGs have more luminous haloes, while no correlation is found for LTGs. These correlations have a sharper increase above a given mass threshold. These points are consistent with the hierarchical paradigm: more massive ETGs have undergone more merger events, triggering more tidal features that eventually contribute to form the extended stellar haloes. This is also consistent with the idea that for the most massive galaxies, mass growth mostly occurs through mergers in comparison to lower-mass galaxies and LTGs where the growth is dominated by smooth accretion of gas. Afterwards, we assessed the impact of the large-scale (field, group and cluster) and the small-scale (presence of a nearby companion) environments. The small-scale environment impacts the fraction of tidal features both in groups and in the Virgo cluster. The large-scale environment affects LTGs, through ram-pressure stripping, and tidal features (less luminous features and less shells in the Virgo cluster). When studying the joint impact of mass and environment, it seems that galaxy mass has a more important impact on tidal features and haloes rather than the environment. The work carried out in this chapter is complementary to the one in Bílek et al. (2022) where we focussed on the relations between internal kinematics of

ETGs and tidal features.

In Chapter 6 we discussed the limitations of our study and their impact on the results. We discussed the reliability of the annotations and the biases related to the visual identification and distinction of tidal features. We quantified the impact of sources of contamination and found that the PSF had an important impact on our results. For instance, haloes are more extended, redder and 15% more luminous because of the artificial amount of light added by the PSF. A careful PSF-deconvolution is needed to obtain accurate photometry. We also noted that the change of MegaCam filters had an important impact on our images, and that the quality of the filters is more important than the depth of a survey in the study of LSB features.

During this thesis, we have brought new insights on the LSB Universe through a detailed quantitative analysis of LSB features. There are many ways to pursue this work: we address some examples in Chapter 7. In particular, a crucial point is to compare observations to the predictions of numerical simulations. This is needed to constrain models of galactic evolution. The annotation tool can be used to compare in a similar and quantitative manner LSB features in observations and simulations, which was previously not possible due to differences in depths, detection methods, or definition of tidal features. We presented preliminary results on the comparison of LSB features between observations and simulations. Such a work would extend the study by Martin et al. (2022) in which I took part, whose goal was to investigate the nature, frequency and visibility of tidal features in simulated images. Another interesting prospect is to use deep learning to automatically detect collisional debris and dwarf galaxies in deep images, as it will no longer be possible to manually annotate them in future surveys covering thousands of square degrees on the sky. Our annotation database can be used as a training set for such deep learning algorithms (see Section 7.2.2 and Richards et al. (2020) with whom I collaborated). We also discussed the possibility to study the density profiles of stellar haloes to obtain more clues on the accretion history of galaxies.

8.2 French version

Le travail effectué pendant ma thèse vise à caractériser les structures de marée à faible brillance de surface autour de galaxies massives et proches, afin de sonder l'histoire de l'assemblage tardif des galaxies. Dans ce chapitre, nous résumons ce qui a été présenté dans ce manuscrit de thèse.

Dans le Chapitre 1, nous avons présenté le contexte général de la formation et de l'évolution des galaxies. Les modèles hiérarchiques de l'évolution galactique prévoient que les galaxies se forment et croissent grâce à des fusions successives et à de l'accrétion de gaz (Section 1.1). L'environnement dans lequel les galaxies résident a également un impact important sur leur évolution. Par exemple, l'environnement des amas a tendance à 'éteindre' les galaxies spirales par de multiples processus qui épuisent leur réservoir de gaz. Les interactions gravitationnelles et les fusions entre galaxies ont tendance à les dépouiller de leurs étoiles, qui forment alors des structures diffuses : elles sont appelées 'structures de marée'. Les propriétés de ces débris de collision dépendent du type de fusion qui s'est produit, ce qui signifie que l'étude des structures de marée donne des indications sur l'histoire passée de l'assemblage tardif des galaxies. L'étude de ces débris est rendue compliquée par leur faible brillance de surface (Section 1.2). Des stratégies d'observation dédiées et des pipelines de réduction de données sont nécessaires pour révéler les structures LSB dans les images profondes.

Une fois les images profondes acquises, la question de la détection et de la caractérisation des débris LSB se pose (Section 1.2.2). Différentes approches peuvent être appliquées, allant de l'inspection visuelle aux méthodes automatisées, en passant par des techniques d'apprentissage automatique. Toutefois, les techniques automatisées n'offrent pas encore la distinction subtile entre les différents types de débris de marée, qui reste essentiellement visuelle. En outre, la reconstitution de l'histoire de l'assemblage des galaxies nécessite des études statistiques détaillées sur un grand nombre de débris de collision, ce qui, à notre connaissance, n'existait pas avant le début de ce doctorat. Par conséquent, l'objectif de ma thèse était de caractériser quantitativement et systématiquement les structures de marée autour d'un échantillon complet de galaxies massives et proches dans des images profondes, couvrant différents environnements et analysées de manière cohérente.

Pour cela, un nouvel outil d'annotation a été développé afin d'obtenir des mesures quantitatives sur les structures LSB (Chapitre 2). Il permet à plusieurs collaborateurs de délimiter précisément les débris LSB, qu'il s'agisse de composantes stellaires ou de sources de contamination, superposés à des images profondes. J'ai contribué à rendre cet outil accessible en le mettant en ligne, en le maintenant et en l'améliorant tout au long de la thèse. J'ai développé des outils d'analyse pour effectuer des calculs sur les débris de marée à partir de la base de données, parmi lesquelles la géométrie, la brillance de surface, le flux, la couleur et la masse stellaire sont des quantités calculées importantes (Section 2.3).

Équipés du serveur d'annotation et des outils d'analyse, l'étape suivante a consisté à déterminer les images profondes à utiliser (Chapitre 3). Nous avons utilisé des images du télescope Canada-France-Hawaii provenant de quatre études différentes : MATLAS, CFIS, VESTIGE (toutes dans la bande r, choisie comme référence) et NGVS (bande g). Nous avons décrit leurs caractéristiques et le pipeline de réduction compatible LSB, ainsi que les quelques étapes de traitement de données nécessaires pour récupérer les images et les intégrer au serveur d'annotation. Ensuite, nous avons déterminé l'échantillon de galaxies sur lequel nous allions effectuer l'analyse statistique des structures LSB (Chapitre 4). À partir de l'échantillon de référence compilé par la collaboration ATLAS^{3D} de galaxies massives et proches, nous avons réalisé l'intersection entre celui-ci et les images profondes dont nous disposions. Notre échantillon final est constitué de 475 galaxies, de type précoce et tardif, situées dans l'amas de la Vierge, dans des groupes de galaxies et dans le champ.

Une fois que nous avions l'échantillon de galaxies, les images et l'outil d'annotation, nous avons annoté les structures LSB autour de nos 475 galaxies. Les résultats sont présentés dans le Chapitre 5 et publiés dans Sola et al. (2022) et Sola et al. (2023) (en préparation). De notre base de données d'annotations, nous avons tiré des informations quantitatives telles que les propriétés géométriques, la brillance de la surface, les couleurs ou le recensement des queues de marée, des courants stellaires, des coquilles et des halos. Grâce aux masques de nos annotations, nous avons calculé la fraction de luminosité puis la masse stellaire dans ces débris LSB. Jusqu'à présent, ces valeurs n'étaient pas très bien contraintes sur de grands échantillons de galaxies. Nous avons commencé par fournir ces propriétés pour les structures individuelles, puis nous avons étudié l'impact de la masse et de l'environnement des galaxies sur ces débris afin d'obtenir des indices sur l'histoire de l'assemblage des galaxies.

Nous avons vu dans la Section 5.1.1 qu'il existe une grande variabilité dans les formes des queues de marée et des courants stellaires, mais des différences statistiquement significatives sont trouvées dans leur morphologie. En effet, les queues de marée sont plus larges et plus brillantes que les courants stellaires, ce qui est attendu d'après les modèles et qui conforte l'approche que nous avons utilisée pour les distinguer. Nous avons trouvé que

bien que ces structures puissent s'étendre sur plusieurs kiloparsecs, elles ne représentent que moins de 2% de la luminosité de la galaxie. Nous avons également calculé les rayons des coquilles et nous n'avons pas retrouvé les coquilles les plus étendues visibles dans les simulations. Nous avons trouvé que 36% de nos galaxies hébergent un type quelconque de débris de marée. En ce qui concerne les halos stellaires étendus (Section 5.2.1), ils représentent environ 10 % de la luminosité totale des galaxies et, comme attendu, les ETGs sont plus rouges que les LTGs.

Dans une seconde partie, nous avons étudié l'impact de la masse et de l'environnement sur les structures de marée (Sections 5.1.2 et 5.2.2). Nous avons trouvé que la fraction de galaxies hébergeant des débris de marée, ainsi que la fraction de luminosité dans les débris, augmente avec la masse de la galaxie. De même, les ETGs les plus massives ont des halos plus lumineux, alors qu'aucune corrélation n'est trouvée pour les LTGs. Ces corrélations augmentent plus fortement à partir d'un certain seuil de masse. Ces points sont cohérents avec le paradigme hiérarchique: les ETGs plus massives ont subi plus de fusions, créant plus de débris de marée qui contribuent finalement à former les halos stellaires étendus. Ceci est également cohérent avec l'idée que pour les galaxies les plus massives, la croissance en masse se produit principalement par des fusions par rapport aux galaxies de plus faible masse et aux LTGs où la croissance est dominée par l'accrétion régulière de gaz. Ensuite, nous avons évalué l'impact de l'environnement à grande échelle (champ, groupe et amas) et à petite échelle (présence d'un compagnon proche). L'environnement à petite échelle a un impact sur la fraction de débris de marée à la fois dans les groupes et dans l'amas de la Vierge. L'environnement à grande échelle affecte les LTGs, par le biais de la pression dynamique, et les débris de marée (moins lumineux et moins de coquilles dans l'amas). Lorsque l'on étudie l'impact conjoint de la masse et de l'environnement, il semble que la masse des galaxies ait un impact plus important sur les structures de marée et les halos que l'environnement. Le travail mené dans ce chapitre est complémentaire à celui de Bílek et al. (2022) où nous nous sommes intéressés aux relations entre la cinématique interne des ETGs et les débris de marée.

Dans le Chapitre 6, nous avons examiné les limites de notre étude et leur impact sur les résultats. Nous avons discuté de la fiabilité des annotations et des biais liés à l'identification visuelle et à la distinction des débris de marée. Nous avons quantifié l'impact des sources de contamination et constaté que la PSF avait un impact important sur nos résultats. Par exemple, les halos sont plus étendus, plus rouges et 15% plus lumineux à cause de la quantité artificielle de lumière ajoutée par la PSF. Une déconvolution minutieuse de la PSF est nécessaire pour obtenir une photométrie précise. Nous avons également noté que le changement des filtres MegaCam a eu un impact important sur nos images, et que la qualité des filtres est plus importante que la profondeur d'un relevé dans l'étude des structures LSB.

Au cours de cette thèse, nous avons apporté de nouvelles connaissances sur l'Univers LSB grâce à une analyse quantitative détaillée des structures LSB. Il existe de nombreuses façons de poursuivre ce travail : nous en donnons quelques exemples au Chapitre 7. En particulier, un point crucial est de comparer les observations aux prédictions des simulations numériques. Ceci est nécessaire pour contraindre les modèles d'évolution galactique. L'outil d'annotation peut être utilisé pour comparer de manière similaire et quantitative les structures LSB dans les observations et les simulations, ce qui n'était pas possible auparavant en raison de différences de profondeur, de méthodes de détection ou de définition des débris de marée. Nous avons présenté quelques résultats préliminaires de la comparaison des débris LSB entre les observations et les simulations. Ce travail étend

celui de Martin et al. (2022) auquel j'ai contribué, dont le but était d'étudier la nature, fréquence et visibilité des débris de marée dans des images simulées. Une autre perspective intéressante est d'utiliser l'apprentissage profond pour détecter automatiquement les débris de collision dans les images profondes, car il ne sera plus possible de les annoter manuellement dans les futurs relevés couvrant des milliers de degrés carrés dans le ciel. Notre base de données d'annotations peut être utilisée comme un ensemble d'entraînement pour ces algorithmes d'apprentissage profond (voir Section 7.2.2 et Richards et al. (2020) avec qui j'ai collaboré). Nous avons également discuté de la possibilité d'étudier les profils de densité des halos stellaires pour obtenir plus d'indices sur l'histoire de l'accrétion des galaxies.

Appendix A Articles

A.1 Article: Characterization of LSB structures in annotated deep images, Sola et al (2022)

Sola et al. (2022)

A&A 662, A124 (2022)
https://doi.org/10.1051/0004-6361/202142675

Astronomy & Astrophysics

Characterization of low surface brightness structures in annotated deep images

Elisabeth Sola[1], Pierre-Alain Duc[1], Felix Richards[2], Adeline Paiement[3], Mathias Urbano[1], Julie Klehammer[1], Michal Bílek[4,5,6], Jean-Charles Cuillandre[7], Stephen Gwyn[8], and Alan McConnachie[9]

[1] Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg (ObAS), UMR 7550, 67000 Strasbourg, France
e-mail: elisabeth.sola@astro.unistra.fr
[2] Department of Computer Science, Swansea University, UK
[3] Université de Toulon, Aix Marseille Univ, CNRS, LIS, Marseille, France
[4] European Southern Observatory, Karl-Schwarzschild-Straße 2, 85748 Garching bei Müchen, Germany
[5] LERMA, Observatoire de Paris, CNRS, PSL Univ., Sorbonne Univ., 75014 Paris, France
[6] Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France
[7] AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, 91191 Gif-sur-Yvette, France
[8] National Research Council, Canadian Astronomy Data Centre, Victoria, BC, V9E 2E7, Canada
[9] NRC Herzberg Astronomy and Astrophysics, 5071 West Saanich Road, Victoria, BC, V9E2E7, Canada

Received 16 November 2021 / Accepted 4 March 2022

ABSTRACT

Context. The identification and characterization of low surface brightness (LSB) stellar structures around galaxies such as tidal debris of ongoing or past collisions is essential to constrain models of galactic evolution. So far most efforts have focused on the numerical census of samples of varying sizes, either through visual inspection or more recently with deep learning. Detailed analyses including photometry have been carried out for a small number of objects, essentially because of the lack of convenient tools able to precisely characterize tidal structures around large samples of galaxies.
Aims. Our goal is to characterize in detail, and in particular obtain quantitative measurements, of LSB structures identified in deep images of samples consisting of hundreds of galaxies.
Methods. We developed an online annotation tool that enables contributors to delineate the shapes of diffuse extended stellar structures with precision, as well as artifacts or foreground structures. All parameters are automatically stored in a database which may be queried to retrieve quantitative measurements. We annotated LSB structures around 352 nearby massive galaxies with deep images obtained with the Canada-France-Hawaii Telescope as part of two large programs: Mass Assembly of early-Type GaLAxies with their fine Structures and Ultraviolet Near Infrared Optical Northern Survey/Canada-France Imaging Survey. Each LSB structure was delineated and labeled according to its likely nature: stellar shells, streams associated with a disrupted satellite, tails that formed in major mergers, ghost reflections, or cirrus.
Results. From our database containing 8441 annotations, the area, size, median surface brightness, and distance to the host of 228 structures were computed. The results confirm the fact that tidal structures defined as streams are thinner than tails, as expected by numerical simulations. In addition, tidal tails appear to exhibit a higher surface brightness than streams (by about 1 mag), which may be related to different survival times for the two types of collisional debris. We did not detect any tidal feature fainter than 27.5 mag arcsec^{-2}, while the nominal surface brightness limits of our surveys range between 28.3 and 29 mag arcsec^{-2}, a difference that needs to be taken into account when estimating the sensitivity of future surveys to identify LSB structures.
Conclusions. We compiled an annotation database of observed LSB structures around nearby massive galaxies including tidal features that may be used for quantitative analysis and as a training set for machine learning algorithms.

Key words. galaxies: interactions – galaxies: evolution – methods: data analysis

1. Introduction

According to hierarchical models of galactic evolution, galaxies assemble through successive mergers, accretion of smaller systems, and smooth accretion of gas (e.g., Kauffmann et al. 1993; Cole et al. 2000; Baugh et al. 2003; Stringer & Benson 2007). These interactions between galaxies leave collisional debris, in particular tidal features such as tidal tails, streams, or shells, that have different morphologies and survival lifetimes (e.g., Arp 1966; Toomre & Toomre 1972; Quinn 1984; Mancillas et al. 2019). Since the different types of features are produced by different types of collisions, their identification and classification give valuable information about the mass assembly history of the host galaxy and about models of galactic evolution in general. However, the detection of collisional debris is complicated by their low surface brightness (LSB), challenging detection by traditional methods.

Detecting LSB tidal features in the Local Group is possible through stellar count, even using ground-based missions, such as the Sloan Digital Sky Survey (SDSS) with which Belokurov et al. (2006) revealed streams and stellar structures in the Sagittarius dwarf spheroidal. The resolved stellar populations photometric Pan-Andromeda Archaeological Survey (PAndAS; Martin et al. 2014) also studied this "field of streams", as well as the surrounding of M31 where signs of galactic interactions were studied (McConnachie et al. 2009). Substructures in M31

halo were detected in other surveys, such as Ibata et al. (2001), or Ferguson et al. (2002) with the *Isaac Newton* Telescope, or Gilbert et al. (2012) with the *Spitzer* Large Area Survey with Hyper-Suprime-Cam (SPLASH) survey. A wealth of new structures have been disclosed in the Milky Way by the *Gaia* mission (e.g., Malhan et al. 2018; Antoja et al. 2020).

However, for systems with increasing distance from the Local Group, stars are less and less resolved individually so that the detection of tidal features relies mainly on the study of the combined diffuse light. Overcoming the observational challenge of detecting faint and extended structures has been made possible by the development of dedicated instruments and/or data reduction pipelines able to produce deep images (i.e., images with a depth sufficient to detect faint structures) with limited undesirable artifacts. To that extent, both professional and small-sized telescopes enabled the discovery of a wealth of LSB structures around nearby galaxies. For instance, Martínez-Delgado et al. (2010) and Javanmardi et al. (2016) have respectively discovered streams and LSB systems around spiral galaxies with several small-size telescopes, while the Dragonfly Telephoto Array (van Dokkum et al. 2014) produced very deep images of nearby galaxies. The Huntsman telescope (Spitler et al. 2019) is based on the same principle as Dragonfly. In addition, LSB studies were conducted using professional telescopes such as the Burrell Schmidt telescope (Mihos et al. 2015, e.g.,), the Canada-France-Hawaii-Telescope (CFHT; e.g., Ferrarese et al. 2012; Duc 2020), the VLT Survey Telescope (VST; e.g., Venhola et al. 2017; Iodice et al. 2021), or the Subaru telescope (e.g., Alabi et al. 2020; Jackson et al. 2021).

Once the deep images are acquired, the LSB structures need to be identified, a task traditionally done with a visual inspection by one or a handful of contributors. This inspection can be performed on the images resulting from the basic data reduction, possibly adjusting the scaling and dynamics of the image intensity to enhance the LSB structures, or on images with an advanced processing, such as residual images obtained by subtracting the light profile model of the target (using softwares like GALFIT (Peng et al. 2002) for example). This technique discloses the inner most tidal debris as asymmetries that are not well fit by symmetric models of galactic light (e.g., Bell et al. 2006; McIntosh et al. 2008; Tal et al. 2009), but generates multiple artifacts.

A number of surveys of tidal features made by a direct visual inspection of the images by a few expert contributors have been carried out. For instance, Jackson et al. (2021) have inspected Subaru Hyper Suprime-Cam images of a sample of 118 low redshift, massive central galaxies; Kluge et al. (2020), Wendelstein Telescope Wide Field Images of the 170 local brightest cluster galaxies; Bílek et al. (2020), CFHT Mass Assembly of early-Type GaLAxies with their fine Structures (MATLAS) images (Duc 2020) of 177 massive nearby early-type galaxies; Morales et al. (2018) SDSS images of 297 Local Volume galaxies; Atkinson et al. (2013), CFHT Legacy Survey (CFHTLS)-Wide images of luminous 1781 galaxies at a redshift of $0.04 < z < 0.2$; Bridge et al. (2010) the 2 square degrees CFHTLS-Deep survey images of 27 000 galaxies.

As the sample size of galaxies increases, the classification by a limited team of experts becomes progressively infeasible. Galaxy features identification may then be done by a crowd of citizen scientists. Their potential lack of expertise is compensated by the higher number of individual annotations per galaxy. The reliability of the classification increases with the number of participants. This is exploited by the Galaxy Zoo project. For instance, Casteels et al. (2013) used the Galaxy Zoo 2 classifications made by 80 000 volunteers of SDSS images to study tidal features of about 150 000 galaxies of similar mass. Similarly, the Galaxy Cruise[1] project aims at classifying the presence of shells, rings, streams, and distorted halos from Subaru-HSC images. However, a method relying on the participation of volunteers with limited science knowledge can only be reliable for simple tasks. Precise classifications require prior knowledge of the field. It is for example the case when trying to disentangle the tidal tails associated with major mergers (i.e., mergers between two similar-mass galaxies) and the streams which trace minor mergers (i.e., mergers with a lower-mass galaxy): just focusing on their shape may not be enough as both collisional debris exhibit only subtle differences, as discussed in this paper. A prior knowledge on galaxy evolution and mergers is required to separate them. Hence, most studies on mergers relying on citizen-science efforts have focused on the presence of tidal distortions or collisional debris, without making any attempt to distinguish subclasses of objects.

As an alternative approach for large samples, fully automated methods have been developed. The level of the morphological asymmetry of galaxies can reveal tidal disturbances associated to ongoing or past mergers (Pawlik et al. 2016). More in general nonparametric methods such as the Gini-M_{20} parameter (e.g., Abraham et al. 2003; Lotz et al. 2004) and/or the CAS system (concentration, asymmetry, smoothness) (e.g., Abraham et al. 1994, 1996; Conselice et al. 2003, 2008; Conselice 2009), that do not assume a particular function for the galactic light distribution, have been often used. However, such parameters are not very sensitive to LSB structures, as they are flux-weighted and dominated by the contribution of the most luminous parts of a galaxy. On the contrary, Wen et al. (2014) developed the $A_O - D_O$ method, which is efficient to select asymmetric galaxies with faint features like tidal tails. In addition, Mantha et al. (2019) proposed a new tool to extract and quantify galactic morphological substructures from residual images, including plausible tidal features, along with a measure of their surface brightness. Automated tidal feature identification can also be performed using algorithms to separate the high and low spatial frequencies in the image (Kado-Fong et al. 2018), respectively corresponding to tidal features and galaxy light, allowing a quantitative analysis of their properties. However, like for the citizen scientist approach, these methods to not allow subtle classification and disentangling between various types of tidal features.

Another promising approach toward identifying faint tidal features on large number of objects is machine learning, as it offers the possibility to work with large samples of galaxies. In particular, convolutional neural networks (CNNs) have been used to classify the morphologies of galaxies (e.g., Huertas-Company et al. 2015; Dieleman et al. 2015; Vega-Ferrero et al. 2021; Domínguez Sánchez et al. 2018; Tohill et al. 2021). CNNs are also able to capture disturbed galactic morphologies that can be the hint of mergers and interactions (e.g., Pearson et al. 2019; Ferreira et al. 2020). They can be used to identify LSB tidal features in observational images: Walmsley et al. (2019) and Bickley et al. (2021) used CNNs to identify tidally-disrupted galaxies and classify tidal features. They obtained high accuracy and low contamination, and in overall performed better than other automated techniques. In addition, Pearson et al. (2019) were able to classify merger features in SDSS observations from a CNN trained on

[1] Galaxy Cruise, `https://galaxycruise.mtk.nao.ac.jp/en/index.html`

snapshots from the Evolution and Assembly of GaLaxies and their Environ- ments (EAGLE) simulation. However, currently the deep learning approach is unable to precisely classify the different types of tidal features, unless it is trained with large sample of images that have been previously precisely annotated. Unsupervised techniques could offer a solution to this problem (e.g., Martin et al. 2020; Uzeirbegovic et al. 2020; Spindler et al. 2021; Cheng et al. 2021), although there may be less control over the output.

The efforts to classify the morphology of galaxies and detecting surrounding tidal perturbations is not restricted to images obtained with telescopes. More and more detailed, realistic, images are produced by numerical simulations. They have the main advantage of providing the ground truth when interpreting the results. Indeed, it is possible to track the merger trees of LSB hosting galaxies and their 3D information to follow the LSB structures wrapping around galaxies (e.g., Hendel & Johnston 2015; Pop et al. 2018; Mancillas et al. 2019; Ebrová et al. 2021; Bílek et al. 2022).

Several types of simulations aim at predicting the formation history of galaxies, including interactions and mergers along with their tidal debris (e.g., Helmi & White 1999; Cooper et al. 2010; Bullock & Johnston 2005; Johnston et al. 2008; Pillepich et al. 2018; Schaye et al. 2015; Crain et al. 2015). In order to be comparable with those observed in the real Universe, simulated galaxies must not be idealized (Bottrell et al. 2019), and should include realistic sky, resolution, contamination sources and surface brightness limit. Identification of the tidal features have been carried out on these realistic mock images: for instance, Mancillas et al. (2019) performed a census of tidal features around galaxies from a hydrodynamical simulation, while Martin et al. (2022) investigated their nature, frequency and visibility around galaxies from the NewHorizon cosmological simulation.

One should note that most of the techniques mentioned above, whether applied to observational images or simulated ones, focus on qualitative aspects such as the probability of presence of one or several tidal features. Quantitative morphological and photometric measurements of tidal debris have so far been made for very limited numbers of galaxies (e.g., Martinez-Delgado et al. 2021), or restricted to ongoing tidally interacting systems. Systematic measurements for larger samples are needed to fully characterize the various types of LSB structures in order to make quantitative comparisons with numerical simulation-based models of galaxy evolution. Such quantitative properties will offer important calibrations toward understanding the physical nature of mergers causing the observed LSB structures and more in general to reconstruct the late assembly histories of individual galaxies. To that end, dedicated tools providing detailed annotations are needed.

In this paper, to systematically characterize tidal structures in deep observations, we have developed a tool that allows users to annotate large samples of galaxies via an accessible and intuitive online interface. The shapes of tidal features may be drawn with precision directly on the displayed images. Although other tools such as the Zooniverse platform also enable citizen scientists to delineate morphological features, such as spiral arms or bars (Masters et al. 2021), our interface provides new functionalities. It offers a larger variety in the annotation shapes used to delineate features, and the flexibility to switch between bands, which are facilities that are well suited for LSB structures annotations.

Given the complexity of the precise annotation task and focus, our tool is more adapted to expert users. The delineated tidal features are stored in a database, from which we can then determine their distribution of shapes, sizes and surface brightness. Delineation is the first step toward exploiting the full 2D profiles of the individual features, which gives additional constraints on the merger that created them.

The paper is organized as follows. In Sect. 2 of this paper, we present the data we used for the deep images. We introduce the annotation server, its features and the annotation process in Sect. 3. In Sect. 4, we detail the analysis tools that were used to retrieve quantitative measurements. Then, in Sect. 5 we present the results obtained from the annotations, and we discuss them in Sect. 6. Finally, we outline the conclusions in Sect. 7.

2. Deep images

The images and surveys we used in this paper are briefly described here. We used data from the 3.6-m Canada-France-Hawaii Telescope (CFHT), with the wide-field optical imager, MegaCam. In particular, we utilised images from two CFHT Large Programs: the Canada-France Imaging Survey (CFIS[2]) and the Mass Assembly of early-Type GaLAxies with their fine Structures survey (MATLAS[3]). MegaCam offers a wide field of view of $1^\circ \times 1^\circ$ with a resolution of 0.18 arcsecond per pixel. Images were processed by the Elixir-LSB dedicated pipeline optimized for the detection of LSB structures (Cuillandre, priv. comm., Duc et al. 2015). The limiting surface brightness reaches 28.3 mag arcsec^{-2} for CFIS (Cuillandre, priv. comm.) and 28.9 mag arcsec^{-2} for MATLAS in the r-band.

While MATLAS targets nearby massive galaxies, especially of early-type (Duc et al. 2015; Duc 2020; Bílek et al. 2020), CFIS is a blind survey that will cover 5000 square degrees in the Northern Hemisphere in the u and r band (Ibata et al. 2017), with additional bands available from observations made with other telescopes as part of the Ultraviolet Near Infrared Optical Northern Survey (UNIONS) project[4].

The galaxies annotated for the work presented in this paper are located within the 3600 square degrees that were covered by CFIS in March 2021. Annotation was done for the r-band images from MATLAS and CFIS.

In these CFHT deep images, we selected massive and nearby galaxies belonging to the reference ATLAS^{3D} (main and parent) samples (Cappellari et al. 2011). They include objects with distances smaller than 42 Mpc, an absolute K-band magnitude brighter than -21.5 mag and a stellar mass higher than 6×10^9 $M_\odot$. In order to study LSB features as a function of the morphology of the host galaxies, we selected two subsamples of comparable sizes consisting of 186 Early-Type Galaxies (ETGs) and 166 Late-Type Galaxies (LTGs) hence a total of 352 galaxies. If a galaxy was present in both surveys, we used the MATLAS image, because this survey is deeper. By doing so, we are biased toward finding faint features in ETGs, because they were the primary targets of MATLAS. Table 1 summarizes the number of galaxies per survey.

The CFIS 30 arcmin wide tiles were combined using SWarp (Bertin et al. 2002) and then cropped in order to center the final

[2] CFIS, https://www.cfht.hawaii.edu/Science/CFIS/

[3] MATLAS, http://obas-matlas.u-strasbg.fr

[4] The UNIONS project is a collaboration of wide field imaging surveys of the Northern Hemisphere. UNIONS consists of the Canada-France Imaging Survey (CFIS), conducted at the 3.6-m CFHT on Maunakea, members of the Pan-STARRS team, and the Wide Imaging with Subaru HyperSuprime-Cam of the Euclid Sky (WISHES) team. CFHT/CFIS is obtaining deep u and r bands, Pan-STARRS is obtaining deep i and moderate-deep z band imaging, and Subaru/WISHES is obtaining deep z band imaging.

Table 1. Number of galaxies surveyed in this work, detailed by their survey and morphological type.

Type	Number of galaxies studied
MATLAS ETGs	179
CFIS ETGs	7
MATLAS LTGs	53
CFIS LTGs	113
Total	352

image on the galaxy of interest. Users annotated structures in images with a field of view (FoV) of 31×31 arcmin. This FoV corresponds to an average physical size of 250×250 kpc and is equivalent to an average size of 50 effective radii around the target galaxies (with a minimum of 50×50 kpc, or 6.5 effective radii), enough to visualize the entire galaxy, its neighborhood and potentially most of its extended tidal features. Inspecting larger areas would have been too time consuming given the number of stellar structures and instrumental artifacts to annotate, and would not have been relevant to this study. The images were downsized by a factor of three (i.e., binned 3×3), both to decrease the size of the files on the web server and to enhance very faint structures.

To further enhance the visual identification of fine structures, a transformation from linear scale to a slightly modified inverse hyperbolic sine, asinh, was applied, with the following formula:

$$\mathrm{asinh(ADU)} = \log\left(\alpha \times (\mathrm{ADU} - b) + \sqrt{\alpha^2 \times (\mathrm{ADU} - b)^2 + 1}\right) \quad (1)$$

where asinh is the value of the pixel in asinh scale, ADU is the pixel value in linear scale, b is the background value and α is a parameter to tune, chosen here as one. The background level was set to a fixed value of zero in our case. This is motivated by the fact that the Elixir-LSB pipeline precisely processes the images in order to achieve a flat background over a given field of view, after correcting for residual instrumental or large scale sky artifacts. However some local contamination of sources, such as star halos or Galactic cirrus, remain after this processing. Therefore the real background may locally be nonzero. Local determination of the background is required to get a precise photometry, but for this paper we fixed it to the standard fixed value as a reasonable approximation.

In addition to asinh scaled images, surface brightness (SB) maps scaled in mag arcsec^{-2} were produced and used to characterize the tidal features. The relation from linear to surface brightness scale is the following:

$$\mu = -2.5 \log_{10}\left(\frac{\mathrm{ADU} - b}{pixsize^2}\right) + 30. \quad (2)$$

where μ is the surface brightness value of the pixel in mag arcsec^{-2} in AB magnitude, ADU is the value of the pixel in the original linear image, b is the value of the background (chosen here as zero), *pixsize* is the size of one pixel in arcsecond and 30 is the value of the zero point.

As it will be explained in Sect. 3.1, the online annotation tool requires the images to be in a particular format (HiPS, see Fernique et al. 2015) in order to display them. Hence, we created the HiPS after the asinh and surface brightness scalings. Having a single image does not enable the user to adjust online the image dynamics[5] such as the contrast or the cuts, however this homogeneity turns out to be an asset since the consistency of the images makes it easier to understand the differences between the annotations of several users.

In addition, $g - r$ colormaps were computed from the surface brightness maps for a subsample of 177 ETGs and 53 LTGs from MATLAS with available g-band images, and exploited to further characterize the LSB strutures, as described in Sect. 4.6.

Finally, we also considered for our annotations shallower true-color images from the Data Release 1 (DR1) of the Panoramic Survey Telescope and Rapid Response System (PanSTARRS)[6] (see Sect. 3), an imaging survey that covered the entire sky north of Dec = −30 deg in five bands (g, r, i, z and y) (Chambers et al. 2016).

3. LSB structures annotation tool

The annotation tool we developed is applicable to any imaging survey. However we focus here specifically on its use for the study of LSB structures.

3.1. General description

Visual classification methods are much faster and efficient with a web browser based tool that provides online facilities such as image visualization, navigation as well as immediate structure identification and labeling. Simple and clear interfaces are needed when numerous collaborators or citizen scientists are asked to review the data, such as the Zooniverse platform[7]. Although the latter enables collaborators to record quantitative information through an annotation tool, most platforms simply offer the possibility to assess the presence of a given feature.

In this paper, we present a web-based annotation tool that enables collaborators (referred to as "users" in this paper) to draw with precision the shapes of LSB structures superimposed on deep images and label them, allowing a quantitative analysis of the LSB structures of various types. In comparison with the Zooniverse interface, our tool offers several drawing options, an easy navigation through the image, the possibility to display images from other surveys and a simple way to verify the annotations once they have been drawn. The annotation process, described in detail in Sect. 3.2, relies on the visualization of astronomical images thanks to an online tool that uses the Aladin Lite[8] facility developed by the Centre de Données astronomiques de Strasbourg (CDS). It enables the visualization of sky regions, overlaid with object information from astronomical databases such as the Set of Identifications, Measurements and Bibliography for Astronomical Data (SIMBAD). Data from various surveys, such as PanSTARRS DR1, can be displayed and explored, but custom images can also be added, provided that the images are in the Hierarchical Progressive Surveys (HiPS, see Fernique et al. 2015) format. The HiPS format enables the representation of large astronomical datasets as the resolution increases when the users zoom on a part of the image. It relies on the hierarchical partitioning of a sphere into smaller and smaller

[5] The original MATLAS server available at `https://obas-matlas.u-strasbg.fr/WP/` allows the user to adjust on line the contrasts and cuts, as discussed in Bílek et al. (2020).
[6] PanSTARRS, `https://panstarrs.stsci.edu/`
[7] Zooniverse `https://www.zooniverse.org/`
[8] Aladin Lite `https://aladin.u-strasbg.fr/AladinLite/`

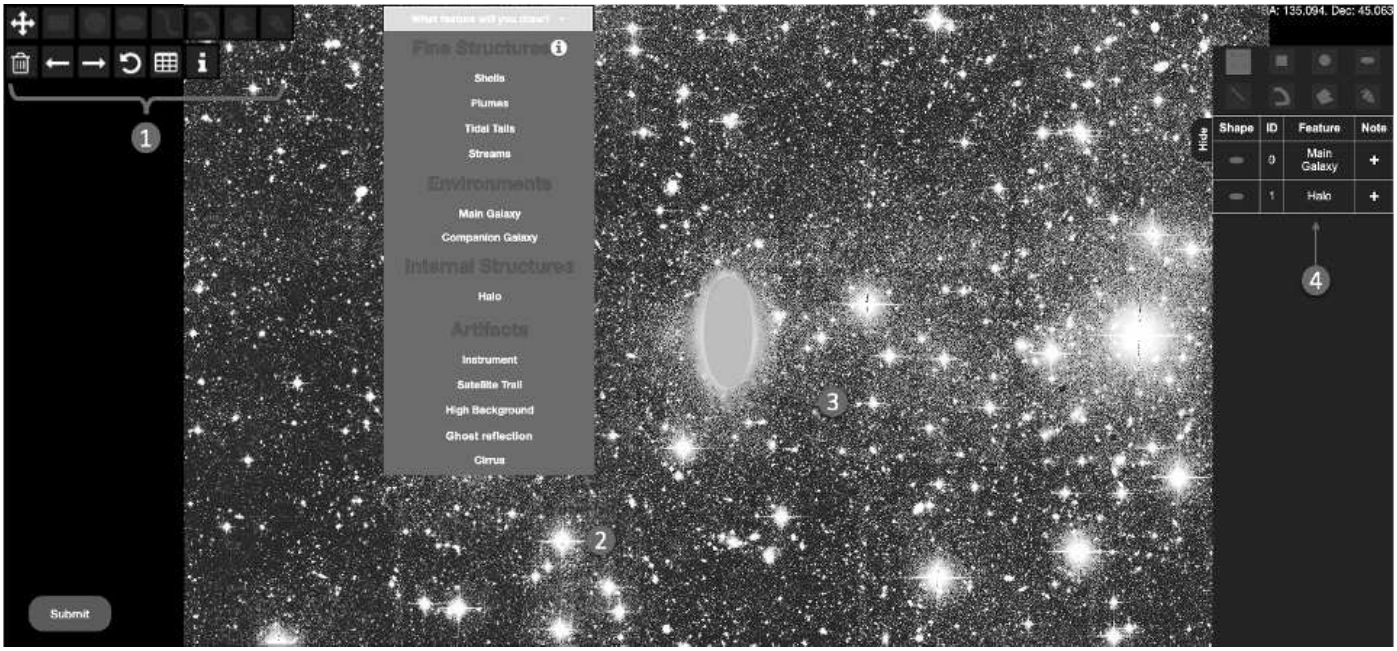

Fig. 1. Annotation interface with its main facilities: drawing buttons (label 1), classification menu (label 2), examples of already drawn annotations (label 3) and summary table (label 4). The background is a CFIS asinh-scaled image.

diamonds as the order of the partitioning increases, each diamond being identified by a unique index and order. Hence, we added our own HiPS images, whose origin is detailed in Sect. 2.

3.2. Annotation process

The annotation process can be divided in several steps: the selection of the galaxy to annotate, the annotation itself, its classification and the verification or modification of the annotation. All users are previously identified and logged in to record their annotations.

The annotation is made directly from the image navigation interface. The latter allows us to zoom in and out in the images, navigate through them and switch between different layers, that is images with different intensity scalings or from various surveys. Catalogs from for instance the SIMBAD database can be overlaid to display pieces of information about the objects in the image, including their velocity when known.

Drawing tool buttons are used to make the annotations. The interface is displayed in Fig. 1 with a CFIS asinh-scaled image, while a Pan-STARRS DR1 color image layer of the same galaxy is shown in Fig. 2. The user is asked to draw the external boundary of the features present in the image, as defined later in this section. To do so, the most appropriate type of shapes among circles, ellipses, rectangles, polygons, or curved lines are selected. Curved lines are cubic Bézier curves defined with four control points. The shapes, superimposed on the images with a semi-transparent red color, may be adjusted with precision. Afterwards, the user needs to associate the drawn shape with a label from a menu. All annotations may be checked and further updated from a Summary table. A tutorial explains how to draw the annotations with specific tools and then how to label them.

The aim is to annotate every feature visible on the image relevant for our study. This means that several annotations are drawn on the same image. It includes the stellar structures of interest and contaminants, among which:

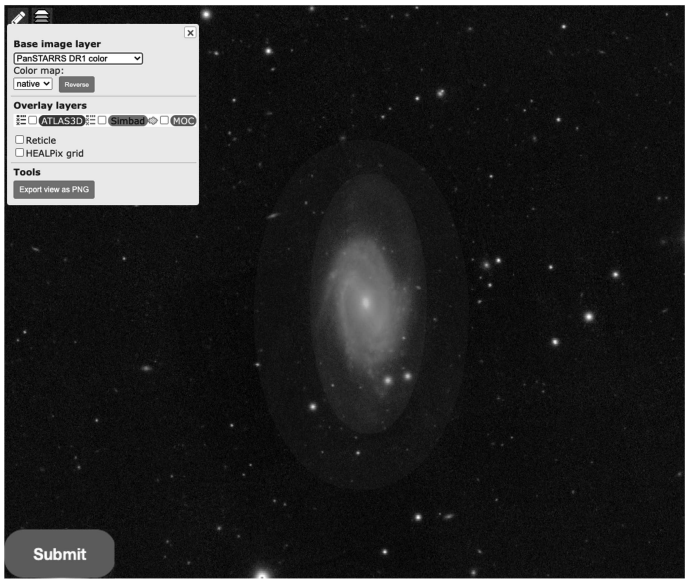

Fig. 2. Zoom on a shallower version of Fig. 1 with the same annotated galaxy. The Pan-STARRS DR1 RGB-color image layer that was used to delineate the "main galaxy" (inner red ellipse) is shown as background.

– Main galaxy: the target galaxy, encompassing its more luminous features, such as spiral arms for late-type galaxies. Such annotation is made on standard shallow images, in particular the PanSTARRS DR1 color images[9]. The outer elliptical

[9] We delineate the main galaxy on shallow images as they best highlight the bright inner structures. Indeed, currently, our Aladin-based tool does not offer the possibility to adjust the contrast and dynamics of the images, but allows us to switch between different sets of images through a convenient layer interface. Among the existing surveys made available through the interface, PanSTARRS DR1 was chosen as it covers the entire sky north of Dec = –30 deg and thus encompasses all our galaxies, contrary to SDSS for instance.

isophote of the galaxy is delineated. A posteriori, with the tool presented later, this isophote roughly corresponds to a surface brightness of 24 mag arcsec^{-2}.

– Halo: the extended low surface brightness stellar halo around the galaxy, as seen on the deep CFHT images. Its outer isophote – generally elliptical but sometimes disturbed – is traced[10].

– Tidal Tail: defined here and in the previous papers of this series (e.g., Duc et al. 2015; Bílek et al. 2020) as elongated stellar features whose stellar material likely comes from the target galaxy, and which then should have formed during major mergers (e.g., Arp 1966; Toomre & Toomre 1972; Mihos 1995). If the shape differs from the standard, antennae-like elongated structure, it is labeled as a "plume".

– Streams: generally thin and elongated tidal features, whose material does not seem to originate from the target galaxy. The stream is either attached to a companion galaxy progenitor whose mass is much lower than that of the target galaxy, or it is defined as an orphan stream if no progenitor is visible. In all cases, such streams likely trace past or ongoing minor mergers (e.g., Bullock & Johnston 2005; Belokurov et al. 2006; Martínez-Delgado et al. 2010).

– Shells: arc-shaped features, that are often present in groups and are concentric. They are typically formed during intermediate mass encounters (with a mass ratio around 1:10) with specific orbital parameters (e.g., Prieur 1990; Ebrova 2013; Duc et al. 2015; Pop et al. 2018).

– Companion: a nearby massive galaxy with a known velocity close to that of the target galaxy. We considered a difference of velocities of about 200 km s^{-1} for the threshold. This value is more conservative than the larger velocity threshold often used to identify companions. Indeed we want to make sure that the two galaxies are currently involved in a tidal interaction able to produce visible tidal debris. The outer envelope of the companion on deep imaging is delineated.

– Ghost reflections: artificial and extended round halos around bright stars caused by internal reflections on the detector and optical elements of the camera.

– Instrument: remaining instrument signature (CCD gap).

– Satellite Trail: trail of any satellite passing in the image.

– Cirrus: dust clouds in our Galaxy, scattering the optical light, and showing up as diffuse but structured (usually filamentary) features on the deep images. Regions likely contaminated by cirrus emission are delineated.

– High Background: regions in the deep images, with background levels higher than the blank sky values, not clearly identified as structured cirrus. These regions may trace dust illuminated by bright objects. They are not flat field defects but are really due to higher foreground emissions in these regions.

The visual classification of tidal features among tails, streams and shells, is necessarily subjective and may be ambiguous. It then makes sense to have several users annotating and making the classification. One of the goals of this paper is to precisely characterize each type of structures and retrospectively assess the relevance of the classification.

The annotation of contaminants such as cirrus, ghost reflections, or high background is essential as they might pollute the stellar structures of interest and either make complicated their detection or skew their annotation. Examples of galaxies with tidal features from CFIS are presented in Fig. 3, while Figs. C.1–C.3 from Bílek et al. (2020) present tidal features identified in MATLAS images.

Once the user has finished annotating a galaxy, the shape parameters, positions (in pixel coordinates and in right ascension and declination) and labels of each annotation are stored in a database hosted by a server. These can be used to redraw the annotations on the images or plot them as thumbnails, as it will be seen later. Examples of annotated galaxies are visible in Fig. 4.

The annotations are made by several users. In this work, all the users are considered to have the same level of expertise in the identification of tidal features. It must be noted however that our tool allows us to take into account possible different levels of expertise. This must be taken into account, as it reflects the degree of confidence one can have in the annotations. To that end, weights can in principle be attributed to the users, this is discussed in Sect. 6.2.

3.3. Thumbnails

In addition to the annotation tool itself, several analysis tools were developed and integrated into the server. One of the main features is the possibility to see the annotations that have been done, through thumbnails. Indeed, quickly visualizing annotations made by a user may be difficult when many features have been annotated. Hence, we developed a fast way to draw the contours of the main features to have an overview of the shapes of the annotations and to detect errors or missing features such as the main galaxy or the halo. The thumbnails only contain the external boundaries of the annotations, not the images themselves. Thumbnails can also be used to visualize annotations of several users on a same page, for comparison purposes.

The Thumbnails page enables users to choose the galaxies to be plotted, the size of the thumbnail box (in arcseconds ($''$), kiloparsec (kpc), or effective radius R_e of the target galaxy) and the type of annotations: for example, main galaxy, halo, tidal tails, streams and/or shells. Furthermore, the possibility to represent all features with thumbnails enables a global visual comparison of their shapes and sizes. Such thumbnails, representing only the shapes of tidal features and free of any contaminant (host galaxy, image artifacts, cirrus), could be used to train an algorithm to classify structures based on their shape. This could complement other machine learning algorithms that would be trained on the original images. Once generated, the thumbnails can be downloaded and displayed as webpages. A few examples are illustrated in Fig. 5.

4. Analysis tools

In this section, we present the analysis tools developed to characterize the LSB structures. As previously mentioned, retrieving quantitative measurements about the LSB tidal features is essential to infer the history and mass assembly of a galaxy. Yet for large samples of galaxies, except for fine structure numerical censuses, detailed analyses have not been systematically performed.

The annotation tool we developed offers new possibilities. Indeed, we now have access to the projected shape of the structures, since the users are invited to delineate the contours with precision. This allows us to retrieve the coordinates of the boundaries of the structures, making possible the determination of the area they cover and their length among other properties. For instance, this will be useful to determine retrospectively whether our criterion to separate streams from tidal tails is relevant.

[10] We note that the extended presumably thick nonstarforming disks that may be present around LTGs are considered here as a halos as there is no way to probe their 3D shape, especially when seen face-on.

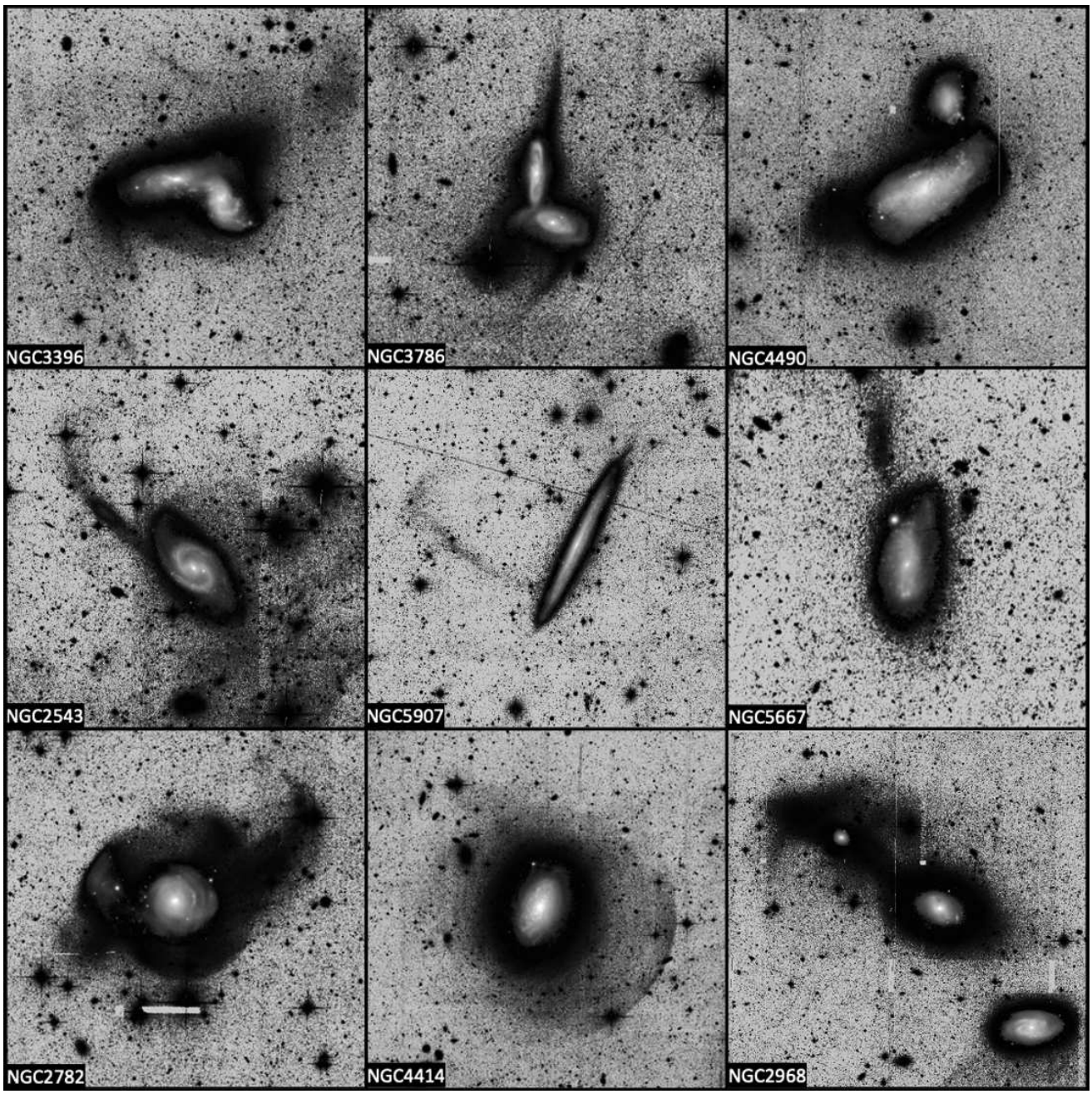

Fig. 3. Examples of tidal features detected in CFIS *r*-band images displayed with a asinh scale. A true color image from the PanSTARRS DR1 survey is overlaid at the center of the target galaxy. The first row shows tidal tails and plumes, the middle row streams and the bottom one shells.

4.1. Area

The first step was to determine what is the area covered by each structure. As mentioned in Sect. 3.2, there are different annotation shapes that must be considered. Since curved lines are cubic Bézier curves, it is not possible to determine their area. For the other shapes, the coordinates of all the points forming the contour of the annotation are retrieved in right ascension (RA) and declination (DEC). The distance between two coordinates was obtained using the on-sky separation[11], which enabled us to compute the areas of circles, rectangles, and ellipses.

To compute areas of simple polygons, we used the shoelace formula which is given by

$$A_{\text{polygon}} = \frac{1}{2}\left|\left(\sum_{i=1}^{n-1} x_i y_{i+1}\right) + x_n y_1 - \left(\sum_{i=1}^{n-1} x_{i+1} y_i\right) - x_1 y_n\right|$$

where A_{polygon} is the area, n is the number of sides of the polygon and $(x_i, y_i), i = 1, ..., n$ are the ordered planar coordinates of the vertices of the polygon.

4.2. Length

In addition to the area, the computation of the projected longest length in tidal structures is also important to characterize them.

[11] Astropy SkyCoord `https://docs.astropy.org/en/stable/api/astropy.coordinates.SkyCoord.html`

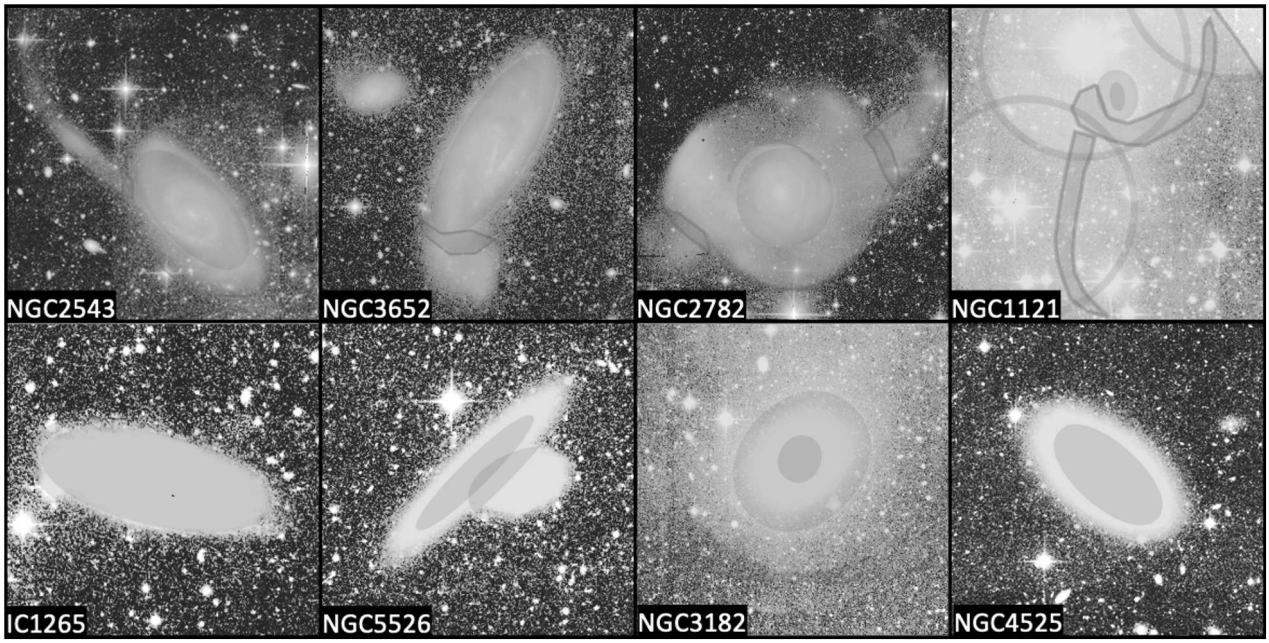

Fig. 4. Examples of annotated galaxies. Some of them present tidal features and pollutants, or a companion galaxy, while others only present the halo and the center of the galaxy.

The definition of the longest length depends on the shape of the annotation. For ellipses, the longest length corresponds to the length of the major axis, for circles it is the diameter, for rectangles and curved lines it is simply the length.

The process is more complicated when dealing with polygons, as they can have various complicated shapes. In these cases, the medial axis was used as the longest length: it can be seen as the topological skeleton and it is defined as the set of points that have at least two closest points on the polygon (i.e., it is the set of points that are equidistant to the contour of the polygon). The medial axes were obtained using scikit-image (van der Walt et al. 2014) through skeletonization. At the end of this step, the skeletons can have several branches but only the longest possible path was kept. To this purpose, the FilFinder package (Koch & Rosolowsky 2015) was used. An example of the previous steps is shown in Fig. 6.

A limit of this method arises for round polygons, that is to say whose shape tends to be circular. In such cases, the topological skeleton tends to be shorter and shorter, up to the limit case of a circle where the skeleton is reduced to the center of the circle. This leads to underestimate the estimated lengths. To tackle this issue, we adapted our method for round polygons and approximate them by the maximum volume inscribed ellipse. The length is then taken as that of the major axis. We also modify the value of the area defined in Sect. 4.1 to keep the area of the computed ellipse. One way to evaluate how round a polygon is is to compute the isoperimetric ratio q:

$$q = 4\pi \frac{\text{area}}{\text{perimeter}^2}. \tag{3}$$

The value of the isoperimetric ratio q is between 0 and 1, it reaches 1 in the case of a disk. In our case, we had to determine a threshold value for q above which a polygon would be considered round enough for our skeleton method to be incorrect. We determined this threshold to be 0.8, based on our measurements and shapes of the polygons (see Sect. 5). Hence, for polygons with $q < 0.8$ we kept the topological skeleton defined above, and for $q \geq 0.8$, we approximated them by an ellipse.

4.3. Distance between shells and the primary galaxy

Wilkinson et al. (1987) and Prieur (1990) divided shells into several types according to their position and orientations with respect to the host galaxy. Type I shells are concentric and centered on the host galaxy, type II shells are circular and randomly distributed around the galaxy, type III shells appear irregular and not concentric, either because there are very few shells or because of their untypical shape.

Hence, the study of the position of a shell relatively to the center of its primary galaxy is important, as comparisons with simulations can give us hints about the merger that triggered the shells, as well as its age (e.g., Ebrová et al. 2021; Bílek et al. 2022). Shell radii are useful also for investigating the gravitational fields of galaxies (e.g., Hernquist & Quinn 1987; Bílek et al. 2013).

With our annotations, we computed the projected distance between the center of the galaxy and the point located in the middle of the curved line defining the shell. Here we make the assumption that the curved lines are circular arcs.

4.4. Progenitor of tidal tails

Our current annotation server does not allow the user to manually assign an annotated feature to a specific galaxy, for instance the primary galaxy or the companion. By default all annotations are tied to the primary galaxy. This is especially an issue for systems made of two interacting target galaxies, present in the same image. In this case, one wishes to attach tidal tails to their real progenitor. There are four different possibilities when determining the progenitor of tidal tails:

– Case 1: the tail has an overlap with the halo of the primary galaxy but not with that of the companion: the progenitor is then the main galaxy.

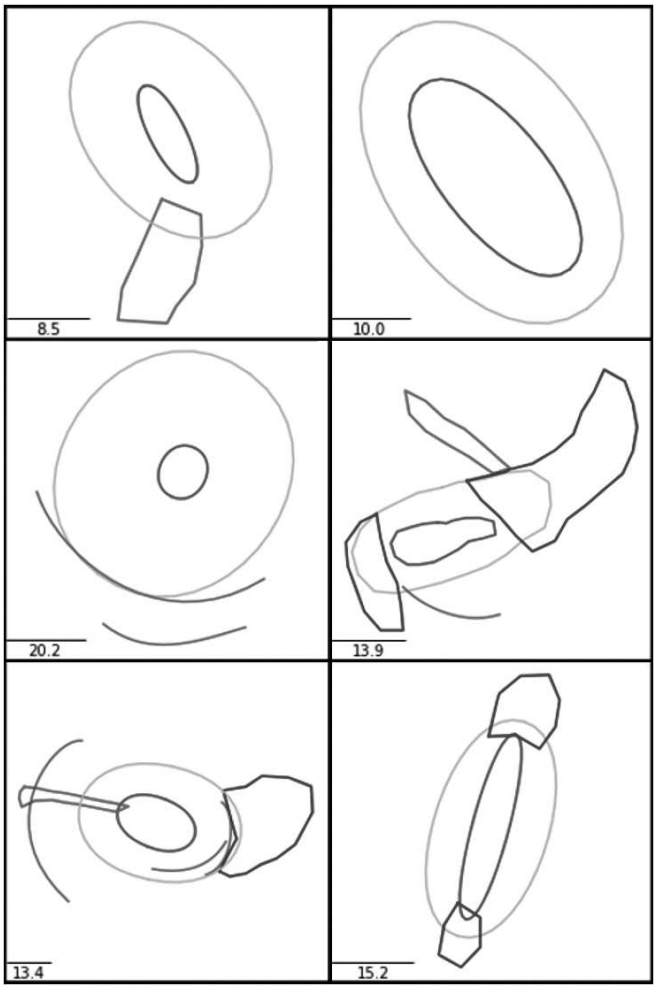

Fig. 5. Examples of annotation thumbnails. The center of the target galaxy is plotted in purple, its halo in yellow, tidal tails in blue, streams in red and shells in green. A scalebar in kiloparsecs is shown at the bottom left of each thumbnail.

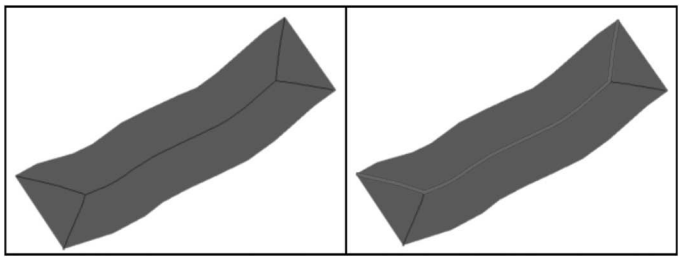

Fig. 6. Example of the longest length in a polygon (in gray). *Left*: medial axis (black line) of the polygon obtained during the skeletonization. *Right*: longest length kept (red line) in the polygon

- Case 2: the tail has an overlap with the halo of the companion but not with that of the primary galaxy: the progenitor is then the companion.
- Case 3: the tail has no apparent overlap with the target galaxy or the companion. In that case, the position of the center of mass of the tail is computed, as well as the positions of the centers of the primary galaxy and its companion. Then, the distance from each galactic center to the center of the tail is computed, and the tail is associated to the galaxy with the shortest distance. If the difference in distances is small and inferior to an arbitrary threshold, then the progenitor is set to "unsure".
- Case 4: the tail has an overlap with both the primary galaxy and its companion: it looks like a bridge. In this case, the area of the intersection of the tail with both galaxies is computed, and the galaxy with the biggest intersection area is kept as the progenitor. If the difference of areas is small and inferior to an arbitrary threshold, then the same computation on distances than for Case 3 is performed.

For the results, only the tidal tails associated to the primary galaxy were kept, in order to count each structure only once. For streams, by definition they are not attached to the primary galaxy so all the annotations of streams are kept.

4.5. Surface brightness measurements

The assessment of the surface brightness (SB) values of each type of tidal feature is important to make comparison with simulations but also determine whether they will be detectable in surveys to come, such as *Euclid* or the *Vera Rubin* Observatory. With our annotation tool, such measurements can be indirectly retrieved.

Indeed, since the coordinates of the contours of the annotations are available, it is possible to retrieve them to create masks that may then be attached to the surface brightness maps. To do so, the RA, Dec coordinates of the contours are converted into pixel coordinates of the Flexible Image Transport System (FITS) image on which we want to create the mask, using Astropy World Coordinate System functions[12]. The interior of that boundary in pixel coordinates is then filled to have the mask, using OpenCV-Python[13] functions. Then, we apply this mask to the SB FITS file and we read the SB values of all the pixels inside the mask[14]. For each annotation of a given type, the median value within the mask area is computed. This way, contribution of light coming from contamination sources such as foreground stars or background galaxies is removed, provided the field is not too crowded. This provides a representative (instead of an average) value of the SB of the structure. A direct aperture photometry would require a proper masking of all contaminant sources, which is beyond the scope of this initial study.

Another potential issue could arise if there is a SB gradient along the structure: deeper imaging will make the median SB value fainter because more LSB pixels are considered. However, since we do not want to get a precise value but to compare trends between tidal features, this issue is not a major one. So our estimate of the SB value of a given structure is quite uncertain. Our measurements are however useful to compare trends between the different classes of structures. For all the annotations of the same type, the median value of the previously computed medians was computed.

Finally, it is also possible to retrieve the SB value along the contour of the annotation (by transforming the RA, Dec coordinates of the contours into into pixel coordinates of the SB FITS files, then reading the corresponding SB values and keeping the median value). This is useful especially for stellar halos (see Sect. 5.3.1), as the corresponding ellipse annotation approximates the outer isophote, whose SB value can be compared to the limiting SB of the survey.

[12] Astropy WCS `https://docs.astropy.org/en/stable/wcs/index.html`
[13] OpenCV-Python `https://pypi.org/project/opencv-python/`
[14] For shells, the SB values of the pixels along the curved line are retrieved.

4.6. Color measurements

As mentioned in Sect. 2, g and r band images were available for all MATLAS galaxies, so we were able to construct the corresponding $g-r$ colormaps. For CFIS, this was not possible as only the r-band was available. We created the MATLAS SB files in the g-band using the same process as the one described in Sect. 2 for the r-band images. The $g-r$ colormaps were computed from the g and r-band SB files.

Afterwards, we applied the masks of our annotations on the $g-r$ colormap FITS file. For each annotation, we retrieved the $g-r$ values of the pixels inside the mask, and we kept the median value to remove the contribution of outlier pixels. Finally, we estimated the median $g-r$ value for each annotation type (tidal feature, halo, or main galaxy). However, these values are only estimates as the colors were computed on polluted images with sources of contamination such as bright reflections or cirri that can overlap with the LSB features of interest.

4.7. Level of contamination

The contamination of the images by pollutants such as ghost reflections, cirrus, high background, satellite trails, or artifacts coming from the instrument can be high in deep images and lead to biased annotations. For instance, if the primary galaxy is embedded in a bright ghost reflection from a nearby star, the user is likely to underestimate the real size of the halo. In order to quantify this degree of contamination, we automatically assigned a reliability index based on the intersection between the halo of the primary galaxy and pollutants as follows. The higher the reliability index, the cleaner the image.

– None: the annotation of the halo is impossible (for instance due to a high contamination by bright sources)
– 1: the entire halo is embedded in a ghost reflection or a high background region
– 2: the halo has an overlap with a ghost reflection or a high background region
– 3: the entire halo is embedded in a companion galaxy
– 4: the entire halo is embedded in a ghost reflection coming from the core of the galaxy
– 5: the halo has an overlap with a companion galaxy
– 6: the entire halo is embedded in cirrus
– 7: the halo has an overlap with cirrus
– 8: the halo has an overlap with a satellite trail
– 9: there is no pollutant overlapping with the halo.

As the reliability index can take several values for a given halo, it is stored in a list. To have an average reliability index per galaxy, it was necessary to compute a weighted average of the values. In order to penalize strong pollutants, the values [1, 2, 3, 4, 5, 6, 7, 9] were associated to the weights [8, 7, 6, 5, 4, 3, 2, 1] respectively. The satellite trails are not taken into account as their impact on the classification is very low. The weighted reliability index in our images is discussed in Sect. 6.3.

In addition, in Sect. 6.4, we define a similarity index that assesses the similarity between two users' annotations of the same physical type (e.g., halos, pollutants). This similarity index is compared as a function of the level of contamination we defined in this section.

4.8. Annotations kept

By construction, several users have annotated the same structures for any given galaxy. We therefore faced the difficulty of keeping the most representative annotations for any specific galaxy. We present in this section our selection process. It is relatively simple for the diffuse halos since they were systematically annotated by all users. We made the basic assumption that more expert users tend to see fainter features in deep images (Bílek et al. 2020). Therefore, we decided, for a given galaxy and structure, to only keep the annotation with the largest area. For the annotations corresponding to the brightest part of the galaxy (and made based on shallow images, like PanSTARRS-DR1), the extent is not a relevant criterion and we chose the annotation that represents the median area as our final annotation.

The process is a bit more complicated for tails and streams, as they were not always identified by all users, or might have been delineated in various ways. For these tidal features, we proceeded in two steps. First, we considered streams and tails together, and we paired the annotations of the two first users. We used the same method as for the reliability index: two paired structures with a percentage of intersection[15] higher than a given threshold (namely 25%) are considered as being the same structure. A same unique identifier is attributed to them. We repeated the process with the annotations of the other users: the new feature was paired with the previous ones, the percentage of intersection was computed and then the feature was either associated with an already existing unique identifier or as a new one. This iterative process is illustrated on the panels a to d of Fig. C.1.

Afterwards, for the features sharing the same unique identifier, the one with the largest area is kept following the procedure used for the stellar halos. Hence, at the end each galaxy will have tidal features with different unique identifiers, as visible on the panel e of Fig. C.1. In the following, only the annotations that were kept after the selection process are taken into account (except if explicitly mentioned otherwise).

5. Results

In this section, we present statistical results based on our annotations of tidal features, including their geometrical properties and surface brightness. We opted for median values rather than mean ones in order to get the most representative values. The database contains 8441 annotations. The number of annotations per feature type is detailed in Table 2. The annotations have been made by four users: two of them have annotated all the galaxies, while 30 and 58.5% of the galaxies were delineated by the two other users.

5.1. Tidal tails and streams

A total of 223 tidal tails and 84 streams in our database were kept after the selection process[16]. Here we present several geometrical and surface brightness analyses of these features. Based on these analyses, we provide a discussion on the criterion used to differentiate tidal tails from streams in Sect. 6.5.

[15] The intersection score is computed as follows: each annotation is defined by its contour coordinates and is considered as a filled polygon (using the shapely package (`https://shapely.readthedocs.io`)). The area of each polygon is computed, as well as the area of the region of the intersection between the two polygons. The intersection score is then the area of the intersection region divided by the area of the larger polygon.

[16] We do not make a distinction between tidal tails (with typical, antennae-like shape) and plumes.

Fig. 7. Thumbnails of all the galaxies having tidal tails or streams. The main galaxy is represented in magenta, the halo in yellow, tidal tails in blue and streams in red. A scalebar representing 30 kpc at the bottom left gives an indication of the size of the structures.

Table 2. Number of annotations stored in the database as a function of their type.

Annotation type	Number
Main galaxy	1013 (352)
Halo	962 (340)
Tidal tails	433 (223)
Streams	171 (84)
Shells	260
Companion galaxy	808
High background	1121
Ghosted halo	3238
Cirrus	283
Satellite trail	30
Instrument	122
Total	8441 (6861)

Notes. In parenthesis is indicated the number of annotations kept after our selection process on the main galaxy, halo, tidal tails and streams.

5.1.1. Qualitative interpretation of global shapes

To have an initial overview of their morphology and of their location with respect to the target galaxy, we inspected their thumbnails as represented in Fig. 7. Their shapes and sizes show a great variability. For a given galaxy, tidal tails often appear in pairs and seem to be rather located near the foci of the ellipse defining the halo, whereas streams seem more isolated. Globally, tidal tails appear rounder and broader than streams, while streams seem more elongated and thinner. It must be noted that the thickest tails were referred as plumes in our study.

To better compare the individual shapes of each type of tidal structure, we present in Appendix A the footprint of the tails and streams without their host galaxy as a function of the morphological type: tidal features for LTGs are shown in Figs. A.1 and in Fig. A.2 for ETGs. All thumbnails have the same physical size (namely 50×50 kpc) and they are sorted by increasing mass of the host galaxy. The mean galaxy mass in each row is detailed in the text of Appendix A.

From these figures, more massive galaxies tend to host larger or more extended tidal tails. For ETGs, tidal tails seem slightly rounder and larger than for LTGs, but ETGs are on average more massive than LTGs. For streams, there is no clear trend neither as a function of the mass of the galaxy nor of its morphological type, which is expected as the material does not originate from the primary galaxy but from a companion.

One important point to note is the fact that tidal tails and streams look relatively similar. Although streams globally seem

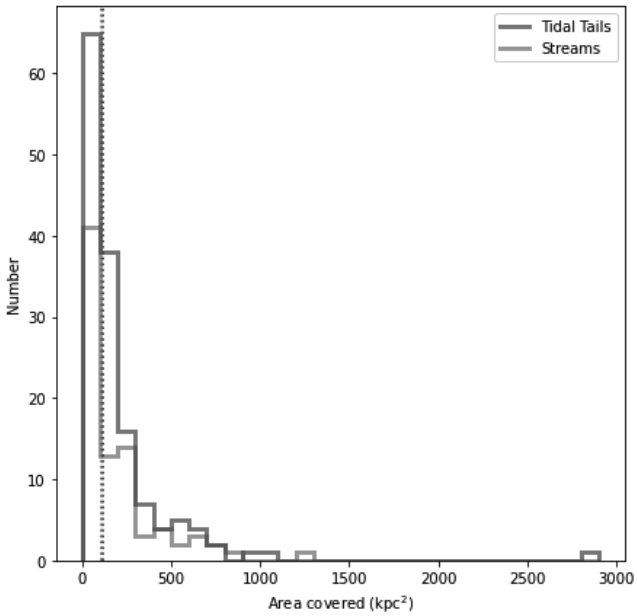

Fig. 8. Histogram of the area (in square kiloparsecs) covered by tidal tails (in blue) and streams (in red), in bins of width 100 kpc^2. The median of each distribution is represented by the dotted lines.

Table 3. Median areas covered by tidal tails and streams for each type of galaxy, associated with their standard deviation.

Galaxy type	Tidal tails		Streams	
	Median area (kpc^2)	std	Median area (kpc^2)	std
All galaxies	116	295	108	227
ETGs	161	445	139	252
LTGs	102	149	86	171

Notes. Values are expressed in square kiloparsecs.

more elongated than tails, there is no obvious visual difference between them: some tails look like streams and conversely, for all mass range.

Though global trends are observed, clearly the large variability of structures (that might be partly due to delineation errors) does not allow us to make a sharp distinction between tails and streams simply based on a visual inspection. In the following sections, we use a more quantitative approach to assess whether a statistically significant distinction between these two types of features can be found.

5.1.2. Quantitative interpretation: area

The distributions of the areas of each type of structure is plotted in Fig. 8. Table 3 summarizes the results, making a distinction between LTGs and ETGs.

The distributions of areas hardly differ for both tidal structures: they are peaked around 115 kpc^2 for tidal tails and 110 kpc^2 for streams. Most structures cover areas smaller than 300 kpc^2, and few between 300 and 800 kpc^2. We note that the structures with the largest areas are associated to systems showing ongoing interactions.

From the table, structures surrounding ETGs seem more extended than structures surrounding LTGs, but this difference is not statistically significant. Indeed, we applied Mood's statistical test (testing the null hypothesis that two samples come from populations with the same median) on the areas of the structures as a function of the morphological type: the p-value is 0.08 (for tidal tails) and 0.18 (for streams), which are higher than 0.05, so we cannot reject the null hypothesis that the medians are the same at a confidence level of 5%.

The values of the standard deviations may be reduced when taking into account trends with the mass of the host and environment. This will be studied in detail in another paper.

Table 4. Median values of the longest length in tidal tails and streams for each type of galaxy, associated with their standard deviation.

Galaxy type	Tidal tails		Streams	
	Median length (kpc)	std	Median length (kpc)	std
All galaxies	22	20	29	29
ETGs	23	27	34	31
LTGs	20	14	25	25

Notes. Values are expressed in kiloparsecs.

Table 5. Median values of the width in tidal tails and streams for each type of galaxy, associated with their standard deviation.

Galaxy type	Tidal tails		Streams	
	Median width (kpc)	std	Median width (kpc)	std
All galaxies	5.7	3.8	4.1	2.1
ETGs	6.6	4.5	4.3	2.1
LTGs	5.0	3.1	3.9	2.0

Notes. Values are expressed in kiloparsecs.

5.1.3. Quantitative interpretation: length and width

Median values of the measured length are summarized in Table 4 while their histograms are shown in Fig. 9. They show that streams are longer than tidal tails (with respectively a median value of 29 kpc and 22 kpc when combining all galaxies), for all galaxy types. The computation of Mood's test gives a p-value of 0.009, which is smaller than 0.05 so we can conclude that the medians of the length for tidal tails and streams are not the same at a significance of 5%. The distribution of the length of streams is more extended and flatter than for tails. A few structures reach a length longer than 80 kpc.

An estimate of the width of tidal tails and streams is obtained by dividing the area they cover by their longest length. This relies on the assumption that these features have ribbon or rectangular shapes. For round polygons, they are approximated by an ellipse (as explained in Sect. 4.2) so the width corresponds to the minor axis (while the length corresponds to the major axis). The results are presented in Table 5 while the histogram of the width of tidal features for all the galaxies is shown in Fig. 9.

From this histogram, one can see that tidal tails are on average wider than streams. This difference is statistically significant, as Mood's test on the medians of the width for tidal tails and streams gives a p-value of 0.0016. Almost all the streams have a width less than 10 kpc with a peak around 4 kpc, while for tails the peak is around 6 kpc and the distribution is more extended. A few tails are very wide (higher than 14 kpc).

The fact that tidal tails are wider than streams was expected from models: indeed, the width of a tail or stream increases with

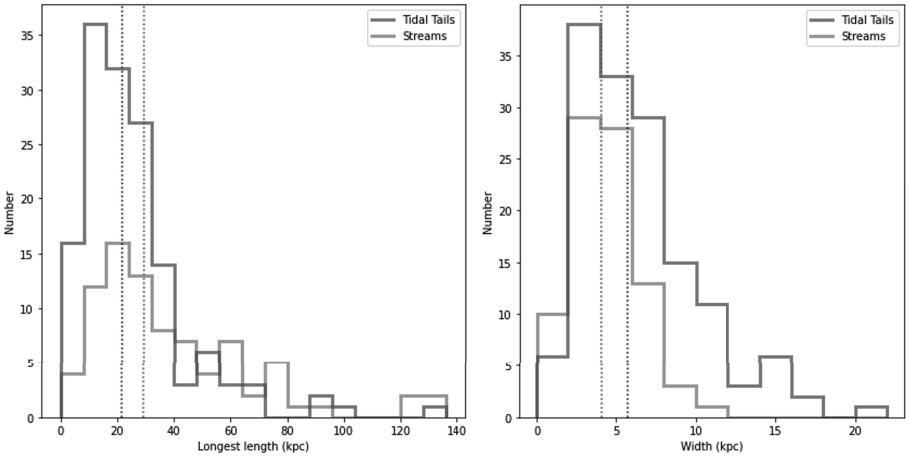

Fig. 9. Histograms of the length in kiloparsecs of tidal tails (in blue) and streams (in red), in bins of width 8 kpc (*left*), and of their width in bins of width 2 kpc (*right*). The median of each distribution is represented by the dotted lines.

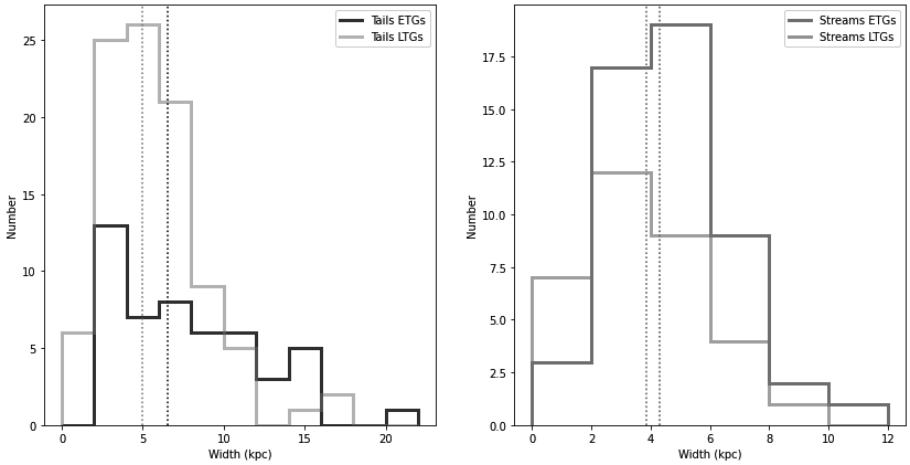

Fig. 10. Histograms of the width in kiloparsecs of tidal tails (*left*) and streams (*right*) in bins of 2 kpc as a function of the morphological type of the galaxies: ETGs are represented by darker shades than LTGs. The median of each distribution is represented by the dotted lines.

the velocity dispersion of the stars that form this structure (e.g., Johnston et al. 1996; Johnston 1998). Yet, the velocity dispersion of a galaxy depends both on its morphological type and of its mass (e.g., Bernardi et al. 2010; Bezanson et al. 2012). The more massive the galaxy, the higher the velocity dispersion and therefore the wider the tail. Hence, since streams originate from low-mass companions, their velocity dispersion is smaller. This contributes to their widths being smaller than for tails. This is consistent with Fig. 6 from Hendel & Johnston (2015) who found in their simulation that if the infalling galaxy satellite had a higher mass, the width of the debris increases. Therefore, our results comfort our approach to make a distinction between tidal tails and streams.

To explore the results even further, the histograms of the width of tidal tails and streams as a function of the morphological type of their host galaxy is shown in Fig. 10. From it, one can see that the distributions of the width of tidal tails are very different for ETGs and LTGs: the distribution is flatter and more extended for ETGs, with a median value of 6.6 kpc, while for LTGs the distribution is peaked around 5 kpc, with few tails having a width between 10 and 18 kpc. This was also expected, as the velocity dispersion is higher for ETGs than for LTGs, producing wider tails. Contrary to the significant difference for tidal tails, there is no real difference for streams between the distributions for ETGs and LTGs, which are relatively similar. This was also expected, as the morphological type of the primary galaxy is not related to the one of its small companion producing streams.

We can also represent the length and width of tails normalized by the effective sizes of their host galaxy, and as a function of the morphological type, as visible in Fig. B.1. The same trends than previously mentioned for tails are visible, so a flatter distribution for the width of ETGs, and a slightly longer length, than for LTGs. We did not normalize the length and width of streams as they do not originate from the primary galaxy.

To summarize this section, we found statistical differences between tidal tails and streams: from our measures, streams are more elongated and thinner than tidal tails, which was already hinted by the visual inspection. The agreement with theoretical arguments gives credibility to our classification based on visual impression.

5.1.4. Surface brightness

In Table 6, we give the overall median inner SB value for each type of structure and host galaxy type. The distributions of these values are visible in Fig. 11.

Having a median SB of 26.2 mag arcsec^{-2}, streams are fainter than tails by 0.9 mag. This difference is statistically significant, as Mood's test applied on the SB values of tidal tails and streams returns a p-value of 8.6×10^{-8}. As seen in Fig. 11, the SB distribution of streams is narrower than that of tidal tails. None of them have SB fainter than 27.5 mag arcsec^{-2}.

Several reasons could explain why tidal tails are on average brighter than streams. First, our sample includes ongoing or

Table 6. Median inner SB measurements for tidal tails and streams, expressed in magnitudes per square arcsecond, associated with their standard deviation.

Galaxy type	Tidal tails		Streams	
	Median inner SB	std	Median inner SB	std
All galaxies	25.3	1.1	26.2	0.7
ETGs	25.4	0.9	26.1	0.7
LTGs	25.1	1.1	26.3	0.7

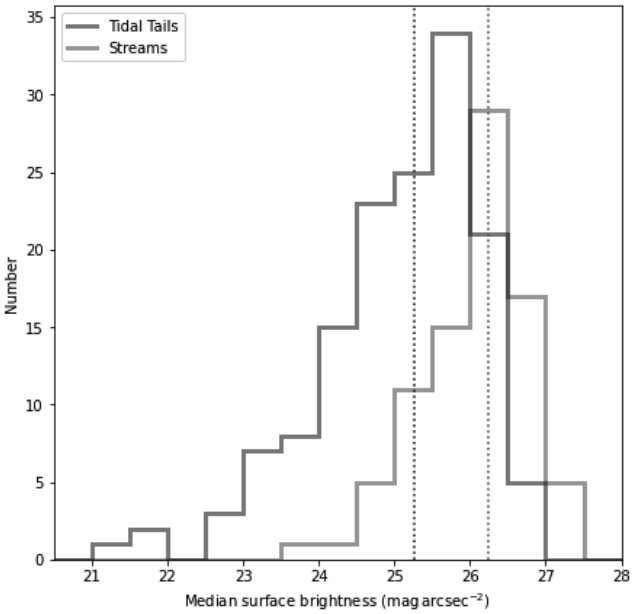

Fig. 11. Histogram of the median surface brightness value in magnitudes per square arcsecond for tidal tails (in blue) and streams (in red), in bins of width 0.5 mag arcsec^{-2}. The median of each distribution is represented by the dotted lines.

recent mergers. Their tails are too young to face the evaporation process that fade collisional debris. Besides, those formed from gas-rich LTGs can contain young and luminous stellar populations (e.g., Elmegreen et al. 1993; Jarrett et al. 2006). In addition, by definition tails form from massive objects and should therefore be more massive and luminous than the streams which emanate from lower-mass satellites. A second explanation could be related to the different survival times of each type of structures (see e.g., Mihos 1995; Ji et al. 2014; Mancillas et al. 2019): streams tend to be visible for a longer time than tails. Tails could be more fragile than streams and therefore could disappear faster, meaning that their typical morphology would start to be lost when the structure orbits the galaxy more than once. The old tail would rather resemble multiple streams. We might be able to detect tails only when they are young enough to keep their typical morphology, hence bright enough.

5.1.5. Overall bending

As seen in numerical simulations of galaxy mergers (e.g., Bullock & Johnston 2005; Cooper et al. 2010; Lux et al. 2013; Gibbons et al. 2014), tidal streams follow approximately the orbit of their progenitors: wrapping around the primary galaxy, they appear as strongly curved. On the other hand, the shape of tails that emanate from the primary galaxy is mainly driven by tidal forces. Depending on their orientation and until their material falls back on the primary, these structures may appear as relatively straight. Therefore curvature may be another criterion to disentangle streams and young tails.

To obtain a basic estimate of the curvature, we fit the topological skeleton (as defined in Sect. 4.2, it is the medial axis and can be seen as a thinner version of the shape that is equidistant to its boundaries) of tidal features by a linear function, using a least-squares regression. It must be noted that we performed this computation only for not-round polygons, as the skeleton for round polygon does not represent properly the shape. We compared the R^2 correlation coefficient determined for streams and tails: the higher R^2, the closer the feature is to a straight curve.

For streams, the mean and median R^2 are respectively 0.62 and 0.79 with an associated standard deviation of 0.37. For tails, the mean and median R^2 values are respectively 0.59 and 0.64, with a standard deviation of 0.34. It appears that there is a large variability, especially for streams where the difference between the mean and the median value of R^2 is more important. In addition, the linear fit for streams is better than for tails. Hence, streams do not appear more curved than tails as we would have expected from simulations.

One possible explanation, besides the projection effects, the uncertainties of the method and/or possible confusions between streams and tails, is that the depth of the survey may not be sufficient to follow the structures over large distances. They must be long enough to get a reliable estimate of the curvature, which is often not the case (see Figs. A.1 and A.2). Besides, the bending of stream might be invisible if it is partly hidden by the host galaxy.

5.1.6. Color

The median $g-r$ color values computed from the colormaps for tidal tails and streams are presented in Table 7. It must be noted that the colors were computed on the images without any cleaning process: pollutants such as bright ghost reflections or high background may affect the color measurements. As pollutants are more visible in the r-band than in the g-band, they will tend to redden all measures. Nevertheless, our measure of the color of streams surrounding ETGs, with a median $g-r$ value of 0.64 mag, is in agreement with Martinez-Delgado et al. (2021) who determined a $g-r$ value between 0.5 and 0.8 mag for 24 streams around local galaxies.

Our analysis tends to show that tidal tails are bluer than streams by around 0.1 mag. This difference is statistically significant, as Mood's test gives a p-value of 0.003. This could be due to the presence of young stars along the tails. In gas-rich mergers, gas is expelled along tidal tails, just like the stars, and may be compressed in the collisional debris and trigger star-formation (e.g., Jarrett et al. 2006; Olson & Kwan 1990). On the contrary, dwarf satellites are usually gas poor, and if they have been stripped the color of their tidal streams will reflect that of their old stellar populations. So the observed difference in colors between tails and streams may be due to age effects: as already argued, tails tend to be observed at a younger age than streams. Taking into account the fact that the color of the old stars of satellites is bluer than that of the primary galaxy, due to their lower metallicity, the age effect may be even stronger.

5.1.7. Stream progenitors

If streams emanate from a disrupted satellite, remnants of the progenitor may still be visible. As matter of fact, the presence of

Table 7. Median $g-r$ value for tidal tails, streams, shells and halos for each type of MATLAS galaxy, associated with their standard deviation.

Galaxy type	Tidal tails		Streams		Shells		Halos	
	Median $g-r$	std	Median $g-r$	std	Median $g-r$	std	Median $g-r$	std
MATLAS all galaxies	0.57 (148)	0.17	0.65 (52)	0.16	0.60 (217)	0.25	0.62 (221)	0.12
MATLAS ETGs	0.57 (95)	0.16	0.64 (49)	0.15	0.60 (198)	0.25	0.63 (170)	0.11
MATLAS LTGs	0.56 (53)	0.19	1.00 (3)	0.21	0.70 (19)	0.20	0.57 (51)	0.15

Notes. The number of structures annotated is indicated in parentheses.

a condensation within a tidal structure was one of our criteria to label it as a possible stream, especially if a sign of an S-shape was present[17].

Hence, it is interesting to check the percentage of streams that have a progenitor from our annotation database. The progenitor of the stream is defined as follows. If there are no companion or dwarf galaxy in the annotations, the stream is considered orphan. If the stream has an overlap (partial or entire) with a companion or dwarf galaxy, then the progenitor is the companion galaxy.

The absence of a progenitor in a genuine stream may indicate that the satellite has been totally destroyed, if the stream was formed long ago, or that it is hidden in the primary galaxy. Statistically, orphan streams should be older than those having a progenitor still visible.

In our results, when taking all CFIS and MATLAS galaxies with streams into account, about 70% of all streams are orphan. In comparison, for the Milky Way and M31, although there is no precise census of the percentage of orphan streams, most of the streams originating from companion galaxies do not present a progenitor, excepted for the Sagittarius stream (Ibata et al. 1994). Trends are similar for our results even though more progenitors are still visible. As mentioned before, this percentage of orphan streams could be related to the age of these structures or to projection effects, but it might also be linked to misclassifications between tidal tails and streams. Indeed, the absence of a progenitor in the structure made the identification more complicated, which in some cases might have mislead users during their classification. We do not see any difference between ETG and LTG hosts, but we did not expect the stream properties or its progenitor to depend on the morphological type of the primary galaxy.

5.2. Shells

A total of 260 shells have been annotated. Geometrical, color and surface brightness measurements of these features are presented here.

5.2.1. Concentricity and radii

Shells have been annotated using curved lines and thus measuring their area is irrelevant. Such an annotation faces a major issue: the selection of the beginning and ending of a shell might be different for an expert or novice user. The former may be aware of the well-shaped circular shells in idealized numerical simulations and consider as a single structure a shell that might be divided into several arcs by the less expert users.

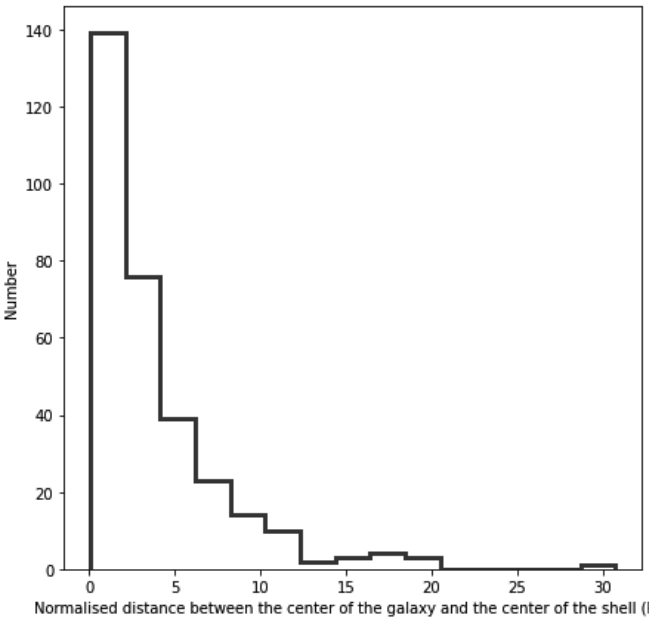

Fig. 12. Concentricity test: histogram of the distance in effective radius between the galactic center and the center of the shell. Higher distances correspond to larger deviations from concentricity.

Nevertheless, interesting metrics can still be computed, such as the concentricity or their radius.

Concentricity. From numerical simulations, shells are usually formed as concentric structures (e.g., Pop et al. 2018; Ebrová et al. 2021; Bílek et al. 2022), a prediction we can directly test with our observations. To compute the concentricity, we assume that the curved line defining the shell is a circular arc, and we compute its center[18]. Since the curved line is a cubic Bézier curve, it may differ from a circular arc if the user did not draw the shell properly. In that case, we only consider the starting, middle and ending point of the curve and we compute the center of the circle passing through these three points. We then compute the distance between the shell center and the center of the host galaxy. The histogram of theses distances is displayed in Fig. 12. A distance equal to zero means the shell is centered on the galaxy (i.e., concentric), while larger distances indicate a higher deviation from concentricity.

Most often the difference between the galactic center and the center of the shell is less than 10 R_e and the median distance is around 2.5 R_e. This means that in general the center of the shell is still located inside the galaxy and we can consider them as relatively concentric. Cases for which the relative distance between

[17] Condensations and tidal dwarf galaxies may be present in tidal tails made in major mergers, but being formed in situ, they do not exhibit the S-shape typical of tidally disrupted dwarfs.

[18] The center of the shell is hence the center of the circle that passes through the circular arc defining the shell.

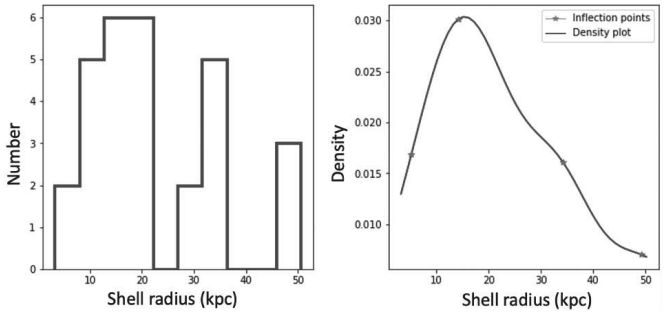

Fig. 13. Example of NGC 0474. *Left*: histogram of the shell radii for this galaxy. *Right*: corresponding density plot with the inflection points of the curve indicated by red stars.

the centers is large may correspond to bad approximations of the shells as circular structures or to real nonconcentric shells such as those of Type III.

Radius. The shell radius is another metric which can easily be compared to predictions from simulations. It is computed as the on-sky separation between the point lying in the middle of the curved line defining the shell and the center of the galaxy host.

Such a computation needs to take into account the fact that for a given galaxy, multiple users might have annotated the same shells. Simply computing the mean shell radius per galaxy averaged over all the users would introduce biases. Indeed, as mentioned earlier, the number of shells annotated depends on the expertise of the user and on the shape of the shell itself. Clearly defined shells will be annotated as one structure, while less-defined ones will be annotated as several shells. This means that the less well-defined ones will have a higher impact and count for more features.

To tackle this issue, we have directly plotted for each galaxy the histogram of the radii of shells annotated by all users, and drawn the corresponding density plot. On that plot, we identified the most representative values, that is to say the radii corresponding to the inflection points, as illustrated in Fig. 13. The inflection points, referred here as "peak radii", are computed using a Scipy function to find peaks in a 1D array using a Ricker wavelet transformation[19]. Their histogram for all galaxies are represented in Fig. 14.

Most of the shells have a radius smaller than 40 kpc with a few extending to 80 kpc. For instance the shells of the prototypical galaxy NGC 0474 range from 10 to 50 kpc Bílek et al. (2022). We do not observe shells beyond 80 kpc, which seems at odds with some simulations. Whereas they extend to 120 kpc in Ebrová et al. (2019) and Karademir et al. (2019), some of the shells in Pop et al. (2018) have a radius reaching 150–200 kpc. Obviously the comparison is not straightforward as the shell orientation (not well constrained from our annotations of real systems) and differences in surface brightness need to be taken into account.

5.2.2. Surface brightness

We measured the SB value along the curved line defining the shell annotation. When considering all galaxies, the median inner SB values for shells is 25.4 mag arcsec^{-2}, it is of 25.3 mag arcsec^{-2} for ETGs and of 25.6 mag arcsec^{-2} for LTGs

[19] Scipy find_peaks_cwt, https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks_cwt.html

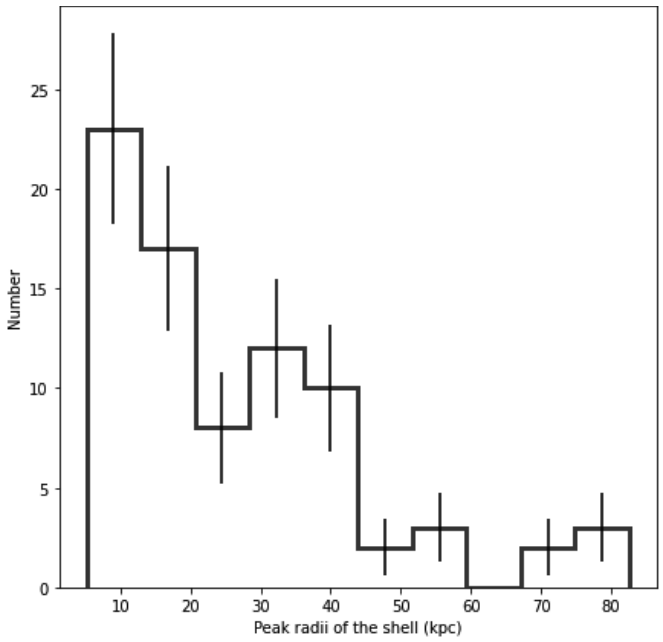

Fig. 14. Histogram of the peak radii of shells (in kiloparsecs) for all galaxies, with Poisson uncertainties plotted as the black error bars in each bin.

hosts. The distribution of SB values, presented in Fig. 15, ranges between 21 and 28 mag arcsec^{-2}. Shells are detected with a maximal surface brightness close to the nominal depth of the surveys. This is linked to their shape, as a circular arc is easy to detect and identify on an image: even very faint shells can be visually recognized, while it is more difficult for complicated shapes like tidal tails or streams (see Sect. 6.5).

We note that we did not subtract the stellar halo of the host before our measurement, explaining why the inner shells (i.e., shells that are overlapping with the halo) are apparently brighter than the external ones (i.e., shells further away from the halo).

5.2.3. Colors

The median $g-r$ color values of the shells computed from the MATLAS colormaps are presented in Table 7. The median color of shells seems close (slightly redder) than that of other tidal features, but again the measurements may be polluted by the stellar halo.

There is a discrepancy of 0.1 mag between the color of the shells of LTGs and ETGs. Mood's test on the medians of the colors of ETGs and LTGs gives a p-value of 1.8×10^{-5}, which is smaller to 0.05 so we can conclude that the medians of shells for ETGs and LTGs are not the same at a significance of 5%.

5.3. Halos

After our selection process, 340 halos were kept. Here, we present the analysis of their surface brightness, radius and color.

5.3.1. Surface brightness and radius

We measured the surface brightness along the external contours of the annotated stellar halos. Figure 15 plots their distribution

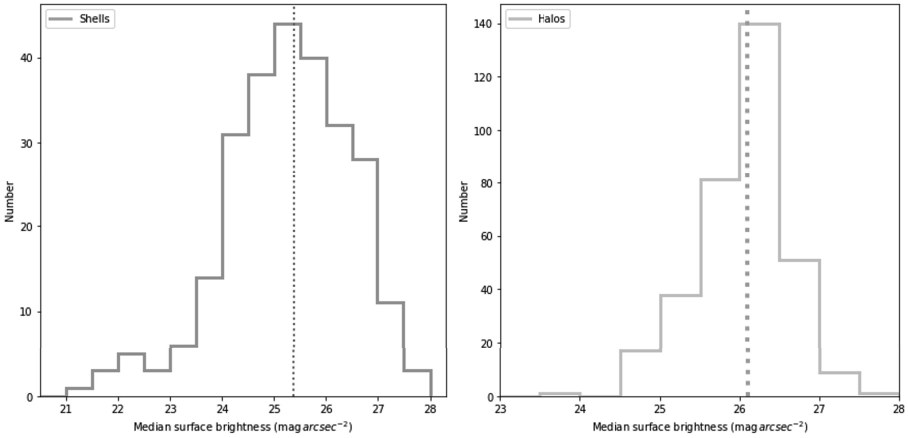

Fig. 15. Histograms of the median surface brightness value in magnitudes per square arcsecond for shells (*left*) and for halos (*right*), in bins of width 0.5 mag arcsec^{-2}. The medians are represented by the dotted lines.

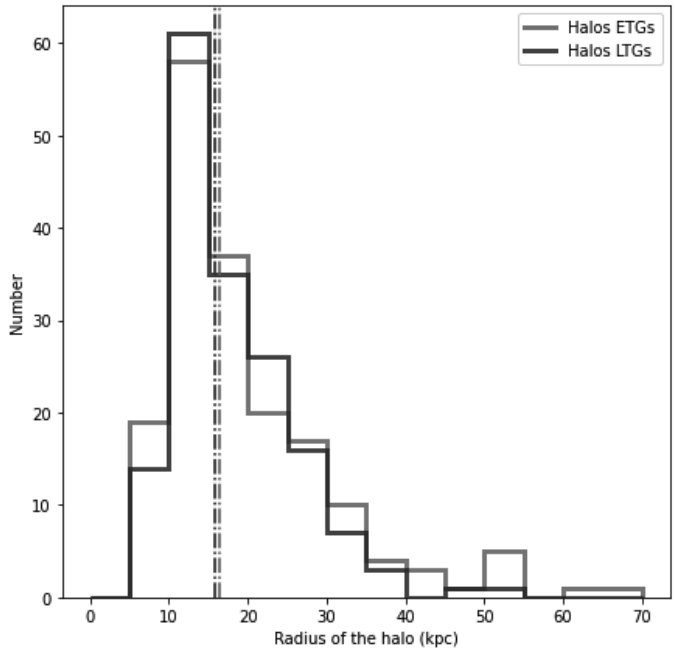

Fig. 16. Histogram of the radius of the halos in kiloparsecs as a function of the morphological type, in bins of width 5 kpc. The median of each distribution is represented by the dotted line.

for the 340 stellar halos kept after our selection process. It peaks at 26.1 mag arcsec^{-2} (median value).

Such a value is clearly much lower than the nominal SB limit of the survey, and that obtained when deriving integrated surface brightness profiles of galaxies. It just reflects the ability of the eye in delineating an external contour on our asinh images.

In addition, from the coordinates of the annotations stored in our database, we can compute the radius of the annotated stellar halos. The histogram of the values measured for all galaxies is shown on Fig. 16. One can see that the median radius is around 16.5 kpc for ETGs and 15.9 kpc for LTGs, with the majority of the halos having a radius smaller than 30 kpc. A few radii extend up to 70 kpc and correspond mostly to ETGs. However, the distributions are relatively similar for the two morphological types.

We would expect ETGs to have larger radii than LTGs. Indeed, there are more satellite companions around ETGs than LTGs for a given mass, and more companions for more massive galaxies (Kawinwanichakij et al. 2014). In addition, the mass growth of present day galaxies seems to be driven by minor mergers (Oser et al. 2010), leading to the formation of streams. Hence ETGs should have more material in their outskirts and so larger radii. The fact that we do not observe this could be explained by the low SB of streams. We may not be able to visually recover all the streams, especially the fainter, so we might underestimate the extent of the faint outskirts of the halo. The study of integrated SB profiles of ETGs and LTGs might reveal differences, but this is behind the scope of this paper.

From our annotation database, we cannot directly infer the stellar mass of the halo, an important parameter to constrain the mass assembly of galaxies (Merritt et al. 2020). This estimate requires a lot of processing, in particular to take into account PSF effects (Karabal et al. 2017). It will be the focus of another paper in this series.

5.3.2. Color

The median $g-r$ color values computed from the colormaps for halos are presented in Table 7. One can see that there is a 0.06 mag difference between the median $g-r$ color for halos of ETGs and LTGs. It is statistically significant (at a significant level of 5%), with a p-value of 0.025 from Mood's test. Not so surprisingly, the stellar halos of LTGs are bluer than for ETGs, due to (low levels of) star formation occurring there.

5.4. Covering factor

As mentioned earlier, our annotation database does not provide an estimate of the stellar mass of the LSB structures surrounding their host galaxies. However one proxy of their relative importance is their "covering factor", defined as the percentage of the pixels belonging to one type of structure with respect to a given field of view. For the latter, we considered boxes of side length 20 R_e centered on the primary galaxy.

For this measurement, we selected the annotations of a given galaxy and a given user. The median values of the covering factor for different types of structures are given in Table 8. A covering factor of 0 (respectively 1) means that the feature is not present on (respectively entirely covers) the given field of view. Results are shown in Fig. 17. Within the selected boxes, the stellar halo has a median covering factor of 0.16: it is computed for the entire outer ellipse annotation of the halo, without subtracting the main galaxy. This value is to be compared to 0.03 for

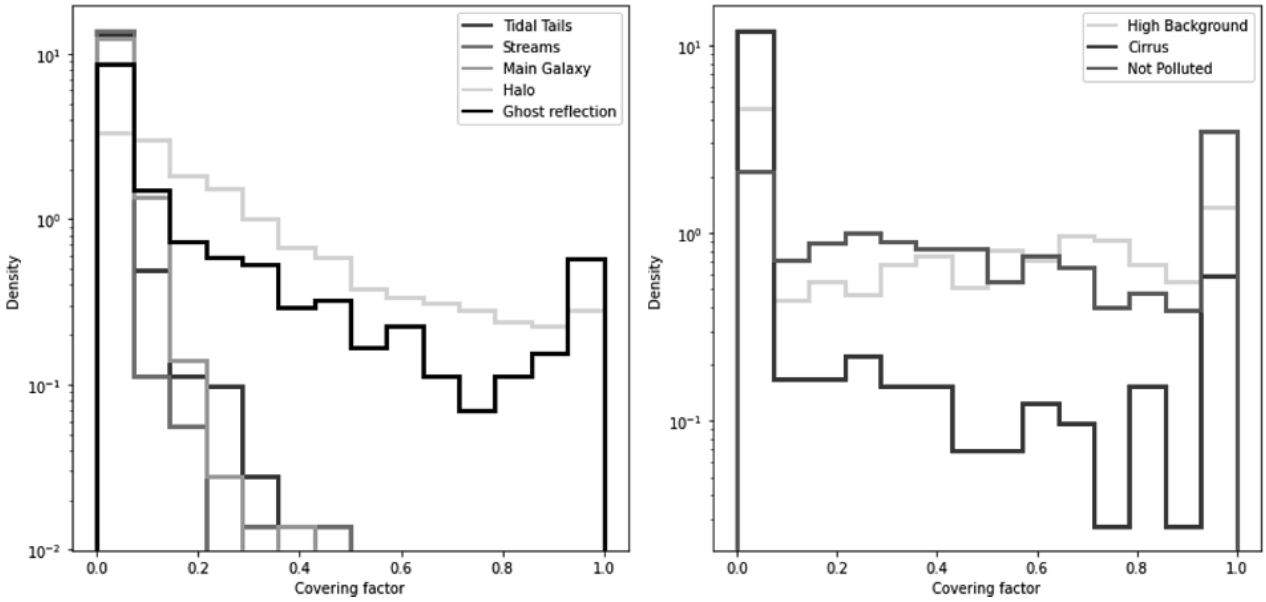

Fig. 17. Histograms of the covering factor. *Left:* covering factor for stellar structures and ghost reflections in a selected box of side length 20 R_e centered on the primary galaxy. *Right:* covering factor for high background, cirrus and clean ("Not Polluted") regions in a field of view of $30.95 \times 30.95'$ around the primary galaxy.

Table 8. Median of the percentage of the covering factor in selected boxes of side length 20 R_e around the primary galaxy.

Feature type	Median covering factor
Main Galaxy	0.03
Halo	0.17
Tidal Tails	0.04
Streams	0.02
Ghost reflections	0.01

Notes. For tidal tails and streams, they are counted only when the galaxy exhibits these features.

the central regions (main galaxy), 0.04 for tidal tails and 0.02 for streams, considering only galaxies that do have tidal tails or streams.

For pollutants such as high background and cirrus, it is more relevant to determine their covering factor for larger fields of view, to get predictions on the contamination levels for other surveys. We considered a field of view of $31 \times 31'$. For regions with high background, the covering factor has two peaks respectively near 0 and 1, with a relatively uniform distribution between these two values. It is higher than 0.9 (fully polluted images) for 11% of our galaxies and lower than 0.1 (clean images) for 34% of our galaxies. Restricting the analysis to the cirrus, the histogram reveals images that are either fully covered or completely absent with very few intermediate cases. Overall, regions free of any contamination sources (excepted the ghosted halos), that is to say with a covering factor higher than 0.9 for the clean pixels, correspond to about one fourth of our images, while 17% of our images are almost fully contaminated, so clean regions cover less than 10%.

6. Discussion

The originality of our approach relies on the use of an annotation database of Low-Surface-Brightness features compiled by a group of users who have precisely delineated a large number of individual structures (representing a total of 8441 annotations) directly on displayed images and classified them. The method raises a number of issues partly posted in the previous sections: (1) the difficulty of matching individual annotations made by different users (2) the presence of overlapping structures, in particular the contaminants which have a large covering factor, which prevent us from determining with precision the boundaries of some structures (3) the ambiguity in the classification of the various types of tidal features (4) the fact that we have used images coming from various CFHT surveys with varying depth and surface brightness limits and (5) the reliability of the annotations when considering users with different levels of expertise. We address all these issues in this discussion.

6.1. Survey sensitivity to identify tidal features

One important point to note is the fact that no annotated tidal tails or streams are fainter than 27.5 mag arcsec^{-2}, even though the nominal depth of the images is at least of 28.3 mag arcsec^{-2}. This discrepancy can be explained by several factors. First, the nominal survey depth was estimated from measures done on boxes of $10'' \times 10''$, while the structures of interest are more extended. Fluctuations of the SB brightness along the most diffuse structures (possibly above or below the SB limit) make it break into several substructures on our images, and its identification and classification as a single genuine stellar feature is very difficult. The presence of artifacts of similar SB as the structures has the same consequence of apparently breaking them into pieces.

This discrepancy of about 1 mag between the faintest tidal structures that may be identified and classified and the nominal SB limit of the survey must be kept in mind when making comparisons with simulations or estimating their visibility with other facilities. Similarly, the outer SB of halos is smaller than the depth of the survey, as the eye is not able to detect the faintest structures compared to what can be obtained with aperture photometry.

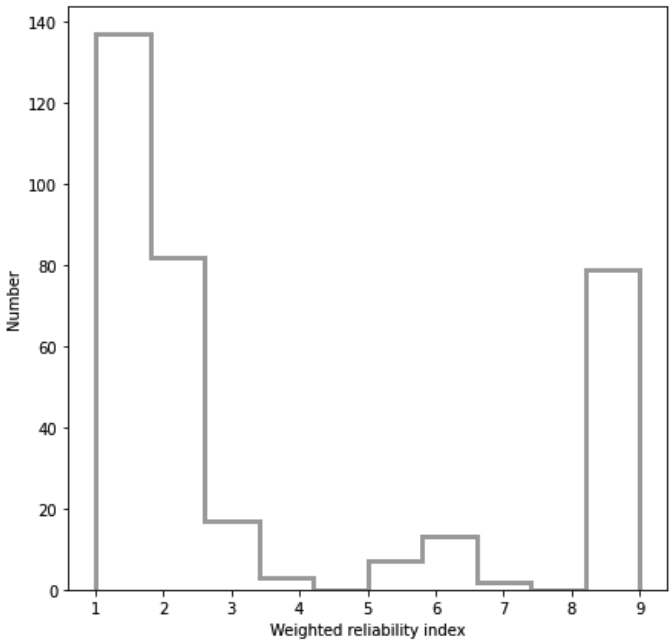

Fig. 18. Histogram of the weighted reliability index for the galaxies. A higher index corresponds to a cleaner image around the halo.

On the contrary, shells are more easily identified because of their characteristic circular shape, even if they are nonlinear, and indeed the SB of the faintest shells are close to the SB limit of the survey (see Sect. 5.2.2).

6.2. Generalization to users with different levels of expertise

In the preliminary study presented in this paper, the annotations were made by four users with a similar level of expertise. However, the annotation tool can be used by anyone and this study can be extended to less expert users. Taking the level of expertise of the user into account in the classification of tidal structures is important, as it reflects the degree of confidence that we can have in the annotation of this user (Bílek et al. 2020). Therefore, it is possible with our annotation tool to attribute weights to the users, and these weights would be inherited by the annotations. In that case, the results should take the weights into account, by computing for instance the weighted median and weighted standard deviation instead of simply the median and standard deviation.

Applying weights to the users could also modify our selection method described in Sect. 4.8. For instance, instead of keeping the largest annotation for halos, tails and streams, one could think of weighting the annotation masks and combining the shapes into a weighted combination of the different annotations.

6.3. Level of contamination

As mentioned in Sect. 4.7, we defined a reliability index to take the pollutant sources in the vicinity of the halo of the galaxy into account, as they might lead to biased annotations. The higher the reliability index, the cleaner the image around the galaxy. The histogram of the weighted reliability index can be seen in Fig. 18. One can see that only 23% of the images are completely free of pollutants around the halo, while 65% are polluted (with an index smaller or equal to 2). This indicates that a majority of our annotations are embedded in a polluted region which might have biased our delineations, but more importantly this could be a major issue for automated classification methods.

6.4. Similarity between annotations

Since several users have annotated the same galaxies, it is necessary to assess their reliability. Following Sect. 6.2, if the users have different levels of expertise, this assessment could be useful to adjust the user weight. Indeed, it is important to characterize whether they annotate like the majority of the group or if their annotations are too different from what is expected (especially for nonexpert users). In that case, the weight associated to that user will be lowered in order to avoid outliers in the results due to nonreliable classifications. This step could be part of a methodological process for future studies. We remind that in this paper, all users have a similar level of expertise and no weight was assigned.

A similarity index ranging from 0 (no similarity) to 1 (total similarity) was computed to assess the similarity between the annotations of two users. It depends on the annotation type and is defined as follows.

First, for halo and main galaxy annotations: the Jaccard index (Hennig 2007) is used. It is defined as the area of the intersection of both structures divided by the area of their union. If an annotation is missing for one of the user the index is set to zero. For shells: since their number is relevant, the similarity index is defined as $\frac{\min(S_i,S_j)}{\max(S_i,S_j)}$, where S_i (respectively S_j) is the number of shells annotated by user$_i$ (respectively user$_j$).

Then, for tails and streams: two metrics are used. The first one, like for shells, is an index computed over the number of features. However, the area covered is more relevant to characterize them (rather than the number of features, as they range from 0, 1, or 2 for tails and generally 0 or 1 for streams). Hence, a second metric was defined. It relies on the pairing of tidal tails and streams as defined in Sect. 4.8 with the unique identifier. The Jaccard index is computed on the structures annotated by user$_i$ and user$_j$ that share the same unique identifier.

Lastly, for contaminants (high background, cirrus): the union of all the contaminant of annotations of the given type is made, as sometimes a user will split a pollutant annotation into two. Then, the Jaccard index is computed between the unions of the contaminants.

We found out that the mean similarity index between two of our users reaches about 0.81 for halos and 0.52 for tidal tails and streams sharing the same unique identifier. The relatively high similarity between halo annotations is an indicator that both users annotated in a comparable way, which is important for our study. The lower similarity index for tidal features was expected, since annotating tidal features is not as clear and easy as annotating halos.

It is also interesting to determine whether users annotate in a comparable way in the presence of pollutants. When we consider only the cleanest annotations, that is to say annotations with as weighted reliability index as defined in Sect. 4.7 equal to 9, we obtain a similarity index of 0.85 for halos and 0.57 for tidal tails and streams. When keeping only the most polluted annotations (with a weighted reliability index smaller or equal to 2), the similarity index for halos is 0.8 and 0.53 for tails and streams. One can see that for halos, the cleaner the image, the higher the similarity index, which is likely related to the fact that users annotate in a more similar manner when the image is less polluted. For tails and streams, the trend is not that clear but this might be due to the fact that it is more complicated to precisely delineate these features in a similar manner. Therefore in general pollutants do

not seem to be the main source of differences in the delineations of tidal debris.

6.5. Disentangling tidal tails from streams

As mentioned in Sect. 3.2, we defined tidal tail as structures with stellar material apparently coming from the primary galaxy, while streams originate from a less-massive smaller companion which may still be visible, hidden, or have been destroyed. Most probably the users have adopted this definition in different ways according to their expertise and used a variety of observables (location, shape, amount of overlap with the closest galaxy, etc.) to assess the classification.

Getting quantitative measurements from our annotation database, we are able to determine whether these 2 classes of tidal objects show different properties, may thus be really distinguished from our images and retrospectively check whether the basic initial criterion for disentangling them was relevant.

As presented in detail in Sect. 5.1.3, we found that streams are narrower than tidal tails, a difference expected from models, as the width of tidal debris largely depends on the velocity dispersion of the progenitor, itself linked with its total mass. Just considering tidal tails, those found around ETGs appear wider than those associated to LTGs, another result at first order consistent with the expectations, since the velocity dispersion of galaxies decreases with their morphological type. In fact this explanation holds only for ETG-ETG collisions which naturally produce plume-like tidal tails. A merger involving one or two LTGs will produce more narrow tails. Conversely, tails coming from late-type galaxies (that have kept their stellar disk) are necessarily relatively thin. Taking into account all configurations, statistically, there should be more wide tidal tails around ETGs, as observed.

It must be noted that the area and length measured (from which we estimated the width) are only based on the projected shape of the structures since it is the only thing that can be annotated. It does not take the inclination and orientation along the line-of-sight of the galaxy into account, so the real intrinsic size of each structure is not known. Some tidal features are likely to be overlooked, especially when the galaxy is seen edge-on or when a tidal tail is hidden behind the galaxy.

We also found a statistically significant difference between the median surface brightness and color of tails and streams. This is also consistent with having progenitors of different masses and ages.

The fact that we see a statistical distinction in the physical properties of tidal tails and streams is a validation of our definitions of these features (see Sect. 3.2). However, from visual inspection of the thumbnails with all the individual tidal features shown in Figs. A.1 and A.2, it can be seen that their shapes can vary a lot from one to the other. Some trends are emerging (tidal tails seem broader than streams while streams seem more elongated), but the great variability observed makes difficult a clear visual separation between the two structures. This might suggest that an automated classification of streams and tails solely based on their individual shape could be very complicated. The human expert classifier might have used a number of criteria to classify streams and tails, in addition to calling his physical intuition on their origin. Thus results might be better when providing the machine all relevant information, including the properties of the primary galaxy, but still requires reliable labels. Deep learning techniques are good candidates to achieve this inclusion of properties thanks to its inherent accounting for visual context.

6.6. Impact of the depth of the survey

In this paper, we used images from two surveys with different depth, CFIS and the deeper MATLAS. Since most of our ETGs are drawn from MATLAS images and most LTGs from CFIS, we are in principle biased toward finding fainter features in ETGs. In this section, we study the impact of the depth of the survey on tidal features properties like the area or the length. To that end, we plot the 2D histogram of the SB values of tails and streams as a function of their area (Fig. 19), for ETGs and LTGs. A similar 2D histogram of the length as a function of the SB is represented in Fig. 20. One must note that we cannot compare these 2D histograms between tidal tails and streams as other processes than the depth are involved (see Sect. 6.5), and we must keep in mind the fact that we do not have the same number of tidal features around ETGs and LTGs.

For the area (Fig. 19) of tidal tails (panels a and b), there seems to be a slight trend that fainter structures are larger. We do not observe important differences between ETGs and LTGs, only three tails are larger and fainter in ETGs compared to LTGs. For the area of streams (panels c and d), the trend is less clear. Some streams have similar areas for a large range of SB values. This was expected as streams originate from a companion galaxy and the morphological type of the primary galaxy should not have an influence. Two streams are fainter than 27.2 mag arcsec^{-2} around ETGs, contrary to LTGs. Hence, for the area, although a few structures are larger and fainter around ETGs compared to LTGs, there is no major difference between ETGs and LTGs that could be attributed to differences in the depth of the surveys.

For the length (Fig. 20) of tidal tails (panels a and b), it seems that fainter structures are longer (the trend is clearer for ETGs than for LTGs). The maximum SB of tails hardly differs: around ETGs it reaches about 26.5 mag arcsec^{-2} for four structures, compared to 26.9 mag arcsec^{-2} for LTGs for two tails. For the length of streams (panels c and d), for ETGs two streams reach 27.5 mag arcsec^{-2}, which is 0.2 mag fainter than the faintest stream around LTGs. The longest streams (with a length $\geq$80 kpc) are not the faintest ones, and the majority of streams have comparable length and SB for ETGs and LTGs. Therefore, although there is a difference in the depth of the surveys, we do not observe a major bias in our results for the area or the length of tidal tails and streams.

6.7. Limits of the study

We remind readers here that all our quantitative measurements relied on the analysis of the annotation database. In particular, we did not perform aperture photometry in our images. This means that we cannot determine the flux of the tidal features or of the halos, and hence their stellar mass. Furthermore, no masking techniques were used to remove the polluting light from foreground or background objects. The SB values that are obtained are useful to compare trends between features, but a more detailed analysis would be needed to precisely assess them.

Similar to any other visual classification process, our annotations would be more reliable if tens or hundreds of people would have participated. In this paper, we rely on the annotations of four users only. However, we realized this process is complicated and the users need to be trained both to recognize LSB structures and to draw their shape with precision, making any citizen science project like Galaxy Zoo difficult to implement. This task is made more complicated by the presence of many pollutants that overlap with the interesting features. The time necessary to annotate is quite important as well (around 10 minutes per galaxy).

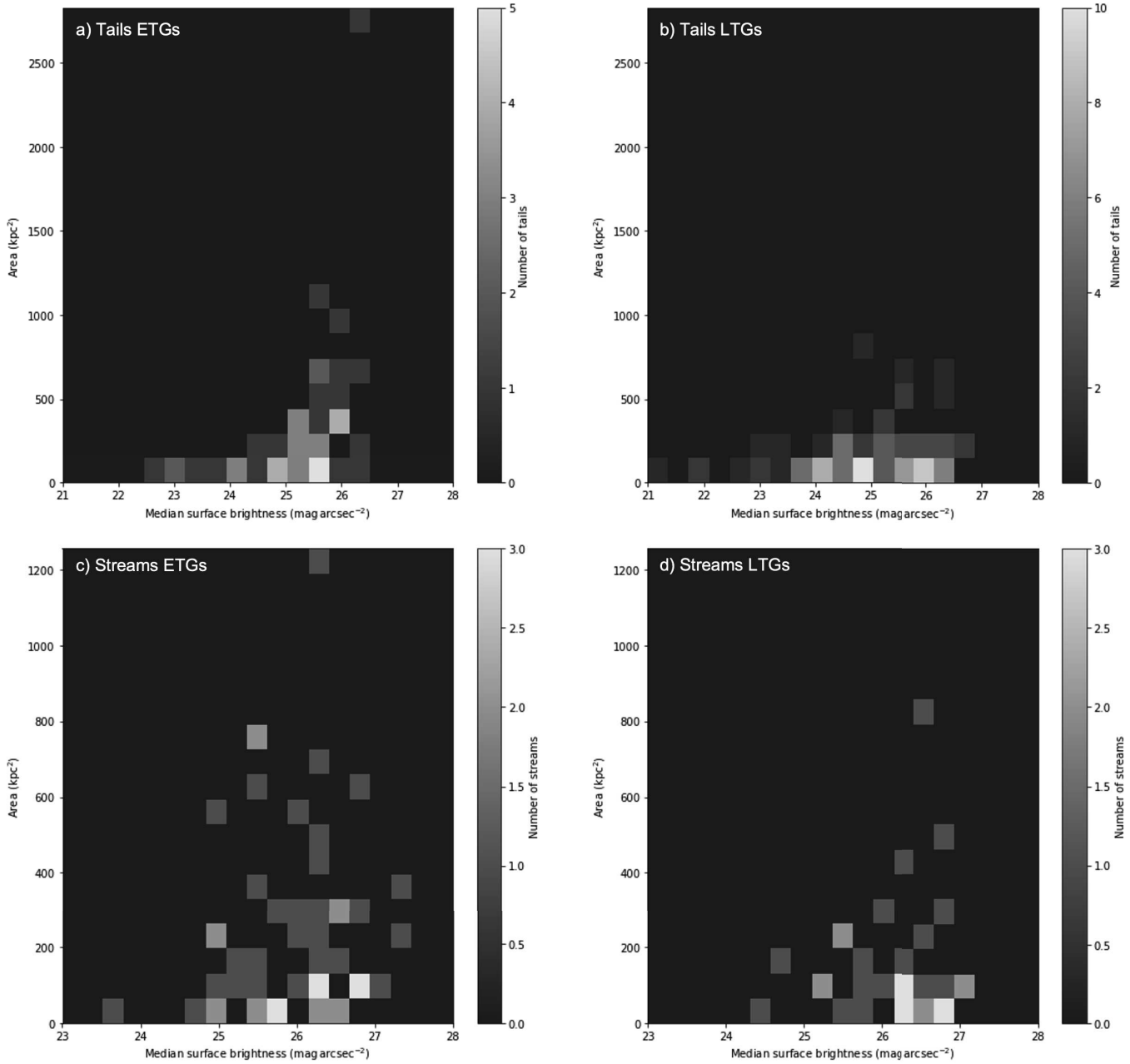

Fig. 19. 2D histograms of the area (in square kiloparsecs) of tidal tails and streams as a function of the median surface brightness (in mag arcsec^{-2}). This 2D histogram for tidal tails around ETGs is presented in *panel a* and for LTGs in *panel b*. For streams, the histogram for ETGs is visible in *panel c* and in *panel d* for LTGs.

Therefore, automated techniques will be needed to analyze larger samples of galaxies with few annotations. The annotations we made could be used to train machine learning algorithms. In fact, it has already been the case with Richards et al. (2020), who used our cirrus annotations to train a new machine learning algorithm to detect cirrus on deep images. Yet, the small number of annotations is a problem as large datasets are needed to train such algorithms. Appropriate data augmentation or annotations of cosmological simulations may be necessary to have large enough training datasets.

7. Conclusions

The detection and classification of tidal features around galaxies is essential as their characterization gives valuable information about the past assembly history of their host galaxy. In this paper, we present an online annotation tool that enables users to draw the shapes of LSB structures with precision in deep images for a large number of galaxies. We also show how we may use our annotation database to estimate some physical quantities, such as the shape, size, surface brightness and colors of tidal tails, streams, shells and stellar halos. Such values were so far not very well constrained on large samples of galaxies. They may be compared to simulations to better understand the type of mergers that took place, and more generally to constrain models of galactic evolution.

Using a customized online tool, we have manually identified, delineated and classified LSB features including contaminants around 352 nearby massive galaxies from the CFHT MATLAS and CFIS surveys. Each field has been inspected by up to

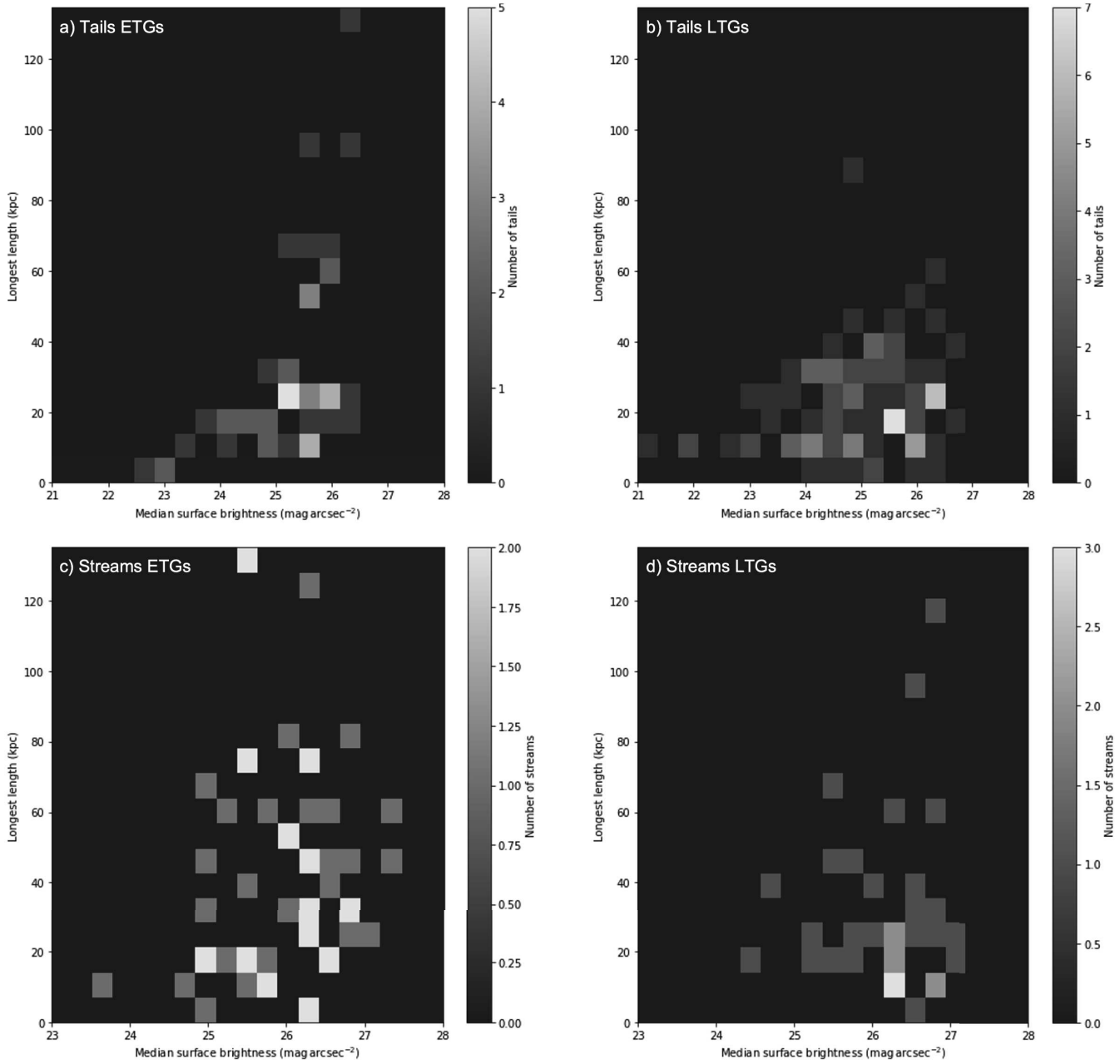

Fig. 20. 2D histograms of the length (in kiloparsecs) of tidal tails and streams as a function of the median surface brightness (in mag arcsec^{-2}). This 2D histogram for tidal tails around ETGs is presented in *panel a* and for LTGs in *panel b*. For streams, the histogram for ETGs is visible in *panel c* and in *panel d* for LTGs.

four different users. A database containing 8441 annotations was compiled (Table 2). We have developed a methodology to associate common annotations between users and link the associations with their host galaxies. With such a project, we aim at differentiating the types of tidal debris and a posteriori determine whether a classification solely based on eye inspection of deep images is reliable.

Not so surprisingly, the annotated structures show an apparent large variety of shapes and sizes when put and compared together (as visible in Figs. A.1 and A.2). There is a large overlap between tails (tidal structures emanating from the primary galaxy and made during major mergers) and streams (tidal debris from disrupted low-mass companions). However, a statistical analysis revealed real differences:

– Tidal tails are wider than streams, with a typical width of 6 kpc against 4 kpc for streams (Fig. 9). This was expected from models: since streams originate from low-mass companions with lower velocity dispersion, their width should be smaller.
– Tidal tails are brighter than streams, with a difference between their median surface brightness of almost 1 mag arcsec^{-2} (Table 6). This may be due to some age effects, with streams having a longer survival (i.e., visibility) time than tails, which are more easily identified as such when they are young.
– Tidal tails are slightly bluer than streams (a difference of 0.1 mag for the $g-r$ color) (Table 7), again likely due to an age bias.

These statistical differences comfort the approach we used to specifically distinguish tidal tails from streams.

- The annotated stellar halos have radii extending mostly between 5 and 30 kpc (Fig. 16) with a similar distribution for ETGs and LTGs in the environments probed here (outside massive clusters).
- Identified shells typically have radii lower than 40 kpc, with a few extending to 80 kpc (Fig. 14): we do not observe the very external extended shells found in some simulations, even though the comparison is not straightforward.

Other tendencies interesting for the analysis for future surveys are worth highlighting:

- The faintest structures identified as tidal tails and streams have a median surface brightness about 1 mag brighter than the nominal limiting surface brightness of the survey. This should also be kept in mind when comparing observations to simulations.
- Artifacts or foreground structures occupy a large fraction of the image pixels (more than 50% in 10 R_e boxes centered on the target galaxy). They may significantly alter our SB and color measurements. Having the possibility to remove them would also be a strong asset for future automatic classifications.

Finally, our annotations were all done manually, which took a non negligible amount of time of a group of participants. Though feasible in samples of hundreds of galaxies, thanks to our dedicated online tool, it cannot be applied to samples of thousands. Machine learning methods will be needed to automatically detect and classify tidal features in large surveys to come, and our annotation database can be used to train such algorithms.

As future prospects, the properties of tidal features with respect to the host galaxy properties will be studied in a future paper. In addition, the amount of stellar mass in the tidal features could be computed. Together with estimates of their lifetime, they can give an estimate of the speed of mass gain of galaxies caused by mergers.

Acknowledgements. The authors thank the anonymous referee for a useful report and Garreth Martin for his feedback on the paper. F.R. was supported by Science and Technology Facilities Council (STFC) grant ST/P006779/1. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France, and of "Aladin sky atlas" developed at CDS, Strasbourg Observatory, France. This work is based on data obtained as part of the Canada-France Imaging Survey, a CFHT large program of the National Research Council of Canada and the French Centre National de la Recherche Scientifique. Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA Saclay, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers (INSU) of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. This research is based in part on data collected at Subaru Telescope, which is operated by the National Astronomical Observatory of Japan. We are honored and grateful for the opportunity of observing the Universe from Maunakea, which has the cultural, historical and natural significance in Hawaii. Pan-STARRS is a project of the Institute for Astronomy of the University of Hawaii, and is supported by the NASA SSO Near Earth Observation Program under grants 80NSSC18K0971, NNX14AM74G, NNX12AR65G, NNX13AQ47G, NNX08AR22G, YORPD20_2-0014 and by the State of Hawaii.

References

Abraham, R. G., Valdes, F., Yee, H. K. C., & van den Bergh, S. 1994, ApJ, 432, 75
Abraham, R. G., van den Bergh, S., Glazebrook, K., et al. 1996, ApJS, 107, 1
Abraham, R. G., van den Bergh, S., & Nair, P. 2003, ApJ, 588, 218
Alabi, A. B., Romanowsky, A. J., Forbes, D. A., Brodie, J. P., & Okabe, N. 2020, MNRAS, 496, 3182
Antoja, T., Ramos, P., Mateu, C., et al. 2020, A&A, 635, L3
Arp, H. 1966, ApJS, 14, 1
Atkinson, A. M., Abraham, R. G., & Ferguson, A. M. N. 2013, ApJ, 765, 28
Baugh, C. M., Benson, A. J., Cole, S., Frenk, C. S., & Lacey, C. 2003, in The Mass of Galaxies at Low and High Redshift, eds. R. Bender, & A. Renzini (Berlin: Springer), 91
Bell, E. F., Naab, T., McIntosh, D. H., et al. 2006, ApJ, 640, 241
Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, ApJ, 642, L137
Bernardi, M., Shankar, F., Hyde, J. B., et al. 2010, MNRAS, 404, 2087
Bertin, E., Mellier, Y., Radovich, M., et al. 2002, ASP Conf. Ser., 281, 228
Bezanson, R., van Dokkum, P., & Franx, M. 2012, ApJ, 760, 62
Bickley, R. W., Bottrell, C., Hani, M. H., et al. 2021, MNRAS, 504, 372
Bílek, M., Jungwiert, B., Jílková, L., et al. 2013, A&A, 559, A110
Bílek, M., Duc, P.-A., Cuillandre, J.-C., et al. 2020, MNRAS, 498, 2138
Bílek, M., Fensch, J., Ebrová, I., et al. 2022, A&A, 660, A28
Bottrell, C., Hani, M. H., Teimoorinia, H., et al. 2019, MNRAS, 490, 5390
Bridge, C. R., Carlberg, R. G., & Sullivan, M. 2010, ApJ, 709, 1067
Bullock, J. S., & Johnston, K. V. 2005, ApJ, 635, 931
Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, MNRAS, 413, 813
Casteels, K. R. V., Bamford, S. P., Skibba, R. A., et al. 2013, MNRAS, 429, 1051
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, ArXiv e-prints [arXiv:1612.05560]
Cheng, T.-Y., Huertas-Company, M., Conselice, C. J., et al. 2021, MNRAS, 503, 4446
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
Conselice, C. J. 2009, MNRAS, 399, L16
Conselice, C. J., Bershady, M. A., Dickinson, M., & Papovich, C. 2003, AJ, 126, 1183
Conselice, C. J., Rajgor, S., & Myers, R. 2008, MNRAS, 386, 909
Cooper, A. P., Cole, S., Frenk, C. S., et al. 2010, MNRAS, 406, 744
Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937
Dieleman, S., Willett, K. W., & Dambre, J. 2015, MNRAS, 450, 1441
Domínguez Sánchez, H., Huertas-Company, M., Bernardi, M., Tuccillo, D., & Fischer, J. L. 2018, MNRAS, 476, 3661
Duc, P.-A. 2020, IAU Symp., 355
Duc, P.-A., Cuillandre, J.-C., Karabal, E., et al. 2015, MNRAS, 446, 120
Ebrova, I. 2013, ArXiv e-prints [arXiv:1312.1643]
Ebrová, I., Bílek, M., & Jungwiert, B. 2019, ArXiv e-prints [arXiv:1909.07393]
Ebrová, I., Bílek, M., Vudragović, A., Yıldız, M. K., & Duc, P.-A. 2021, A&A, 650, A50
Elmegreen, B. G., Kaufman, M., & Thomasson, M. 1993, ApJ, 412, 90
Ferguson, A. M. N., Irwin, M. J., Ibata, R. A., Lewis, G. F., & Tanvir, N. R. 2002, AJ, 124, 1452
Fernique, P., Allen, M. G., Boch, T., et al. 2015, A&A, 578, A114
Ferrarese, L., Côté, P., Cuillandre, J.-C., et al. 2012, ApJS, 200, 4
Ferreira, L., Conselice, C. J., Duncan, K., et al. 2020, ApJ, 895, 115
Gibbons, S. L. J., Belokurov, V., & Evans, N. W. 2014, MNRAS, 445, 3788
Gilbert, K. M., Guhathakurta, P., Beaton, R. L., et al. 2012, ApJ, 760, 76
Helmi, A., & White, S. D. M. 1999, MNRAS, 307, 495
Hendel, D., & Johnston, K. V. 2015, MNRAS, 454, 2472
Hennig, C. 2007, Comput. Stat. Data Anal., 52, 258
Hernquist, L., & Quinn, P. J. 1987, ApJ, 312, 1
Huertas-Company, M., Gravet, R., Cabrera-Vives, G., et al. 2015, ApJS, 221, 8
Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, Nature, 370, 194
Ibata, R., Irwin, M., Lewis, G., Ferguson, A. M. N., & Tanvir, N. 2001, Nature, 412, 49
Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017, ApJ, 848, 128
Iodice, E., Spavone, M., Capaccioli, M., et al. 2021, The Messenger, 183, 25
Jackson, T. M., Pasquali, A., La Barbera, F., More, S., & Grebel, E. K. 2021, ArXiv e-prints [arXiv:2102.02241]
Jarrett, T. H., Polletta, M., Fournon, I. P., et al. 2006, AJ, 131, 261
Javanmardi, B., Martinez-Delgado, D., Kroupa, P., et al. 2016, A&A, 588, A89
Ji, I., Peirani, S., & Yi, S. K. 2014, A&A, 566, A97
Johnston, K. V. 1998, ASP Conf. Ser., 136, 365
Johnston, K. V., Hernquist, L., & Bolte, M. 1996, ApJ, 465, 278
Johnston, K. V., Bullock, J. S., Sharma, S., et al. 2008, ApJ, 689, 936
Kado-Fong, E., Greene, J. E., Hendel, D., et al. 2018, ApJ, 866, 103
Karabal, E., Duc, P. A., Kuntschner, H., et al. 2017, A&A, 601, A86
Karademir, G. S., Remus, R.-S., Burkert, A., et al. 2019, MNRAS, 487, 318
Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, MNRAS, 264, 201
Kawinwanichakij, L., Papovich, C., Quadri, R. F., et al. 2014, ApJ, 792, 103
Kluge, M., Neureiter, B., Riffeser, A., et al. 2020, ApJS, 247, 43
Koch, E., & Rosolowsky, E. 2015, MNRAS, 452, 3435
Lotz, J. M., Primack, J., & Madau, P. 2004, AJ, 128, 163
Lux, H., Read, J. I., Lake, G., & Johnston, K. V. 2013, MNRAS, 436, 2386
Malhan, K., Ibata, R. A., & Martin, N. F. 2018, MNRAS, 481, 3442
Mancillas, B., Duc, P.-A., Combes, F., et al. 2019, A&A, 632, A122

Mantha, K. B., McIntosh, D. H., Ciaschi, C. P., et al. 2019, MNRAS, 486, 2643
Martin, N. F., Ibata, R. A., Rich, R. M., et al. 2014, ApJ, 787, 19
Martin, G., Kaviraj, S., Hocking, A., Read, S. C., & Geach, J. E. 2020, MNRAS, 491, 1408
Martin, G., Bazkiaei, A., Spavone, M., et al. 2022, MNRAS, 513, 1459
Martínez-Delgado, D., Gabany, R. J., Crawford, K., et al. 2010, AJ, 140, 962
Martinez-Delgado, D., Cooper, A. P., Roman, J., et al. 2021, ArXiv e-prints [arXiv:2104.06071]
Masters, K. L., Krawczyk, C., Shamsi, S., et al. 2021, MNRAS, 507, 3923
McConnachie, A. W., Irwin, M. J., Ibata, R. A., et al. 2009, Nature, 461, 66
McIntosh, D. H., Guo, Y., Hertzberg, J., et al. 2008, MNRAS, 388, 1537
Merritt, A., Pillepich, A., van Dokkum, P., et al. 2020, MNRAS, 495, 4570
Mihos, J. C. 1995, ApJ, 438, L75
Mihos, J. C., Durrell, P. R., Ferrarese, L., et al. 2015, ApJ, 809, L21
Morales, G., Martínez-Delgado, D., Grebel, E. K., et al. 2018, A&A, 614, A143
Olson, K. M., & Kwan, J. 1990, ApJ, 361, 426
Oser, L., Ostriker, J. P., Naab, T., Johansson, P. H., & Burkert, A. 2010, ApJ, 725, 2312
Pawlik, M. M., Wild, V., Walcher, C. J., et al. 2016, MNRAS, 456, 3032
Pearson, W. J., Wang, L., Trayford, J. W., Petrillo, C. E., & van der Tak, F. F. S. 2019, A&A, 626, A49
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Pillepich, A., Springel, V., Nelson, D., et al. 2018, MNRAS, 473, 4077
Pop, A.-R., Pillepich, A., Amorisco, N. C., & Hernquist, L. 2018, MNRAS, 480, 1715
Prieur, J. L. 1990, Status of Shell Galaxies, ed. R. Wielen (Berlin: Springer), 72
Quinn, P. J. 1984, ApJ, 279, 596
Richards, F., Paiement, A., Xie, X., Sola, E., & Duc, P.-A. 2020, ArXiv e-prints [arXiv:2011.11734]
Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
Spindler, A., Geach, J. E., & Smith, M. J. 2021, MNRAS, 502, 985
Spitler, L. R., Longbottom, F. D., Alvarado-Montes, J. A., et al. 2019, ArXiv e-prints [arXiv:1911.11579]
Stringer, M. J., & Benson, A. J. 2007, MNRAS, 382, 641
Tal, T., van Dokkum, P. G., Nelan, J., & Bezanson, R. 2009, AJ, 138, 1417
Tohill, C., Ferreira, L., Conselice, C. J., Bamford, S. P., & Ferrari, F. 2021, ApJ, 916, 4
Toomre, A., & Toomre, J. 1972, ApJ, 178, 623
Uzeirbegovic, E., Geach, J. E., & Kaviraj, S. 2020, MNRAS, 498, 4021
van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., et al. 2014, https://doi.org/10.7717/peerj.453
van Dokkum, P. G., Abraham, R., & Merritt, A. 2014, ApJ, 782, L24
Vega-Ferrero, J., Domínguez Sánchez, H., Bernardi, M., et al. 2021, MNRAS, 506, 1927
Venhola, A., Peletier, R., Laurikainen, E., et al. 2017, A&A, 608, A142
Walmsley, M., Ferguson, A. M. N., Mann, R. G., & Lintott, C. J. 2019, MNRAS, 483, 2968
Wen, Z. Z., Zheng, X. Z., & An, F. X. 2014, ApJ, 787, 130
Wilkinson, A., Sparks, W. B., Carter, D., & Malin, D. A. 1987, IAU Symp., 127, 465

Appendix A: Individual shapes of tidal tails and streams

In this section, we present the individual thumbnails of tidal tails (in blue) and streams (in red) identified around LTGs (Figure A.1) and ETGs (Figure A.2). The galaxies are sorted by increasing mass, from the top-left for the less massive one to the bottom-right for the most massive one.

For the streams of LTGs (Figure A.1), the mean galaxy mass per each row is respectively 1.5×10^{10}, 4.6×10^{10} and $1.1\times10^{11}\,M_{\odot}$ for the first, second and third rows. For the tails of LTGs, the mean mass per row is: 8.5×10^{9}, 1.4×10^{10}, 2.3×10^{10}, 3.2×10^{10}, 4.9×10^{10}, 6.9×10^{10}, 7.8×10^{10} and finally $1.1\times10^{11}M_{\odot}$.

Likewise, the mean galaxy mass per row for the streams of ETG (Figure A.2) are 1.3×10^{10}, 3.1×10^{10}, 6.2×10^{9}, 1.0×10^{11} and $1.8\times10^{11}M_{\odot}$. For the tails of ETGs, this progression per row is: 1.9×10^{10}, 3.3×10^{10}, 5.3×10^{10}, 1.1×10^{11} and $1.5\times10^{11}M_{\odot}$.

Appendix B: Normalized length and width of tidal tails

In this section, we present the histograms of the length and width of tidal tails normalized by the effective radius of their host galaxy, and as a function of the morphological type. One can see that the trends are similar to the ones presented in section 5.1.3, so the distribution of the width is flatter for ETGs than for LTGs. The length of the tidal tails for ETGs is slightly longer than for LTGs.

Appendix C: Precision on the annotations kept

In this appendix, we describe in more detail the selection process for tidal tails and streams seen in section 4.8. The aim is to keep the most representative features by attributing a unique identifier to the annotations, and then, for annotations sharing the same unique identifier, keeping the one with the largest area. This iterative process is illustrated in Figure C.1 for one galaxy. First, all the tidal tails and streams from all users are considered (panel *a)*. The first iteration starts (panel *b)*, and the annotations of User 2 and User 1 are compared. The red arrows outline which annotations are currently compared. If their intersection score is high enough, they are paired and share the same unique identifier (e.g., here the magenta annotations), otherwise a new unique identifier is given (e.g., here the green annotation). During the second iteration (panel *c)*, the annotations of User 3 and User 2 are compared with the same method. During the third iteration (panel *d)*, the annotations of User 3 and User 1 are compared. At the end of this step, all the annotations have been attributed a unique identifier. Finally, for the annotations sharing the same unique identifier, only the one with the largest area is kept (panel *e)*.

Appendix D: The Next Generation Virgo Cluster Survey

In addition to the images from CFIS and MATLAS, we have also annotated the LSB structures visible in the deep images from the Next Generation Virgo Cluster Survey (NGVS[20]). This CFHT Large program surveyed 104 squared degrees in the Virgo cluster in four bands, with a depth of 29 mag arcsec^{-2} in the g-band (Ferrarese et al. 2012). Unfortunately the r-band image is not available for this survey, and no direct comparison could be directly done with CFIS and MATLAS. We therefore decided not to include the NGVS annotations in our survey but to summarize our analysis in this Appendix.

Two users annotated a total of 2217 features (among which 1898 have been kept after our selection process) around 58 ETGs and 65 LTGs from the Virgo Cluster. Figure D.1 displays the median SB of tails and streams, and Figure D.2 shows the median value of the outer contour of the annotated stellar halos.

The maximal outer SB value (28.5 mag arcsec^{-2}) in the g-band is closer to the nominal depth of the survey than the maximal outer SB value in the r-band for MATLAS. One possible reason is that the g-band is less sensitive to artifacts such as ghost reflections or high background regions and therefore less contaminated, which enables to detect fainter isophotes than in the r-band. Figure D.3 displays the radius of the annotated halos for the Virgo LTGs and ETGs.

[20] NGVS, http://astrowww.phys.uvic.ca/~lff/NGVS/Home.html

Fig. A.1. Thumbnails of the streams (*top, in red*) and tidal tails (*bottom, in blue*) identified around the LTGs, plotted in boxes of 50×50 kpc. They are sorted by increasing mass of the host galaxy, starting from the top-left for the lightest LTG to the bottom-right for the most massive LTG.

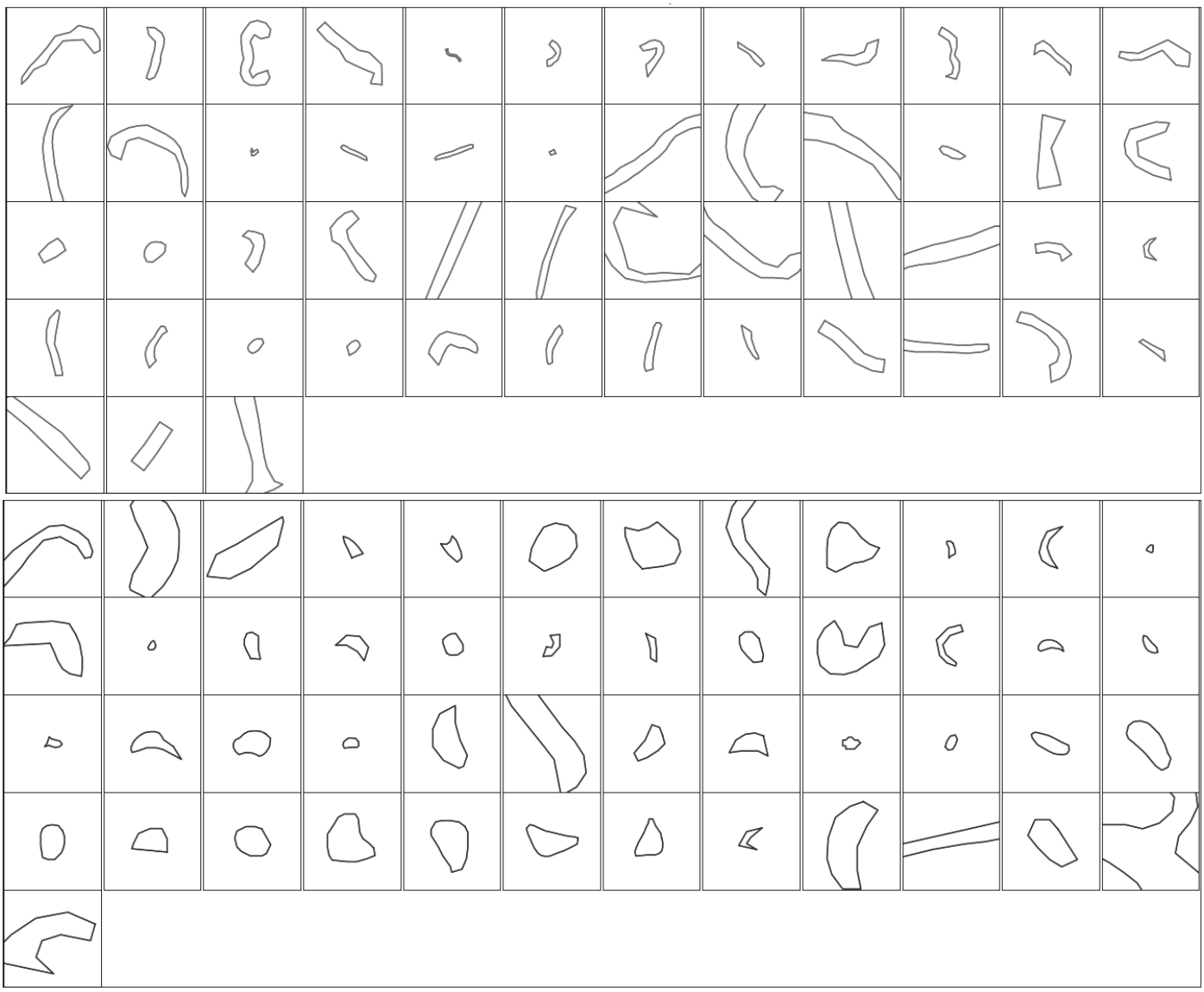

Fig. A.2. Thumbnails of the streams (*top, in red*) and tidal tails (*bottom, in blue*) identified around the ETGs, plotted in boxes of 50×50 kpc. They are sorted by increasing mass of the host galaxy, starting from the top-left for the lightest ETG to the bottom-right for the most massive ETG.

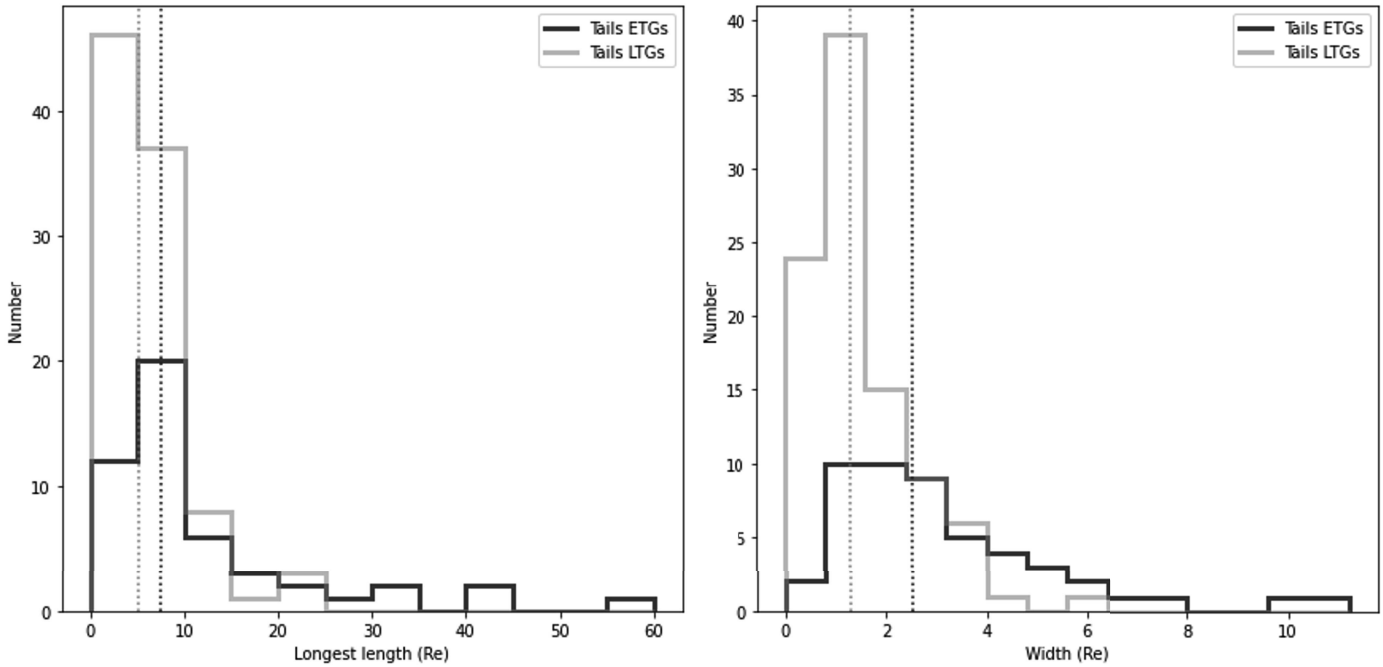

Fig. B.1. Histogram of the length (*left*) and of the width (*right*) of tidal tails normalized by the effective radius of the host galaxy, as a function of the morphological type. ETGs are represented by darker shades than LTGs. The median of each distribution is represented by the dotted lines. The histogram of the length has bins width of 5 R_e and of 0.8 R_e for the width.

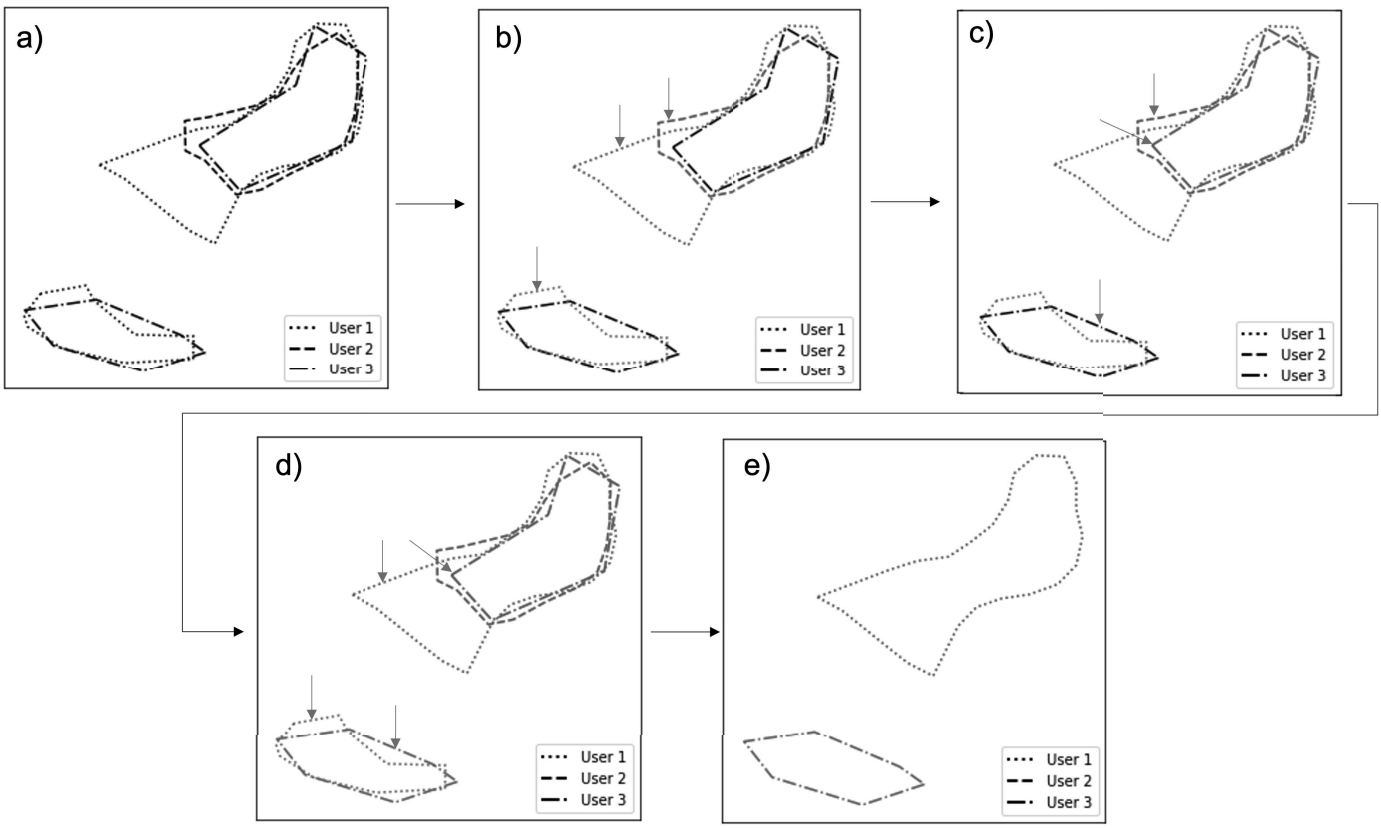

Fig. C.1. Illustration of the selection process for tidal tails and streams for one galaxy. Black annotations indicate that the unique identifier has not been attributed yet, while each color corresponds to one unique identifier. The linestyles correspond to the different users. The red arrows indicate which annotations are currently compared. *a)* Initialization: all the annotations of tidal tails and streams are considered. *b)* First iteration: the annotations of User 2 and User 1 are compared. *c)* Second iteration: comparison of the annotations of User 3 and User 2. *d)* Third iteration: comparison of the annotations of User 3 and User 1. *e)* Final step: for the annotations sharing the same unique identifier, only the one with the largest area is kept.

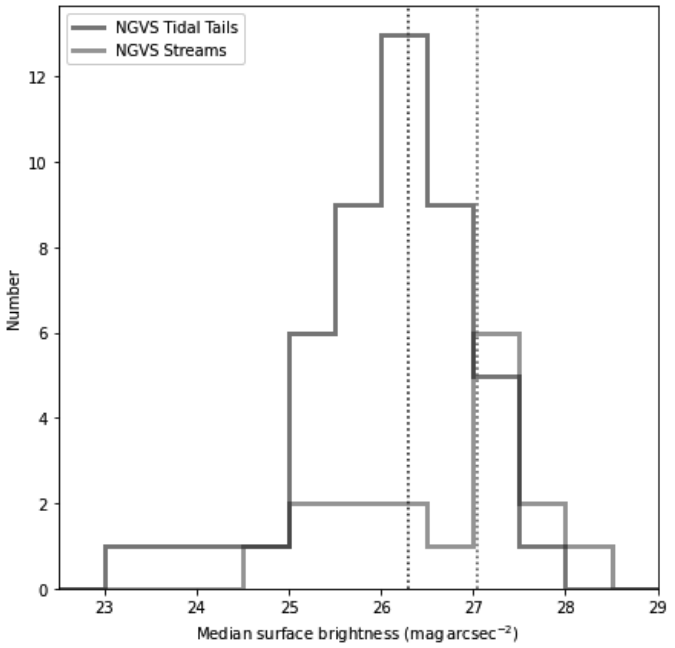

Fig. D.1. For NGVS only: histogram of the median surface brightness values in magnitudes per square arcsecond for tidal tails (in blue) and streams (in red), in bins of 0.5 mag arcsec^{-2}. The median of each distribution is represented by the dotted lines.

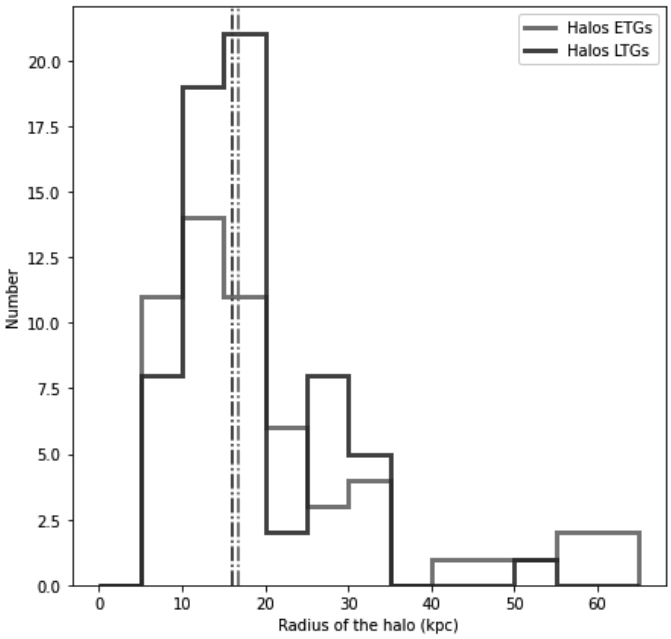

Fig. D.3. For NGVS only: histogram of the radius of the halos, in bins of 5 kpc, as a function of the morphological type. The median of each distribution is represented by the dotted line.

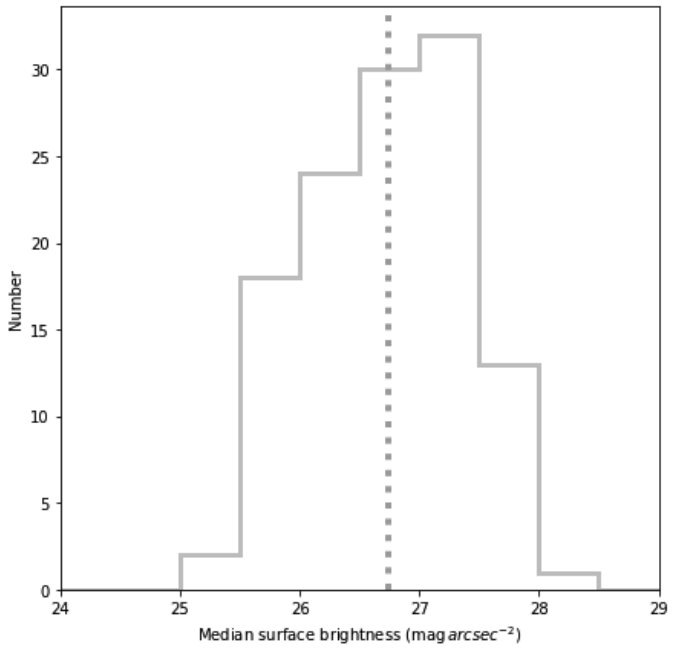

Fig. D.2. For NGVS only: histogram of the outer median surface brightness values in magnitudes per square arcsecond for halos, in bins of 0.5 mag arcsec^{-2}. The median is represented by the dotted line.

A.2 Article: Origin of the differences in rotational support among early-type galaxies: The case of galaxies outside clusters, Bílek et al (2023)

Bílek et al. (2022)

A&A 672, A27 (2023)
https://doi.org/10.1051/0004-6361/202244749

Astronomy & Astrophysics

Origin of the differences in rotational support among early-type galaxies: The case of galaxies outside clusters⋆

M. Bílek[1,2,3], P.-A. Duc[3], and E. Sola[3]

[1] LERMA, Observatoire de Paris, CNRS, PSL Univ., Sorbonne Univ., 75014 Paris, France
e-mail: michal.bilek@obspm.fr
[2] Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France
[3] Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg (ObAS), 11 rue de l'Université, Strasbourg, France

Received 15 August 2022 / Accepted 4 October 2022

ABSTRACT

Context. Early-type galaxies (ETGs) are divided into slow and fast rotators (FRs and SRs) according to the degree of ordered rotation of their stellar populations. Cosmological hydrodynamical simulations indicate that galaxies form as FRs before their rotational support decreases, usually because of mergers.
Aims. We aimed to investigate this process observationally for galaxies outside of clusters.
Methods. We made use of the fact that different merger types leave different traces that have different lifetimes. We statistically analyzed multiple characteristics of galaxies that are expected to be influenced by mergers, such as tidal features, kinematically distinct cores, and stellar ages. They were taken from the MATLAS and ATLAS^{3D} databases. Through multilinear regression we identified the quantities that, at a fixed mass and environmental density of the galaxy, significantly correlate with a measure of the ordered rotation of the galaxy, $\lambda^{\mathrm{N}}_{R_{\mathrm{e}}}$.
Results. We found a negative correlation of the rotational support with the occurrence of tidal disturbances and kinematic substructures, and a positive correlation with metallicity and metallicity gradients. For massive galaxies, the rotational support correlates negatively with the abundance of α-elements, and for the galaxies in low-density environments, it correlates negatively with the central photometric cuspiness. These and additional literature observational constraints are explained the easiest if the mergers that decreased the rotational support of ETGs were typically minor, wet, and happening at $z \approx 2$. They did not form the currently observed tidal features. The observed frequency of tidal features implies a merging rate of 0.07–0.2 per Gyr. This is insufficient to explain the observed growth of the radii of ETGs with redshift by mergers.

Key words. galaxies: elliptical and lenticular, cD – galaxies: kinematics and dynamics – galaxies: interactions – galaxies: evolution – methods: statistical – methods: observational

1. Introduction

The formation of early-type galaxies (ETGs) seems to have proceeded in two phases (Oser et al. 2010; Naab et al. 2014). The first phase, a "wet clumpy collapse" (Thomas et al. 1999), is a period of intensive in situ star formation. Galaxies were assembled through numerous gas-rich mergers or a smooth accretion of gas from cosmic filaments. This formation stage is witnessed by the chemical composition of ETGs, which is best explained by the monolithic-collapse models (Thomas et al. 1999; Kobayashi 2004; Pipino et al. 2010, 2013). These models reproduce the observed values of metallicity and abundance of α-elements, and the correlations of these quantities with the masses of the galaxies. The end of the first phase of formation of ETGs is characterized by the termination of star formation in the galaxies. It happens approximately at a redshift of two, but this limit is not universal. Observations suggest that some ETGs, particularly the most massive ones or those with the highest stellar surface densities, were already quenched as soon as at $z = 7-9$, while others continued forming stars much longer, up to $z = 0.5-2$ (McDermid et al. 2015; González Delgado et al. 2017; Forrest et al. 2020; Estrada-Carpenter et al. 2020; Carnall et al. 2020; Costantin et al. 2021; Tacchella et al. 2022).

There are many ways for a galaxy to become quenched, the relative importance of which has not yet been fully clarified. For the most massive galaxies, the mass quenching mechanism seems plausible (Kereš et al. 2005; Dekel & Birnboim 2006): when the falling interagalactic gas reaches the halos of galaxies that are massive enough, it is shock-heated so much that the cooling time exceeds the age of the Universe. The infalling clouds of cold gas are not able to reach the galaxy and get dissolved in the hot gas halo of the galaxy (Afruni et al. 2019). The filaments of cold interagalactic gas are able to penetrate the hot circumgalactic gaseous halos only before $z = 1.5-2$. Therefore, the massive galaxies are currently mostly passive. A number of mechanisms have been identified for the lighter galaxies. They include, for galaxies falling to galaxy clusters, the mechanisms of starvation (Larson et al. 1980) or gas shocks (Bitsakis et al. 2016, 2019; Ardila et al. 2018). Other spirals get quenched in centers of clusters by ram-pressure striping. Theoretical arguments suggest that even if a galaxy has a substantial gas content, the bare presence of a spheroidal component can postpone or prevent star formation; this is the so called "morphological quenching" (Martig et al. 2009, 2013). The role of the activity of galactic nuclei in the quenching of galaxies is still debated (see Harrison 2017 for a review). In any case, the chemical composition of

⋆ Table of the parameters of the galaxies is only available at the CDS via anonymous ftp to cdsarc.cds.unistra.fr (130.79.128.5) or via https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/672/A27

ETGs indicates that the dominant quenching mechanism has to cut the ETGs progenitors off from the inflow of fresh intergalactic gas (Peng et al. 2015; Trussler et al. 2020), which is called "strangulation". The same data also show signs of a removal of the cold gas from the galaxies, particularly from the less massive ones, but this is an effect of the secondary importance (Trussler et al. 2020).

The monolithic collapse phase and quenching are still not sufficient to explain the observational data. The passive galaxies at redshift around two are more compact (Daddi et al. 2005; Trujillo et al. 2006; van Dokkum et al. 2009) and disky (Toft et al. 2007; Hill et al. 2019) than the local ETGs. The simplest explanation of galaxy expansion involves a bare mass loss of the galaxy by stellar evolution and gaseous outflows caused by active galactic nuclei (Fan et al. 2008; Damjanov et al. 2009; Fan et al. 2010). However, cosmological simulations and some observational evidence (Naab et al. 2009; Trujillo et al. 2011) rather suggest that primarily minor mergers are responsible for the expansion. This is supposed to happen from about $z = 2$ until today, and it is called the second phase of ETG formation. This would explain several other observations, such as the existence of tidal features (e.g., Malin & Carter 1983; Atkinson et al. 2013; Duc et al. 2015; Bílek et al. 2020), the properties of globular cluster systems (Côté et al. 1998), and the flattening of metallicity gradients at the outskirts of ETGs (Zibetti et al. 2020; Oyarzún et al. 2019). Finally, it has been observed that star formation can be renewed in ETGs after they have been quenched (Thomas et al. 2010; Serra et al. 2014; Gavazzi et al. 2018; Mancini et al. 2019; Yıldız et al. 2020). The new stars then form in a disk.

Some ETGs were likely formed outside of the two-phase scenario by binary mergers of spirals (Toomre 1977). While this mechanism was popular in the past, it turns out that ETGs, particularly the massive ones, are formed only rarely through this mechanism. (Thomas et al. 1999; Naab & Ostriker 2009; Harris 2001; Krajnović et al. 2011).

The unsolved problems of ETG formation in Λ-cold-dark-matter (ΛCDM) cosmological models are generally related to a too gradual formation of ETGs: the main signs are the frequent observations of very massive quenched galaxies at very large redshifts (Hill et al. 2017; Schreiber et al. 2018; Merlin et al. 2019; Faisst et al. 2019; Forrest et al. 2020; Stevans et al. 2021; Carnall et al. 2023), and the problems with reproducing the values and correlations of the abundance of the α-elements (Thomas et al. 1999, 2002, 2003; Nagashima et al. 2005; Pipino et al. 2009; De Lucia et al. 2017; Okamoto et al. 2017; Vincenzo et al. 2018).

Early-type galaxies seem to form bimodal statistical distributions in the space of their properties. This is why they are divided into slow and fast rotators (SRs and FRs, see Cappellari 2016 for a review). Fast rotators show a regular rotational pattern in the kinematic maps of their inner stellar populations (Emsellem et al. 2011). Their kinematic axes are aligned well with the minor photometric axes of the galaxies. Fast rotators turn out to be intrinsically oblate axisymmetric ellipsoids (Weijmans et al. 2014; Foster et al. 2017). Kinematic maps of SRs instead either do not show any rotation, or show complex features, such as kinematically distinct components (Emsellem et al. 2007, 2011). Slow rotators are rounder than FRs and are weakly triaxial (Weijmans et al. 2014; Foster et al. 2017; Li et al. 2018b), and their kinematic and photometric axes do not align well (Krajnović et al. 2011; Ene et al. 2018). Slow rotators are also typically more massive than FRs (Emsellem et al. 2011). The vast majority of ETGs are FRs, but SRs prevail among the most massive ETGs (Emsellem et al. 2011). This is because the degree of rotation is probably primarily a function of the mass of the galaxy (Veale et al. 2017; Brough et al. 2017; Greene et al. 2017) and the most massive galaxies are located in the centers of their groups or clusters, even though some works suggest that the rotator type is also influenced independently by the density of the environment of the galaxy (Graham et al. 2019a,b; van de Sande et al. 2021a). While the primary motivation for introducing the FRs and SRs was the distinct appearance of their kinematic maps, many works rely on quantitative definitions, which are supposed to be roughly equivalent to the morphological definition. Here we build on the widely used parametric separation criterion by Emsellem et al. (2011): ETGs having their $\lambda/\sqrt{\epsilon}$ below the value of 0.31 are classified as SRs while the rest are classified as FRs. Here ϵ stands for the apparent ellipticity of the galaxy within the half-light radius, and the parameter λ quantifies the relative importance of the ordered rotation and velocity dispersion in supporting the galaxy.

Simulations provide us with insights into the formation of FRs and SRs. The 44 high-resolution zoom-in simulations by Naab et al. (2014) showed that there are many ways to form an SR or an FR in terms of the number of mergers, their mass ratios, and gas fractions. The slowest rotators, however, were formed by many minor mergers. The early work by Naab et al. (2014) had the disadvantage that it could not indicate how important the individual formation channels of SRs and FRs are. Ideally, this would be shown by cosmological hydrodynamical simulations. Such a route was followed by Penoyre et al. (2017), who inspected the Illustris simulation. They found that at very high redshifts, all galaxies are FRs, with a high rotational support. The progenitors of the current SRs and FRs are still nearly indistinguishable at $z = 1$ in terms of the distributions of stellar masses and of the λ parameter. The rotation support was found to decrease after major mergers but very massive galaxies decreased it even in periods without mergers. Lagos (2018) undertook a similar approach with the EAGLE (Schaye et al. 2015) and HYDRANGEA (Bahé et al. 2017) simulations. They again found a link between mergers and a decrease in the degree of rotational support. They found that galaxies that experienced dry mergers in the simulations, either minor or major, usually ended up with a lower degree of rotation than galaxies that underwent wet or no mergers. Some mergers can increase the rotational support, but such mergers are in the minority. A small fraction of SRs did not experience any mergers and such galaxies inhabited low-spin halos. Nevertheless, in most cases, the transformation from FRs to SRs happens through mergers. The formation of SRs from FRs in the EAGLE simulation was confirmed by Lagos (2022) and in the MAGNETICUM simulation by Schulze et al. (2018).

Cosmological hydrodynamic simulations still have relatively low resolution. This might be the reason why the stellar kinematics of the simulated galaxies do not fully reproduce the reality. For example, Ebrová et al. (2020) noted that the kinematically distinct components in Illustris are too large, Lagos (2022) pointed out the nonrealistic radial profiles of velocity dispersion of the galaxies in EAGLE, and Schulze et al. (2018) reported a population of overly flattened SRs in MAGNETICUM.

In this paper we investigate the details of the decrease in the rotational support of ETGs observationally for galaxies outside of galaxy clusters. In order to quantify how much the transformation has progressed in a given galaxy, we heuristically exploit the parameter $\lambda^{\mathrm{N}}_{R_\mathrm{e}} = \lambda/\sqrt{\epsilon}$ (Emsellem et al. 2011), which we call the rotational support. We note that more recent works use

more elaborate criteria than $\lambda^{\rm N}_{R_{\rm e}} <$ or > 0.31 to classify galaxies into SRs and FRs. The newer criteria are supposed to capture the bimodality of ETGs better. For example, Cappellari (2016) define SRs as satisfying the conditions $\lambda < 0.08 + \epsilon/4$ and $\epsilon < 0.4$. Here we build on the older SR and FR separation criterion that uses the $\lambda^{\rm N}_{R_{\rm e}}$ parameter because it seems obvious how to use it to continuously quantify the stage of the kinematical transformation. It is possible that future studies will find a more suitable quantification of the transformation stage than the $\lambda^{\rm N}_{R_{\rm e}}$ parameter.

The main idea of this paper is the following. We assume, inspired by the simulations, that all ETGs form initially with a relatively high and a relatively universal value of rotational support, and then the rotational support is decreased by mergers. Next, we assume that if the mergers have been very important in the evolution of the galaxy until now, the galaxy would, in most cases, be observed to have a low value of the rotational support at the current cosmic epoch. The mergers change various parameters of the galaxy. The traces of the mergers depend on whether the merger is gas rich or gas poor, and on the mass ratio of the merging galaxies. In addition, the different merger signs have different lifetimes. We make use of that and investigate the correlations of various merger-sensitive indicators with the rotational support in order to study the mergers that are responsible for the decrease in the rotational support. In order to reduce confounding effects, we studied the correlations at a fixed mass and environmental density through mulitilinear regression.

The paper is organized as follows. Our data sources are presented in Sect. 2. In Sect. 3, we list the merger-sensitive parameters that we exploit to derive the characteristics of the mergers that caused the decrease in the rotational support of ETGs. For each parameter, we explain how it is expected to change after different types of mergers. Section 4 is devoted to the estimation of how long after a merger a galaxy appears morphologically disturbed. The main method of the paper is described in Sect. 5, where we explain how we determined whether and how different merger sensitive parameters correlate with the $\lambda^{\rm N}_{R_{\rm e}}$ parameter for galaxies at a fixed mass and environmental density, and present the results. These results and additional literature findings are then used in Sect. 6 to deduce what types of mergers were typically decreasing the rotational support of galaxies and when they happened. Given that we a get different time of the kinematic transformation than predicted by cosmological simulations, we verify our conclusion by independent methods in Sect. 7. As a by-product, we estimate in Sect. 7.2 the current merging rate of our galaxies from the incidence of tidal disturbances. In Sect. 8, we show that many observational findings about the occurrence of FRs and SRs can be explained as simple consequences of the fact that galaxy mergers usually decrease the rotational support of galaxies. We synthesize our findings in Sect. 9, where we propose how typical FRs, SRs, and spiral galaxies form. The paper is summarized in Sect. 10. In Appendix A we provide the correlations of the various merger-sensitive parameters with galaxy mass, environment density, and rotational support. We also compare the merger-sensitive parameters in FRs and SRs.

For the conversion between look-back time and redshift, we used Ned Wright's cosmology calculator[1] (Wright 2006) with the cosmological parameters $H_0 = 69.6\,{\rm km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\rm M} = 0.286$ and $\Omega_{\rm vac} = 0.714$ (Bennett et al. 2014).

[1] http://www.astro.ucla.edu/~wright/CosmoCalc.html

2. Data

Our work is based on the public data provided by the ATLAS$^{\rm 3D}$ (Cappellari et al. 2011a)[2] and MATLAS (Duc et al. 2015; Bílek et al. 2020)[3] surveys. The ATLAS$^{\rm 3D}$ survey targets nearby (<42 Mpc) massive ($M_K < -21.5$ mag) ETGs (galaxies lacking spiral arms) and is volume complete. The survey strives to collect all possible information about its targets. All of the data we used in this paper was taken from ATLAS$^{\rm 3D}$, except for those on the photometric irregularities (tidal features, irregular outer isophotes) and dust lanes; that information was taken from the catalog of Bílek et al. (2020) extracted from the MATLAS survey. The MATLAS survey took very deep ($\sim$28.5 mag arcsec^{-2}) wide-field ($1 \times 1^\circ$) optical (the u^*, g', r', and i' bands) images of all ATLAS$^{\rm 3D}$ ETGs, except for those in the Virgo Cluster, with the 3.6 m Canada-France-Hawaii Telescope, employing observing and data processing strategies optimized for detecting large-scale low-surface-brightness structures. We used the MATLAS sample (Bílek et al. 2020) where we excluded the two galaxies without mass measurement (PGC 056772 and PGC 061468). In total, the sample considered here contains 175 objects. The list is available in electronic form at the CDS. Importantly, the MATLAS survey avoided the cluster environments, and therefore our results pertain only to galaxies in low- to medium-density environments such as galaxy groups.

In this paper, we characterize the mass of a galaxy by $M_{\rm JAM}$, adopted from Cappellari et al. (2013). It is defined as $M_{\rm JAM} = (M/L)_{\rm e} L$, where L is the total luminosity of the galaxy and $(M/L)_{\rm e}$ is the effective dynamical mass-to-light ratio of the galaxy within the galactocentric distance of one $R_{\rm e}$ derived using Jeans Anisotropic Modeling (Cappellari 2008). Cappellari et al. (2013) showed that the median fraction of dark matter mass within a sphere of radius $R_{\rm e}$ is 13%; thus the stellar mass of the galaxy can be estimated as $0.87\,M_{\rm JAM}$, or, in the logarithmic scale, the stellar mass is lower by 0.06 dex. The 15th, 50th, and 85th percentiles of the distribution of $\log(M_{\rm JAM}/M_\odot)$ in our sample are 10.2, 10.6, and 11.0, respectively.

The environmental density of a galaxy in this paper is quantified by the ρ_{10} parameter of Cappellari et al. (2011b), defined as the mean density of galaxies inside a sphere that is centered on the target galaxy and contains the ten nearest neighbors. The 15th, 50th, and 85th, percentiles of the distribution of $\log(\rho_{10}/{\rm Mpc^{-3}})$ in our sample are -2.2, -1.6, and -1.0, respectively. The highest ρ_{10} reaches $1.66\,{\rm Mpc^{-3}}$ for NGC 4623.

In our work, we quantify the rotational support through the parameter

$$\lambda^{\rm N}_{R_{\rm e}} = \frac{\lambda}{\sqrt{\epsilon}}, \tag{1}$$

introduced by Emsellem et al. (2011). Here λ quantifies the degree of ordered rotation of the galaxy, as it is observed projected on the sky, and ϵ the apparent ellipticity, see Emsellem et al. (2011) for details. Both of the quantities are measured within one $R_{\rm e}$ from the center of the galaxy. The values were taken from Emsellem et al. (2011). The 15th, 50th, and 85th percentiles of the distribution of $\lambda^{\rm N}_{R_{\rm e}}$ in our sample are 0.32, 0.77, and 1.0, respectively. The distribution of our sample in the space of galaxy mass – environmental density – rotational support is shown in Fig. B.1. The used values of $M_{\rm JAM}$, ρ_{10} and $\lambda^{\rm N}_{R_{\rm e}}$ are given in the table available at the CDS.

[2] http://www-astro.physics.ox.ac.uk/atlas3d/

[3] http://obas-matlas.u-strasbg.fr

3. Traces of past mergers

In this section, we describe the merger-sensitive parameters that we exploit later in the paper. It shall be kept in mind, as we discuss in Sect. 6, that none of them are totally reliable – each can be influenced by other processes than mergers. Figure 1 shows the distribution of these merger indicators in our galaxy sample in the space of rotational support versus mass of the galaxy. Figure B.2 shows the same but the mass is replaced by the environmental density. In Appendix A, we inspect the correlations of mass, environmental density and the rotational support with the different merger-sensitive parameters. In that appendix we also compare the values of these parameters between FRs and SRs. The used values of the merger-sensitive parameters listed in the following subsections are given in the table available at the CDS. It summarizes also the corresponding data sources.

3.1. Tidal disturbances

Galaxy mergers leave morphological irregularities in the images of the galaxies. They can remain observable for several gigayears before the system relaxes or before the existing tidal features are destroyed by another merger. The survivability depends, for example, on the mass ratio of the merged galaxies, their orbital configuration, their morphological types, and the density of the environment. An attempt to quantify the lifetimes of tidal features was made by Mancillas et al. (2019), and we discuss the issue further in Sect. 4. A catalog of various types of tidal disturbances was presented in Bílek et al. (2020). It was constructed by a visual inspection of deep optical images of the galaxies in the MATLAS survey. We consider the following types of tidal disturbances.

Shells. These are arc-like sharp photometric jumps whose center of curvature most often coincides with the core of the galaxy. The conditions of their formation are still under debate but generally radial mergers are preferred (Hendel & Johnston 2015; Amorisco 2015), but see, e.g., the simulation by Ebrová et al. (2020). While in older works they were attributed to minor mergers (see the review in Bílek 2016), recent studies have suggested that shells often form also in intermediate or even major mergers (Pop et al. 2018; Kado-Fong et al. 2018).

Streams. They appear as thin, long structures, sometimes going through the center of the host galaxy, but most of time wrapping around it. Generally, they can be attributed to non-radial minor mergers (Hendel & Johnston 2015; Amorisco 2015) while even major mergers involving disks can produce similar features (Wang et al. 2012). In that case, typically there are several streams, or they are accompanied by other tidal disturbances.

Tails. Tails are morphologically similar to streams but they have a higher thickness (Sola et al. 2022), up to the size of the host galaxy, and they are always attached to it. The presence of a tail or tails in a galaxy indicates that the galaxy is just in the process of disruption by a massive neighbor, or that the galaxy is a remnant a of past major merger.

Disturbed outer isophotes. In a relaxed ETG, outer isophotes are axially symmetric about the minor and major photometric axes. An interaction breaks this symmetry, making the isophotes lopsided or irregular. In many mergers, the irregularities would take the form of some of the tidal features above. However, once the tidal features become too old, they lose their distinct form and they appear just as irregular isophotes. Disturbed outer isophotes can also signify an ongoing or distant interaction.

Shells, streams, and tails are collectively called the "tidal features". In this paper we call the union of tidal features and disturbed outer isophotes "tidal disturbances". To quantify whether a given tidal disturbance is present in the galaxy, we adopted the rating from Bílek et al. (2020). They list, for each of the tidal disturbances, its rating expressing the visual prominence of the given feature in the galaxy. A value of zero indicates that the feature is not present, a value of one that it is possibly present, and a value of two that it is certainly present in the galaxy. We also took the frequency of each type of tidal feature in a galaxy, which indicates how many tidal features of the given type are present in the galaxy, from Bílek et al. (2020). Several types of tidal disturbances can be present in a galaxy at the same time. Whenever we wanted to quantify whether any tidal disturbance is present in the galaxy, we took the maximum of the ratings of the individual types of tidal disturbances.

3.2. Dust lanes

Dust is common in spiral galaxies but less so in ETGs. This suggests that dust gets into ETGs by the accretion of spirals. Indeed, ETGs with tidal disturbances often show prominent dust lanes. We took the information about the occurrence of dust lanes from the catalog by Bílek et al. (2020), who rate the presence of dust lanes similarly to tidal disturbances: ranging from zero for no dust lanes to two for prominent dust lanes.

3.3. Kinematic substructures

The majority of ETGs shows regular ordered rotation around the photometric minor axis (Krajnović et al. 2011; Emsellem et al. 2011). The are, however, exceptions[4]: "counter-rotating cores" (CRCs) rotate around the same axis as the rest of the galaxy below $\sim 1R_e$ but in the opposite sense; "Kinematically distinct cores" (KDCs) do not share the rotation axis with the rest of the galaxy; and finally "kinematic twists" (KTs) are characterized by a gradual change of the kinematic position angle through the volume of the galaxy below $\sim 1R_e$. The so-called double σ (or 2σ) galaxies show, in their maps of velocity dispersion, two peaks that lie on the major photometric axis of the galaxy and their separation is greater than half of the effective radius of the galaxy. The peaks likely arise because there are two counter-rotating stellar disks in these galaxies (Krajnović et al. 2011). Together, we call KDCs, CRCs, KTs, and the 2σ features "kinematic substructures". We quantified the presence of a given type of kinematic substructures by a value of one, and its absence by zero. Kinematic substructures can form in mergers, even if alternative formation channels exist (Ebrová et al. 2021; Young et al. 2020).

3.4. Effective radius

At the redshifts over about 1–2, passive galaxies of any fixed mass are around five times smaller than the local ETGs (Daddi et al. 2005; Trujillo et al. 2006; van Dokkum et al. 2009). The growth of the effective radius is usually attributed to galaxy mergers: the initial potential energy of the two separated galaxies is transferred into the internal energy of the merger remnant. It was shown that a larger expansion is achieved if a given mass is accreted through minor mergers than through

[4] We follow here the classification by Krajnović et al. (2011).

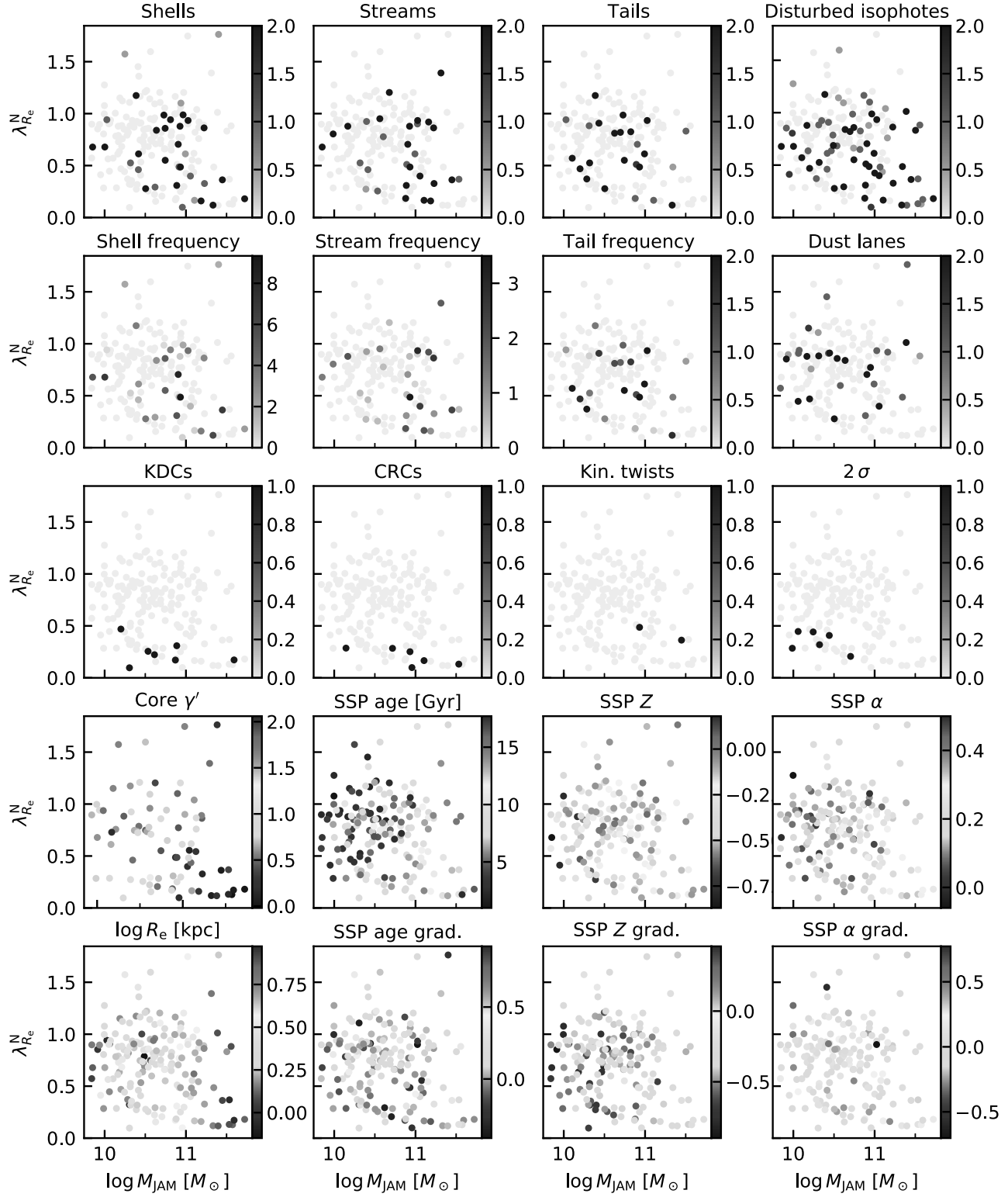

Fig. 1. Investigated merger-sensitive parameters as functions of the rotational support (vertical axis in each tile) and the $M_{\rm JAM}$ mass of the galaxy (horizontal axis in each tile). Each point represents one galaxy. The colors of the points indicate the values of the quantities stated in the tiles of the figures, coded according to the color bars given to the right of each tile. For shells, streams, tails, disturbed isophotes, and dust lanes, the values of zero, one, or two indicate that the feature is not present, is likely present, or is certainly present, respectively. The frequency of shells, streams, and tails indicates the number of these features in the galaxy. For KDCs, CRCs, KTs, and the 2σ features, the values of zero or one mean that the kinematic substructure is not or is present in the galaxy, respectively.

major mergers (Naab et al. 2009). The number of mergers that caused the expansion is expected to be relatively low. For example, Trujillo et al. (2011) calculated that the growth of radius since $z \approx 0.8$ can be achieved by 3 ± 1 mergers of the mass ratio of 1:3 or by 8 ± 2 mergers of the mass ratio of 1:10. Some local ETGs might even have coincidentally avoided any

substantial mergers since $z = 1-2$, as their properties suggest (Martín-Navarro et al. 2018; Beasley et al. 2018; Yıldırım et al. 2017). Therefore, if SRs and FRs had different merger histories, we expect a difference in their effective radii at a fixed stellar mass. For the effective radii, we took the R_e^{maj} parameters from Cappellari et al. (2013). They were obtained as the major axes of multi-Gaussian fits of the galaxies. They were published in the angular scale. We converted them into kiloparsecs assuming the distances of the galaxies given in Cappellari et al. (2011a).

3.5. Inner photometric profile

The photometric profiles of ETGs are generally described well by a Sérsic profile. However, high-resolution images, mainly those taken by the *Hubble* Space Telescope, reveal that there are deviations from it in the centers of ETGs (Lauer et al. 1995, see Graham 2013 for a review). Some inner photometric profiles follow a power law, creating an excess of light with respect to the Sérsic profile fitted to a wider radial range. These are called the cusps. On the contrary, the derivative of the photometric profile can nearly reach zero in galaxy centers, such that there is deficit of light with respect to a Sérsic profile. Such features are called the cores. Cored profiles are found typically in bright galaxies but, interestingly, stellar densities are higher in the low-luminosity galaxies with cuspy profiles (Faber et al. 1997). The formation of cores is usually explained in the following way. After a merger of two galaxies, there are two supermassive black holes in the merger remnant galaxy. They experience dynamical friction and sink toward the center of the galaxy. Their potential energy transforms into the kinetic energy of stars, which are subsequently ejected from the center of the galaxy, so that a central core forms (e.g., Faber et al. 1997; Milosavljević et al. 2002). However, gas-rich mergers tend to form new stars in the center, so that a remnant of a gas-rich merger can eventually become more cuspy than the progenitor galaxies (e.g., Hopkins et al. 2009b; Haan et al. 2013). We use here the parameter γ' tabulated by Krajnović et al. (2020) to quantify the shape of the inner photometric profile. It comes from fits by the Nuker law. Higher values of γ' indicate more cuspy profiles.

3.6. Chemical composition

The stellar metallicity of galaxies, $[Z/H]$, is known to increase with the mass of the galaxy and the density of its environment (Peng et al. 2010; Maiolino & Mannucci 2019). In addition, star-forming galaxies tend to have lower metallicities than passive galaxies of the same mass (e.g., Thomas et al. 2010; Peng et al. 2015; Maiolino & Mannucci 2019). Similarly, the abundance of α-elements, $[\alpha/\mathrm{Fe}]$, is generally higher for more massive ETGs (Thomas et al. 2010; McDermid et al. 2015) while the α-abundance is lower in spirals (Proctor & Sansom 2002; Peletier et al. 2007; Ganda et al. 2007; Scott et al. 2017; Parikh et al. 2021). The chemical composition of galaxies can thus be expected to depend on their merger history. A high value of α-enhancement signifies rapid star formation (e.g., Thomas et al. 1999). The metallicity of ETGs grows toward their centers (e.g., Spolaor et al. 2009; Kuntschner et al. 2010; Scott et al. 2013; Li et al. 2018a), while the α-elements show nearly no gradient (Rowlands et al. 2018).

Dry major mergers tend to flatten the preexisting metallicity gradients simply because of the mixing of stellar populations (Di Matteo et al. 2009b). On the other hand, minor mergers tend to steepen metallicity gradients, since minor mergers tend to deposit the low-metallicity material of the small galaxies at high radii (Amorisco 2017; Karademir et al. 2019). Simulations indicate that gas-rich major mergers decrease both the central metallicity and the metallicity gradient, since tidal torques bring the metal-poor gas in galaxy centers (Kobayashi 2004; Montuori et al. 2010; Perez et al. 2011; Navarro-González et al. 2013; Hirschmann et al. 2015; Taylor & Kobayashi 2017, but see Hopkins et al. 2009a), as also indicated observationally (Rupke et al. 2010; Maiolino & Mannucci 2019). This means that the central starburst is not powerful enough to increase the central metallicity. On the other hand, the central starburst is able to increase the central enrichment by α-elements, which is hinted at both by simulations (Perez et al. 2011) and observations (Pierce et al. 2005). The values of metallicity and α-abundance we used here are the measurements by McDermid et al. (2015) within one effective radius, based on single-stellar-population (SSP) models. We checked that our results did not change if we used the star-formation-history ages published in the same work, which are based on full spectral fitting. Stellar population gradients were taken from Krajnović et al. (2020) and again are derived from SSP models. A higher value of the metallicity gradient means a less negative, and therefore a flatter, gradient. Chemical properties derived from SSP models are known to preferentially trace the properties of the old stellar populations present in the galaxy (Serra & Trager 2007).

Early-type galaxies can also form stars in situ during the second phase of assembly. There are indeed observational indications that star formation can be renewed in ETGs after some period of passivity (Thomas et al. 2010; Bílek et al. 2020; Yıldız et al. 2020). Simulations show that the stars formed in situ can constitute a substantial fraction of the stars of the resulting ETG too (e.g., Penoyre et al. 2017; Lagos 2018). One can argue that these in situ stars are formed by freshly accreted gas from the intergalactic medium, and therefore the new stars decrease the total metallicity of the galaxy. However, here we rely on the hypothesis that the stars formed in the second phase of ETG assembly do not cause a decrease in the metallicity of the galaxy. The first reason for this is the fact that the rejuvenated ETGs in the sample of Thomas et al. (2010), which are about 2 Gyr old, show the same or rather somewhat higher metallicities than the standard, old ETGs of the same velocity dispersion. Second, according to the calculations of Trussler et al. (2020), a typical star-forming galaxy transforms into a typical passive galaxy with the typical metallicity relatively quickly, on the characteristic time scale of around 2 Gyr. An ETG that experienced a temporal burst of star formation would likely turn back to the standard mass-metallicity relation even faster.

3.7. Stellar ages

Ages of galaxies generally grow toward higher masses, earlier morphologies, and denser environments (e.g., Thomas et al. 2010; McDermid et al. 2015; González Delgado et al. 2017). Stellar populations inside of galaxies generally become older toward the centers of galaxies (Spolaor et al. 2009; Kuntschner et al. 2010; Li et al. 2018a). Dry mergers lead to population mixing, while wet mergers can give rise to completely new stars in the merger remnant. Both dry and wet major mergers tend to flatten age gradients, while minor mergers, which deposit material far from the center, can increase them.

Here we consider the stellar ages, expressed in gigayears, measured by the SSP method from line indices. The

measurements were taken from McDermid et al. (2015) and pertain to the stellar population in one effective radius. The ages are defined as the time passed since half of the stellar population was formed. The SSP age is know to be biased toward the younger stellar populations in the galaxy (Serra & Trager 2007). Uncertainties in age measurements grow quickly with the age. The age gradients we use here were taken from Krajnović et al. (2020).

4. Survivability of tidal disturbances

Estimating the distribution of the lifetimes of tidal disturbances is a difficult task. Once a tidal disturbance is induced by a merger, it will have a morphology of a tidal feature, that is shells, streams, tails, or their combination. Mancillas et al. (2019) did the first step toward the estimation of the lifetimes of different types of tidal features by inspecting one zoom-in hydrodynamic simulation of a galaxy in a ΛCDM Universe. With this approach, they included the destruction of tidal features by subsequent galaxy interaction. Of specific interest for this work, they investigated whether the features would be detected by the MATLAS survey. The tails were found to have the shortest lifetimes, namely below 1 Gyr. The longest lifetimes were for the shells, which lived for up to 4 Gyr. As the tidal disturbance evolves, it will eventually lose its characteristic morphology of a shell, tail, or a stream. After that it will be observable as disturbed outer isophotes. Mancillas et al. (2019) did not discuss the lifetimes of disturbed outer isophotes. In our paper, we assume that it will take at least 4 Gyr before any tidal disturbance disappears completely, regardless of its initial morphology.

We make an analytic upper estimate of the survivability of tidal disturbances in our galaxies. We ignore the fact that tidal features can be destroyed by new galaxy interactions (Mancillas et al. 2019). Then tidal features disappear because of the phase mixing mechanism (e.g., Mo et al. 2010), and therefore the characteristic time scale of the dissolving of the tidal disturbances is the orbital period of a star at the position of the tidal disturbance.

We measured the sizes of the irregularities in the galaxies with disturbed isophotes in MATLAS images and found that they are usually ten times larger than the effective radii. Therefore, we approximated the radii of the outermost isophotes as $R_{\rm OI} = 10R_{\rm e}$. We estimated the orbital period at $R_{\rm OI}$ as

$$T_{\rm orb} = \frac{2\pi R_{\rm OI}}{\sqrt{-a\,R_{\rm OI}}}, \tag{2}$$

where a stands for the gravitational acceleration at $R_{\rm OI}$. That was determined from the empirical radial acceleration relation (Lelli et al. 2017, see also Milgrom 1983) and the total stellar mass of the galaxy, which is $0.87M_{\rm JAM}$. The results of Bílek et al. (2019a) indicate that the gravitational acceleration in ETGs might be even stronger than expected from the radial acceleration relation, in line with our effort to give an upper estimate for the survivability time. We assumed that tidal disturbances live at most for $10T_{\rm orb}$ after the merger. We then obtained that mean maximal survival time for our galaxy sample is 9 ± 3 Gyr, nearly independently of galaxy mass. Therefore, we adopted 9 Gyr as the maximum lifetime of tidal disturbances. This suggests that most of the tidal disturbances must have been formed before $z = 1.4$, which means during the second phase of ETG assembly.

5. Correlations of merger-sensitive parameters with $\lambda^{\rm N}_{R_{\rm e}}$ at a fixed mass and environmental density

Here we describe our methodology. As we explained in Sect. 1, we assume that galaxies gradually lose rotational support primarily by mergers. At the present cosmic epoch, different galaxies are observed in different stages of the transformation. Some of the galaxies would still be classified as FRs according to their value of $\lambda^{\rm N}_{R_{\rm e}}$, while others would already be classified as SRs.

Later in this paper, we deduce when these mergers were primarily happening and what type they typically were (i.e., wet or dry, and minor or major). This is found by determining how the merger-sensitive parameters change with the rotational support. However, it is necessary to take into account that the merger-sensitive parameters would correlate with the rotational support even if mergers did not influence the rotational support at all, for the reasons described below.

The typical number of mergers that a galaxy experiences is expected to be an increasing function of the current mass of the galaxy, because more massive galaxies have stronger gravitational fields and are more extended, and therefore dynamical friction can be effective up to larger distances. Therefore, their effective cross-section for mergers is expected to be larger than for galaxies of a lower mass. Next, the typical number of mergers that a galaxy experiences is expected to increase with the density of its environment because there are more galaxies in the vicinity available for merging[5]. The typical amount of stellar mass formed in situ is expected to be an increasing function of the current galaxy mass, again because a more massive galaxy is able to attract the accreted gas from a larger distance than a less massive galaxy. The typical amount of stellar mass formed in situ is expected to decrease with the density of environment because in these environments the effects of starvation, strangulation, ram pressure striping, and shock heating take place. As a result, the correlations of the rotational support with mass and environmental density would induce correlations of the rotational support with merger-sensitive parameters, even if the mergers did not influence the rotational support. In statistical literature, this is called the confounding effect. In Appendix A, we check that the rotational support correlates significantly with galaxy mass and partly also with environmental density, and thus the necessary condition for the confounding effect to happen is satisfied. The way to eliminate the confounding effect is to inspect the correlations of the quantities of interest at fixed values of the confounding quantities. In our case, this means inspecting the correlations of merger-sensitive parameters with the rotational support for galaxies of a fixed mass and environmental density.

The main method of this paper is based on the assumption that if we consider galaxies of a fixed mass and environmental density, then mergers typically have a larger importance for galaxies with a lower rotational support than for those with a higher rotational support. The word "typical" is important here because, for example, some mergers with a specific orbital configuration and gas content can spin up the galaxy (Di Matteo et al. 2009a; Qu et al. 2010; Naab et al. 2014; Penoyre et al. 2017), or because the rotational support depends also on the orientation of the galaxy with respect to the direction to the observer. The trends of the importance of mergers and rotational support are thus expected to be valid only in the statistical sense.

[5] Actually, in galaxy clusters mergers can be prevented because of high relative velocities of the galaxies (Ghigna et al. 1998; Mihos 2003), but we do not have cluster galaxies in our sample.

Table 1. Predictability of the investigated merger-sensitive parameters from the rotational support, $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$.

Parameter	All		$\log M_{\mathrm{JAM}} < 11$		$\log M_{\mathrm{JAM}} \geq 11$		$\log \rho_{10} < -2$		$\log \rho_{10} \geq -2$	
	sign	p [%]	sign	p [%]	sign	p [%]	sign	p [%]	sign	p [%]
Shells	−	19	−	53	−	20	+	79	−	13
Streams	−	34	−	29	−	59	+	88	−	31
Tails	−	22	−	24	−	46	−	43	−	35
Disturbed isophotes	−	**0.64**	−	6.0	−	10	−	29	−	**1.1**
Any TF	−	22	−	30	−	46	−	95	−	21
Any TF or DI	−	13	−	29	−	33	−	36	−	22
Shell fr.	−	16	−	30	−	39	+	59	−	11
Stream fr.	−	40	−	14	−	88	−	78	−	46
Tail fr.	−	20	−	15	−	65	−	29	−	38
Dust	+	49	+	97	+	18	+	33	+	74
KDC	−	**0.0072**	−	**0.0037**	−	47	−	9.1	−	**0.027**
CRC	−	**0.014**	−	**0.051**	−	13	−	**2.7**	−	**0.20**
KT	−	36	−	43	−	85	N/A	N/A	−	43
2σ	−	**0.13**	−	**0.058**	N/A	N/A	−	26	−	**0.46**
No kin. feature	+	**0.14**	+	**0.0091**	−	69	+	**0.82**	+	6.8
$\log R_\mathrm{e}$	−	8.6	−	30	−	8.4	−	21	−	23
Core γ'	+	35	+	92	+	19	−	**1.9**	+	10
SSP age	−	87	−	86	−	76	−	60	−	97
SSP age gr.	−	82	+	73	−	70	+	66	−	78
SSP Z	+	**0.0025**	+	**0.0075**	+	15	+	**1.7**	+	**0.089**
SSP Z gr.	+	**3.2**	+	6.2	+	13	+	**1.5**	+	33
SSP α	+	83	+	32	−	**3.5**	−	30	+	26
SSP α gr.	+	59	+	31	−	25	−	34	+	11

Notes. The sign sub-column means the sign of the correlation, while the p subcolumn shows the probability that there is actually no correlation.

Ideally, one should sort the galaxy sample into narrow bins of mass and environmental density, and inspect in them the correlations between the rotational support and the merger-sensitive parameters. This is not possible for our sample because it is too small. We thus used another method that is able to eliminate or mitigate the confounding effect, namely the multilinear regression.

In particular, we made multilinear fits of each merger-sensitive parameter as a function of galaxy mass, environmental density, and rotational support:

$$[\text{parameter}] = b + a_M \log M_{\mathrm{JAM}} + a_\rho \log \rho_{10} + a_{\mathrm{KS}} \lambda^{\mathrm{N}}_{R_\mathrm{e}}. \quad (3)$$

The fitted coefficients are listed in Appendix C. The coefficient at the rotational support in Eq. (3), a_{KS}, can be used for assessing the correlation between the rotational support and the investigated merger-sensitive parameter without the confounding effects of mass and environmental density, provided that the relation between the quantities is well described by Eq. (3). In this method, one essentially substitutes the actual value of the merger-sensitive parameter of a given galaxy by the value predicted by the formula in Eq. (3) on the basis of the M_{JAM}, ρ_{10}, and $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$ of the galaxy. It then remains to be ascertained whether the correlation of the merger-sensitive parameter and the rotational support at a fixed mass and environmental density is statistically significant. To this end, we made another multilinear fit for the merger-sensitive parameter, but only as a function of the galaxy mass and environmental density:

$$[\text{parameter}] = \tilde{b} + \tilde{a}_M \log M_{\mathrm{JAM}} + \tilde{a}_\rho \log \rho_{10}. \quad (4)$$

The statistical significance of the correlation of the given merger-sensitive parameter with the rotational support at a fixed mass and environmental density was evaluated through an F-test applied to the residuals of the fitting by Eqs. (3) and (4). The F-test indicated whether the addition of the rotational support among the independent quantities improved the quality of the fit significantly. In other words, we found in this way whether the rotational support provides any information on the given merger-sensitive parameter if the mass and environmental density of the galaxy are already known. We accepted the significance level of 5% for the F-test. This means that the probability that the test will indicate a significant correlation between the quantities, if the quantities actually do not correlate, is 5%.

The results are presented in Table 1. The second main column represents the results for our whole galaxy sample. The third and fourth main columns represent the results of our test when applied only to galaxies that have an M_{JAM} mass lower or a higher than $10^{11}\,M_\odot$, respectively. This limit is motivated by the theoretical expectations described in Sect. 1. Similarly, in the last two main columns of Table 1, we divided our galaxy sample by the environmental density at $\rho_{10} = -2$. We are not aware of any past works that would show that galaxy properties abruptly change at a particular value of environmental density (e.g., separating filaments from galaxy groups). We chose our separating value on the basis of Fig. B.2 because, at this value, the incidences of tidal disturbances, dust lanes, and KDCs seem to change abruptly. We explored whether the results changed if we used other quantifications of the environmental density instead of $\log \rho_{10}$, namely the parameters Σ_3 and ν_{10} of Cappellari et al. (2011b), and we found no substantial difference.

In Table 1, the sign subcolumns specify whether the given merger indicator typically increases or decreases toward a higher value of the rotational support for galaxies at a fixed mass and environmental density. This is the sign of the fitted parameter a_{KS} in Eq. (3). The p subcolumns of the table give the p-value of the F-test – that is the probability that the inclusion of the

rotational support in the multilinear fit of the given merger-sensitive parameter actually does not improve the fit substantially, even if the F-test indicates so. Thus, the lower the p-value, the more significant the correlation of the rotational support with the given parameter. The emphasized values signalize that including the $\lambda^{\mathrm{N}}_{R_{\mathrm{e}}}$ term in the multilinear model improves the fit of the merger indicators at the significance level of 5% (bold font) or 1% (bold and larger font).

The table tells us that the rotational support indeed helps us to predict some of the merger-sensitive parameters, even if we already know the galaxy mass and environmental density. Namely, in the sample as whole, we find that for galaxies at a fixed mass and environment density, the rotational support correlates negatively with the incidence of disturbed isophotes, as well as with the incidences of KDCs, CRCs, and the 2σ features. On the other hand, the correlation is positive with a regular kinematic appearance. Next, we find that for galaxies at a fixed mass and environment density, the rotational support correlates positively with the values of metallicity and metallicity gradients (i.e., the galaxies of a higher rotational support usually have flatter metallicity gradients, since the gradients are usually negative). The most significant correlations of the rotational support are those with metallicity and with the presence of a KDC. Again, we point out that not detecting a significant correlation does not imply that the correlation does not exist. It can just be too weak to be detected in the current data.

Moving to the group of galaxies with $\log M_{\mathrm{JAM}} \geq 11$, for galaxies at a fixed mass and environmental density, we detected only one significant correlation: galaxies of low rotational support are more likely to have a relatively high abundance of α-elements. On the contrary, for the less massive part of our sample, we basically obtained the same correlations as for the whole sample. The exceptions are the correlations of the rotational support with disturbed isophotes and metallicity gradients, which slightly miss our significance threshold.

In the low-density subsample ($\log \rho_{10} < -2$), we noted two substantial changes with respect to the full sample. First of all, at a fixed mass and environmental density, the central photometric cuspiness significantly decreases toward higher values of rotational support. In the high-density part of the sample, the trend is opposite, yet insignificant: the cuspiness tends to increase with increasing rotational support. The second substantial difference is the loss of a significant correlation between the rotational support and the incidence of disturbed isophotes with respect to the complete galaxy sample. In the high-density subsample, we detect most of the significant correlations detected in the full sample. The first exception is the one between the rotational support and regular rotation, whose significance misses somewhat our limit. The second change is the lack of correlation of the rotational support with the metallicity gradient.

It is known that metallicity correlates strongly with the stellar velocity dispersion of galaxies (Thomas et al. 2010; McDermid et al. 2015). We thus tested if the strong correlation of the rotational support and metallicity disappeared if we fixed not only the mass and environmental density, but also the stellar velocity dispersion taken from Cappellari et al. (2013), namely the velocity dispersion within the isophote counting half of the galaxy luminosity. We made a multilinear fitting and an F-test as before. The sense of the correlation of the rotational support and metallicity remained as in Table 1 and the significance did not change much.

It is useful to define the cleaned merger-sensitive parameters as:

$$[\text{cleaned parameter}] = [\text{parameter}] - (b + a_M \log M_{\mathrm{JAM}} + a_\rho \log \rho_{10}), \quad (5)$$

where the values of b, a_M, and a_ρ were obtained by fitting Eq. (3) to the data. The cleaning allowed us to compare the values of the merger-sensitive parameters of galaxies without being affected by the confounding effects of mass and environmental density. The values of the cleaned parameters are provided in the table available at the CDS. Figure B.3 shows our galaxy sample in the space of the cleaned parameters versus the rotational support for all investigated merger-sensitive parameters.

6. Deducing the mechanism of decreasing the rotational support of ETGs from observations

Cosmological simulations of galaxy formation predict that ETGs decrease the level of their rotational support with time and that mergers play a substantial role in that (Sect. 1), which we take as a basic assumption in this paper. In this section, we exploit the correlations of the rotational support with merger-sensitive parameters at a fixed galaxy mass and environment density, as found in Sect. 5, to learn more about this transformation process from observations. We make use the fact that the different merger-sensitive parameters have different lifetimes and are sensitive to different types of mergers. We add further constraints based on findings from the literature. Relying on a constraint provided by a single parameter can be misleading since the merger-sensitive parameters can be influenced by other mechanisms than by mergers. It turned out to be advantageous that we are working here with many merger-sensitive parameters, because this allowed us to confirm some of our conclusions by several independent pieces of evidence. This allowed us to reduce the probability of a mistake caused by measurement errors or by misinterpreting the data. The individual pieces of evidence we found are summarized in Table 2. In the following paragraphs, we aim to find, for every constraint, not only an interpretation based on mergers, but also on other mechanisms. It turns out that every observation requires another alternative mechanism, which is often speculative or not underpinned by quantitative models. On the other hand, mergers appear as a more solid and universal explanation of all observations.

In Sect. 5 we found that, at a fixed mass and environmental density, the metallicity of a galaxy generally decreases with the decreasing rotational support. This is the most significant correlation we found. A similar result was found by Bernardi et al. (2019). This observational finding agrees with our assumption that the rotational support decreases because of mergers of smaller galaxies, because metallicity generally decreases with galaxy mass. Nevertheless, as with the other merger indicators, the metallicity of a galaxy is not influenced only by mergers, but also by other mechanisms or factors. These include the balance between the energy of gas outflow and the depth of the potential well of the galaxy (Pipino et al. 2010), the star formation history of the galaxy and its initial mass function (e.g., Matteucci 2014), a removal of gas cold gas from the galaxy, or an interruption of the inflows of intergalactic gas to the galaxy (Peng et al. 2015; Trussler et al. 2020). One can speculate that during the monolithic collapse phase, the formation of galaxies with a low rotational support is faster, and thus some of the alternative mechanisms are more effective than when the galaxy has a lower rotational support.

We detected that, at a fixed mass and environmental density, galaxies with a lower rotational support have a significantly higher incidence of disturbed outer isophotes. With a decreasing rotational support, the incidences of the different types of tidal features

Table 2. Summary of the constraints on the typical mergers that caused the decrease in the ordered rotation of ETGs.

Observation	Time since transformation [Gyr]	Indicates minor mergers?	Indicates wet mergers?	From our sample?
Tidal disturbances	>4	n	n	Y
KDCs	<12?	n	n	Y
KDCs + tidal disturbances	>4	n	n	Y
Stellar age + α-elements (massive ETGs)	>10	n	Y	Y
Galaxy counts at high z	>8–9	n	n	n
Ellipticities of massive quenched galaxies at high z	>10	n	n	n
Sérsic indices of massive quenched galaxies at high z	>12	n	n	n
Luminosities, surface brightnesses and effective radii of brightest cluster galaxies at high z	>10	n	n	n
Inner phot. profiles (low environmental density)	>10	n	Y	Y
Dust lanes	>1	n	n	Y
Effective radii	>12	n	Y	Y
Metallicity gradients		Y	n	Y

Notes. Observation: The observable providing constraints on the properties of mergers that caused the kinematic transformation. The order is as discussed in the text. Time since transformation: The typical time of the kinematic transformation implied by the given observable. A question mark indicates a speculative estimate. Indicates minor mergers? and Indicates wet mergers?: The columns show whether the given observable implies that the mergers causing the kinematic transformation were minor or wet, respectively. The symbols "Y" and "n" indicate yes and no, respectively. None of the observing facts considered in this study indicate that the mergers were major or dry. From our sample?: Indicates whether the given observational data are those analyzed in this paper or come from literature, with the meaning of symbols as before. The data from literature sources might be less telling for the galaxy sample investigated in this paper.

increase too, even if these correlations are not statistically significant. All types of tidal disturbances become particularly rare above the rotational support of 1.0, as Figs. 1 and B.3 show. This indicates that even at a fixed mass and environmental density, the galaxies with a lower rotational support experience more mergers. The tidal disturbances had to be younger than the maximum lifetime of tidal disturbances, which we estimated in Sect. 4 as 9 Gyr. On the other hand, it is important to note that even among the galaxies with the lowest rotational support, most galaxies do not show tidal disturbances. Indeed, out of the 25 SRs in our sample (i.e., their $\lambda^{\mathrm{N}}_{R_{\mathrm{e}}} < 0.31$), only eight have their cleaned indicator of disturbed isophotes higher than zero. Thus, if we ask for the typical cosmic epoch when mergers were decreasing the rotational support, it had to be before the minimum lifetime of disturbed isophotes, that is, before 4 Gyr (Sect. 4). There is no guarantee that the observed tidal disturbances were caused by the mergers that contributed most to the decreasing the rotational support. The findings suggest the possibility that galaxies with a lower rotational support live in environments where galaxy interactions are more common compared to those that have the same mass and environment density but a higher rotational support. As for the alternative explanations that do not rely on mergers, tidal disturbances could have been caused by non-merging galaxy flybys, which are expected to be common by alternative theories of gravity (Bílek et al. 2018, 2019b). If the rotational support of ETGs was set by internal processes, then we have to speculate that galaxies with a lower rotational support experience more non-merging flybys than galaxies with a higher rotational support. Tidal disturbances might also dissolve faster if the gravitational potential of the galaxy is flatter because the stars would not continue orbiting in the plane of collision, as in a spherical potential. It has indeed been found that FRs are oblate ellipsoids while SRs are closer to being spherical (Cappellari et al. 2007; Weijmans et al. 2014; Foster et al. 2017; Li et al. 2018b).

Next, we detected that, at a fixed mass and environmental density, galaxies with a lower rotational support possess KDCs, CRCs, and the 2σ features more often. This suggests that these kinematic substructures formed in the same mergers that were responsible for decreasing the rotational support of the galaxies. This would likely mean that the decrease in rotational support did not happen during the era when galaxies were mostly gaseous, that is, say before the redshift of three (more than 12 Gyr ago), because a streaming of gas flows in opposing directions in the same system is difficult. This maximum age estimate agrees well with the result of Ebrová et al. (2021), who found that the oldest KDC in the cosmological hydrodynamical simulation Illustris is 11.4 Gyr old at the current epoch. This observational constraint is, however, not very strong, since only relatively few galaxies have a KDC or CRC – the typical epoch of decreasing the rotational support might have been earlier. On the other hand, one should keep in mind that kinematic substructures can have a different origin than mergers, such as non-merging galaxy flybys (Hau & Thomson 1994; Young et al. 2020), projection effects (Statler 1991), and sequential accretion of gas from different cosmic filaments (Algorry et al. 2014, see also Ebrová et al. 2021). Also, we can speculate that a kinematic substructure might be more difficult to observe if its host galaxy has a strong net rotation, because the substructure would have a lower contrast in the kinematic map. Yet another speculative possibility is that during the monolithic collapse phase, the angular momentum of the inflowing gas changed its direction at some point.

An interesting constraint on the typical time of the kinematic transformation appears once we note that, at a fixed galaxy mass and environmental density, the incidence of KDCs, CRCs, and the 2σ features does not correlate significantly with the presence of tidal disturbances. We found this by the combination of the multilinear fitting and an F-test, just as we did when looking for correlations of our merger-sensitive parameters with $\lambda^{\mathrm{N}}_{R_{\mathrm{e}}}$. Taken in another way, out of the 19 galaxies with KDCs, CRCs, or the 2σ features, only five have a positive cleaned parameter of tidal disturbances (that is more prominent tidal disturbances than typical for galaxies of the given mass and environment density), which is $26 \pm 12\%$ (Poisson error assumed). If we count only the galaxies with KDCs or CRCs, only five of 13 ($38 \pm 17\%$) have a positive cleaned parameter of tidal disturbances. This is not an excess compared to the whole investigated galaxy sample, where the fraction of galaxies with a positive cleaned parameter of tidal disturbances is $36 \pm 5\%$. This suggests that the transformation

of the rotational support typically occurred before the minimum lifetime of the tidal disturbances, which is before 4 Gyr, otherwise most galaxies with kinematic substructures would show a positive cleaned parameter of tidal disturbances.

We did not detect any significant correlation between the rotational support and the stellar age for galaxies at a fixed mass and environmental density. This indicates that the decrease in the rotational support did not typically occur by the means of recent wet mergers that would form a lot of new stars. For the massive part of our galaxy sample ($\log M_{\rm JAM} > 11$), we can combine this with additional constraints to get a more complete picture. For them, we detected that at a fixed mass and environmental density, a lower rotational support implies a higher abundance of α-elements. This is a signature of wet mergers, not dry. For these massive galaxies, we do not expect substantial recent in situ star formation that would be responsible for the high α-abundance. This is suggested both by the reconstructed observed star formation histories (McDermid et al. 2015) and cosmological simulations (Penoyre et al. 2017). To explain these constraints simultaneously, we propose that the mergers that lowered the rotational support of the massive galaxies had to be wet, but occurring at high redshifts, so that the difference in stellar age with respect to the galaxies with a higher rotational support cannot be measured. The mergers preferably had to happen when ETGs were still forming stars, which is at about $z > 2$, (about 10 Gyr ago). This view is supported by the observations of quiescent galaxies at $z = 1.6$ (9.6 Gyr ago) by Onodera et al. (2015). Their galaxies mostly had stellar masses over $10^{11}\,M_\odot$, and most of these galaxies are SRs (Emsellem et al. 2011). Onodera et al. (2015) found that if their galaxies evolved passively since the redshift of the observation, their metallicity and α-abundance would agree excellently with that of the ETGs in the nearby Universe. One could perhaps explain the negative correlation of the α-abundance with rotational support without mergers. For example if the progenitor gas cloud had a higher angular momentum, it settled more slowly, such that the star formation was less bursty, and that resulted in a stellar population that is poor in α-elements.

Another interesting constraint on the time of the lowering of the rotational support by mergers comes from the evolution of the cosmic spatial density of quiescent galaxies with redshift. Recent results suggest that the cosmic number density of quiescent galaxies with stellar masses over $10^{11}\,M_\odot$ (i.e.those that are mostly SRs in the nearby Universe), have not evolved at least since $z = 1.4$ (9.1 Gyr ago, Kawinwanichakij et al. 2020). This suggests that such galaxies do not experience any substantial mass growth by mergers. The lighter galaxies continued growing to a later time: the cosmic number density for the quiescent galaxies with stellar masses over $10^{10}\,M_\odot$ has not evolved since $z = 1$ (7.8 Gyr ago). We do not have lighter galaxies in our sample. The mergers that would cause the decrease in the rotational support bring some material and increase the stellar mass of the galaxies. Altogether, the observations show that galaxies of the masses investigated in this paper stopped evolving 8–9 Gyr ago. During the earlier epochs, the galaxies could change their masses either by mergers or in situ star formation. Thus, we can conclude that the mergers happened more than 8–9 Gyr ago. One might argue against this constraint: in the late cosmic times, galaxies grow more by minor mergers than by major mergers and the minor mergers tend to deposit their material at the outskirts of galaxies (Amorisco 2017; Karademir et al. 2019). This material is difficult to detect because of its low surface brightness, but it can comprise a non-negligible fraction of the stellar mass of the galaxy (Huang et al. 2018). However, it is then questionable if mergers that deposit material at the outskirts of galaxies can decrease the rotational support of the galaxy that is measured within one effective radius of the galaxy. In addition, the sample of Huang et al. (2018) consisted of extremely massive galaxies (logarithmic stellar masses over 11.4); for the MATLAS sample, Duc et al. (2015) found the halos of our galaxies to contain, on average, about 5% of the total luminosities of the galaxies (confirmed by another method in Sola et al. in prep).

Observations of massive galaxies at high redshifts provide yet further clues as to the time of formation of massive SRs. It turns out that quenched galaxies at higher redshifts are generally flatter than the local ETGs. The exceptions are the galaxies with logarithmic stellar masses over 11.3 that appear always round in projection, at least to $z = 2$ (Chang et al. 2013). If we assume that galaxies form first as rotating disks and then they transform to pressure-supported spheroids by mergers, then the most massive galaxies, which are usually are SRs, had to be formed before that redshift.

In addition, the flat quiescent galaxies at high redshifts also have lower Sérsic indices, resembling the spiral galaxies in the local Universe in this regard. The increase in the typical Sérsic index of quiescent galaxies in time can be attributed to mergers (Schweizer 1982; Hilz et al. 2013). Lustig et al. (2021) nevertheless found that the quiescent galaxies with stellar masses around $10^{11}\,M_\odot$ already had high Sérsic indices around 12 Gyr ago ($z = 3$), suggesting that the mergers happened before that time. The flat objects with low Sérsic indices, which prevail among quiescent galaxies at high redshifts, can be the progenitors of the local FRs, which prevail also in the population of the local ETGs. In addition, it is observed that effective radii, surface brightnesses, and luminosities of brightest cluster galaxies have not evolved at least for 10 Gyr ($z = 1.8$) (Chu et al. 2021, 2022).

Another constraint on the time of the transformation of the rotational support is provided by the inner photometric profiles. For our sample as a whole, we did not detect any trend of the γ' parameter with the rotational support at a fixed mass and environmental density. We only found that the galaxies belonging to the low-density subsample ($\log \rho_{10} < -2$) with a lower rotational support have more cuspy profiles. This disfavors dry mergers as the cause of the decrease in the rotational support, because such mergers should make the profiles more cored (this was already suggested for low-mass SRs by Krajnović et al. 2020). This suggests that mergers that decreased the rotational support were gas-rich, such that the increase in cores by the merging of the central black holes was balanced by the formation of new cusps, as explained in Sect. 3. This fits in our picture where the decrease in the rotational support happens primarily at high redshifts, when ETGs were still forming stars. Combining the lack of correlation between the rotational support and central photometric profile with the old ages of stellar populations, we expect that the mergers happened typically before $z = 2$ (10 Gyr ago). The higher cuspiness of galaxies with a low rotational support in the low-density environments suggests that in these environments the mergers were particularly gas-rich, in agreement with the environment-morphology relation. This was already suggested by Krajnović et al. (2020). We note that cores in ETGs can also be induced during the stage of the monolithic collapse (Nipoti et al. 2006), or as a result of repeated removals of gas from centers of galaxies by activity of galactic nuclei (van der Vlugt & Costa 2019). In order to explain the correlation found without mergers, these two alternative core-forming processes would have to be stronger for galaxies with a higher rotational support, which is again a speculative possibility. Our results could be somewhat biased by the fact that we had the information about the central photometric profile for a much

lower number of galaxies than for the other investigated merger-sensitive parameters; namely, the slope of the inner photometric profile is available for 80 out of our 175 galaxies.

We did not detect any statistically significant correlation between the incidence of dust lanes and the rotational support for galaxies at a fixed mass and environmental density. Given that the lifetime of dust is expected to be under 1 Gyr (Patil et al. 2007), the mergers that decreased the rotational support likely typically happened before that time. The same result would be expected if the mergers were dry, which is, however, disfavored by some of the previous pieces of evidence.

We also did not detect a statistically significant correlation of the effective radius with the rotational support at fixed galaxy mass and environmental density. Here we assume that the decrease in the rotational support happens primarily by mergers, which implies that at a fixed galaxy mass and environment density, the galaxies with a lower rotational support should have larger effective radii if the mergers were dry, as explained in Sect. 3. This suggests that the decrease in the rotational support happened primarily when the galaxies were still mostly gaseous, say at $z > 3$, because the expansion by mergers works only for dissipationless systems. Alternatively, one might argue that the growth of the radius of the galaxies with a low rotational support by mergers was approximately the same as the growth of radius of the galaxies with a higher rotational support by in situ star formation.

Finally, we found that at a fixed galaxy mass and environmental density, the galaxies with a lower rotational support have steeper metallicity gradients[6]. As explained in Sect. 3, this signifies that mergers that decreased the rotational support were preferably minor. The gradients we used, being derived by the SSP method, are biased toward the old stellar populations. This further supports the hypothesis that the transformation happened a long time ago. The correlation is stronger in the low-density environment subsample. This can be explained by the accreted galaxies being more metal poor, as can be expected due to the fact that the metallicity of galaxies decreases toward low-density environments at a fixed galaxy mass. The steeper metallicity gradient of the galaxies with a low $\lambda^{\mathrm{N}}_{R_{\mathrm{e}}}$ with respect to the galaxies with a high $\lambda^{\mathrm{N}}_{R_{\mathrm{e}}}$ could also be explained by the fact that the latter routinely experience major mergers. This, however, goes against our assumptions and several observational findings stated above.

Another option that remains to be discussed is that the galaxies that currently have a low rotational support first experienced the ancient wet mergers at $z > 2$, which changed the values of the merger-sensitive parameters, but these mergers were not those that decreased the rotational support. The rotational support was instead changed at a lower redshift by mergers that did not influence the merger-sensitive parameters. We dismiss this option because it seems too fine-tuned and opposes Occam's razor. In addition, it contradicts the abovementioned observed constancy of volume density of quenched galaxies since $z = 1{-}1.4$, and because in Sect. 7 we find that mergers in the current Universe are very rare.

To summarize this section, the observations can be explained easiest if the mergers that decreased the rotational support were wet and minor. They had to happen when the ETG progenitors were not completely gaseous, but still contained a substantial fraction of gas. This is at around $z = 2$ or even before, depending probably on the mass and surface density of the galaxy (see Sect. 1). The individual constraints on the time and way of the transformation described above are summarized in Table 2. This way of galaxies forming with a low rotational support resembles the pictures of the monolith collapse from the simulations of Kobayashi (2004), where the galaxy is rapidly assembled through many wet minor mergers. A wet clumpy collapse is what was also assumed in the successful chemical model of Thomas et al. (1999). A similar scenario was claimed to explain the chemical properties of KCDs (Bender & Surma 1992). Also, the simulations of Naab et al. (2014) showed that the galaxies with the lowest λ had to be formed by multiple minor mergers. The higher incidence of tidal disturbances observed in the galaxies with a low rotational support suggests that such galaxies live in environments where galaxy interactions are more common, such as the intersections of cosmic filaments. The late interactions, however, do not seem to be determinant for the properties of the galaxies, which we check again below in Sect. 7.

The observational evidence does not seem to agree with the prediction of cosmological simulations that the transformation of the kinematic structure of ETGs happened after the redshift of one (8 Gyr ago) (Penoyre et al. 2017; Lagos 2018). This seems to be another manifestation of the problem of the too gradual formation of ETGs in cosmological simulations (see Sect. 1). The ideal solution to determine when the transformation occurred would be to observe spatially resolved stellar kinematics at high redshifts. Fist observations of this type start appearing nowadays (Newman et al. 2018; Cole et al. 2020) but we have to wait until data become available for a statistically meaningful sample of galaxies.

It is interesting to put the typical time of the transformation that we determined, the redshift of about two, in the context of the other events happening in the Universe at the same epoch. The redshift of two is the time when the global star formation in the Universe, dominated by massive galaxies, started to drop (Madau & Dickinson 2014; Liu et al. 2018; Wilkins et al. 2019). At the same time, the frequency of galaxy interactions dropped too (Huertas-Company et al. 2015; Ventou et al. 2017, 2019). Before $z = 2$, the interstellar medium of star-forming galaxies was moving randomly, under the effect of turbulence, and the gas was forming a lot of giant gas clumps, which could hold a substantial fraction of the baryonic mass of the galaxies. At $z \sim 2$, ordered motions in star-forming galaxies started to prevail (Simons et al. 2017). At the same time, galaxies attained the standard *Hubble* morphologies known from the local Universe, instead of showing prominent giant gas clumps (Mortlock et al. 2013; Lee et al. 2013). The redshift of about two thus seems to be a plausible point at which the kinematic morphology settles. Indeed, it has been proposed that the kinematic morphology is more fundamental than the *Hubble* photometric morphology (Cappellari et al. 2011b), and therefore settling both morphologies at the same time appears logical.

7. Sanity check: The role of the late galactic interactions in the decrease in the rotational support

In the previous section, we found multiple independent pieces of evidence that the rotational support is decreased by mergers typically before the redshift of two. It might be concerning that we

[6] We noted the apparent contradiction with the results of Krajnović et al. (2020) who reported less steep metallicity gradients in SRs compared to FRs. There are several reasons for this: 1) Krajnović et al. (2020) did not compare the γ' parameter for FRs and SRs of the same mass and environmental density; 2) while they analyzed the ATLAS^{3D} sample just as we did, our sample contains only its MATLAS subset, which avoids the Virgo cluster; 3) Krajnović et al. (2020) applied a binary separation of ETGs to FRs and SRs.

also found that galaxies with a lower rotational support possess tidal disturbances more often at a fixed mass and environmental density. This observation alone could also be interpreted to mean that the rotational support is decreased before the maximum lifetime of tidal features, which is 9 Gyr (while this contradicts the other findings from the previous section). In this section we therefore inspected in more detail the role of the late interactions, whose signatures can still be observed as tidal features, in decreasing the rotational support. It seems that the late interactions do not significantly influence galaxies.

7.1. Recent major mergers do not contribute to the correlations with rotational support

In Sect. 5, we detected correlations between the rotational support and some of the merger-sensitive galaxy parameters. We used the correlations to deduce what mergers caused the decrease in the rotational support and found that they happened at high redshifts. In this section, we perform a sanity check on the early drop in the rotational support. According to the Illustris simulation, at $z = 1$ (8 Gyr ago) the progenitors of the current-day FRs and SRs had indistinguishable distributions of the λ parameter (Penoyre et al. 2017). Galaxy merging was relatively rare in the Universe in these late epochs. We expect every galaxy to have had relatively few of these recent mergers (e.g., Rodriguez-Gomez et al. 2015, and see Sect. 7.2) since $z = 1$ (7.8 Gyr ago), typically about one or five, depending on the mass ratio and whether we trust rather observations or simulations. These few mergers had to set the current λ of galaxies. There would be many galaxies that did not receive any substantial merger, and such objects would not show any correlation between the merger sensitive parameters and $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$. The correlations that we detected would be caused only by the galaxies that experienced substantial merging. Galaxies showing large morphological disturbances (i.e., probable remnants of recent major mergers), would then contribute a large portion of the points that induce the detected correlations. We test this in this section and find that that it does not seem to be the case.

We identified the following 14 galaxies as the most serious candidates for major merger remnants, because of their morphology in the deep MATLAS images[7]: NGC 0474, NGC 1222, NGC 2764, NGC 3414, NGC 3610, NGC 3619, NGC 3640, NGC 4382, NGC 4636, NGC 4753, NGC 5485, NGC 5493, NGC 5557, and NGC 5866. These candidates are identified in the galaxy list available at the CDS. They show either strong morphological disturbances, even close to the centers of the galaxies, or their stellar halos are strongly offset from the central parts of the galaxies. Their indicator of disturbed isophotes is two (13 cases) or one (1 case). The galaxies often show complex tidal features and unrelaxed dust patches. Such galaxies underwent relatively strong interactions recently (before the lifetime of tidal features, see Sect. 4). We avoided the galaxies that seem to be involved in an ongoing interaction with their neighbors, or the galaxies whose morphology is difficult assess, for example because of pollution by the light scattered from a nearby star, or because the galaxies overlap with neighbors in projection. Our selection is likely biased in favor of mergers involving spiral or lenticular galaxies. Such mergers typically produce more distinct photometric irregularities than major mergers of pressure-supported galaxies. Our sample of major merger candidates spans both mass bins and both the density bins defined in Sect. 5.

We compared the properties of the major merger candidates to the less-disturbed galaxies in our sample in Fig. 2. It shows the galaxies in the space of the cleaned merger-sensitive parameters versus the rotational support. The figure shows only the parameters that were identified in Sect. 5 to correlate statistically significantly with the rotational support at a fixed mass and environment density of the galaxy. If the statistically significant correlation was detected only for one of the considered galaxy subsamples (that is the low- and high-mass subsamples, and the low and high environmental density subsamples), then the figure shows only the galaxies belonging to that subsample. In addition, to stress our arguments (see below), we also show the galaxies in the plane of the cleaned SSP age versus the rotational support, where the galaxies do not form a significant correlation. Every galaxy in Fig. 2 is represented by a point, and the major merger candidates are highlighted by crosses. The median values of the cleaned parameters and of the rotational support for all the depicted galaxies are marked by the dashed lines. We did not include among the plots in this figure the disturbed outer isophotes because their presence was the main criterion to select the major merger candidates – the major merger candidates would be offset from the other galaxies in the sample by definition. In agreement with what was found in Sect. 5, Fig. 2 shows that the major merger candidates have relatively low rotational support, compared to other galaxies.

The figure shows that the major merger candidates do not seem to contribute to the correlations. For example, it is striking that almost none of the candidates possess a KDC or other kinematic substructures. For the central photometric slope and the abundance of α-elements, there does not seem to be any cloud of points corresponding to the galaxies that did not experience any substantial recent merging. The data are rather explained better, such that another process defines the distribution of points in the space of cleaned parameters versus $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$, and that the major merger candidates are just randomly selected points at the given value of $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$. This other process could be, in our interpretation, the numerous wet mergers at high redshift. In contrast, the distribution of the major merger candidates in the space of the cleaned SSP age versus $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$ is clearly biased. This is how we expect the major merger candidates to be distributed in the other plots, if the late mergers were responsible for forming the correlations.

We also note that the gradients of metallicity for the major merger candidates seem to be more positive (i.e. flatter) than for the other galaxies. At the same time, the major merger candidates have a low $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$ compared to the others. Together, this means that the major merger candidates actually make the detected correlation weaker. One interpretation is that these recent major mergers erase a correlation that was induced by the ancient numerous minor mergers.

The interactions that created the recent major merger candidates thus do not seem to be important for shaping the correlations that we used to infer how the rotational support of galaxies is lowered. This suggests that the mergers that lowered the rotational support had to be more major or numerous in order to induce the observed trends.

7.2. Events forming observable tidal disturbances are rare

Here we provide an estimate of how many interactions the galaxies in our sample experience per gigayear, on the basis of the

[7] Images of all MATLAS galaxies are available at `http://obas-matlas.u-strasbg.fr`.

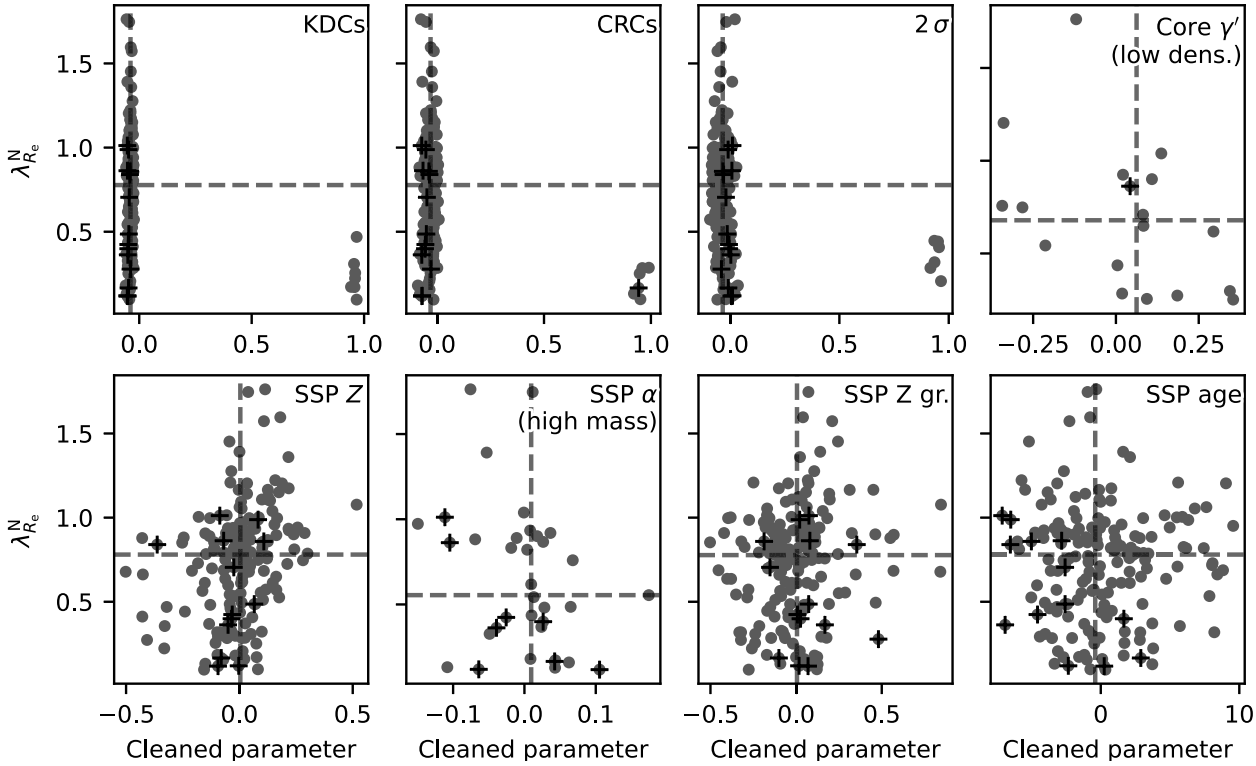

Fig. 2. Demonstration that the significant correlations between the rotational support and the merger-sensitive parameters detected above were not caused by recent major mergers. All galaxies from the sample are shown as points, while the candidates for the recent major merger remnants are highlighted by crosses. The horizontal coordinate of each panel shows the cleaned value of the given parameter. The vertical axis shows the rotational support. The dashed lines indicate the median values calculated for all available data.

observed incidence of tidal disturbances. The interactions appear to be relatively rare.

We assume that every tidal disturbance has the same life time, T_L = 4–9 Gyr (see Sect. 4), that the frequency of the disturbance-forming events is the same for all the galaxies in the sample under consideration, and that the frequency is constant in time.We count only the independent interactions. This means, for example, that the interaction of a given galaxy with another galaxy and its satellite is counted as one interaction. Then the number of interactions that created tidal disturbances in a given galaxy, n, follows a Poisson probability distribution:

$$P(n \text{ interactions}) = \frac{\mu^n e^{-\mu}}{n!}. \quad (6)$$

One can estimate the parameter of the distribution, μ, from the fraction of galaxies that are observed to have a tidal disturbance, f, by making use of the fact that

$$f = P(n > 0) = 1 - P(n = 0) = 1 - e^{-\mu}. \quad (7)$$

This gives $\mu = -\ln(1 - f)$. The mean of a Poisson distribution is the parameter μ itself, and therefore the mean number of events that formed tidal disturbances in a galaxy in the considered sample can be estimated as

$$\langle n \rangle = -\ln(1 - f). \quad (8)$$

If galaxies typically experience a large number of events forming tidal disturbances in T_L, then we expect tidal disturbances to be present in a large fraction of galaxies.

In the census of the tidal disturbances in the MATLAS sample, Bílek et al. (2020) found f for the total sample to be 41%, if we count both the certain and likely detections. The likely detections include the galaxies where the tidal disturbances were too faint to be detected with certainty, or the galaxies whose images had an inferior quality. This value of f is thus rather the upper limit of the true value. With this fraction of galaxies with tidal disturbances, we found that every galaxy typically experienced only 0.51 independent interactions in the last T_L. For massive galaxies with $\log M_{\rm JAM} > 11$, Bílek et al. (2020) found that f increases to 60%, meaning that such galaxies experienced on average 0.91 disturbing events in T_L. Some tidal disturbances might had been undetected because of unfavorable projection effects. Mancillas et al. (2019) investigated this issue in simulations and found that streams and tails are not sensitive to projection effects, but shells are. Nevertheless, because several shells are usually present in a galaxy, it is possible to detect at least one of them from any line of sight. Moreover, that work did not consider galaxies with disturbed isophotes. Doing so would decrease the sensitivity to the projection effects. Even the most massive galaxies, therefore, have experienced just about one interaction in the last 4–9 Gyr. This also includes the minor mergers forming stellar streams. For the high-mass galaxies, we obtained the rate of interactions forming tidal disturbances of 0.23 per Gyr for the minimum T_L, or 0.1 for the maximum one. This is in reasonable agreement with the merging rates estimated from the spatial density of close galactic pairs (Man et al. 2016). Out of the galaxies with $M_{\rm JAM} < 10^{11}\,M_\odot$ in MATLAS, 36% have likely or certain detections of tidal disturbances. Equation (8) then implies that only 0.44 interactions happened in these lighter galaxies in the time T_L, meaning that the merging rate is at most 0.11 per Gyr.

We were able to estimate the frequency of major mergers from the number of major merger candidates from the last section. The 14 galaxies constitute 8% of the whole MATLAS

sample. From here we obtained $\langle n \rangle = 0.08$, or at most 0.02 major mergers per Gyr, for the minimum lifetime of tidal disturbances. This agrees with the major merger rates found by the counts of galaxy pairs (Mundy et al. 2017). This shows that major mergers are extremely rare in the current Universe.

Thus, we again find evidence against the decrease in the rotational support within the lifetime of tidal disturbances. The increased incidence of tidal disturbances in galaxies with a low rotational support thus rather reflects that galaxies with a low rotational support have continued to receive more mergers even until today, but the late mergers do not usually change the properties of the galaxies substantially.

The above estimates of merging frequencies can be used to test the hypothesis that quenched galaxies increase their radius with redshift because of mergers. Trujillo et al. (2011) estimated the number of mergers necessary for explaining the observed evolution of radii of quenched galaxies with redshift. The masses of their galaxies were similar to ours. For the redshift of 0.4, which corresponds to the look-back time of about the minimum lifetime of tidal disturbances, they give 1.4 ± 0.3 mergers for the mass ratio of 1:3 or 3.0 ± 0.7 mergers for the mass ratio 1:10. The value for the major mergers is somewhat higher than the number of interactions that we estimated for the massive galaxies from the frequency of tidal disturbances, but still in the 2σ uncertainty limit. Nevertheless, we have to remember that many of the disturbances, especially the tidal streams, likely come from minor mergers and others are false detections. For the low-mass galaxies, the discrepancy is even higher. If we instead adopt the maximum lifetime of tidal disturbances, which corresponds to $z = 0.8$, the estimates of Trujillo et al. (2011) predict even a 2.5 times higher number of mergers than since the redshift of 0.4. This all suggests that the expansion of ETGs with redshift must be partly attributed to other mechanisms than mergers (Fan et al. 2008, 2010; Damjanov et al. 2009; Ishibashi et al. 2013; van der Vlugt & Costa 2019). Man et al. (2016) and Newman et al. (2012) arrived to the same conclusion from the counts of ongoing merger candidates. The need for alternative mechanisms is particularly strong for the low-mass galaxies. Indeed, the chemical models by Trussler et al. (2020) predict that gas outflows become more and more important toward ETGs with lower masses.

Further, it is worth noting that in the simulation inspected by Mancillas et al. (2019), at least one tidal stream in the reach of the MATLAS survey was detected for any time and for any line of sight. This contrasts with the large fraction of observed galaxies without any tidal disturbances; only 16% of galaxies in MATLAS have certain or likely detection of streams. This again points to a too extended formation of ETGs in cosmological simulations.

8. Corollaries of the assumption of the decrease in the rotational support by mergers

This paper is based on the assumption that galaxies are formed with a high rotational support, which is then decreased primarily in mergers. In this section we point out that this assumption alone can explain several observational facts.

If a galaxy gained a substantial fraction of mass through mergers, it will probably end with a low rotational support. Many massive galaxies and the central galaxies of clusters and galaxy groups are expected to have formed in this way (Dubinski 1998; Mihos 2003; De Lucia & Blaizot 2007; Penoyre et al. 2017). It follows from here that these galaxies can be expected to preferably be SRs. Indeed, the fraction of SRs increases with galaxy mass and the central galaxies are usually SRs (Emsellem et al. 2011; Houghton et al. 2013; D'Eugenio et al. 2013; Brough et al. 2017; Veale et al. 2017; van de Sande et al. 2021a).

Rotational support cannot be lower than zero. If galaxy mergers usually decrease the rotational support, then the galaxies that experienced many mergers will have their $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$ parameters clustered toward zero. This explains the finding that the distribution of galaxies in the λ plane versus stellar mass is bimodal, with one of the peaks near $\lambda = 0$ (van de Sande et al. 2021b) (even if similar studies can be affected by the sample selection bias, as demonstrated in Graham et al. 2018). Next, it is observed that for galaxies of a fixed stellar mass, the prominence of the peak near $\lambda = 0$ increases with the stellar mass, which can be explained by the greater fraction of material gained by accretion by the massive galaxies.

A galaxy that has gained a large fraction of mass through many mergers from different directions can be expected to lose any sign of the initial rotation. They would acquire a new kinematic morphology. Cappellari (2016) divided galaxies with a low value of rotational support into misclassified and genuine SRs. The two classes have different kinematic structures. The orbits of stars in the genuine SRs are randomly oriented, while the misclassified SRs seem to consist of two counterrotating disks. The genuine SRs probably correspond to the galaxies that experienced a large number of mergers. The misclassified SRs might have formed from FRs that experienced one merger that coincidentally had a suitable orbital configuration. Indeed, the massive galaxies that experience many mergers, are typically observed to be genuine SRs (Cappellari 2016).

It is worth noting that the galaxy sample investigated here is expected to contain only a few genuine SRs. Such galaxies have a $\lambda^{\mathrm{N}}_{R_\mathrm{e}} < 0.31$ and $M_{\mathrm{JAM}} > 2 \times 10^{11}\, M_\odot$ (Cappellari 2016). Even if our sample does not contain many genuine SRs, its members still experienced mergers. The final stage of this process is the formation of genuine SRs. Our approach here resembles the studies of the formation of quenched galaxies that make use of the green valley galaxies that have not quenched completely yet. (e.g., Trussler et al. 2020; Carleton et al. 2020; Noirot et al. 2022).

9. Final picture of the formation of SRs, FRs, and massive spirals

In this section, we address the typical formation paths undertaken by spiral galaxies, and fast and slow rotator ETGs. The proposed scenarios are outlined in Fig. 3. Our picture for the formation of the different galaxy types is based on the assumption that the most predictive factor of the evolution of a galaxy is its final mass. The primary dependence on mass agrees, for example, with the findings from Appendix A and with the fact that the morphology of galaxies correlates strongly with their mass: spirals are typically less massive than ETGs (Read & Trentham 2005; Blanton & Moustakas 2009; Kelvin et al. 2014) and FRs are less massive than SRs (Fig. B.1, Emsellem et al. 2011; Cappellari 2016; Veale et al. 2017; Brough et al. 2017; Greene et al. 2017; Graham et al. 2018). The first three rows of Fig. 3 depict the typical time evolution of a standard galaxy. In the last row, we included some of the nontypical ways the discussed galaxy types are formed.

We begin with the proposed formation scenario of a typical SR, which is depicted in the right column of Fig. 3. It is based mostly on our findings for galaxies outside of galaxy clusters,

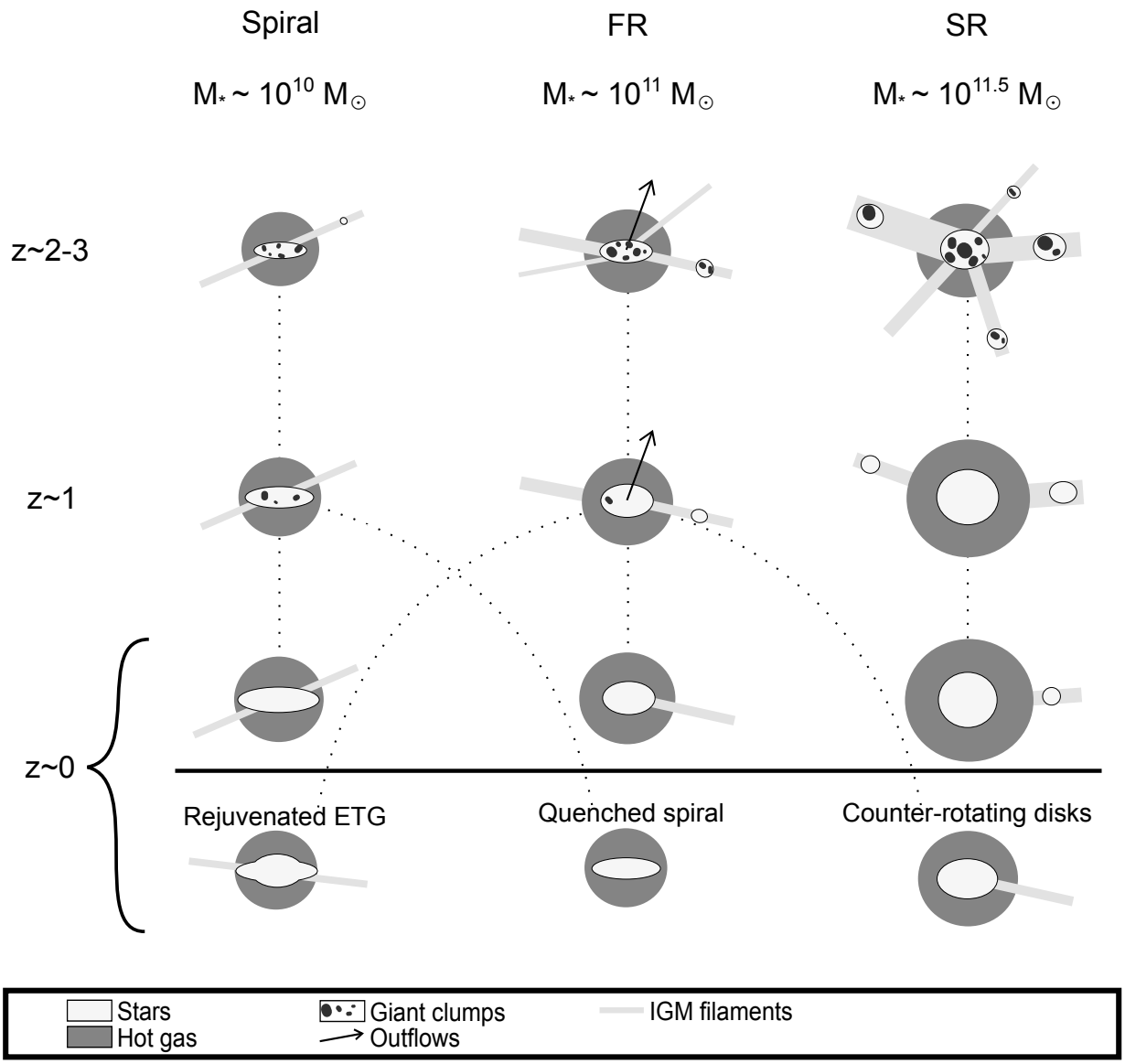

Fig. 3. Proposed scenario for the formation of massive galaxies outside of galaxy clusters. The columns from left to right correspond to the formation of a typical spiral, FR, and SR. At the top of each column, we state the typical current stellar mass of the respective galaxy type. The last row of the figure below the thick black line shows some nontypical ways of forming the different galaxy types. The dotted lines mark the evolutionary tracks of the different galaxy types.

but we argue later that even the cluster SRs follow a similar formation path. Before the redshift of about two, the galaxy rapidly grows by in situ star formation (Sect. 1). It experiences numerous wet mergers. This leads to a high enrichment of the stellar population by α-elements. The present-day slow rotation suggests that the accreted material was arriving from multiple directions, a process that is favored if the galaxy was located at the intersection of the multiple cosmic filaments that bring in the galaxy intergalactic medium (IGM) and other, smaller, galaxies. The location of SRs at the intersections of cosmic filaments does not necessarily imply that they live in regions of extremely high concentrations of galaxies. Some filaments may indeed host a low number of galaxies that are sparsely distributed. Part of the inflowing IGM is shock-heated, forming a halo of hot gas around the galaxy (Sect. 1). At the redshifts above two, the cold IGM is still able to penetrate the hot halo of the massive galaxies and feed the star formation in the galaxies. According to the successful chemical models of monolithic collapse by Pipino et al. (2010), the galaxy quenches because of supernova winds. The quenching proceeds outside-in.

The mergers experienced by the galaxy contribute to the decrease in the rotation support of the galaxy. Nevertheless, there are some additional mechanisms that might contribute as well. We consider two equal gas clouds that have the same apocentric distance with respect to the center of a spherical galaxy, but one cloud is on a circular orbit and the other is on a radial orbit. The binding energy of the gas cloud on the circular orbit is lower. The supernova winds that are supposed to quench the galaxy will thus preferentially remove the gas cloud from the galaxy on the circular orbit. The α-enhancement identified for the galaxies with a low rotational support suggests that SRs experience a stronger feedback than FRs of the same mass. Similarly, the non-detection of a correlation between the effective radius and the rotational support at a fixed mass and environment density indicates that the original potential energy of the accreted galaxies must be removed from the system, either in the form of

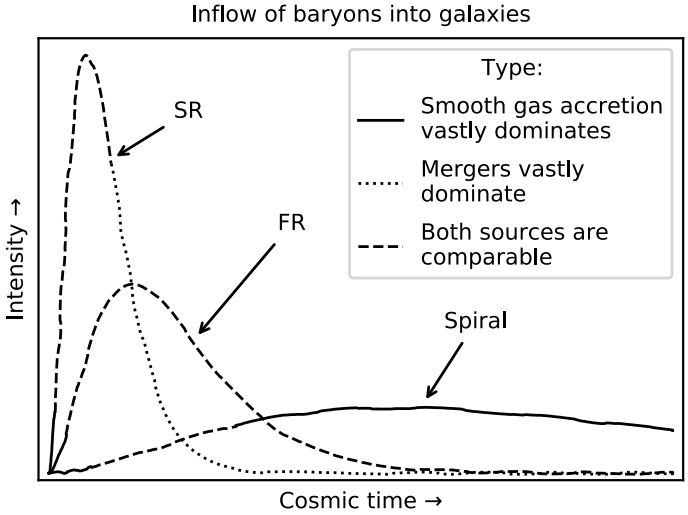

Fig. 4. Proposed assembly of different types of massive non-cluster galaxies. The curves schematically show the intensity and type of inflow of baryonic material into the galaxies over time.

radiation or gas outflows, a process that will preferentially remove the gas clouds on originally circular trajectories.

It is relevant to note that the gas-rich galaxies that we identified as the progenitors of SRs contain many giant gas clumps at high redshifts (Conselice et al. 2008; Elmegreen et al. 2009; Elmegreen & Elmegreen 2014; Cava et al. 2018). Their masses can reach up to $10^9\,M_\odot$ and are probably gravitationally self-bound (Bournaud et al. 2014; Tadaki et al. 2018; Fensch & Bournaud 2021). Their sizes are often comparable to the sizes of the most distant quiescent galaxies (∼1 kpc). Once their original host galaxy is accreted by the future SR and turns into tidal debris, these gas clouds move on randomly oriented orbits and contribute to the turbulent nature of gas around the future SRs.

At the redshift of around one, the formation of the SRs becomes much quieter. By the redshift of around two, the IGM filaments can no longer reach the center of the galaxy (Sect. 1). All arriving IGM is shock-heated so that it contributes to the hot halo. The frequency of mergers also decreases. The mergers become drier, because generally all massive galaxies become gas poorer in later cosmic epochs, and partly because of the quenching of the satellites in the hot halo of the main galaxy. Toward the redshift of zero, the frequency of mergers decreases even further. The possibility that SRs reside at the intersections of several filaments is strengthened by the fact that we found that with the decrease in the rotational support of the galaxies, the incidence of disturbed isophotes increases at a fixed mass and environmental density.

Here we come to the question of the formation of the cluster SRs. For the non-cluster SRs, we explained above the observational constraints by the SRs residing on intersections of cosmic filaments. But this is exactly the location of galaxy clusters and massive groups. The reasoning above thus hold true also for the central galaxies of these structures. Our assumption on the transformation of the kinematic structure by mergers alone has an interesting implication: the most massive galaxies are expected to experience the highest number of mergers partly because of the strength of their gravitational field and partly because they tend to be the central galaxies of their environments. In the end, the galaxy gains most of its mass by mergers. This object type differs from FRs and spirals by a relatively low fraction of stellar mass formed in situ. As galaxy groups and clusters merge together, the central SRs of the original structures can become noncentral galaxies of the new structures. The central SRs of groups and clusters are expected to experience, compared to the non-cluster SRs, many dry mergers (Dubinski 1998; Mihos 2003; De Lucia & Blaizot 2007). These produce the observed central photometric cores (e.g., Krajnović et al. 2020). One has to remember that even the central cluster galaxies were gas rich at early cosmic epochs and their rotational support might have already decreased to the current value by that time.

The formation of spirals seems to be the reverse of the formation of SRs. Most spiral galaxies do not have classical bulges and many do not contain detectable stellar halos (Kormendy et al. 2010; Peebles & Nusser 2010; Fisher & Drory 2011; Merritt et al. 2016). Such features are supposed to form mostly by mergers (Naab & Burkert 2003; Bournaud et al. 2005, 2007). This suggests that spirals form much more smoothly than ETGs: either completely by smooth accretion of the IGM (Sancisi et al. 2008), with a small contribution of minor mergers, or possibly by very gas-rich mergers at high redshifts. The low abundance of classical bulges in the observed spirals and the non-detection of the stellar halos are difficult to reproduce in current cosmological simulations of galaxy formation (Brooks 2016; Peebles 2020; Merritt et al. 2020). Spiral galaxies possess hot gas halos at least down to the stellar mass of $10^9\,M_\odot$ (Li et al. 2014, 2017) but cold gas streams are able to go through the halos (Dekel & Birnboim 2006).

The formation of FRs seems to be an interpolation between the formation of spirals and SRs. At the redshift of two to three, a typical FR faced fewer mergers than a typical SR and it grew more by in situ star formation. If the FR formed in a node of the cosmic web, then only a few filaments were joining in the node, or some of the filaments conducted much more material into the node than the other filaments. As the galaxy contained a lot of gas, it became gravitationally unstable (Tadaki et al. 2018) and formed massive giant clumps, which helped to dynamically heat the disk, make it thicker (Bournaud et al. 2014; Clarke et al. 2019), and produce a bulge (Noguchi 1999; Elmegreen et al. 2008; Hopkins et al. 2012). Mergers contribute to the thickening of the disk and to the formation of the bulge. Another source of thickening of the disk might be the strong bars that are hinted at by observations of submillimeter galaxies (Gullberg et al. 2019; Hodge et al. 2019; Dudzevičiūtė et al. 2020). Similarly to SRs, at $z = 1$ the merging activity has decreased, just as the inflow of fresh intergalactic gas. Because of the typically lower masses of the FRs, the gas is partly able to reach the galaxies. This is suggested by the presence of molecular gas in many galaxies in the given mass range (Young et al. 2011), or the non-negligible star formation implied by the reconstructed star formation histories (McDermid et al. 2015). In addition to mass quenching, star formation is further suppressed by morphological quenching and possibly other mechanisms. Our results from Sect. 7.2 suggest that the frequency of mergers is too low to explain the growth of radii of quenched galaxies with redshift and the discrepancy is the strongest for the galaxies with $M_{\rm JAM} < 10^{11}\,M_\odot$. Therefore, an additional mechanism is needed in these galaxies, such as gas outflows (Sect. 1). The need for outflows is hinted at also by the chemical composition of these galaxies (Trussler et al. 2020).

Figure 4 schematically shows curves of the proposed intensity and source of the accreted material by different types of galaxies as functions of the cosmic time. It is inspired by studies of stellar populations (Thomas et al. 2010; Pipino et al. 2013; McDermid et al. 2015; González Delgado et al. 2017) and the merging histories deduced above. The shapes of

the curves reflect the "galaxy downsizing" phenomenon (e.g., Neistein et al. 2006); in other words, that the high-mass galaxies form first and the less massive galaxies form later.

Not all galaxies follow the typical paths described above. For example, a spiral galaxy can be quenched by starvation and turn into an ETG after entering the hot gas halo of a galaxy cluster. The opposite transition is also possible. There are signatures that formerly passive galaxies acquire star-forming disks and become bulged spirals (de la Rosa et al. 2016). Mergers of special relative orbital configurations might lead to the formation of galaxies with two counter-rotating stellar populations. They would then be classified as SRs according to our accepted parametric criterion.

10. Summary and conclusions

Cosmological hydrodynamical simulations suggest that galaxies are formed with a high degree of rotational support, which decreases later. Mergers play a substantial role in the transformation. According to these simulations, this transition happened less than about 8 Gyr ago (redshift of one). In this paper we have investigated this transformation observationally, for ETGs outside of galaxy clusters, initially assuming that mergers are solely responsible for the transformation. To quantify how much the transformation proceeded in a given galaxy until now, we primarily made use of the parameter $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$ introduced by Emsellem et al. (2011) and called the rotational support here. By definition, SRs have a rotational support lower than 0.31, and FRs have a rotational support higher than this. We exploited multilinear regressions on data from the ATLAS$^{3\mathrm{D}}$ and MATLAS surveys in order to study the correlations of various parameters that are expected to be sensitive to mergers with the rotational support for galaxies of a fixed mass and environmental density. The results were then interpreted making use of the facts that every merger indicator has a different lifetime and is sensitive to a different type of merger (minor or major, wet or dry). It was crucial to study the correlations at a fixed mass and environmental density in order to account for the so-called confounding effect, that is, the fact that $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$ correlates with the galaxy mass and with the environmental density, and that different merger-sensitive parameters correlate with these quantities as well. We also considered other observational results from the literature, mostly from observations of the high-redshift Universe. We can summarize our results as follows:

1. The observations indeed agree with the hypothesis that the rotational support of galaxies is initially high and then it decreases due to successive mergers. At a fixed stellar mass and environmental density, the galaxies with a low rotational support more often contain tidal disturbances and kinematic substructures compared with the galaxies with a higher rotational support. In addition, the metallicity of the galaxies with a lower $\lambda^{\mathrm{N}}_{R_\mathrm{e}}$ is typically lower.
2. We found many pieces of evidence showing that the decrease in the rotational support happened preferentially toward the end of the first phase of the ETG assembly, when the galaxies still contained a substantial fraction of gas. This means around $z = 2$ or perhaps before, depending on the mass or surface density of the galaxy (Sect. 1). This is supported by the fact that the majority of present-day SRs do not show tidal disturbances, and that, at a fixed mass and environmental density, the galaxies with a lower rotational support do not have substantially different effective radii than galaxies with a higher rotational support. The mergers had to be typically wet, for example because at a fixed mass and environmental density, the galaxies with low rotational support have a higher abundance of α-elements, as if the mergers decreasing the kinematic stage were causing starbursts. The mergers also had to be minor to account for the fact that, at a fixed mass and environment density, galaxies with a low rotational support tend to have steeper metallicity gradients. The evidence for the early time of kinematic transformation is particularly strong for high-mass galaxies ($M_{\mathrm{JAM}} > 10^{11}\,M_\odot$). For them, observations by other authors of their probable progenitors at high redshifts indicate that they already had spherical shapes and their current stellar masses at $z = 1{-}2$. The timing of the establishment of the kinematic morphology coincides with that of the photometric morphology (i.e., the *Hubble* sequence). At that same epoch, ordered motions of gas prevailed over chaotic motions in star-forming galaxies.
   We found that the kinematic transformation of ETGs happened earlier than simulations predict. This agrees with the independent findings of other works that the formation of ETGs in cosmological simulations is too extended.
3. Galaxies with a lower rotational support still experience more galaxy interactions today than the galaxies with a higher rotational support at a fixed mass and environmental density. They seem to live in environments where interactions are more frequent. Such recent interactions, however, are rare and do not greatly influence the properties of the galaxies.
4. From the measured frequency and estimated survival time of the tidal disturbances observed in deep images, we estimated a merger rate of at most $0.23\,\mathrm{Gyr}^{-1}$ for galaxies with $M_{\mathrm{JAM}} > 10^{11}\,M_\odot$, or $0.11\,\mathrm{Gyr}^{-1}$ for galaxies with $M_{\mathrm{JAM}} < 10^{11}\,M_\odot$. We inferred a frequency of major mergers of $0.02\,\mathrm{Gyr}^{-1}$. These numbers agree with what has been estimated from the frequency of close galaxy pairs in the literature. The frequency of mergers seems to be insufficient to explain the growth of the effective radius of ETGs with time, and therefore additional mechanisms are necessary to account for it.
5. We have proposed a picture of formation of FRs, SRs, and massive spirals that seems to account for all observational constraints considered in this paper (Sect. 9). The formation of a typical SR is fast and involves numerous mergers. Spirals typically assemble gradually and smoothly, with most material gained through the accretion of the IGM. The formation of typical FRs lies between these two extremes.

Acknowledgements. We thank the anonymous referee for a constructive report. We thank D. Krajnović, M. Cappellari and I. Ebrová for valuable comments and discussions. MB is grateful for the financial support by Cercle Gutenberg. MB acknowledges the support from the Polish National Science Centre under the grant 2017/26/D/ST9/00449.

References

Afruni, A., Fraternali, F., & Pezzulli, G. 2019, A&A, 625, A11
Algorry, D. G., Navarro, J. F., Abadi, M. G., et al. 2014, MNRAS, 437, 3596
Amorisco, N. C. 2015, MNRAS, 450, 575
Amorisco, N. C. 2017, MNRAS, 464, 2882
Ardila, F., Alatalo, K., Lanz, L., et al. 2018, ApJ, 863, 28
Atkinson, A. M., Abraham, R. G., & Ferguson, A. M. N. 2013, ApJ, 765, 28
Bahé, Y. M., Barnes, D. J., Dalla Vecchia, C., et al. 2017, MNRAS, 470, 4186
Beasley, M. A., Trujillo, I., Leaman, R., & Montes, M. 2018, Nature, 555, 483
Bender, R., & Surma, P. 1992, A&A, 258, 250
Bennett, C. L., Larson, D., Weiland, J. L., & Hinshaw, G. 2014, ApJ, 794, 135
Bernardi, M., Domínguez Sánchez, H., Brownstein, J. R., Drory, N., & Sheth, R. K. 2019, MNRAS, 489, 5633
Bílek, M. 2016, ArXiv e-prints [arXiv:1601.01240]
Bílek, M., Thies, I., Kroupa, P., & Famaey, B. 2018, A&A, 614, A59

Bílek, M., Samurović, S., & Renaud, F. 2019a, A&A, 625, A32
Bílek, M., Thies, I., Kroupa, P., & Famaey, B. 2019b, ArXiv e-prints [arXiv:1908.07537]
Bílek, M., Duc, P.-A., Cuillandre, J.-C., et al. 2020, MNRAS, 498, 2138
Bitsakis, T., Dultzin, D., Ciesla, L., et al. 2016, MNRAS, 459, 957
Bitsakis, T., Sánchez, S. F., Ciesla, L., et al. 2019, MNRAS, 483, 370
Blanton, M. R., & Moustakas, J. 2009, ARA&A, 47, 159
Bournaud, F., Jog, C. J., & Combes, F. 2005, A&A, 437, 69
Bournaud, F., Jog, C. J., & Combes, F. 2007, A&A, 476, 1179
Bournaud, F., Perret, V., Renaud, F., et al. 2014, ApJ, 780, 57
Brooks, A., & Christensen, C. 2016, in Bulge Formation via Mergers in Cosmological Simulations, eds. E. Laurikainen, R. Peletier, & D. Gadotti, 418, 317
Brough, S., van de Sande, J., Owers, M. S., et al. 2017, ApJ, 844, 59
Cappellari, M. 2008, MNRAS, 390, 71
Cappellari, M. 2016, ARA&A, 54, 597
Cappellari, M., Emsellem, E., Bacon, R., et al. 2007, MNRAS, 379, 418
Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011a, MNRAS, 413, 813
Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011b, MNRAS, 416, 1680
Cappellari, M., Scott, N., Alatalo, K., et al. 2013, MNRAS, 432, 1709
Carleton, T., Guo, Y., Nayyeri, H., et al. 2020, MNRAS, 491, 2822
Carnall, A. C., Walker, S., McLure, R. J., et al. 2020, MNRAS, 496, 695
Carnall, A. C., McLeod, D. J., McLure, R. J., et al. 2023, MNRAS, 520, 3974
Cava, A., Schaerer, D., Richard, J., et al. 2018, Nat. Astron., 2, 76
Chang, Y.-Y., van der Wel, A., Rix, H.-W., et al. 2013, ApJ, 762, 83
Chu, A., Durret, F., & Márquez, I. 2021, A&A, 649, A42
Chu, A., Sarron, F., Durret, F., & Márquez, I. 2022, A&A, 666, A54
Clarke, A. J., Debattista, V. P., Nidever, D. L., et al. 2019, MNRAS, 484, 3476
Cole, J., Bezanson, R., van der Wel, A., et al. 2020, ApJ, 890, L25
Conselice, C. J., Rajgor, S., & Myers, R. 2008, MNRAS, 386, 909
Costantin, L., Pérez-González, P. G., Méndez-Abreu, J., et al. 2021, ApJ, 913, 125
Côté, P., Marzke, R. O., & West, M. J. 1998, ApJ, 501, 554
Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, ApJ, 626, 680
Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, ApJ, 695, 101
Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2
de la Rosa, I. G., La Barbera, F., Ferreras, I., et al. 2016, MNRAS, 457, 1916
De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2
De Lucia, G., Fontanot, F., & Hirschmann, M. 2017, MNRAS, 466, L88
D'Eugenio, F., Houghton, R. C. W., Davies, R. L., & Dalla Bontà, E. 2013, MNRAS, 429, 1258
Di Matteo, P., Pipino, A., Lehnert, M. D., Combes, F., & Semelin, B. 2009a, A&A, 499, 427
Di Matteo, P., Jog, C. J., Lehnert, M. D., Combes, F., & Semelin, B. 2009b, A&A, 501, L9
Dubinski, J. 1998, ApJ, 502, 141
Duc, P.-A., Cuillandre, J.-C., Karabal, E., et al. 2015, MNRAS, 446, 120
Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2020, MNRAS, 494, 3828
Ebrová, I., Bílek, M., Yıldız, M. K., & Eliášek, J. 2020, A&A, 634, A73
Ebrová, I., Łokas, E. L., & Eliášek, J. 2021, A&A, 647, A103
Elmegreen, D. M., & Elmegreen, B. G. 2014, ApJ, 781, 11
Elmegreen, B. G., Bournaud, F., & Elmegreen, D. M. 2008, ApJ, 688, 67
Elmegreen, D. M., Elmegreen, B. G., Marcus, M. T., et al. 2009, ApJ, 701, 306
Emsellem, E., Cappellari, M., Krajnović, D., et al. 2007, MNRAS, 379, 401
Emsellem, E., Cappellari, M., Krajnović, D., et al. 2011, MNRAS, 414, 888
Ene, I., Ma, C.-P., Veale, M., et al. 2018, MNRAS, 479, 2810
Estrada-Carpenter, V., Papovich, C., Momcheva, I., et al. 2020, ApJ, 898, 171
Faber, S. M., Tremaine, S., Ajhar, E. A., et al. 1997, AJ, 114, 1771
Faisst, A. L., Capak, P. L., Emami, N., Tacchella, S., & Larson, K. L. 2019, ApJ, 884, 133
Fan, L., Lapi, A., De Zotti, G., & Danese, L. 2008, ApJ, 689, L101
Fan, L., Lapi, A., Bressan, A., et al. 2010, ApJ, 718, 1460
Fensch, J., & Bournaud, F. 2021, MNRAS, 505, 3579
Fisher, D. B., & Drory, N. 2011, ApJ, 733, L47
Forrest, B., Annunziatella, M., Wilson, G., et al. 2020, ApJ, 890, L1
Foster, C., van de Sande, J., D'Eugenio, F., et al. 2017, MNRAS, 472, 966
Ganda, K., Peletier, R. F., McDermid, R. M., et al. 2007, MNRAS, 380, 506
Gavazzi, G., Consolandi, G., Pedraglio, S., et al. 2018, A&A, 611, A28
Ghigna, S., Moore, B., Governato, F., et al. 1998, MNRAS, 300, 146
González Delgado, R. M., Pérez, E., Fernand, Cid, es, R., , et al. 2017, A&A, 607, A128
Graham, A. W. 2013, Planets, Stars and Stellar Systems, eds. T. D. Oswalt, & W. C. Keel (Springer Science+Business Media Dordrecht), 6, 91
Graham, M. T., Cappellari, M., Li, H., et al. 2018, MNRAS, 477, 4711
Graham, M. T., Cappellari, M., Bershady, M. A., & Drory, N. 2019a, ArXiv e-prints [arXiv:1911.06103]
Graham, M. T., Cappellari, M., Bershady, M. A., & Drory, N. 2019b, ArXiv e-prints [arXiv:1910.05139]
Greene, J. E., Leauthaud, A., Emsellem, E., et al. 2017, ApJ, 851, L33
Gullberg, B., Smail, I., Swinbank, A. M., et al. 2019, MNRAS, 490, 4956
Haan, S., Armus, L., Surace, J. A., et al. 2013, MNRAS, 434, 1264
Harris, W. E. 2001, Globular Cluster Systems, 28 (Berlin Heidelberg): (Springer-Verlag). 223
Harrison, C. M. 2017, Nat. Astron., 1, 0165
Hau, G. K. T., & Thomson, R. C. 1994, MNRAS, 270, L23
Hendel, D., & Johnston, K. V. 2015, MNRAS, 454, 2472
Hill, A. R., Muzzin, A., Franx, M., & Marchesini, D. 2017, ApJ, 849, L26
Hill, A. R., van der Wel, A., Franx, M., et al. 2019, ApJ, 871, 76
Hilz, M., Naab, T., & Ostriker, J. P. 2013, MNRAS, 429, 2924
Hirschmann, M., Naab, T., Ostriker, J. P., et al. 2015, MNRAS, 449, 528
Hodge, J. A., Smail, I., Walter, F., et al. 2019, ApJ, 876, 130
Hopkins, P. F., Cox, T. J., Dutta, S. N., et al. 2009a, ApJS, 181, 135
Hopkins, P. F., Bundy, K., Murray, N., et al. 2009b, MNRAS, 398, 898
Hopkins, P. F., Kereš, D., Murray, N., Quataert, E., & Hernquist, L. 2012, MNRAS, 427, 968
Houghton, R. C. W., Davies, R. L., D'Eugenio, F., et al. 2013, MNRAS, 436, 19
Huang, S., Leauthaud, A., Greene, J. E., et al. 2018, MNRAS, 475, 3348
Huertas-Company, M., Pérez-González, P. G., Mei, S., et al. 2015, ApJ, 809, 95
Ishibashi, W., Fabian, A. C., & Canning, R. E. A. 2013, MNRAS, 431, 2350
Kado-Fong, E., Greene, J. E., Hendel, D., et al. 2018, ApJ, 866, 103
Karademir, G. S., Remus, R.-S., Burkert, A., et al. 2019, MNRAS, 487, 318
Kawinwanichakij, L., Papovich, C., Ciardullo, R., et al. 2020, ApJ, 892, 7
Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2014, MNRAS, 444, 1647
Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2
Kobayashi, C. 2004, MNRAS, 347, 740
Kormendy, J., Drory, N., Bender, R., & Cornell, M. E. 2010, ApJ, 723, 54
Krajnović, D., Emsellem, E., Cappellari, M., et al. 2011, MNRAS, 414, 2923
Krajnović, D., Ural, U., Kuntschner, H., et al. 2020, A&A, 635, A129
Kuntschner, H., Emsellem, E., Bacon, R., et al. 2010, MNRAS, 408, 97
Lagos, C., d. P., Schaye, J., Bahé, Y., , et al. 2018, MNRAS, 476, 4327
Lagos, C., d. P., Emsellem, E., van de Sande, J., , et al. 2022, MNRAS, 509, 4372
Larson, R. B., Tinsley, B. M., & Caldwell, C. N. 1980, ApJ, 237, 692
Lauer, T. R., Ajhar, E. A., Byun, Y. I., et al. 1995, AJ, 110, 2622
Lee, B., Giavalisco, M., Williams, C. C., et al. 2013, ApJ, 774, 47
Lelli, F., McGaugh, S. S., Schombert, J. M., & Pawlowski, M. S. 2017, ApJ, 836, 152
Li, J.-T., Crain, R. A., & Wang, Q. D. 2014, MNRAS, 440, 859
Li, J.-T., Bregman, J. N., Wang, Q. D., et al. 2017, ApJS, 233, 20
Li, H., Mao, S., Cappellari, M., et al. 2018a, MNRAS, 476, 1765
Li, H., Mao, S., Cappellari, M., et al. 2018b, ApJ, 863, L19
Liu, D., Daddi, E., Dickinson, M., et al. 2018, ApJ, 853, 172
Lustig, P., Strazzullo, V., D'Eugenio, C., et al. 2021, MNRAS, 501, 2659
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Maiolino, R., & Mannucci, F. 2019, A&A Rev., 27, 3
Malin, D. F., & Carter, D. 1983, ApJ, 274, 534
Man, A. W. S., Zirm, A. W., & Toft, S. 2016, ApJ, 830, 89
Mancillas, B., Duc, P.-A., Combes, F., et al. 2019, A&A, 632, A122
Mancini, C., Daddi, E., Juneau, S., et al. 2019, MNRAS, 489, 1265
Martig, M., Bournaud, F., Teyssier, R., & Dekel, A. 2009, ApJ, 707, 250
Martig, M., Crocker, A. F., Bournaud, F., et al. 2013, MNRAS, 432, 1914
Martín-Navarro, I., Vazdekis, A., Falcón-Barroso, J., et al. 2018, MNRAS, 475, 3700
Matteucci, F. 2014, Saas-Fee Adv. Course, 37, 145
McDermid, R. M., Alatalo, K., Blitz, L., et al. 2015, MNRAS, 448, 3484
Merlin, E., Fortuni, F., Torelli, M., et al. 2019, MNRAS, 490, 3309
Merritt, A., van Dokkum, P., Abraham, R., & Zhang, J. 2016, ApJ, 830, 62
Merritt, A., Pillepich, A., van Dokkum, P., et al. 2020, MNRAS, 495, 4570
Mihos, C. 2003, ArXiv e-prints [arXiv:astro-ph/0305512]
Milgrom, M. 1983, ApJ, 270, 365
Milosavljević, M., Merritt, D., Rest, A., & van den Bosch, F. C. 2002, MNRAS, 331, L51
Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution (UK: Cambridge University Press)
Montuori, M., Di Matteo, P., Lehnert, M. D., Combes, F., & Semelin, B. 2010, A&A, 518, A56
Mortlock, A., Conselice, C. J., Hartley, W. G., et al. 2013, MNRAS, 433, 1185
Mundy, C. J., Conselice, C. J., Duncan, K. J., et al. 2017, MNRAS, 470, 3507
Naab, T., & Burkert, A. 2003, ApJ, 597, 893
Naab, T., & Ostriker, J. P. 2009, ApJ, 690, 1452
Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, ApJ, 699, L178
Naab, T., Oser, L., Emsellem, E., et al. 2014, MNRAS, 444, 3357
Nagashima, M., Lacey, C. G., Okamoto, T., et al. 2005, MNRAS, 363, L31
Navarro-González, J., Ricciardelli, E., Quilis, V., & Vazdekis, A. 2013, MNRAS, 436, 3507
Neistein, E., van den Bosch, F. C., & Dekel, A. 2006, MNRAS, 372, 933
Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, ApJ, 746, 162

Newman, A. B., Belli, S., Ellis, R. S., & Patel, S. G. 2018, ApJ, 862, 126
Nipoti, C., Londrillo, P., & Ciotti, L. 2006, MNRAS, 370, 681
Noguchi, M. 1999, ApJ, 514, 77
Noirot, G., Sawicki, M., Abraham, R., et al. 2022, MNRAS, 512, 3566
Okamoto, T., Nagashima, M., Lacey, C. G., & Frenk, C. S. 2017, MNRAS, 464, 4866
Onodera, M., Carollo, C. M., Renzini, A., et al. 2015, ApJ, 808, 161
Oser, L., Ostriker, J. P., Naab, T., Johansson, P. H., & Burkert, A. 2010, ApJ, 725, 2312
Oyarzún, G. A., Bundy, K., Westfall, K. B., et al. 2019, ApJ, 880, 111
Parikh, T., Thomas, D., Maraston, C., et al. 2021, MNRAS, 502, 5508
Patil, M. K., Pandey, S. K., Sahu, D. K., & Kembhavi, A. 2007, A&A, 461, 103
Peebles, P. J. E. 2020, MNRAS, 498, 4386
Peebles, P. J. E., & Nusser, A. 2010, Nature, 465, 565
Peletier, R. F., Falcón-Barroso, J., Bacon, R., et al. 2007, MNRAS, 379, 445
Peng, Y.-J., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193
Peng, Y., Maiolino, R., & Cochrane, R. 2015, Nature, 521, 192
Penoyre, Z., Moster, B. P., Sijacki, D., & Genel, S. 2017, MNRAS, 468, 3883
Perez, J., Michel-Dansac, L., & Tissera, P. B. 2011, MNRAS, 417, 580
Pierce, M., Brodie, J. P., Forbes, D. A., et al. 2005, MNRAS, 358, 419
Pipino, A., Devriendt, J. E. G., Thomas, D., Silk, J., & Kaviraj, S. 2009, A&A, 505, 1075
Pipino, A., D'Ercole, A., Chiappini, C., & Matteucci, F. 2010, MNRAS, 407, 1347
Pipino, A., Calura, F., & Matteucci, F. 2013, MNRAS, 432, 2541
Pop, A.-R., Pillepich, A., Amorisco, N. C., & Hernquist, L. 2018, MNRAS, 480, 1715
Proctor, R. N., & Sansom, A. E. 2002, MNRAS, 333, 517
Qu, Y., Di Matteo, P., Lehnert, M., van Driel, W., & Jog, C. J. 2010, A&A, 515, A11
Read, J. I., & Trentham, N. 2005, Phil. Trans. R. Soc. London Ser. A, 363, 2693
Rodriguez-Gomez, V., Genel, S., Vogelsberger, M., et al. 2015, MNRAS, 449, 49
Rowlands, K., Heckman, T., Wild, V., et al. 2018, MNRAS, 480, 2544
Rupke, D. S. N., Kewley, L. J., & Barnes, J. E. 2010, ApJ, 710, L156
Sancisi, R., Fraternali, F., Oosterloo, T., & van der Hulst, T. 2008, A&A Rev., 15, 189
Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
Schreiber, C., Glazebrook, K., Nanayakkara, T., et al. 2018, A&A, 618, A85
Schulze, F., Remus, R.-S., Dolag, K., et al. 2018, MNRAS, 480, 4636
Schweizer, F. 1982, ApJ, 252, 455
Scott, N., Cappellari, M., Davies, R. L., et al. 2013, MNRAS, 432, 1894
Scott, N., Brough, S., Croom, S. M., et al. 2017, MNRAS, 472, 2833
Serra, P., & Trager, S. C. 2007, MNRAS, 374, 769
Serra, P., Oser, L., Krajnović, D., et al. 2014, MNRAS, 444, 3388
Simons, R. C., Kassin, S. A., Weiner, B. J., et al. 2017, ApJ, 843, 46
Sola, E., Duc, P.-A., Richards, F., et al. 2022, A&A, 662, A124
Spolaor, M., Proctor, R. N., Forbes, D. A., & Couch, W. J. 2009, ApJ, 691, L138
Statler, T. S. 1991, ApJ, 382, L11
Stevans, M. L., Finkelstein, S. L., Kawinwanichakij, L., et al. 2021, ApJ, 921, 58
Tacchella, S., Conroy, C., Faber, S. M., et al. 2022, ApJ, 926, 134
Tadaki, K., Iono, D., Yun, M. S., et al. 2018, Nature, 560, 613
Taylor, P., & Kobayashi, C. 2017, MNRAS, 471, 3856
Thomas, D., Greggio, L., & Bender, R. 1999, MNRAS, 302, 537
Thomas, D., Maraston, C., & Bender, R. 2002, Ap&SS, 281, 371
Thomas, D., Bender, R., Hopp, U., Maraston, C., & Greggio, L. 2003, Ap&SS, 284, 599
Thomas, D., Maraston, C., Schawinski, K., Sarzi, M., & Silk, J. 2010, MNRAS, 404, 1775
Toft, S., van Dokkum, P., Franx, M., et al. 2007, ApJ, 671, 285
Toomre, A. 1977, in Evolution of Galaxies and Stellar Populations, eds. B. M. Tinsley, D. C. Larson, & R. B. Gehret, 401
Trujillo, I., Förster Schreiber, N. M., Rudnick, G., et al. 2006, ApJ, 650, 18
Trujillo, I., Ferreras, I., & de La Rosa, I. G. 2011, MNRAS, 415, 3903
Trussler, J., Maiolino, R., Maraston, C., et al. 2020, MNRAS, 491, 5406
van de Sande, J., Croom, S. M., Bland-Hawthorn, J., et al. 2021a, MNRAS, 508, 2307
van de Sande, J., Vaughan, S. P., Cortese, L., et al. 2021b, MNRAS, 505, 3078
van der Vlugt, D., & Costa, T. 2019, MNRAS, 490, 4918
van Dokkum, P. G., Kriek, M., & Franx, M. 2009, Nature, 460, 717
Veale, M., Ma, C.-P., Greene, J. E., et al. 2017, MNRAS, 471, 1428
Ventou, E., Contini, T., Bouché, N., et al. 2017, A&A, 608, A9
Ventou, E., Contini, T., Bouché, N., et al. 2019, A&A, 631, A87
Vincenzo, F., Kobayashi, C., & Taylor, P. 2018, MNRAS, 480, L38
Wang, J., Hammer, F., Athanassoula, E., et al. 2012, A&A, 538, A121
Weijmans, A.-M., de Zeeuw, P. T., Emsellem, E., et al. 2014, MNRAS, 444, 3340
Wilkins, S. M., Lovell, C. C., & Stanway, E. R. 2019, MNRAS, 490, 5359
Wright, E. L. 2006, PASP, 118, 1711
Yıldırım, A., van den Bosch, R. C. E., van de Ven, G., et al. 2017, MNRAS, 468, 4216
Yıldız, M. K., Peletier, R. F., Duc, P. A., & Serra, P. 2020, A&A, 636, A8
Young, L. M., Bureau, M., Davis, T. A., et al. 2011, MNRAS, 414, 940
Young, L. M., Krajnović, D., Duc, P.-A., & Serra, P. 2020, MNRAS, 495, 1433
Zibetti, S., Gallazzi, A. R., Hirschmann, M., et al. 2020, MNRAS, 491, 3562

Appendix A: Correlations between the merger-sensitive parameters, rotational support, galaxy mass, and environmental density

In this section we explore whether the galaxy mass $M_{\rm JAM}$, environmental density ρ_{10}, and the rotational support $\lambda^{\rm N}_{R_e}$ correlate in our sample with each other, and with the investigated merger-sensitive parameters. We find evidence of such correlations, which is a warning that one should be aware of the confounding effect when investigating the influence of mergers on the rotational support through the merger-sensitive parameters. In addition, we also investigate the differences in the merger sensitive parameters for SRs and FRs.

We explored how the mentioned quantities correlate exploiting Pearson's correlation coefficient. The results are listed in Table A.1. In the first three main columns, the sub-column "sign" signifies the sign of the correlation. The sub-column "p" shows the p-value of every correlation, (i.e., the probability that there is actually no correlation). All other relevant numbers pertaining the correlations are provided in Appendix C.

To summarize the results in Table A.1, we found the following statistically significant correlations between the merger-sensitive parameters and the rotational support. With increasing rotational support, a galaxy has a decreasing probability of possessing disturbed outer isophotes, containing some form of tidal disturbances, hosting a KDC or a CRC, and having a smaller effective radius. With an increasing rotational support, a galaxy has an increasing probability of not containing any kinematical substructures, of having a more cored central photometric profile, and a higher metallicity.

We compare the merger-sensitive parameters of FRs and SRs in the fifth main column of Table A.1. The first sub-column indicates the average value of the given parameter for SRs. The second sub-column shows the analogous values for FRs. The numbers in parenthesis indicate the uncertainty in the last digit of the mean value. The last column is the probability that the statistical distributions of a given parameter are the same between the two groups of galaxies. This was obtained through the two-sample KS test. The SRs show, with respect to FRs, a higher incidence and frequency of shells, a higher incidence of KDCs and CRCs, a lower incidence of regular kinematic fields, and they have larger effective radii. All of these differences can be expected simply by the fact that SRs are, on average, more massive than FRs.

The correlations with the mass of the galaxy can be read in the third main column of Table A.1. The more massive a galaxy is, the more it is probable that it will host tidal features of any type, or that it will possess disturbed outer isophotes. It will also be more probable that it will host a higher number of shells and streams. More massive galaxies also tend to have a larger effective radius, older stellar populations, a higher metallicity, and a higher content of α-elements. On the contrary, more massive galaxies tend to have more negative age gradients.

Table A.1. Correlations of the investigated merger-sensitive parameters with different properties of the galaxies.

	$\lambda^{\rm N}_{R_e}$		$\log M_{\rm JAM}$		$\log \rho_{10}$		SR/FR		
Parameter	sign	p [%]	sign	p [%]	sign	p [%]	Mean SR	Mean FR	p [%]
Shells	-	8.7	+	**1.0**	+	19	0.5(2)	0.26(5)	**3.1**
Streams	-	16	+	**0.53**	+	16	0.4(1)	0.25(5)	≥ 25
Tails	-	18	+	60	+	13	0.2(1)	0.22(5)	≥ 25
Disturbed isophotes	-	**0.090**	+	**0.029**	+	7.8	0.8(2)	0.52(6)	24
Any TF	-	10	+	**1.2**	+	**4.6**	0.7(2)	0.52(7)	22
Any TF or DI	-	**2.8**	+	**6.3E − 3**	+	**4.4**	0.9(2)	0.69(7)	22
Shell fr.	-	8.7	+	**4.8**	+	29	1.0(4)	0.5(1)	**1.4**
Stream fr.	-	19	+	**0.45**	+	20	0.3(1)	0.20(5)	≥ 25
Tail fr.	-	15	+	50	+	14	0.11(7)	0.16(4)	≥ 25
Dust	+	46	-	43	-	82	0.12(9)	0.26(5)	15
KDC	-	**6.1E − 3**	+	60	-	92	0.24(9)	0.007(7)	$\leq$ **0.10**
CRC	-	**6.4E − 3**	+	12	-	92	0.24(9)	0(0)	$\leq$ **0.10**
KT	-	19	+	6.1	+	11	0(0)	0.013(9)	≥ 25
$2\,\sigma$	-	**0.44**	-	**4.7**	+	67	0.08(5)	0.03(1)	11
No kin. feature	+	**0.047**	-	6.0	-	17	0.32(9)	0.80(3)	$\leq$ **0.10**
$\log R_e$	-	**1.3**	+	**1.5E − 27**	+	33	0.60(4)	0.43(2)	**0.11**
Core γ'	+	**4.1**	-	**0.88E − 6**	-	59	0.59(8)	0.65(5)	8.4
SFH age	-	14	+	**0.035E − 15**	+	33	10.9(4)	9.9(3)	13
SSP age	-	22	+	**2.2E − 12**	+	42	9.6(7)	8.5(4)	22
SSP age gr.	+	68	-	**1.1E − 3**	+	20	0.09(5)	0.18(2)	5.1
SSP Z	+	**0.39**	+	**0.25E − 6**	+	40	-0.24(3)	-0.23(1)	≥ 25
SSP Z gr.	+	6.9	+	14	-	30	-0.36(4)	-0.35(2)	≥ 25
SSP α	-	87	+	**1.0**	-	31	0.19(2)	0.214(8)	≥ 25
SSP α gr.	+	48	-	37	+	59	-0.00(3)	0.05(1)	13

Notes. The sign sub-columns show the sign of the correlation, while the p subcolumns show the probability that there is actually no correlation. The following columns give the mean value of the merger indicator for galaxies sorted into SRs and FRs. The number in parenthesis indicates the uncertainty in the last digit of the mean. The last sub-column gives the probability that the statistical distributions of the given parameter is the same for SRs and FRs.

The correlations of merger-sensitive parameters with the density of environment are shown in the fourth main column of Table A.1. We remind the reader that we do not have galaxies in clusters in our sample. We found two correlations that are statistically significant. Galaxies in denser environments have a higher probability of hosting at least some type of tidal features or a tidal disturbance.

For our sample, we found a significant correlation of the rotational support with $\log M_{\rm JAM}$ having a Pearson coefficient of -0.16 with a p-value of 3.6%. This agrees with the dedicated studies (see Sect. 1). For the correlation of $\log \rho_{10}$ with the rotational support, we found a Pearson coefficient of -0.06 with a p-value of 41%. In agreement with a previous study of the ATLAS$^{\rm 3D}$ sample (Cappellari et al. 2011b), we thus found a very weak correlation between kinematics and environment within our subsample of mostly group galaxies. A significant trend between kinematics and environment was previously reported for the Virgo cluster galaxies (see Figs. 6-7 of Cappellari et al. 2011b). For our sample, the correlation of $\log M_{\rm JAM}$ and $\log \rho_{10}$ has a Pearson coefficient of 0.01 and a p-value of 90%.

Appendix B: Supplementary figures and data

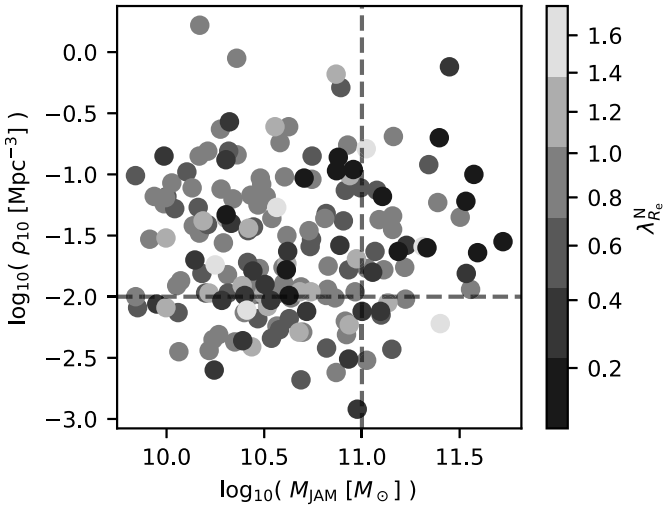

Fig. B.1. Rotational support as a function of $M_{\rm JAM}$ and ρ_{10}. The dashed lines indicate the cuts that we used to define the low-(high-) mass or density subsamples.

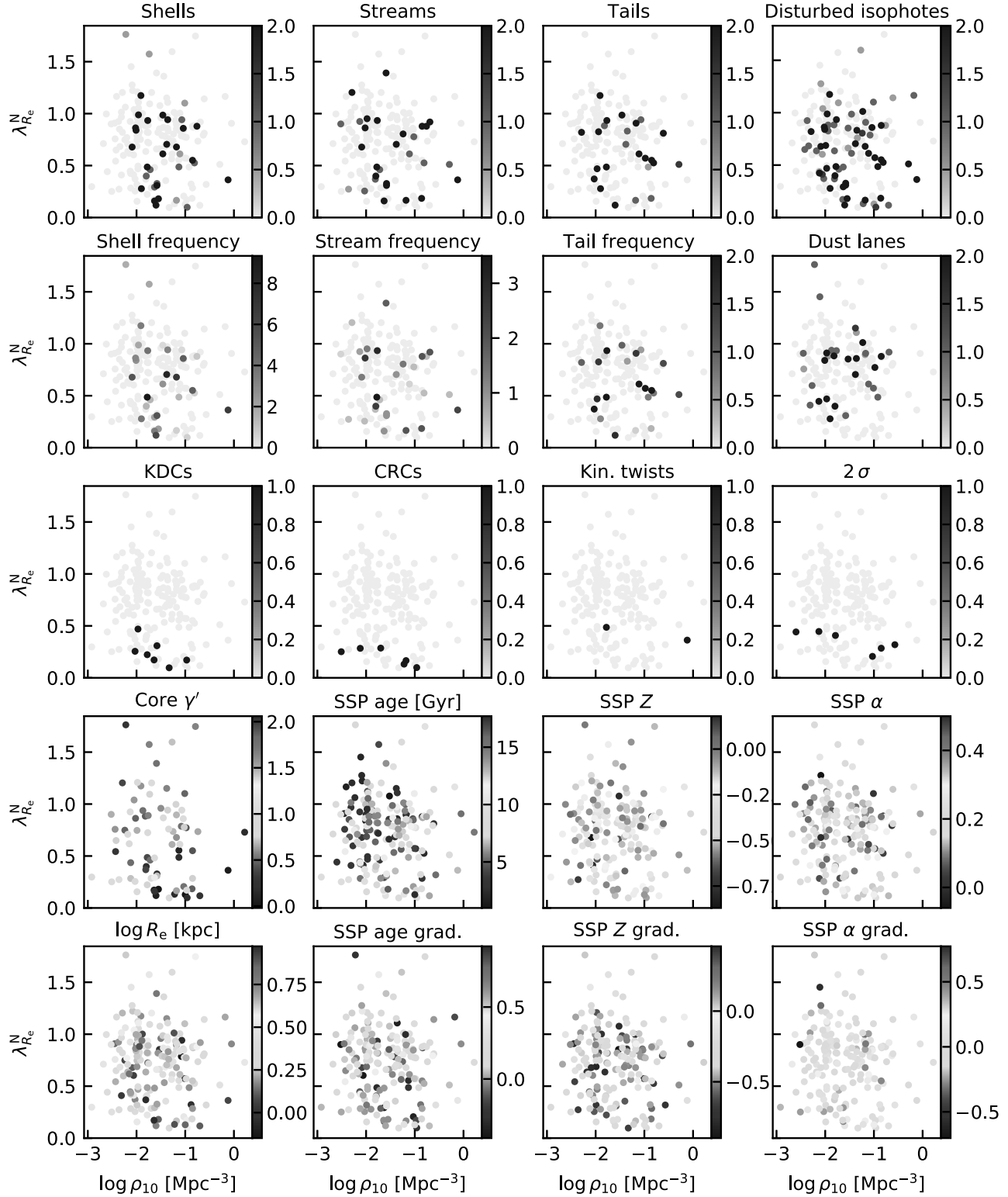

Fig. B.2. Investigated merger-sensitive parameters as functions of the rotational support (vertical axis in each tile) and the ρ_{10} environmental density (horizontal axis in each tile). We point out again that our sample does not contain galaxies in clusters. See Fig. 1 for an explanation of the meaning of the colors of the points.

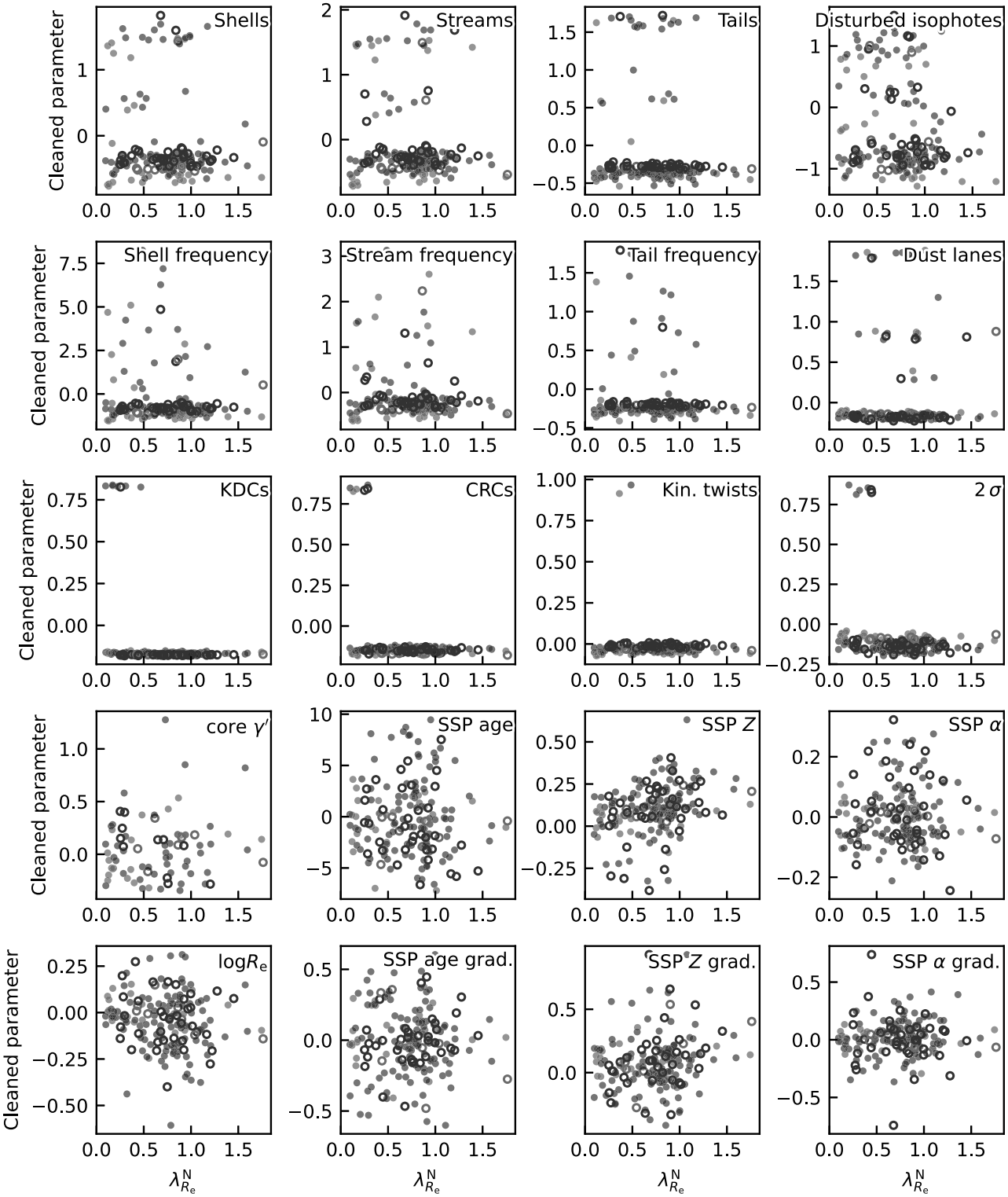

Fig. B.3. Relation of the cleaned merger-sensitive parameters (vertical axis in each tile) and the rotational support (horizontal axis in each tile). Red points show galaxies with $\log M_{\rm JAM} \geq 11$. Blue points show galaxies with $\log M_{\rm JAM} < 11$. Full points show galaxies in environmental densities $\log \rho_{10} \geq -2$. Empty points show galaxies in environmental densities $\log \rho_{10} < -2$.

Appendix C: Linear fits

Table C.1. Linear fits of the characteristics of the galaxies in the form *[parameter]* = $b + ax$, where x is either the rotational support, $\log M_{\rm JAM}$, or $\log \rho_{10}$.

	Rotational support			$\log M_{\rm JAM}$			$\log \rho_{10}$		
Parameter	b	a	r	b	a	r	b	a	r
Shells	0.5(1)	−0.3(1)	-13	−3(1)	0.3(1)	19	0.5(1)	0.11(9)	9.9
Streams	0.4(1)	−0.2(1)	-11	−3(1)	0.3(1)	21	0.5(1)	0.12(8)	11
Tails	0.4(1)	−0.2(1)	-10	−0(1)	0.1(1)	4.0	0.4(1)	0.12(8)	11
Disturbed isophotes	1.0(1)	−0.6(2)	-25	−5(1)	0.5(1)	27	0.8(2)	0.2(1)	13
Any TF	0.8(2)	−0.3(2)	-12	−3(2)	0.4(1)	19	0.9(2)	0.2(1)	15
Any TF or DI	1.0(2)	−0.4(2)	-17	−6(2)	0.6(1)	30	1.1(2)	0.2(1)	15
Shell fr.	1.0(3)	−0.6(4)	-13	−5(3)	0.6(3)	15	0.9(4)	0.2(2)	8.0
Stream fr.	0.3(1)	−0.2(1)	-9.9	−3(1)	0.3(1)	21	0.4(1)	0.10(8)	9.7
Tail fr.	0.27(8)	−0.2(1)	-11	−0.4(9)	0.06(8)	5.1	0.3(1)	0.09(6)	11
Dust	0.2(1)	0.1(1)	5.6	1(1)	−0.1(1)	-6.0	0.2(1)	−0.02(8)	-1.8
KDC	0.17(3)	−0.17(4)	-30	−0.2(4)	0.02(4)	4.0	0.04(4)	−0.00(3)	-0.78
CRC	0.15(3)	−0.16(4)	-30	−0.5(3)	0.05(3)	12	0.03(4)	−0.00(2)	-0.74
KT	0.03(2)	−0.03(2)	-9.9	−0.4(2)	0.04(2)	14	0.05(2)	0.02(1)	12
$2\,\sigma$	0.12(3)	−0.12(4)	-21	0.7(3)	−0.06(3)	-15	0.05(4)	0.01(2)	3.2
No kin. feature	0.48(8)	0.3(1)	26	2.3(8)	−0.15(8)	-14	0.6(1)	−0.08(6)	-10
$\log R_e$	0.54(4)	−0.12(5)	-19	−3.4(3)	0.36(3)	72	0.50(5)	0.03(3)	7.4
Core γ'	0.50(8)	0.2(1)	23	5.9(8)	−0.49(8)	-59	0.6(1)	−0.04(7)	-6.2
SFH age	10.8(5)	−1.0(7)	-11	−36(5)	4.3(4)	61	10.7(7)	0.4(4)	7.3
SSP age	9.5(8)	−1(1)	-9.3	−49(7)	5.4(6)	54	9(1)	0.5(6)	6.2
SSP age gr.	0.15(4)	0.02(6)	3.1	2.2(4)	−0.19(4)	-33	0.24(5)	0.04(3)	9.7
SSP Z	−0.31(3)	0.11(4)	22	−2.0(3)	0.17(3)	43	−0.21(4)	0.02(2)	6.4
SSP Z gr.	−0.42(4)	0.09(5)	14	−1.0(4)	0.06(4)	11	−0.40(5)	−0.03(3)	-7.8
SSP α	0.21(2)	−0.00(2)	-1.2	−0.3(2)	0.05(2)	19	0.19(2)	−0.01(1)	-7.6
SSP α gr.	0.02(3)	0.03(4)	5.4	0.3(3)	−0.03(3)	-6.8	0.06(4)	0.01(2)	4.1

Notes. The number in parenthesis indicates the uncertainty in the last decimal place of the measured value. The r columns indicate the Pearson correlation coefficient multiplied by one hundred.

Table C.2. Linear fits of characteristics of the galaxies in the form *[parameter]* = $b + a_M \log M_{\rm JAM} + a_\rho \log \rho_{10} + a_{\rm KS} \lambda^{\rm N}_{R_{\rm e}}$ for our sample in its completeness.

	All				
Parameter	b	a_M	a_ρ	$a_{\rm KS}$	$N_{\rm data}$
Shells	−2(1)	0.3(1)	0.10(9)	−0.2(1)	175
Streams	−3(1)	0.3(1)	0.11(8)	−0.1(1)	175
Tails	0(1)	0.0(1)	0.11(8)	−0.2(1)	175
Disturbed isophotes	−3(2)	0.4(1)	0.2(1)	−0.5(2)	174
Any TF	−3(2)	0.3(1)	0.2(1)	−0.2(2)	175
Any TF or DI	−5(2)	0.6(1)	0.2(1)	−0.3(2)	174
Shell fr.	−4(3)	0.5(3)	0.2(2)	−0.5(4)	175
Stream fr.	−3(1)	0.3(1)	0.09(8)	−0.1(1)	175
Tail fr.	−0.0(9)	0.04(8)	0.08(6)	−0.1(1)	175
Dust	1(1)	−0.1(1)	−0.01(8)	0.1(1)	175
KDC	0.2(4)	−0.00(3)	−0.01(3)	−0.18(4)	175
CRC	−0.2(3)	0.03(3)	−0.01(2)	−0.16(4)	175
KT	−0.3(2)	0.03(2)	0.02(1)	−0.02(2)	175
$2\,\sigma$	1.0(3)	−0.08(3)	0.00(2)	−0.13(4)	175
No kin. feature	1.6(8)	−0.11(8)	−0.06(6)	0.3(1)	175
$\log R_e$	−3.3(3)	0.36(3)	0.01(2)	−0.06(3)	175
Core γ'	5.6(9)	−0.47(8)	−0.05(6)	0.08(8)	79
SSP age	−48(7)	5.4(7)	0.4(5)	−0.1(8)	174
SSP age gr.	2.3(5)	−0.19(4)	0.04(3)	−0.01(5)	175
SSP Z	−2.3(3)	0.18(3)	0.02(2)	0.14(3)	174
SSP Z gr.	−1.2(4)	0.07(4)	−0.02(3)	0.11(5)	175
SSP α	−0.3(2)	0.05(2)	−0.01(1)	0.00(2)	174
SSP α gr.	0.3(3)	−0.02(3)	0.01(2)	0.02(4)	175

Notes. The number in parenthesis indicates the uncertainty in the last decimal place of the measured value. The column $N_{\rm data}$ indicates the number of the available data points.

Table C.3. Linear fits of characteristics of the galaxies in the form *[parameter]* $= b + a_M \log M_{\rm JAM} + a_\rho \log \rho_{10} + a_{\rm KS} \lambda^{\rm N}_{R_e}$ for our sample divided into two mass bins.

	$\log M_{\rm JAM} < 11$					$\log M_{\rm JAM} \geq 11$				
Parameter	b	a_M	a_ρ	$a_{\rm KS}$	$N_{\rm data}$	b	a_M	a_ρ	$a_{\rm KS}$	$N_{\rm data}$
Shells	−4(2)	0.4(2)	0.12(9)	−0.1(2)	140	−1(8)	0.1(7)	0.0(3)	−0.4(3)	35
Streams	−2(2)	0.2(2)	0.09(8)	−0.2(2)	140	14(9)	−1.1(8)	0.3(3)	−0.2(4)	35
Tails	−2(2)	0.2(2)	0.14(9)	−0.2(2)	140	3(6)	−0.3(5)	0.0(2)	−0.2(2)	35
Disturbed isophotes	−2(2)	0.3(2)	0.1(1)	−0.4(2)	140	−7(9)	0.7(8)	0.1(3)	−0.7(4)	34
Any TF	−4(2)	0.5(2)	0.2(1)	−0.2(2)	140	0(0)	−0.1(9)	0.2(3)	−0.3(4)	35
Any TF or DI	−4(2)	0.5(2)	0.2(1)	−0.2(2)	140	−5(9)	0.6(8)	0.2(3)	−0.4(4)	34
Shell fr.	−6(5)	0.7(4)	0.2(2)	−0.4(4)	140	−0(0)	0(2)	0.5(6)	−0.6(7)	35
Stream fr.	−1(1)	0.1(1)	0.05(6)	−0.2(1)	140	20(0)	−1.2(9)	0.4(3)	−0.1(4)	35
Tail fr.	−1(1)	0.2(1)	0.11(7)	−0.2(1)	140	2(5)	−0.2(4)	0.0(2)	−0.1(2)	35
Dust	2(2)	−0.2(2)	−0.03(8)	0.0(2)	140	−1(6)	0.1(5)	0.1(2)	0.3(2)	35
KDC	−0.1(6)	0.03(5)	−0.01(3)	−0.23(5)	140	−3(2)	0.2(2)	−0.04(6)	−0.05(7)	35
CRC	−0.5(5)	0.05(5)	−0.02(2)	−0.16(4)	140	1(2)	−0.0(2)	0.06(8)	−0.2(1)	35
KT	−0.3(3)	0.03(2)	−0.00(1)	−0.02(2)	140	−0(2)	0.0(1)	0.14(5)	−0.01(6)	35
$2\,\sigma$	1.2(6)	−0.10(5)	0.01(3)	−0.19(5)	140	N/A	N/A	N/A	N/A	35
No kin. feature	3(1)	−0.2(1)	0.01(6)	0.5(1)	140	5(4)	−0.4(4)	−0.5(1)	−0.1(2)	35
$\log R_e$	−2.5(5)	0.28(4)	0.01(2)	−0.05(4)	140	−5.0(9)	0.51(8)	−0.00(3)	−0.06(3)	35
Core γ'	5(2)	−0.4(2)	−0.03(8)	0.0(1)	53	9(3)	−0.8(2)	−0.04(9)	0.1(1)	26
SSP age	−60(0)	7(1)	0.6(6)	−0(1)	139	−50(0)	6(2)	−0.6(9)	−0(1)	35
SSP age gr.	3.2(7)	−0.29(7)	0.02(3)	0.02(6)	140	−1(2)	0.1(2)	0.07(7)	−0.03(8)	35
SSP Z	−2.8(4)	0.23(4)	0.03(2)	0.17(4)	139	−0.6(9)	0.03(8)	−0.01(3)	0.05(4)	35
SSP Z gr.	−0.3(7)	−0.02(6)	−0.03(3)	0.12(6)	140	−1(2)	0.1(1)	−0.01(5)	0.10(6)	35
SSP α	−0.4(3)	0.06(3)	−0.02(2)	0.03(3)	139	1.0(6)	−0.06(6)	0.01(2)	−0.06(3)	35
SSP α gr.	0.1(5)	−0.00(5)	0.01(3)	0.05(5)	140	0.3(9)	−0.02(8)	0.02(3)	−0.04(4)	35

Notes. The number in parenthesis indicates the uncertainty in the last decimal place of the measured value. The $N_{\rm data}$ columns indicate the number of the available data points.

Table C.4. Linear fits of characteristics of the galaxies in the form *[parameter]* $= b + a_M \log M_{\rm JAM} + a_\rho \log \rho_{10} + a_{\rm KS} \lambda^{\rm N}_{R_e}$ for our sample divided into two environmental density bins.

	$\log \rho_{10} < -2$					$\log \rho_{10} \geq -2$				
Parameter	b	a_M	a_ρ	$a_{\rm KS}$	$N_{\rm data}$	b	a_M	a_ρ	$a_{\rm KS}$	$N_{\rm data}$
Shells	1(2)	0.0(2)	0.8(4)	0.1(2)	43	−3(2)	0.3(1)	−0.0(1)	−0.3(2)	131
Streams	2(3)	−0.0(2)	0.5(4)	0.0(3)	43	−3(1)	0.4(1)	0.1(1)	−0.2(2)	131
Tails	1(2)	−0.1(2)	0.3(3)	−0.2(2)	43	−0(1)	0.1(1)	0.0(1)	−0.2(2)	131
Disturbed isophotes	4(3)	−0.0(3)	1.3(5)	−0.4(3)	42	−5(2)	0.5(2)	0.1(1)	−0.5(2)	131
Any TF	3(3)	−0.0(3)	1.0(5)	−0.0(3)	43	−4(2)	0.4(2)	0.1(2)	−0.3(2)	131
Any TF or DI	5(3)	−0.1(3)	1.5(6)	−0.3(4)	42	−6(2)	0.7(2)	0.1(2)	−0.3(2)	131
Shell fr.	4(5)	−0.1(4)	1.3(8)	0.3(5)	43	−5(4)	0.6(4)	−0.1(3)	−0.7(5)	131
Stream fr.	1(2)	0.1(2)	0.6(4)	−0.1(2)	43	−3(1)	0.3(1)	0.1(1)	−0.1(2)	131
Tail fr.	2(2)	−0.1(1)	0.3(3)	−0.2(2)	43	−0(1)	0.1(1)	0.03(9)	−0.1(1)	131
Dust	3(2)	−0.2(2)	0.3(4)	0.2(2)	43	1(1)	−0.0(1)	−0.1(1)	0.1(2)	131
KDC	0.4(7)	0.00(6)	0.2(1)	−0.12(7)	43	0.2(5)	−0.01(4)	−0.05(4)	−0.20(5)	131
CRC	−0.7(9)	0.09(8)	0.0(2)	−0.22(9)	43	−0.1(4)	0.02(3)	0.01(3)	−0.14(4)	131
KT	N/A	N/A	N/A	N/A	43	−0.4(3)	0.04(3)	0.03(2)	−0.03(3)	131
$2\,\sigma$	1.5(9)	−0.15(9)	−0.1(2)	−0.1(1)	43	0.9(4)	−0.07(3)	0.04(3)	−0.13(4)	131
No kin. feature	0(2)	0.1(2)	0.3(3)	0.5(2)	43	2(1)	−0.17(9)	−0.12(8)	0.2(1)	131
$\log R_e$	−2.9(6)	0.33(6)	0.1(1)	−0.08(7)	43	−3.3(3)	0.36(3)	0.03(3)	−0.05(4)	131
Core γ'	3(2)	−0.2(1)	−0.1(2)	−0.3(1)	18	6(1)	−0.47(9)	0.01(9)	0.2(1)	61
SSP age	−60(0)	7(2)	0(3)	−1(2)	43	−44(8)	5.0(8)	−0.3(7)	−0(1)	130
SSP age gr.	3.0(9)	−0.28(9)	−0.0(2)	0.0(1)	43	2.1(5)	−0.17(5)	0.04(4)	−0.02(6)	131
SSP Z	−4.1(7)	0.30(6)	−0.2(1)	0.18(7)	43	−1.9(3)	0.15(3)	0.01(2)	0.12(4)	130
SSP Z gr.	−1(1)	0.02(9)	−0.0(2)	0.3(1)	43	−1.3(5)	0.09(4)	0.01(4)	0.05(5)	131
SSP α	0.5(5)	0.00(5)	0.12(9)	−0.06(5)	43	−0.6(2)	0.07(2)	−0.01(2)	0.03(3)	130
SSP α gr.	0(1)	−0.02(9)	0.1(2)	−0.1(1)	43	0.3(3)	−0.03(3)	−0.01(3)	0.06(4)	131

Notes. The number in parenthesis indicates the uncertainty in the last decimal place of the measured value. The $N_{\rm data}$ columns indicate the number of the available data points.

A.3 Article: Preparing for low surface brightness science with the Vera C. Rubin Observatory: Characterization of tidal features from mock images, Martin et al (2022)

Martin et al. (2022)

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

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

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

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

Affiliations are listed at the end of the paper

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

ABSTRACT

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

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

1 INTRODUCTION

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

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

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

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

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

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

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

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

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

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

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

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

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

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

(iv) In Section 5 we summarize our results.

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

2 METHOD

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

2.1 The NEWHORIZON simulation

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

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

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

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

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

2.1.1 Numerical resolution limit for detecting tidal features

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

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

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

2.2 Galaxy sample

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

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

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

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

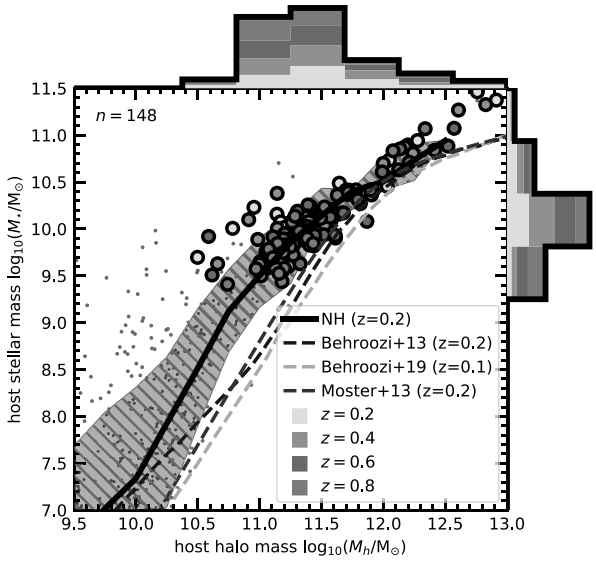

Figure 1. Scatter plot and stacked histograms showing the distribution of host galaxy halo masses and stellar masses selected at $z = 0.2$ along with their progenitors at $z = 0.4$, 0.6, and 0.8. There are a total of 148 objects (37 unique objects at four different snapshots). Objects in our sample are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

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

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

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

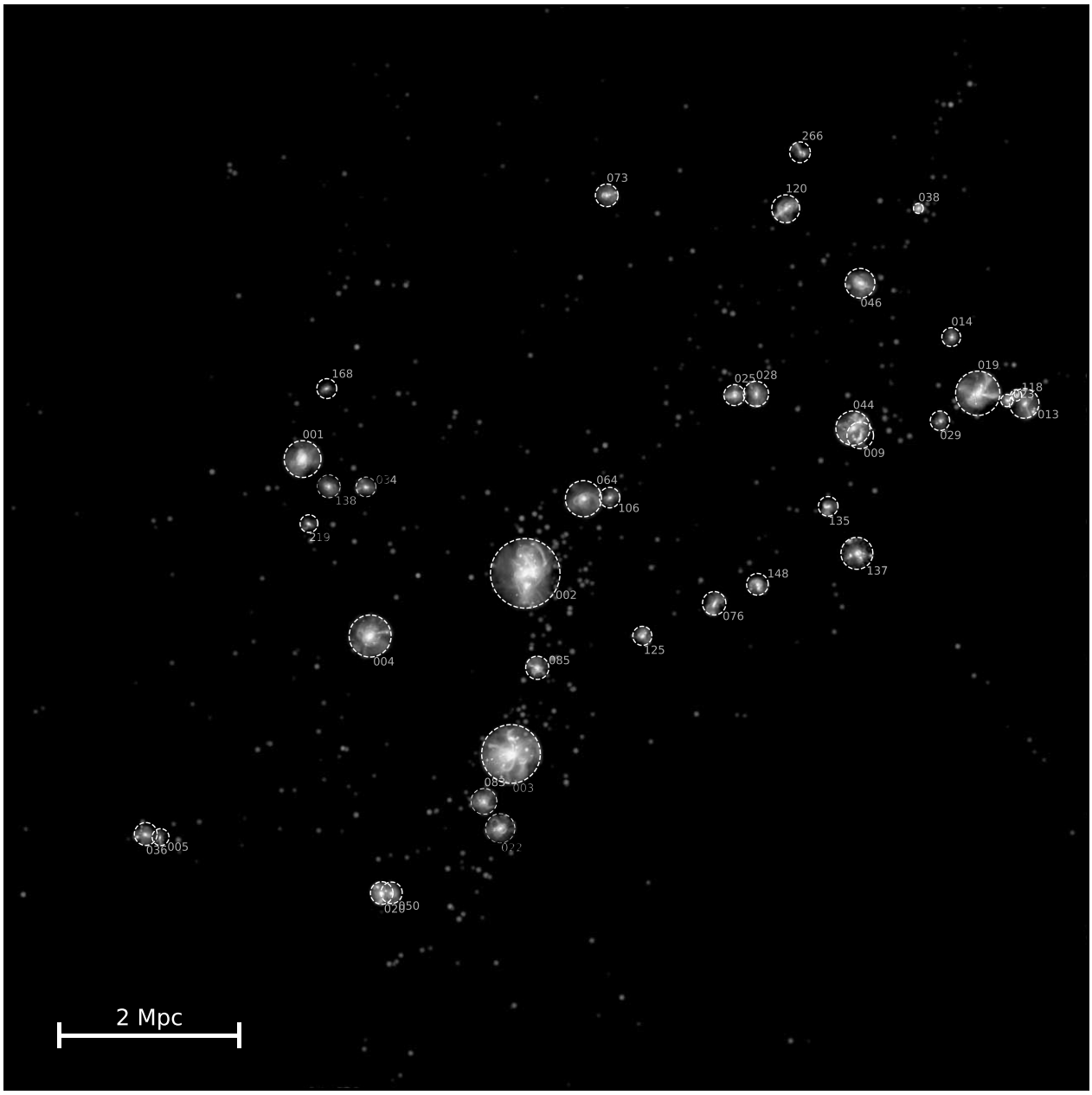

Figure 2. *g*, *r*, *i* false colour mock image showing the distribution of light within 1 $R_{\rm vir}$ of each of the massive galaxies in our sample in the context of the $\sim (16\,{\rm Mpc})^3$ simulation volume at a single simulation snapshot corresponding to $z = 0.4$. The dashed circles enclose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

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

2.3 Mock images

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

2.3.1 Star particle fluxes

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

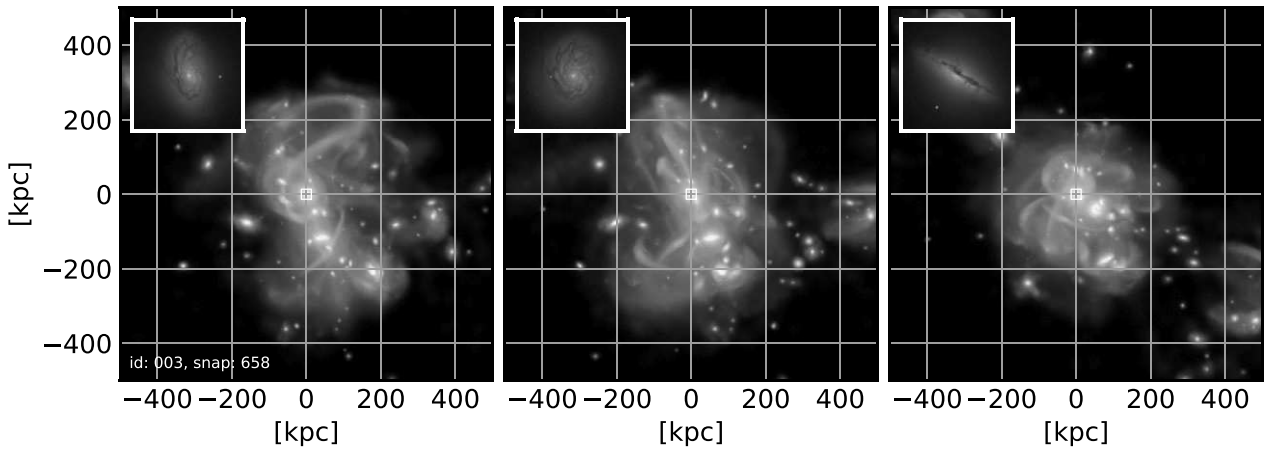

Figure 3. *g*, *r*, *i* false colour mock images of the same 1 Mpc field. From left to right, the panels show projections in *xy*, *xz*, and *yz*. The total stellar mass enclosed in the image is 3.2×10^{11} $M_\odot$ with the host accounting for 2.9×10^{11} $M_\odot$ (90 per cent). Images are produced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

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

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

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

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

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

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

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

2.3.2 *Smoothing*

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

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

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

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

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

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

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

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

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

2.3.3 *Mock observations*

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

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

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

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

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

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

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

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

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

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

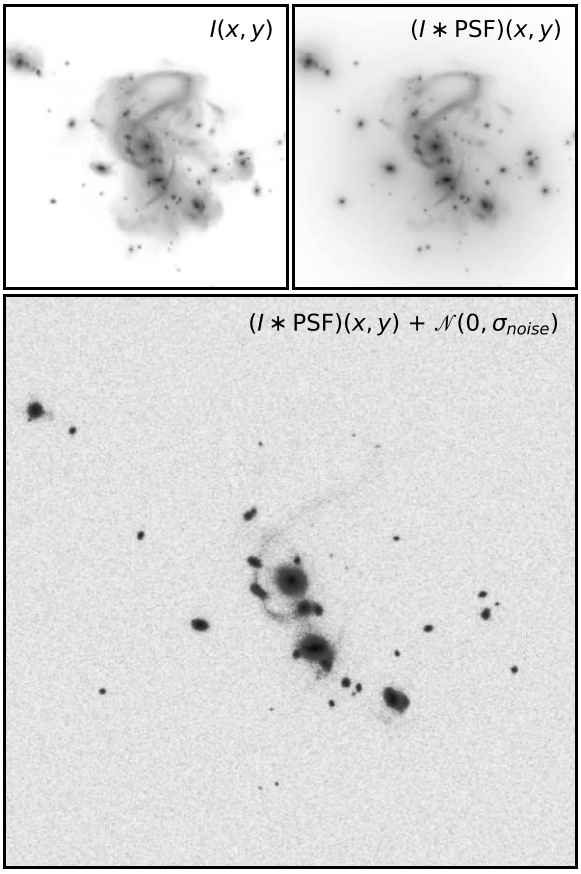

Figure 4. Plot demonstrating the steps taken to produce the mock images used for classification. As in Fig. 3, white corresponds to surface brightnesses fainter than $\sim$35 mag arcsec^{-2}. Top left-hand panel: *r*-band mock image of a galaxy at $z = 0.2$. Top right-hand panel: the image is convolved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

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

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

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

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

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

2.4 Measuring tidal features

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

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

Using these definitions we identify the following regions:

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

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

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

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

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

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

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

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

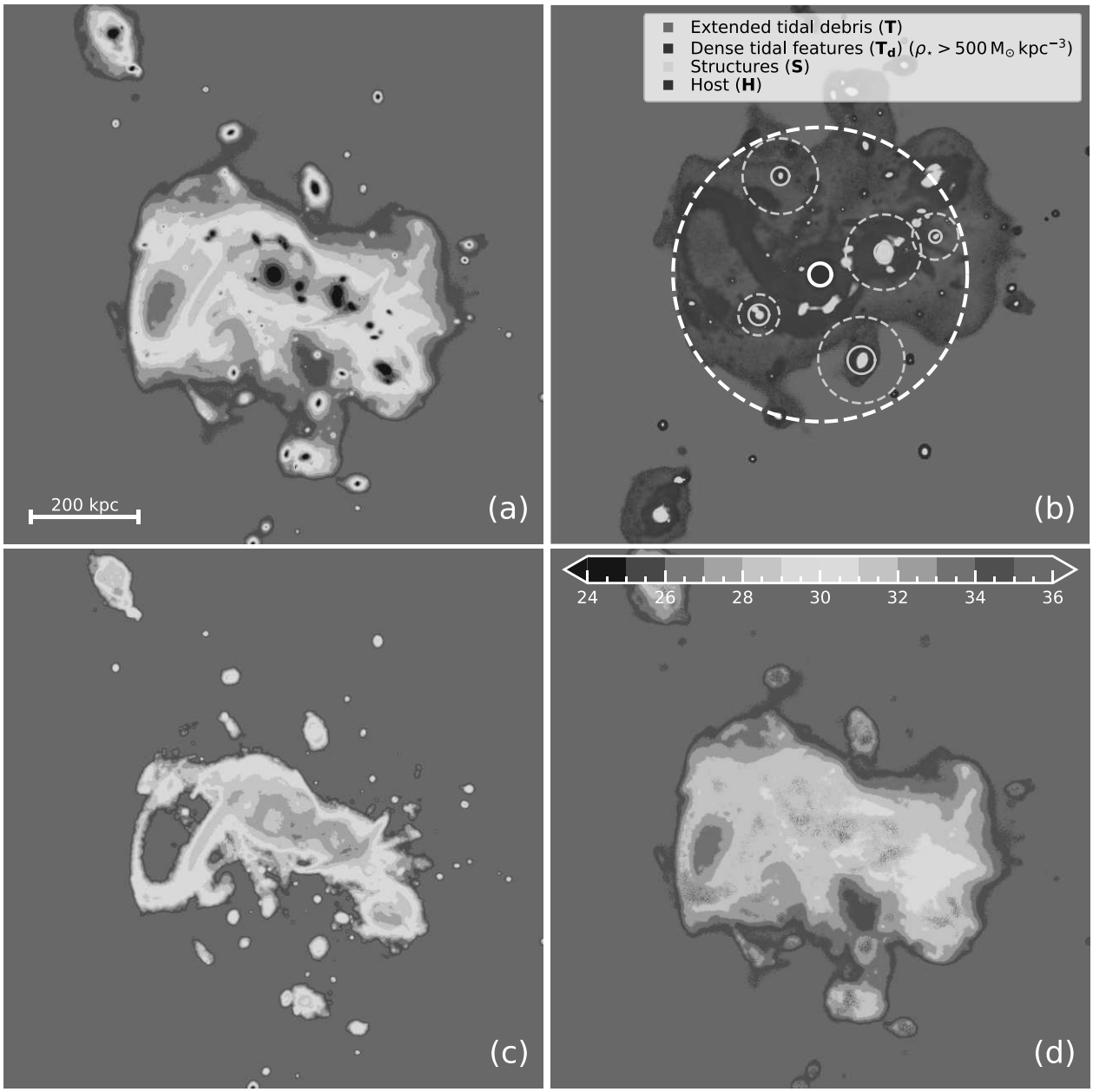

Figure 5. (a) *r*-band surface brightness map of one of the same mock fields as shown in Fig. 3. (b) A plot showing the same field with colours corresponding to tidal debris (T, red), dense tidal features ($\mathbf{T_d}$, purple), structures identified by ADAPTAHOP (S, green) and the central host galaxy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

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

2.5 Merger tree

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

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

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

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

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

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

2.6 Visual classification

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

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

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

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

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

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

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

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

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

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

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

3.1 Quantifying the distribution and strength of tidal features

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

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

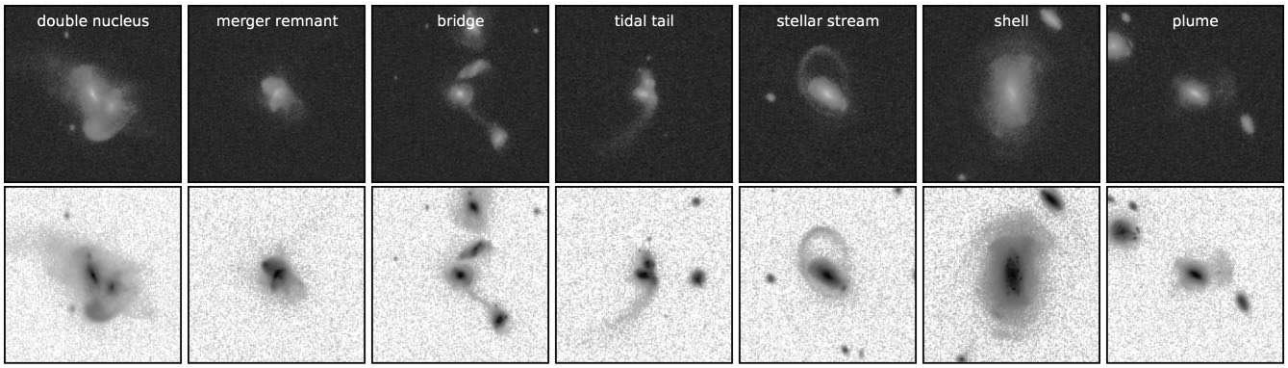

Figure 6. Example 100 Mpc ×100 Mpc thumbnail images for each category. For each object, we show the arcsinh stretched *g*, *r*, *i* false colour image (top row) and greyscale surface brightness map (bottom row).

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

3.1.1 Distribution of light in extended structures

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

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

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

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

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

3.1.2 Proportion of galaxy flux in tidal features

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

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

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

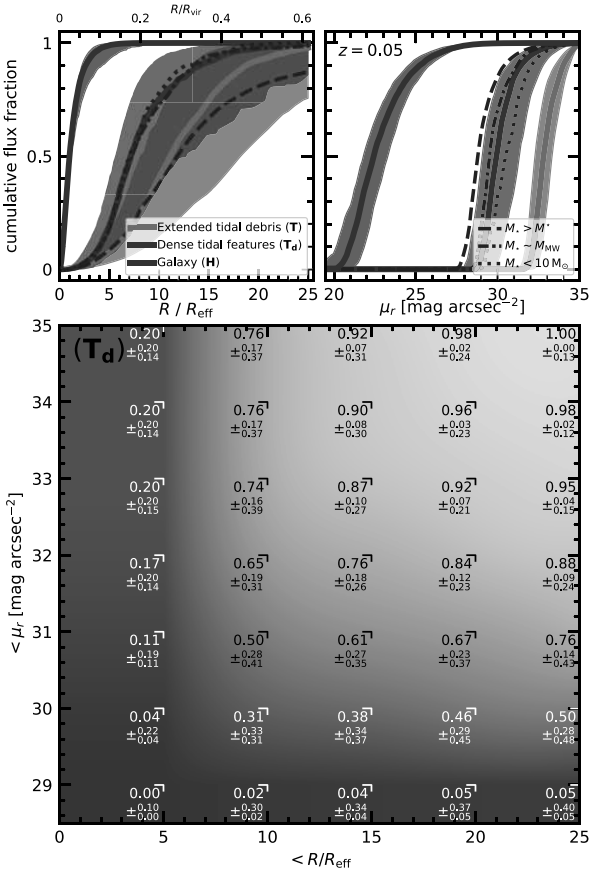

Figure 7. Top left-hand panel: The median cumulative fraction of flux as a function of projected radius ($\mu_r < 35$, $R < R_{\rm vir}$) for the host galaxy (**H**, blue), dense tidal features ($\mathbf{T_d}$, purple), and extended tidal debris (**T**, red). The coloured filled regions indicate the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

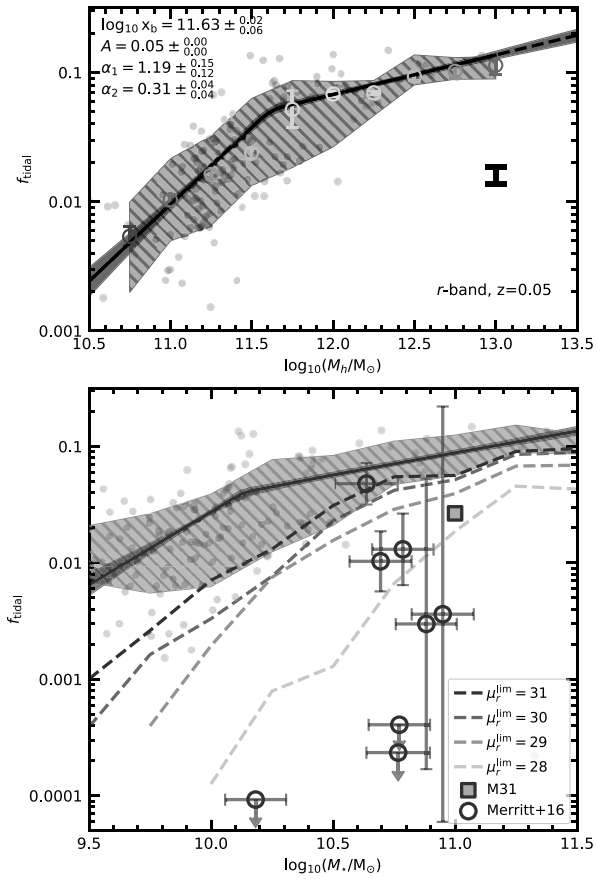

Figure 8. Top: Scatter plot showing the tidal flux fraction ($f_{\rm tidal}$, equation 5) in the r-band as a function of halo mass. The black line and grey shaded region show a broken power-law fit to the grey points representing individual NEWHORIZON galaxies and the associated 1σ uncertainty obtained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

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

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

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

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

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

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

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

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

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

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

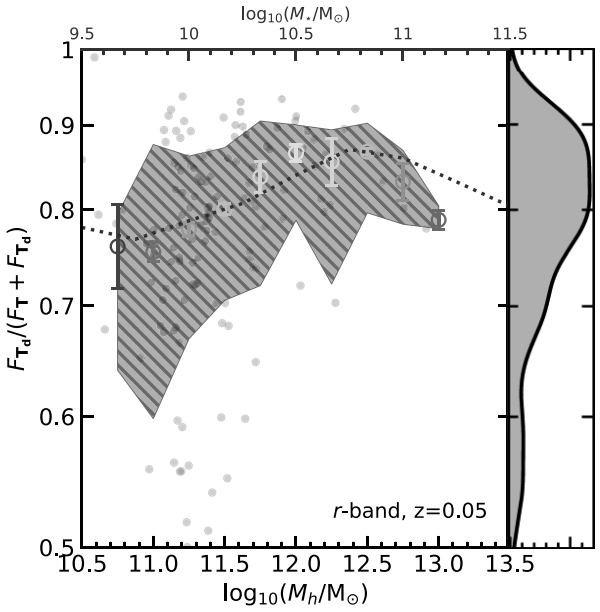

Figure 9. Scatter plot showing the fraction of tidal flux that is found in dense ($\rho_\star > 500$ M$_\odot$ kpc^{-3}) tidal features ($F_{\mathbf{T}_d}/(F_{\mathbf{T}_d}+F_{\mathbf{T}})$) versus the total flux in all tidal features) as a function of halo mass. Coloured error bars show the median value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

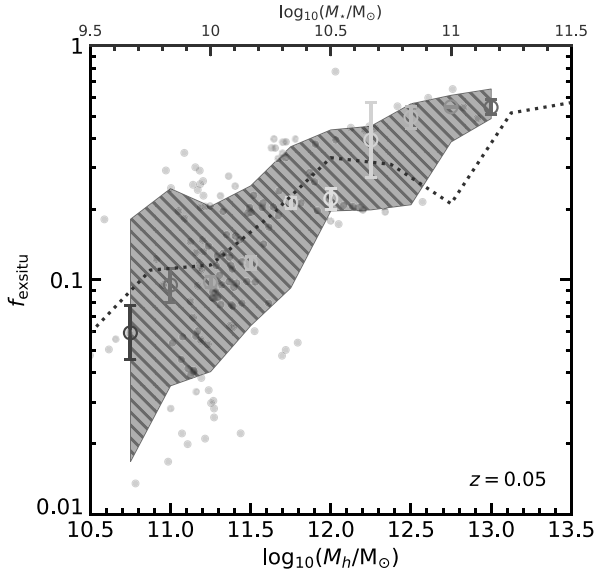

Figure 10. Scatter plot showing ex-situ mass fraction ($f_{\rm exsitu}$, equation 4) as a function of halo mass. The shaded and hatched region indicates the 1σ scatter of the grey points. Coloured error bars show the median value in overlapping halo mass bins with 1σ error bars determined from 100 000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

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

3.2 Tidal features and accreted mass

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

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

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

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

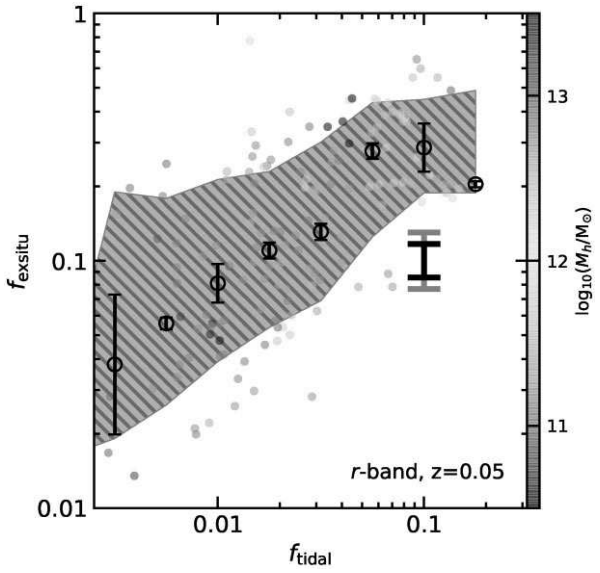

Figure 11. Scatter plot showing ex-situ mass fraction ($f_{\rm exsitu}$, equation 4) as a function of tidal flux fraction ($f_{\rm tidal}$, equation 5) in the *r*-band colour coded by halo mass. Black open circles with error bars indicate the median value in overlapping bins, with 1σ error bars determined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

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

3.3 Tidal feature detection

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

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

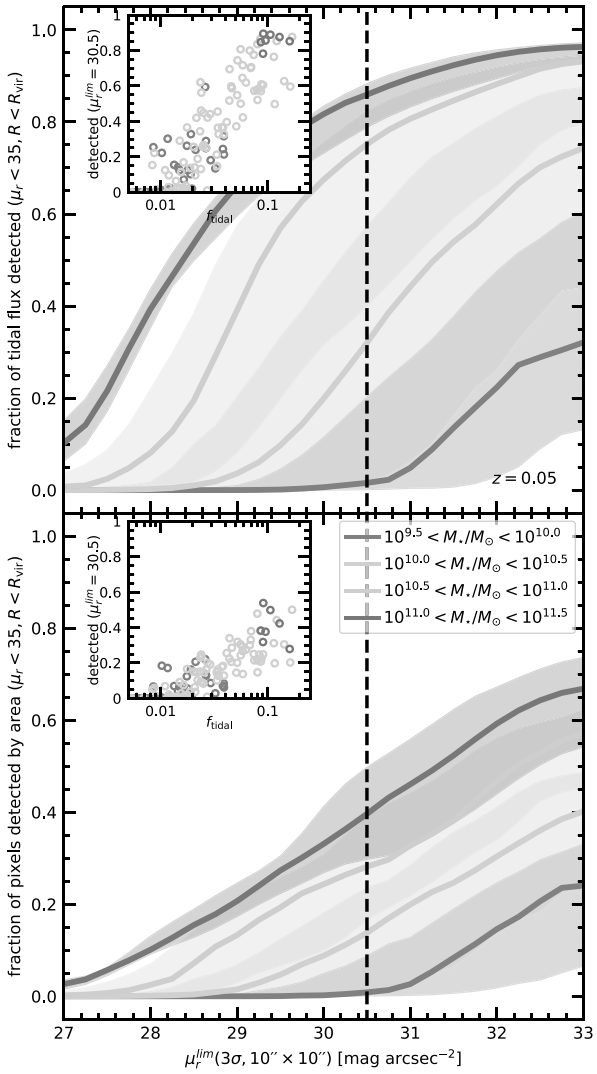

Figure 12. Top: the fraction of tidal flux found in detected structures as a function of *r*-band limiting surface brightness. Coloured lines indicate the median fraction of flux in detected structures for different stellar mass bins, and coloured regions indicate the central 68th percentile (1σ). The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

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

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

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

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

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

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

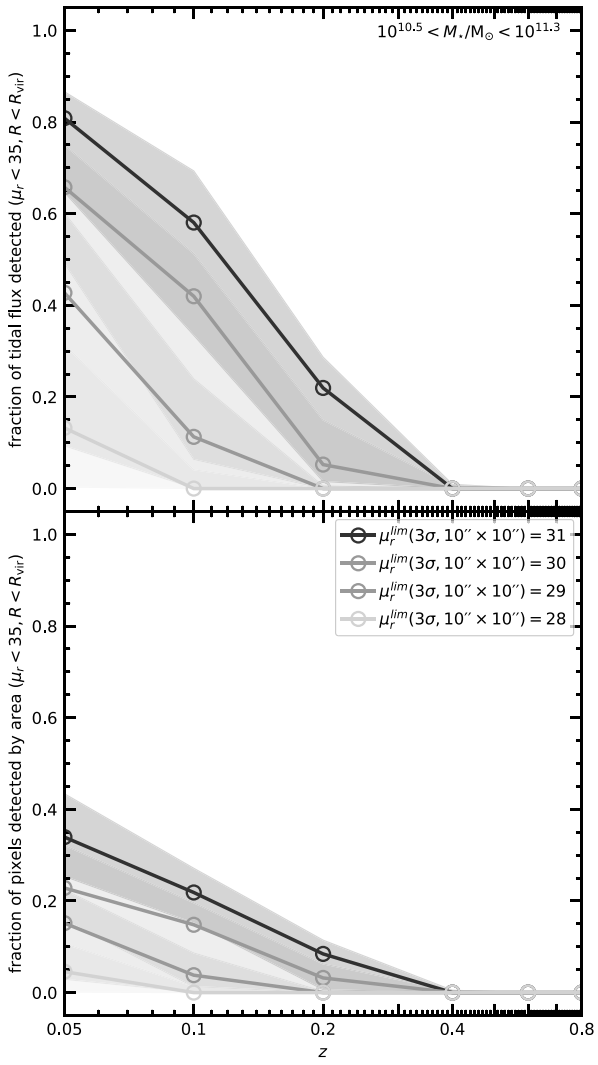

Figure 13. Top the fraction of tidal flux found in detected structures as a function of redshift in $\sim M^\star$ galaxies for different r-band limiting surface brightnesses. The coloured lines indicate the median fraction of flux in detected structures for a given redshift and coloured regions indicate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

4 CHARACTERIZATION OF TIDAL FEATURES BY HUMAN CLASSIFICATION

4.1 Visual classification of mock images

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

4.1.1 Collated feedback

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

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

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

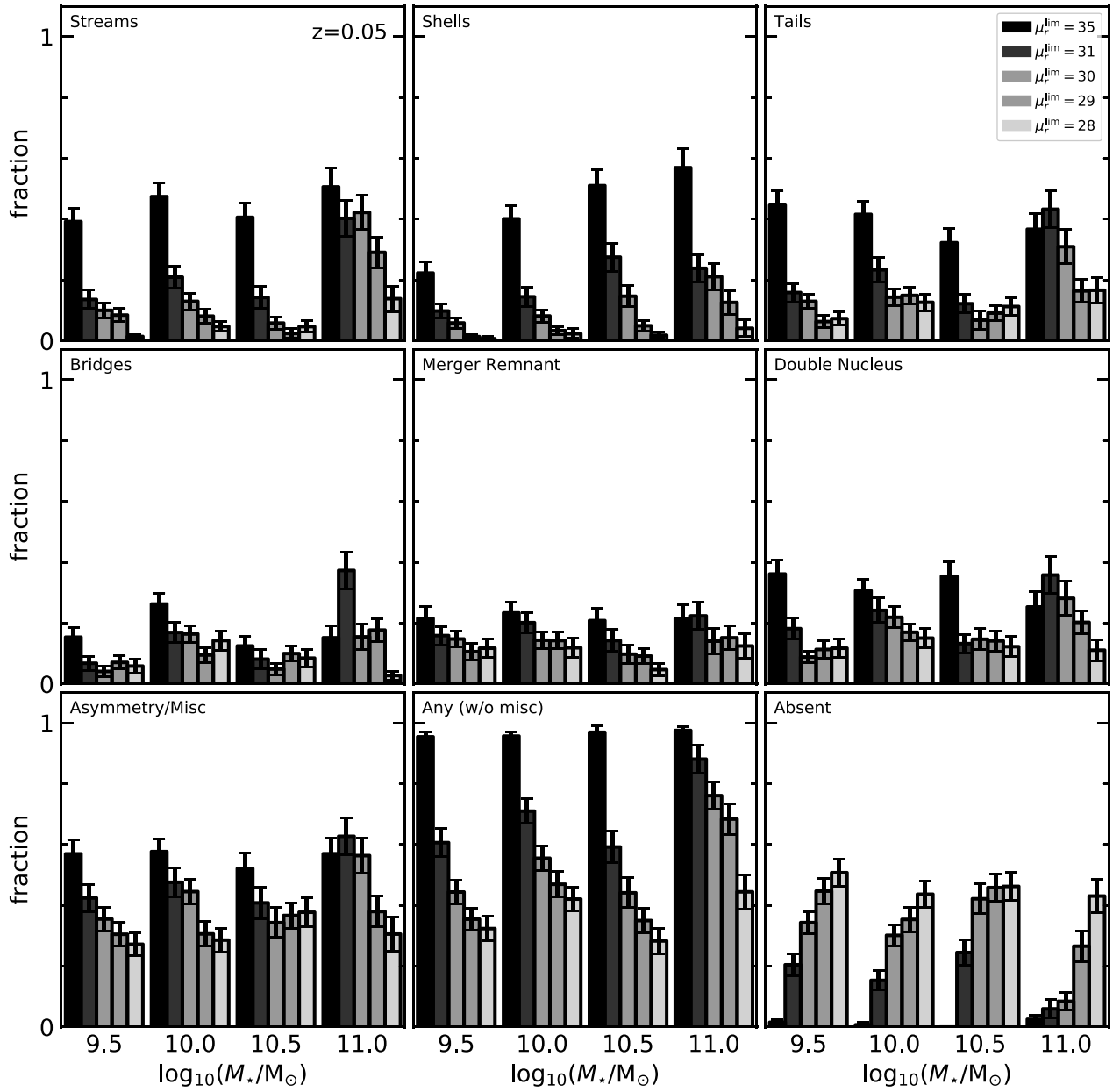

Figure 14. Histograms indicating the prevalence of different classes of tidal feature as a function of stellar mass and limiting surface brightness. The bars indicate the fraction of galaxies that exhibit at least one instance of a given class of tidal feature in each mass bin. The four mass bins have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

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

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

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

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

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

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

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

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

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

4.1.2 Census of tidal features by class

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

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

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

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

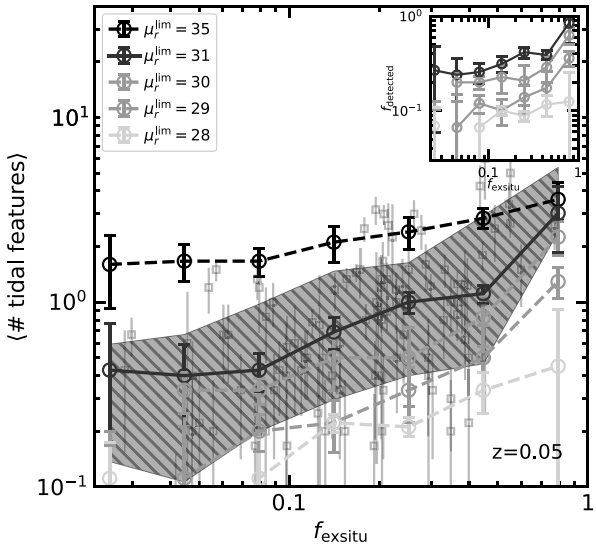

Figure 15. The average number of tidal features identified by classifiers as a function of galaxy ex-situ mass fraction. The solid red line indicates the average number of distinct tidal features identified per galaxy (excluding the merger, double nuclei, and asymmetry / misc category) at $z = 0.05$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

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

4.2 Visual biases

4.2.1 Effect of redshift and limiting surface brightness

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

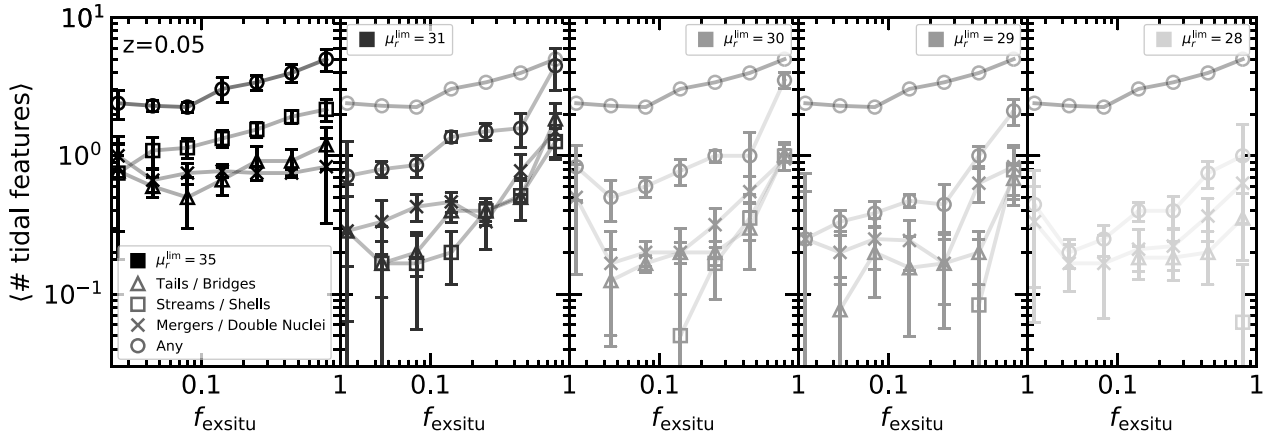

Figure 16. The average number of tidal features identified by classifiers as a function of galaxy ex-situ mass fraction at $z = 0.05$ split by limiting surface brightness and tidal feature class. Tidal tails and bridges are indicated by triangular markers, tidal streams, and shells by squares, mergers,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

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

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

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

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

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

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

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

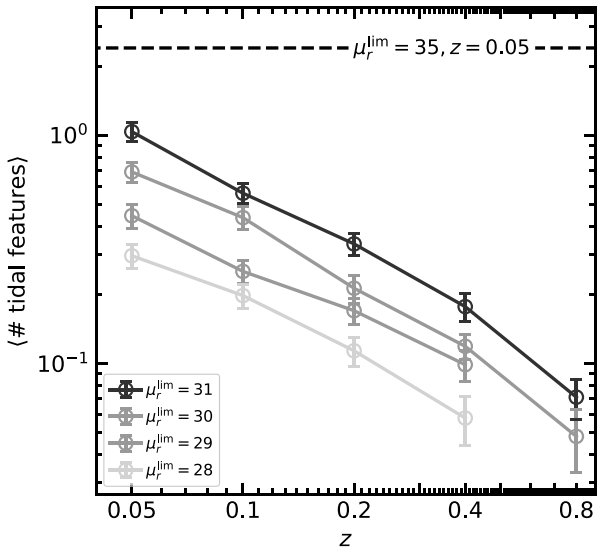

Figure 17. The average number of tidal features identified by classifiers as a function of redshift. The solid coloured lines indicate the average number of distinct tidal features identified per galaxy (excluding the asymmetry / misc category) as a function of redshift and for the different limiting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

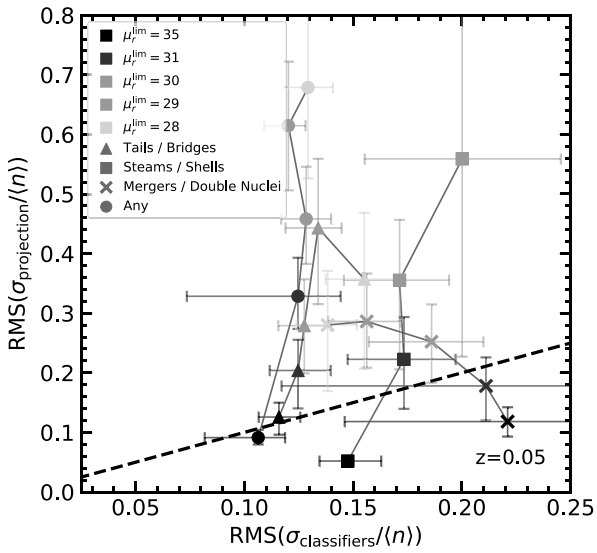

Figure 18. Evolution of the RMS fractional $\sigma_{\rm projection}$ and $\sigma_{\rm classifiers}$ and as a function of limiting surface brightness at $z = 0.05$. Tidal tails and bridges are indicated by triangular markers, tidal streams and shells by squares, mergers or double nuclei by crosses and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

4.2.2 *Effect of projection*

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

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

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

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

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

5 SUMMARY

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

(i) Distribution of tidal flux:

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

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

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

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

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

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

(ii) Reliability of human classification:

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

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

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

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

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

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

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

ACKNOWLEDGEMENTS

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

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

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

DATA AVAILABILITY

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

REFERENCES

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

Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189
Fensch J. et al., 2020, A&A, 644, A164
Ferreras I. et al., 2014, MNRAS, 444, 906
Fiacconi D., Feldmann R., Mayer L., 2015, MNRAS, 446, 1957
Fliri J., Trujillo I., 2016, MNRAS, 456, 1359
Foster C. et al., 2014, MNRAS, 442, 3544
Fry A. M., Morrison H. L., Harding P., Boroson T. A., 1999, AJ, 118, 1209
Gargiulo I. D. et al., 2019, MNRAS, 489, 5742
Garrison-Kimmel S. et al., 2019, MNRAS, 489, 4574
Gentile G., Famaey B., Combes F., Kroupa P., Zhao H. S., Tiret O., 2007, A&A, 472, L25
Gómez F. A. et al., 2017, MNRAS, 472, 3722
Graham A. W., 2002, ApJ, 568, L13
Graham A. W., Dullo B. T., Savorgnan G. A. D., 2015, ApJ, 804, 32
Greco J. P. et al., 2018a, PASJ, 70, S19
Greco J. P. et al., 2018b, ApJ, 857, 104
Grillmair C. J., Freeman K. C., Irwin M., Quinn P. J., 1995, AJ, 109, 2553
Gwyn S. D. J., 2012, AJ, 143, 38
Harmsen B., Monachesi A., Bell E. F., de Jong R. S., Bailin J., Radburn-Smith D. J., Holwerda B. W., 2017, MNRAS, 466, 1491
Hendel D., Johnston K. V., 2015, MNRAS, 454, 2472
Hendel D., Johnston K. V., Patra R. K., Sen B., 2019, MNRAS, 486, 3604
Hood C. E., Kannappan S. J., Stark D. V., Dell'Antonio I. P., Moffett A. J., Eckert K. D., Norris M. A., Hendel D., 2018, ApJ, 857, 144
Huertas-Company M., Aguerri J. A. L., Tresse L., Bolzonella M., Koekemoer A. M., Maier C., 2010, A&A, 515, A3
Ibata R., Thomas G., Famaey B., Malhan K., Martin N., Monari G., 2020, ApJ, 891, 161
Iodice E. et al., 2016, ApJ, 820, 42
Iodice E. et al., 2019, A&A, 623, A1
Ivezić Ž. et al., 2019, ApJ, 873, 111
Jackson R. A., Martin G., Kaviraj S., Laigle C., Devriendt J. E. G., Dubois Y., Pichon C., 2019, MNRAS, 489, 4679
Ji I., Hasan I., Schmidt S. J., Tyson J. A., 2018, PASP, 130, 084504
Johnston K. V., Sigurdsson S., Hernquist L., 1999, MNRAS, 302, 771
Johnston K. V., Sackett P. D., Bullock J. S., 2001, ApJ, 557, 137
Johnston K. V., Choi P. I., Guhathakurta P., 2002, AJ, 124, 127
Johnston K. V., Bullock J. S., Sharma S., Font A., Robertson B. E., Leitner S. N., 2008, ApJ, 689, 936
Jonsson P., 2006, MNRAS, 372, 2
Junais et al., 2020, A&A, 637, A21
Kado-Fong E. et al., 2018, ApJ, 866, 103
Kadowaki J., Zaritsky D., Donnerstein R. L., 2017, ApJ, 838, L21
Karunakaran A., Spekkens K., Zaritsky D., Donnerstein R. L., Kadowaki J., Dey A., 2020, ApJ, 902, 39
Kaviraj S., 2014, MNRAS, 440, 2944
Kaviraj S., 2020, preprint (arXiv:2001.01728)
Kaviraj S., Tan K.-M., Ellis R. S., Silk J., 2011, MNRAS, 411, 2148
Kaviraj S., Darg D., Lintott C., Schawinski K., Silk J., 2012, MNRAS, 419, 70
Kaviraj S., Devriendt J., Dubois Y., Slyz A., Welker C., Pichon C., Peirani S., Le Borgne D., 2015, MNRAS, 452, 2845
Kaviraj S. et al., 2017, MNRAS, 467, 4739
Kaviraj S., Martin G., Silk J., 2019, MNRAS, 489, L12
Kawinwanichakij L. et al., 2017, ApJ, 847, 134
Kesden M., Kamionkowski M., 2006, Phys. Rev. D, 74, 083007
Kim T. et al., 2012, ApJ, 753, 43
Kniazev A. Y., Grebel E. K., Pustilnik S. A., Pramskij A. G., Kniazeva T. F., Prada F., Harbeck D., 2004, AJ, 127, 704
Komatsu E. et al., 2011, ApJS, 192, 18
Kroupa P., 2002, Science, 295, 82
Kuijken K. et al., 2002, The Messenger, 110, 15
Laine S. et al., 2018, preprint (arXiv:1812.04897)
Lambrides E. L. et al., 2021, ApJ, 919, 43
Lauer T. R., 1988, ApJ, 325, 49
Lee N. et al., 2015, ApJ, 801, 80
Lintott C. J. et al., 2008, MNRAS, 389, 1179
Lofthouse E. K., Kaviraj S., Conselice C. J., Mortlock A., Hartley W., 2017, MNRAS, 465, 2895
Lotz J. M., Primack J., Madau P., 2004, AJ, 128, 163
Lupton R., Blanton M. R., Fekete G., Hogg D. W., O'Mullane W., Szalay A., Wherry N., 2004, PASP, 116, 133
Malhan K., Valluri M., Freese K., 2021, MNRAS, 501, 179
Malin D. F., Carter D., 1983, ApJ, 274, 534
Mancillas B., Duc P.-A., Combes F., Bournaud F., Emsellem E., Martig M., Michel-Dansac L., 2019, A&A, 632, A122
Mantha K. B. et al., 2019, MNRAS, 486, 2643
Martin G., Kaviraj S., Devriendt J. E. G., Dubois Y., Laigle C., Pichon C., 2017, MNRAS, 472, L50
Martin G., Kaviraj S., Devriendt J. E. G., Dubois Y., Pichon C., 2018a, MNRAS, 480, 2266
Martin G., Kaviraj S., Devriendt J. E. G., Dubois Y., Pichon C., Laigle C., 2018b, MNRAS, 474, 3140
Martin G. et al., 2019, MNRAS, 485, 796
Martin G., Kaviraj S., Hocking A., Read S. C., Geach J. E., 2020, MNRAS, 491, 1408
Martin G. et al., 2021, MNRAS, 500, 4937
Martínez-Delgado D., Pohlen M., Gabany R. J., Majewski S. R., Peñarrubia J., Palma C., 2009, ApJ, 692, 955
Martínez-Delgado D. et al., 2012, ApJ, 748, L24
Martinez-Delgado D. et al., 2021, preprint (arXiv:2104.06071)
Merritt A., van Dokkum P., Danieli S., Abraham R., Zhang J., Karachentsev I. D., Makarova L. N., 2016, ApJ, 833, 168
Merritt A., Pillepich A., van Dokkum P., Nelson D., Hernquist L., Marinacci F., Vogelsberger M., 2020, MNRAS, 495, 4570
Mihos J. C., 2019, preprint (arXiv:1909.09456)
Mihos J. C., Hernquist L., 1996, ApJ, 464, 641
Mihos J. C., Harding P., Feldmeier J., Morrison H., 2005, ApJ, 631, L41
Mihos J. C. et al., 2015, ApJ, 809, L21
Mihos J. C., Harding P., Feldmeier J. J., Rudick C., Janowiecki S., Morrison H., Slater C., Watkins A., 2017, ApJ, 834, 16
Miville-Deschênes M. A., Duc P. A., Marleau F., Cuillandre J. C., Didelon P., Gwyn S., Karabal E., 2016, A&A, 593, A4
Miyazaki S. et al., 2002, PASJ, 54, 833
Miyazaki S. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 9
Moffat A. F. J., 1969, A&A, 3, 455
Monachesi A., Gómez F. A., Grand R. J. J., Kauffmann G., Marinacci F., Pakmor R., Springel V., Frenk C. S., 2016, MNRAS, 459, L46
Monachesi A. et al., 2019, MNRAS, 485, 2589
Montes M., 2019, preprint (arXiv:1912.01616)
Montes M., Infante-Sainz R., Madrigal-Aguado A., Román J., Monelli M., Borlaff A. S., Trujillo I., 2020, ApJ, 904, 114
Montes M., Brough S., Owers M. S., Santucci G., 2021, ApJ, 910, 45
Morales G., Martínez-Delgado D., Grebel E. K., Cooper A. P., Javanmardi B., Miskolczi A., 2018, A&A, 614, A143
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Müller O., Vudragović A., Bílek M., 2019, A&A, 632, L13
Naab T. et al., 2014, MNRAS, 444, 3357
Nelson D. et al., 2019, MNRAS, 490, 3234
Olivier S. S., Seppala L., Gilmore K., 2008, in Atad-Ettedgui E., Lemke D., eds, Proc. SPIE Conf. Ser. Vol. 7018, Advanced Optical and Mechanical Technologies in Telescopes and Instrumentation. SPIE, Bellingham, p. 70182G
Olsen K. P. et al., 2021, ApJ, 922, 88
Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, ApJ, 725, 2312
Park M.-J. et al., 2019, ApJ, 883, 25
Pawlik M. M., Wild V., Walcher C. J., Johansson P. H., Villforth C., Rowlands K., Mendez-Abreu J., Hewlett T., 2016, MNRAS, 456, 3032
Pawlowski M. S., 2021, Galaxies, 9, 66
Pearson S., Clark S. E., Demirjian A. J., Johnston K. V., Ness M. K., Starkenburg T. K., Williams B. F., Ibata R. A., 2022, ApJ, 926, 166

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

APPENDIX A: EFFECTIVE RESOLUTION LIMIT FOR DETECTING SHELLS

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

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

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

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

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

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

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

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

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

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

$$\mathrm{SNR}_{\rm shell} = \frac{n_{\rm shell}}{\sigma_{\rm shell}} = \frac{n_{\rm shell}}{\sqrt{2\,n_{\rm galaxy} + n_{\rm shell}}}, \tag{A1}$$

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

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

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

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

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

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

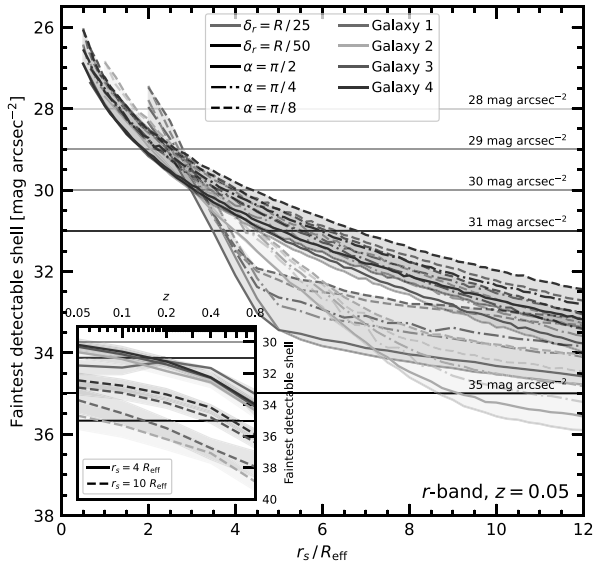

Figure A1. The main plot shows numerical lower limits for the average surface brightness of the faintest detectable shell as a function of galactocentric radius for different galaxy and shell models. Different coloured lines indicate different galaxy Sérsic models, while solid, dot-dashed, and dashed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

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

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

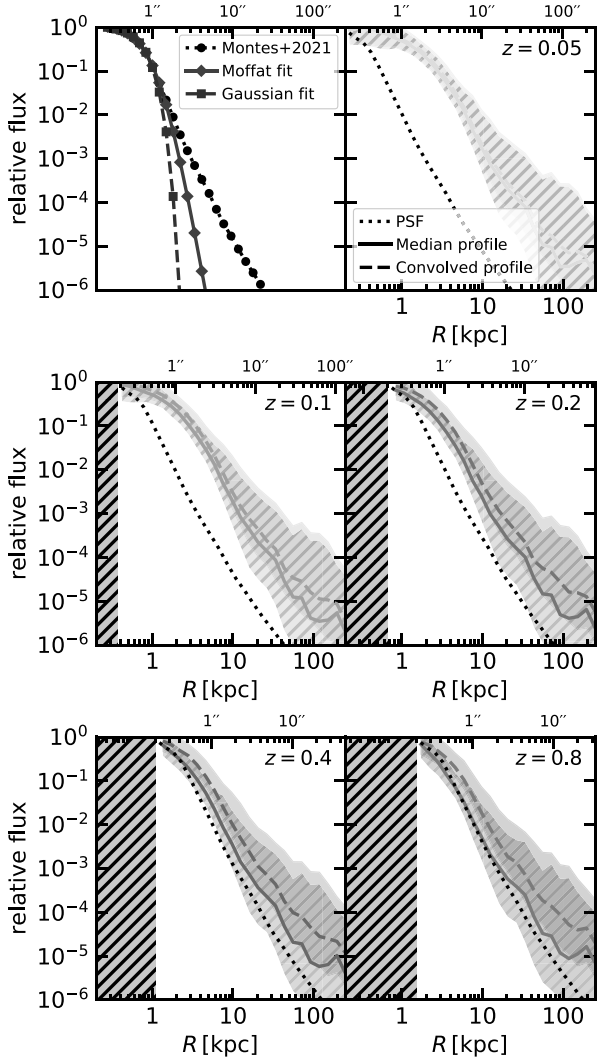

Figure B1. Top left-hand panel: Comparison of the measured 1D PSF from Montes et al. (2021) with a the best fit Moffat (1969) and Gaussian distributions. Other panels: surface brightness profiles compared to the PSF for different redshifts (indicated in the top right-hand corner of each panel). The solid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

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

APPENDIX C: SELECTING THE EX-SITU MASS TIME INTERVAL

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

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

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

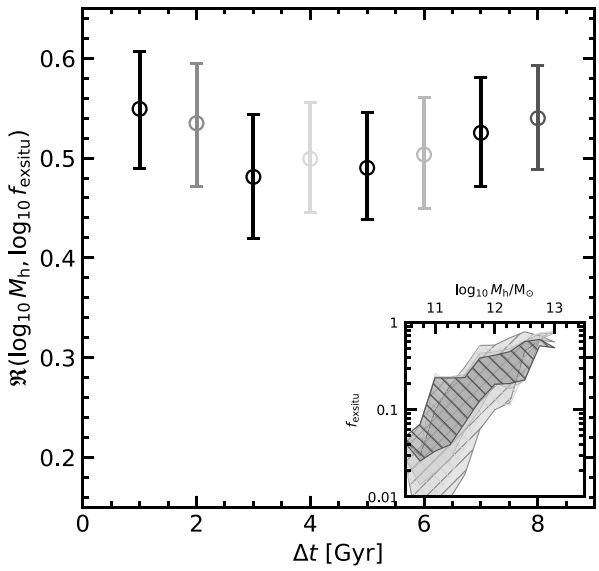

Figure C1. Correlation coefficient for different time intervals with 1σ uncertainties indicated by error bars. The inset plot shows a region enclosing the 1σ scatter for the M_h–$f_{\rm exsitu}$ relation for $\Delta t = [2, 4, 6, 8]$ where the colour of the region corresponds to the value Δt indicated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

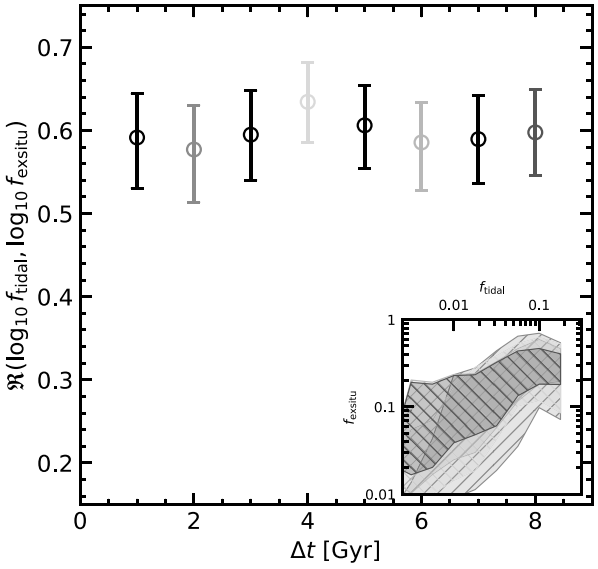

Figure C2. Correlation coefficient for different time intervals with 1σ uncertainties indicated by error bars. The inset plot shows a region enclosing the 1σ scatter for the M_h–$f_{\rm exsitu}$ relation for $\Delta t = [2, 4, 6, 8]$ where the colour of the region corresponds to the value Δt indicated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

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

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

APPENDIX D: REDSHIFT EVOLUTION

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

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

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

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

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

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

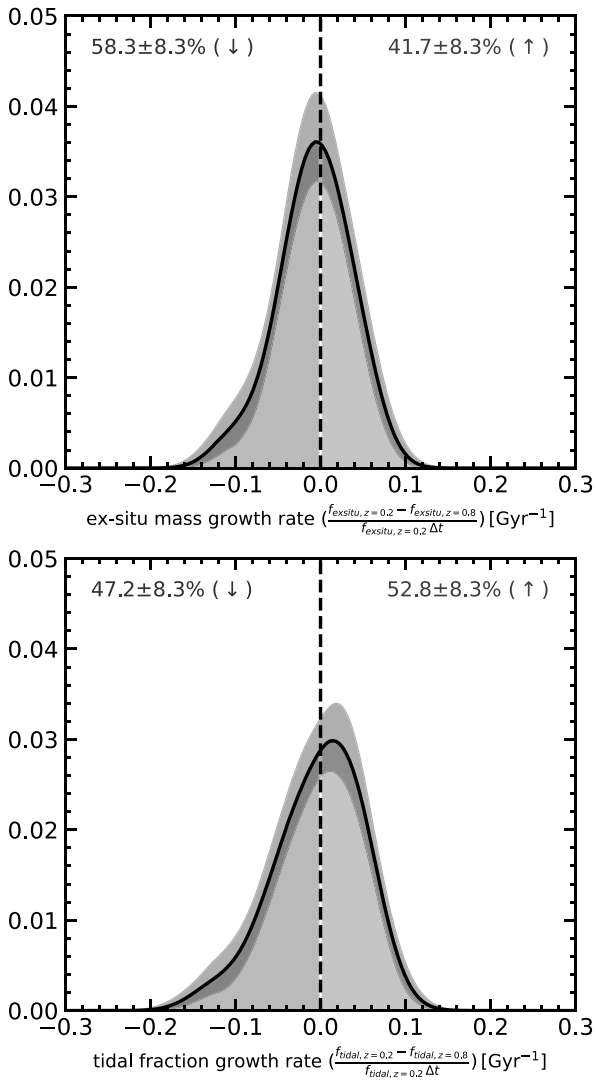

Figure D1. Top: the distribution of the $f_{\rm exsitu}$ growth rate ($\Gamma_{\rm exsitu}$) from Gaussian kernel density estimates from 10 000 bootstraps. Bottom: the distribution of the $f_{\rm tidal}$ growth rate ($\Gamma_{\rm tidal}$) from Gaussian kernel density estimates from 10 000 bootstraps.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

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

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

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

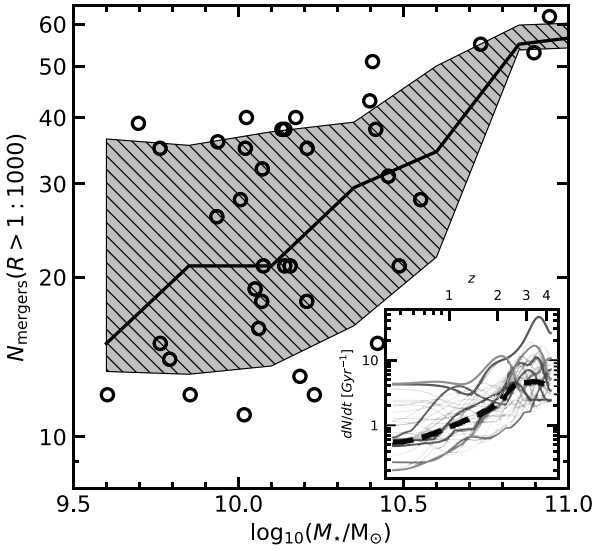

Figure D2. Number of mergers with mass ratio $R < 1{:}1000$ undergone per galaxy since $z = 5$ as a function of galaxy stellar mass. The hatched region encloses the 1σ scatter and the black line indicates the median. The inset plot shows the merger rate ($R < 1{:}1000$) per Gyr of individual galaxies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

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

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

APPENDIX E: TIDAL FEATURE DETECTION METHOD

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

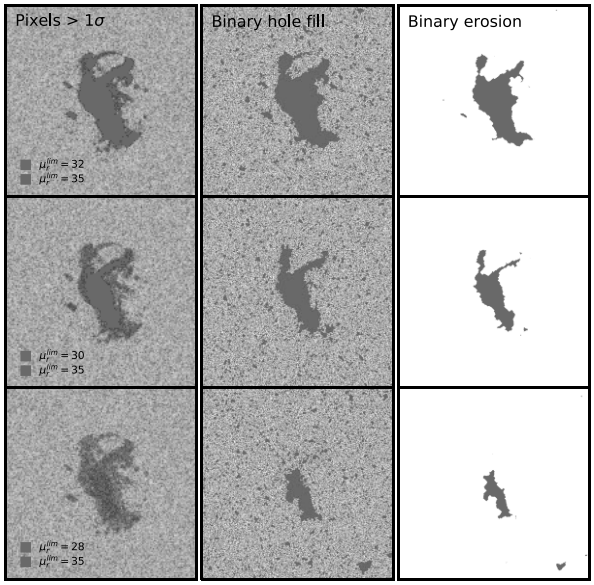

Figure E1. Example images showing the process of locating detected structures for limiting surface brightnesses of $\mu_r^{\rm lim}(3\sigma, 10\,{\rm arcsec} \times 10\,{\rm arcsec}) = 32$ mag arcsec^{-2}, 30 mag arcsec^{-2}, and 28 mag arcsec^{-2}. Each panel of the first column shows a map of pixels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

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

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

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

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

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

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

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

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

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

A.4 Article: Learnable Gabor modulated complex-valued networks for orientation robustness, Richards et al (2020)

Richards et al. (2020)

Learnable Gabor modulated complex-valued networks for orientation robustness

Felix Richards[a,*], Adeline Paiement[b], Xianghua Xie[a], Elisabeth Sola[c], Pierre-Alain Duc[c]

[a]*Department of Computer Science Swansea University Swansea, UK*
[b]*Université de Toulon, Aix Marseille Univ, CNRS, LIS Marseille, France*
[c]*Université de Strasbourg CNRS, Observatoire astronomique de Strasbourg, France*

Abstract

Robustness to transformation is desirable in many computer vision tasks, given that input data often exhibits pose variance. While translation invariance and equivariance is a documented phenomenon of CNNs, sensitivity to other transformations is typically encouraged through data augmentation. We investigate the modulation of complex valued convolutional weights with learned Gabor filters to enable orientation robustness. The resulting network can generate orientation dependent features free of interpolation with a single set of learnable rotation-governing parameters. By choosing to either retain or pool orientation channels, the choice of equivariance versus invariance can be directly controlled. Moreover, we introduce rotational weight-tying through a proposed cyclic Gabor convolution, further enabling generalisation over rotations. We combine these innovations into Learnable Gabor Convolutional Networks (LGCNs), that are parameter-efficient and offer increased model complexity. We demonstrate their rotation invariance and equivariance on MNIST, BSD and a dataset of simulated and real astronomical images of Galactic cirri.

Keywords: rotation, invariance, equivariance, astronomy, bsd, segmentation, analytical, gabor, robustness, complex space

*Corresponding author
Email addresses: felixarichards@gmail.com (Felix Richards), adeline.paiement@univ-tln.fr (Adeline Paiement), x.xie@swansea.ac.uk (Xianghua Xie), elisabeth.sola@astro.unistra.fr (Elisabeth Sola), pierre-alain.duc@astro.unistra.fr (Pierre-Alain Duc)

1. Introduction

We enable learning of approximate orientation invariance and equivariance in convolutional neural networks (CNN). Datasets in various domains often exhibit a range of pose variation (e.g. scale, translation, orientation, reflection). CNNs are inherently equipped to handle translation invariance, but remedies for other symmetries often involve large models and datasets with plenty of augmentation. This inability to properly adapt to transformations such as local/global rotations is a major limitation in CNNs.

An important distinction is that of equivariance versus invariance. For a network to be equivariant, it should be robust to variation in pose and be able to carry over transformations of the input to transformed features and output. For tasks where output is dependent on these transformations, network invariance alone is suboptimal as transformation information is discarded, by definition. For example, in ultra deep astronomical imaging, the scattered light from foreground Galactic cirrus contaminates and occludes interesting Low-Surface Brightness (LSB) extragalactic objects. These cirrus clouds exhibit orientation dependent features: segmenting cloud regions is a problem requiring invariance, as orientation of cloud streaks does not necessarily affect the geometry of the cloud's envelope. On the other hand, removing occluding clouds, which is crucial to studying background LSB galaxies, is a denoising problem that requires robust and descriptive equivariant features.

Numerous works have been published alongside CNN research attempting to integrate forms of rotation invariant and equivariant feature learning in an a-priori fashion. Approaches typically generate rotation dependent responses by one of the following strategies: 1) learning orientations by constructing filters from a steerable basis [1, 2, 3], 2) rotating convolution filters/input by preset angles [4, 5, 6], or 3) introducing orientation information through analytical filters [7, 8]. A significant drawback of the former type is that it introduces significant computational overhead [9]. In the second category, the rotation process imposes the use of interpolation which results in artefacts for any rotation outside of the discrete sampling grid. This is overcome in the latter category by using analytical filters with an inherent rotation parameter. Orientations are static in [7, 8], similarly to the second category, however there is no inherent limitation of

analytical filters preventing them from having learnable orientation parameters. There is thus a need for a dynamic orientation sensitive architecture that can accurately adapt to the input's transformation. We address this need in this work using Gabor filters, analytical filters that are parameterised by orientation, scale and frequency among other variables. Furthermore, Gabor filters are differentiable with respect to their parameters, meaning that these parameters can be learned through steepest descent style algorithms.

Contributions - In this paper we propose Learnable Gabor Convolutional Networks (LGCN), a complex-valued CNN architecture highly sensitive to rotation transformations. We utilise adjustable Gabor modulation of convolutional weights to generate dynamic orientation activations. By learning Gabor parameters alongside convolutional filters we achieve features that are dependent on exact angles with no interpolation artefacts. Moreover, there is no explicit constraint on convolutional filters, allowing a diverse feature space that adapts to the degree of rotation equivariance required. We extend the modulation approach used in [7] to complex space, enabling use of the full complex Gabor filter and exploiting the inherent descriptive power of complex neurons. Further, we build on this and propose a convolutional operator where Gabor filter modulation is cyclically shifted, inspired by group theory CNNs [10, 1, 11, 12, 3], allowing propagation of orientation information throughout a forward pass in an equivariant manner.

2. Previous works

Methods have been developed in an attempt to integrate transformation invariance in an a-priori fashion. Prior to CNN popularity, the use of hand-crafted features such as SIFT [13] and Gabor filters [14, 15] was explored to generate rotation/scale invariant representations. A widely adopted technique in deep learning is to augment transformations into a dataset [16, 17, 18]. This brute force approach introduces new samples to prompt the model to learn this new range of transformations. Models with learned invariance through augmentation require a very large parameter space to capitalise on data augmentation, and still may generalise poorly to unseen transformations.

There has been much work recently on encoding symmetries into CNN architectures.

Early efforts utilised pooling over transformed responses, e.g. siamese networks [19], training-time augmentation [20, 21], parallel convolutional layers [22, 11], kernel-based affine pooling [23], and image warping [24, 25, 26, 27]. Specifically in the last few years there has been a surge of interest in rotation equivariant architectures. Authors have been able to formulate CNNs entirely from principles of group theory and thus construct modified operators and/or constrain filters [10, 1, 28]. Bekkers et al. [12] employ bi-linear interpolation to enable any regular sampling of the continuous group of 2D rotation. Similarly [5] and [4] utilise copied and rotated filters, but pool over the produced activation maps. While interpolation allows rotation by exact angles (as in [12, 5, 4]) it introduces artefacts for angles outside of the discrete sampling grid. In [29] residual blocks are combined with principles of steerable bases to learn approximate equivariance. Worrall et al. [2] allow exact orientation representations while overcoming dependence on interpolation by constraining filters to the family of complex circular harmonics: there is a clear demonstration of complex neurons encoding rotational information which justifies our usage of complex CNNs for rotation equivariance. Finzi et al. [30] are able to construct group equivariance without steerable filters by constructing filters as parameterisations of Lie algebra . Similarly, Weiler et al. [3] present a CNN architecture with learnable steerable filters, and derive a generalised weight initialisation method for steerable basis coefficients. Using a formulation of steerable filter architectures, [31] proposes a general framework for equivariant networks under any combination of rotation, reflection or translation. We draw inspiration from the cyclic shifting group convolutions commonly used in group theory based CNNs [10, 1, 11, 12, 3], and propose a similar operation for rotation generalisation without requiring derivation from group theory and reducing computational overhead.

Analytical filters have made a resurgence in many deep learning contexts. Specific to transformation invariance, analytical filters parameterised by rotation are fast and can extract orientational features dependent on exact angles, overcoming interpolation artefacts. Several approaches replace convolutional weights with wavelet filters [32, 33, 34]. Wavelets are also applied to inputs of standard convolutional layers in a preprocessing fashion [35]. In [36] authors present a framework for convolutional weight modulation, achieving enhanced filters with binarised weights. Zhou et al. [8]

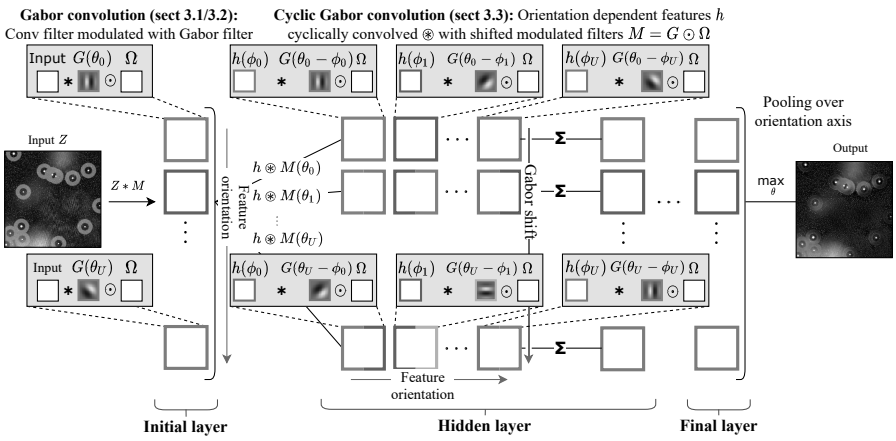

Figure 1: Overview of LGCN, with illustration of the filter modulation and cyclic convolution concepts. Input channel indices are omitted to highlight the rotational aspect of the proposed method. Colour denotes individual feature orientations. To achieve orientation invariance, an additional pooling operation (not represented here) may be added at the end of each hidden layer to pool over orientations. In this equivariant problem example, orientation information is retained across layers. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this article.)

exploit rotation parameterisation of discrete Fourier transforms to extract orientation information, modulating standard convolutional filters with a filter bank of rotated analytical filters. Luan et al. [7] implement a similar approach but opt to use Gabor filters, demonstrating that they are more robust to rotation and scale transformations. In [37] wavelet filter hyperparameters are learned in an end to end fashion for spectral decomposition through wavelet deconvolutions. We combine lessons learned from [37] with [7] to construct Gabor filter modulation with learnable parameters.

3. Methodology

Our LGCN achieves sensitivity to rotation transformations through adjustable Gabor modulation of convolutional weights. In the architecture defined below, modulation parameters are learned alongside convolutional filters. Having separate modulation and convolution parameters keeps backpropagation simple. Given that convolution filters are not explicitly constrained as in other methods attempting to overcome transformations,

the result is a larger space of possible features. Fig. 1 illustrates the general structure of LGCNs, providing an overview of the concepts proposed throughout this section. An important development of this approach is that parameters belong to complex space, allowing both real and imaginary parts of analytical filters to be utilised. Given that frequency response filters are often designed over complex space, this enables a variety of modulation choices.

LGCNs can consider several orientations simultaneously, which are finely tuned to the task being solved. LGCNs are able to achieve activations dependent on arbitrary continuous rotations with no interpolation artifacts and without using steerable filter bases. With modulation, LGCNs increase model complexity at little cost to the parameter size. In our case Gabor filters are calculated with orientation θ and wavelength λ, meaning they can generate a new feature channel with only two parameters. Finally, LGCNs utilise a novel convolutional operator where Gabor filters are cyclically shifted during modulation, enabling propagation of orientation information across layers and thus facilitating learned invariance and equivariance.

3.1. Filter Modulation of Complex-Valued Networks

In order to enable compatibility with a wide range of analytic frequency response filters, we construct complex-valued CNN layers. As described in [38] we construct complex numbers by encoding real and imaginary parts as separate real valued elements. A complex convolutional weight tensor can be written as $\Omega = A + iB$, where A and B are stored internally as real tensors. Complex arithmetic is then simulated with appropriate real operations acting on these elements. For a complex valued input $H = X + iY$, convolution is computed as:

$$Z = \Omega * H = (A * X - B * Y) + i(B * X + A * Y) \tag{1}$$

where $*$ denotes the convolution operator. For nonlinearities, we use the complex ReLU proposed in [39], $\mathbb{C}$-ReLU(Z) = ReLU$(|Z| + b)\frac{Z}{|Z|}$ with b a real-valued bias term. We also implement complex analogues of batch normalisation, given in [38], and average spatial pooling, trivially given by considering the average of real and imaginary parts separately.

For an analytical filter $\Phi^P = \Phi^P_{\text{Re}} + i\Phi^P_{\text{Im}}$ with D parameters $p_d \in P = \{p_0, ..., p_{D-1}\}$ we extend the convolutional modulation presented in [36] to complex space. The modulation of convolutional kernel $\Omega^c = A^c + iB^c$ of channel c with discretised filter Φ^P is given by $M^{c,P} = \Phi^P \odot \Omega^c$, where $\odot$ represents complex element-wise multiplication. Output from convolution with the modulated filter is then given by $Z^{c,P} = M^{c,P} * H$, or for each pixel at coordinates (s, t):

$$z^{c,P}_{s,t} = \sum_{k=1}^{K} \sum_{l=1}^{K} m^{c,P}_{k,l} \, h_{s+k,t+l} \tag{2}$$

This construction of modulated filters can be viewed as a collection of filter banks, where both the underlying kernels (via convolutional filter Ω) and frequency sub-bands (via Φ) are learnt. The complete filter bank has dimensions $2 \times C_{\text{out}} \times C_{\text{in}} \times U \times K \times K$, with C_{out} and C_{in} the number of output and input convolutional channels respectively, U the number of modulating filters, and K the convolution kernel size. Any given filter is obtained by modulating a convolutional filter W^c of channel c with analytical filter Φ^{P^u}. A significant advantage of this formulation is that a filter bank of U filters is created from a single canonical filter, meaning that encoding of transformation representations requires little computational overhead. Adjusting parameters through backpropagation requires calculating the gradient of a differentiable loss function L with respect to p_d:

$$\frac{\partial L}{\partial p_d} = \frac{\partial L}{\partial M^{c,P}} \frac{\partial M^{c,P}}{\partial p_d} = \sum_{k=1}^{K} \sum_{l=1}^{K} \frac{\partial L}{\partial m^{c,P}_{k,l}} \frac{\partial m^{c,P}_{k,l}}{\partial p_d} \quad = \sum_{k=1}^{K} \sum_{l=1}^{K} \frac{\partial m^{c,P}_{k,l}}{\partial p_d} \sum_{s=1}^{N} \sum_{t=1}^{N} \frac{\partial L}{\partial z^{c,P}_{s,t}} h_{s+k,t+l}. \tag{3}$$

Thus the only constraint on choice of analytical filter Φ is that it is differentiable with respect to parameter p_d. In the following subsection we compute the above derivative in the scenario where only a subset of parameters are learned.

3.2. Learnable Gabor filters

In this paper we modulate with Gabor filters similarly to [7], which are feature detectors characterised by rotation sensitivity and frequency localisation: $G(\lambda, \theta, \psi, \sigma, \gamma)_{k,l} = e^{-\frac{k'^2 + \gamma^2 l'^2}{2\sigma^2}} e^{i(\frac{2\pi}{\lambda} k' + \psi)}$ with $k' = k\cos\theta + l\sin\theta$ and $l' = l\cos\theta - k\sin\theta$. Two major differences with [7] is that we work with complex-valued networks, and we learn the

parameters of the filters while they were fixed to static orientations in [7]. A significant advantage of Gabor filters in comparison to Fourier related methods such as DCT is that they are not constructed from a sinusoidal basis, meaning that discontinuous patterns, such as edges, can more easily be represented. We fix (hyper)parameters other than orientation θ and wavelength λ: $G(\lambda, \theta, 0, \frac{1}{\sqrt{2}}, 1)$ as in [7] which demonstrated that this provides sufficient expressivity while simplifying computation. Though we choose to modulate with the well-documented Gabor filters due to orientation and frequency parameterisation, it is possible to modulate with a variety of complex analytical filters with this approach.

Thus the modulated filter $M^{c,P}$ can be written as $M^{c,P} = G^P \odot \Omega^c$. We evaluate $\partial m^{c,P}_{k,l}/\partial p_d$ at pixel k, l in the context of Gabor filter modulation for both parameters. Given that $\theta, \lambda \in \mathbb{R}$ we treat $M^{c,P}$ as a function of the real and imaginary parts separately (k, l indices omitted for readability):

$$\frac{\partial m^{c,P}}{\partial \theta} = a'^c \frac{\partial G_{\text{Re}}}{\partial \theta} + b'^c \frac{\partial G_{\text{Im}}}{\partial \theta} = \frac{2\pi}{\lambda} e^{-(k^2+l^2)} l' [-a'^c \sin(\frac{2\pi}{\lambda} k') + b'^c \cos(\frac{2\pi}{\lambda} k')] \quad (4)$$

$$\frac{\partial m^{c,P}}{\partial \lambda} = a'^c \frac{\partial G_{\text{Re}}}{\partial \lambda} + b'^c \frac{\partial G_{\text{Im}}}{\partial \lambda} = \frac{2\pi}{\lambda^2} e^{-(k^2+l^2)} k' [a'^c \sin(\frac{2\pi}{\lambda} k') - b'^c \cos(\frac{2\pi}{\lambda} k')], \quad (5)$$

where $a'^c = a^c + b^c$, $b'^c = a^c - b^c$. Backpropagation $\frac{\partial L}{\partial p_d} = \frac{\partial L}{\partial M^{c,P}} \frac{\partial M^{c,P}}{\partial p_d}$ can now be calculated, enabling learning of Gabor filters' parameters alongside convolutional weights. Accordingly, parameters are updated by $\theta' = \theta - \eta \frac{\partial L}{\partial \theta}$ and $\lambda' = \lambda - \eta \frac{\partial L}{\partial \lambda}$, with η denoting learning rate.

3.3. Cyclic Gabor Convolutions

In intermediate layers we implement cyclic convolutions to further increase rotation information without increasing parameter size and utilise the additional feature channels generated by Gabor modulation. We exploit the cyclic property of finite subgroups of 2D rotation transformations to create convolutional filters based on all permutations of orientation and canonical filters. By sharing all weights across every orientation, the underlying canonical filters further generalise over rotations. This is analogous to how filters are exposed to all translations in standard CNNs to encourage generalisation over translations. This propagation of rotation dependence directly facilitates equivariance,

in contrast to non-cyclic Gabor convolutions which must pool over the orientation axis per layer.

Note that this cyclic framework does not require analytical filters that are steerable, only that filters can be parameterised by rotation. That is to say, it is not a requirement that filters meet the criteria of linear steerability according to Freeman et al. [40]. In particular, we demonstrate that rotational weight sharing through cyclic shifting can be achieved with Gabor filters, which are not steerable. Networks constructed from steerable filters where basis coefficients are learned in place of convolutional kernels inherently and explicitly limit the filter space – whether this is a downside or an optimal regularisation to achieve rotation equivariance is yet to be shown. In comparison, learnable modulation with Gabor filters implicitly regularises filter space.

The cyclic convolution design we propose takes inspiration from group convolutions presented in [10]. Specifically, cyclic Gabor convolutions utilise the shifting operation used in group convolutions defined over 2D roto-translations [10, 3]. As the modulation transformation cannot be used to form a symmetry group, we do not derive computation using the group framework. However, using the orientation sensitivity of the Gabor filter we implement a similar resulting feature composition, enabling rotational weight sharing without requiring a proof for strict equivariance.

With $h^c(\theta)$ denoting channel c and orientation θ of the previous layer's activation map, for a single orientation and output channel, cyclic convolution $\circledast$ is computed as:

$$z^{\hat{c}}(\theta) = \sum_{c=1}^{C_{\text{in}}} \left[h^c \circledast M^{\hat{c}c} \right](\theta) = \sum_{c=1}^{C_{\text{in}}} \sum_{\phi \in P} \left[h^c(\phi) * (G(\theta - \phi) \odot \Omega^{\hat{c}c}) \right]. \tag{6}$$

P is the set of U orientations that are used to generate Gabor filters. Note this formulation allows any size of P. In order to keep implementation efficient and avoid recalculating Gabor filters for all permutations of learned orientations we keep $\phi \in P$ as the original angles. This choice allows filters to be reused with a cyclic shift of the orientation components per different output orientation θ.

3.4. Learnable Gabor Convolutional Networks

The framework presented above allows learnable modulation to be added into any convolutional layer, making the method very versatile. There are some considerations

to take into account however, which we discuss in this section.

Complex weight initialisation - LGCNs operate over complex space, requiring weight initialisation to be rethought. Principles of He weight initialisation [41] no longer hold given that $\text{Var}(\Omega) \neq \text{Var}(A) + i\text{Var}(B)$, i.e. real and imaginary parts cannot be initialised independently. We use Trabelsi's generalisation of He's strategy over complex space [38], setting $\text{Var}[|\Omega|] = \frac{4-\pi}{2n_{\text{in}}}$ with n_{in} denoting the number of input units. The phase is then uniformly distributed around the circle. It is worth noting that He's derivation is specific to the traditional ReLU, using the result that for a given input X_l to a layer l, and previous output Y_{l-1}: $\text{E}[X_l^2] = \frac{1}{2}\text{Var}[Y_{l-1}]$. This holds for traditional ReLU, $X_l = \max(0, Y_{l-1})$, as Y_{l-1} has zero mean and a symmetric distribution which is essentially split along its axis of symmetry. However with $\mathbb{C}$-ReLU, for $b < 0$, Y_{l-1} is no longer divided along the axis of symmetry. For this reason we simply initialise the biases of $\mathbb{C}$-ReLU layers to zero.

The choice of initialisation for modulation parameters is largely dependent on the choice of analytical filter, and should be influenced by the function's domain and the roles of individual variables. For initialisation of Gabor parameters, as discussed in Section 3.2, we fix phase shift ψ, aspect ratio γ and scale σ in order to simplify computation. Given that wavelength is a non-negative quantity we initialise λ with mean $3\sqrt{U}$ and variance $\frac{\sqrt{U}}{4}$ as per [41], and verified that training is stable. This choice of initialisation also avoids spatial aliasing of the Gabor filter for all kernel sizes (i.e. 3×3 or larger) at network initialisation. As the filter is sampled more than twice per phase, the signal is adequately captured, as per the Nyquist-Shannon sampling theorem. For orientation θ, in the real case there is no benefit of using the full interval of rotations due to evenness, however in the complex case the oddness of the imaginary part causes orthogonal filters for θ with differing sign. For this reason we initialise θ uniformly around the full circle.

Gabor axis considerations - Though the ability to create enhanced filters from a single canonical filter has advantages of parameter efficiency and weight-tying, it leaves the network prone to dimensionality explosion. This can be controlled using one or more of three approaches depending on the problem at hand: adjusting the number of convolutional channels C depending on the dataset's feature complexity; adjusting the

number of modulating filters U based on the dataset's pose variation; and max pooling along the orientation axis i.e. over the modulating filters for each pixel of the bank of modulated feature maps. The latter operation has the additional advantage of focusing the attention of the network on (local) dominant orientations, which is a particularly useful feature for orientation invariance.

Invariance vs equivariance - There is a clear relation between pooling technique and invariance versus equivariance. Preserving only the strongest orientation response discards low response representations and disentangles features, this is however at the cost of encouraging invariance to local rotations rather than equivariance. In practice, invariance is achieved through pooling after each hidden layer over the feature orientation dimension or the Gabor shift dimension, for convolutions and cyclic convolutions, respectively – see Fig. 1.

Projection between $\mathbb{C}$ and $\mathbb{R}$ - Finally, since data used in this paper is real, we set the imaginary part of inputs to zero. Some works [38, 42] opt to include a preprocessing step to estimate the imaginary part though we found this had a detrimental effect on performance. For real classification, final complex feature maps must be projected back onto real space. We experimented with several projection methods such as complex linear layers and using magnitudes, but found that simply concatenating real and imaginary values into fully connected linear layers performed best.

4. Experiments

In this section we validate our learnable modulation formulation, showing that learning analytical filter parameters leads to improved accuracy on both artificial and real data. Initially, LGCNs are evaluated on variants of MNIST [43] containing rotated samples, where we evaluate the network's learned invariance. In the next section we compare invariance and equivariance in both a standard CNN and a learnable Gabor modulated CNN, where networks process synthesised samples of galactic cirri. All experiments throughout this section were run using a single NVIDIA GTX 1080 Ti.

4.1. Orientation invariance on MNIST

MNIST [43] (CC BY-SA 3.0 license) is a standard benchmark for transformation invariance because of its simplicity, interpretability and vast array of variants. We apply a random rotation between $[0, 2\pi)$ to yield a rotated MNIST, and train with 5-fold validation. Our baseline classification architecture is similar to that used in [3, 10, 2], with three blocks of increasing channels, representing a hierarchy of feature complexity. Each block contains two learnable Gabor modulated convolutional layers with a kernel size of 3×3 followed by max pooling along the orientation axis and average spatial pooling. We use no cyclic Gabor convolutional layers, but these may be included in future experiments. In the final block, features are pooled globally so that a given activation contains one complex value per feature channel. We then concatenate real and imaginary parts into a single vector and use three (real valued) fully connected layers for classification. The Adam optimiser [44] is used for network training, starting at a learning rate of 0.001 and then decaying with an exponential schedule by 0.9 every epoch. L2 weight regularisation is also enforced with a penalty of 10^{-7}.

Exploration of rotation invariance in the feature maps – The number of modulation filters U has a direct effect on the network's ability to capture rotation dependent features. We vary this parameter and investigate its effect on network's performance and the learned features of the first layer, which are the most directly affected by low-level geometrical transformations of the input image. For this first experiment, we train networks with $U \in \{1, 2, 4, 8, 16\}$. We measure and compare response magnitudes (measured as the ratio between the average magnitudes of input and output activations) between original and rotated samples, for all rotations in the (discrete) range [0,360°], for each network (Fig. 2 right). Though response magnitude varies slightly, this may be largely due to interpolation artefacts caused by rotation of the input samples. Nonetheless, the pattern remains predictable throughout the rotation interval with decreasing amplitude for increasing U, indicating that the number of modulating filters has a direct impact on rotation invariance. We also measure classification accuracy as a function of rotation for 1000 samples from the MNIST test set for each network (Fig. 2 left). The small difference in accuracy between $U = 1$ and others indicates that even a little orientation information is helpful in generating intra-class rotation-invariant features that

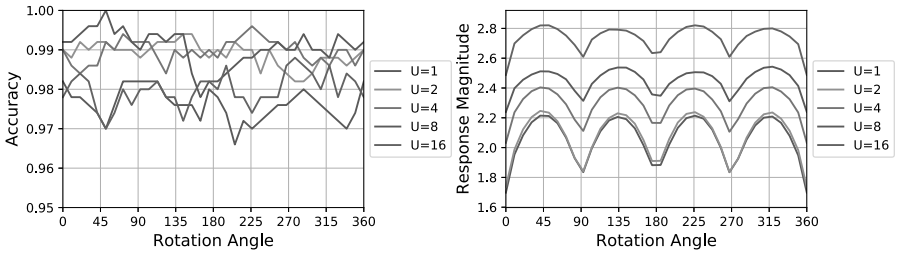

Figure 2: Effect of input rotation on MNIST classification accuracy (left) and magnitude of activations in the first modulated layer of the network (right), for different numbers of modulating filters and orientations, and on a subset of 1000 testing samples of MNIST.

remain inter-class separable. At $U = 16$ there is a detrimental saturation of orientations possibly due to the model becoming too complex for the dataset size and task. Optimal performances are reached for U between 2 and 8, with LGCN being not very sensitive to the exact value of this hyperparameter.

Evaluation of the individual modifications to [7] – We evaluate the performance improvements from our two modifications to [7] individually, namely the use of complex-valued filters and of learnable Gabor orientation parameters. In these experiments we apply these modifications both in turn and jointly to the model of [7]. The channel sizes for each LGCN variant were adjusted so that the total parameter size is at most equal to all of the compared models: for complex models this required halving the number of feature channels. The final results, shown in Tab. 1, show that both modifications improve classification accuracy, demonstrating the additional feature expressivity afforded in comparison to standard CNNs. All variants also outperform GCNs which use real and static Gabor modulation, with an absolute error difference of 0.6 for LGCNs, showing the benefit of our method's changes over the previous work. The combination of modifications leads to a large performance increase that may indicate a synergy between the two approaches. One possible explanation for this is that the complex Gabor filter provides smoother gradients with respect to θ and λ, as opposed to only the real part. We will test this hypothesis in future work.

Effect of adjusting learning strategies of Gabor parameters – We investigate the effect of learning Gabor parameters other than θ, and study how changing initialisation

Table 1: Classification accuracy on randomly rotated MNIST images.

LGCN (proposed)	Complex static	Real learnable	GCN4 [7]	ORN8 [8]	CNN
0.9950	0.9915	0.9911	0.9890	0.9888	0.9718

methods impacts the model's performance. The default training configuration for LGCNs uses fixed σ set at π, and learnable λ initialised using a normal distribution with mean $3\sqrt{U}$ and variance $\frac{\sqrt{U}}{4}$. We experiment with fixing or learning wavelength λ and scale σ and study how different configurations affect LGCNs. In addition we compare additional weight initialisation strategies for these variables. For wavelength initialisation we apply: fixed $\lambda = 3$; normal distribution with unit mean and unit variance (not adjusting for U); and uniform distribution between $[-1.5, 1.5]$. For scale, in addition to fixed $\sigma = \pi$, we initialise with a normal distribution with mean equal to π and quarter variance, and enable backpropagation. Finally, we repeat these experiments with only one λ, σ for all modulating Gabor filters per layer.

We train parameter restricted models with varying Gabor parameter learning strategies on rotated MNIST for 30 epochs with 5-fold validation, and record the average performance over all splits. Results are shown in Table 2. While initialising with a normal distribution $\lambda = \mathcal{N}(3\sqrt{U}, \frac{\sqrt{U}}{4})$ and fixing $\sigma = \pi$ achieves the highest average performance, there is no clear strategy for either variable that remains best with the other variable strategy changed. Notably, there is a performance decrease when aliasing of the modulating Gabor filters is forcibly introduced by initialising wavelength λ from a uniform distribution with bounds $[-1.5, 1.5]$. In further tests it was noticed that in this scenario λ values do not recover from this range of aliasing even after training for >100 epochs. For this experiment we conclude that given parameters are not in an aliasing range, LGCNs are not particularly sensitive to learning strategy of wavelength λ and scale σ.

4.2. Invariance and equivariance to the dominant orientation of galactic cirri

We validate the benefit of modulation by applying LGCNs to a domain demanding robust orientation-sensitive features. We demonstrate that modulation not only enables

Table 2: Classification accuracy on rotated MNIST averaged over 5 splits for different learning strategies of Gabor parameters wavelength λ and scale σ. Rows are divided in the centre to denote whether a single λ and σ is used for all U modulating Gabor filters, or λ and σ are separate for each modulating Gabor filter.

		$\lambda = \mathcal{U}(-1.5, 1.5)$	Fixed $\lambda = 3$	$\mathcal{N}(3, \frac{1}{4})$	$\mathcal{N}(3\sqrt{U}, \frac{\sqrt{U}}{4})$
Separate λ, σ	Fixed $\sigma = \pi$	0.9672	0.9702	0.9686	0.9692
	$\sigma = \mathcal{N}(\pi, \frac{1}{4})$	0.9690	0.9704	0.9693	0.9678
Single λ, σ	Fixed $\sigma = \pi$	0.9684	0.9707	0.9698	**0.9713**
	$\sigma = \mathcal{N}(\pi, \frac{1}{4})$	0.9673	0.9699	0.9707	0.9685

the network to learn invariance and equivariance, but aids the network's ability to generate features unaffected by local disturbances. For these experiments we analyse samples of galactic cirrus clouds – astronomical objects with striped quasi-textures exhibiting clear dominant orientations, as shown in Fig. 3 – as they allow the design of experiments that assess both orientation invariance and equivariance separately and the comparative robustness between different models. These images are very challenging, exhibiting overlapping semi-transparent objects, including foreground cirrus with oriented patterns, background objects (e.g. galaxies) with vastly different textures and intensities, and telescope artefacts. In the initial experiment we evaluate performance on various datasets composed from synthesised images of cirrus structures. Following this, we extend to a real world problem using low surface brightness telescope images containing cirrus contamination.

Generation of synthesised cirrus images – In order to create images exhibiting discriminative features similar to real images of galactic cirrus, multiple noise patterns are combined. All synthesised images are formed from at least three parts, background B, cirrus C, and bright regions R. Pixels to form the background B are drawn from a Gaussian distribution and inverted. The cirrus image C contains textured cloud shapes with smooth boundaries. The cloud shapes are produced by a 2D Gaussian mixture model (GMM) with 4 to 6 randomly located components with random standard deviations. The cirrus texture is created by combining a cloud texture and a streak texture, both generated from Perlin gradient noise of varying frequencies. A binary mask extracted from the GMM that forms the cirrus segmentation target. Bright regions

R are smooth isotropic bright regions resembling regions of diffuse light, and are created from a GMM with a similar process to the cirrus case. Finally, all parts are combined according to $\gamma B + \gamma C + R$ with the denoising target set to $\gamma B + R$, before normalisation between 0 and 1, where $\gamma = 0.4$ balances bright regions vs. background and cirrus regions.

We design the dataset to have three variations of increasing realness. The first variation possesses only cirrus clouds with constant orientation and bright regions; the second randomises cirrus orientation; finally the third introduces star-like objects with telescope halo artefacts (i.e. bright transparent halos around each bright spots simulating stars). These star-like objects are created from a sharp Gaussian profile approximating a point source, where the standard deviation of each star's Gaussian profile is randomly slightly varied to ensure variation. A synthetic halo resembling a telescope artefact is then added around each star, and is created from a circle of fixed radius and width and with a uniform brightness proportional to the star's associated Gaussian standard deviation.

Our synthesised dataset contains 300 samples: 160 for training, 40 for validation (for 5-fold validation) and 100 for testing.

Comparing LGCN against a traditional CNN on synthesised cirrus images – We create a U-Net [45] style architecture in both standard form and with Gabor modulated convolutional layers, where skip connections are combined via summation (as in [46]) rather than concatenation. To enable comparison, we create four variants of this network: one with plain convolutions; one with complex-valued convolutions, denoted $\mathbb{C}$-CNN; one with static real Gabor filter modulation as in [7]; and one with learnable complex Gabor modulation with cyclic convolutions. These networks are tasked with first segmenting the cirrus clouds, and secondly removing clouds and artefacts (if applicable). The complex filters of $\mathbb{C}$-CNN and LGCN naturally require twice the convolutional filter parameters. We ensure a fair comparison by adjusting channel sizes accordingly, thus keeping total parameter size of the two networks roughly equal. For the denoising task we do not utilise orientation pooling so that orientation information is preserved and equivariance is encouraged, as per discussion in 3.4, and experimental verification. Results are presented in Tab. 3 with IoU metric for

segmentation and peak signal to noise ratio (PSNR) for denoising.

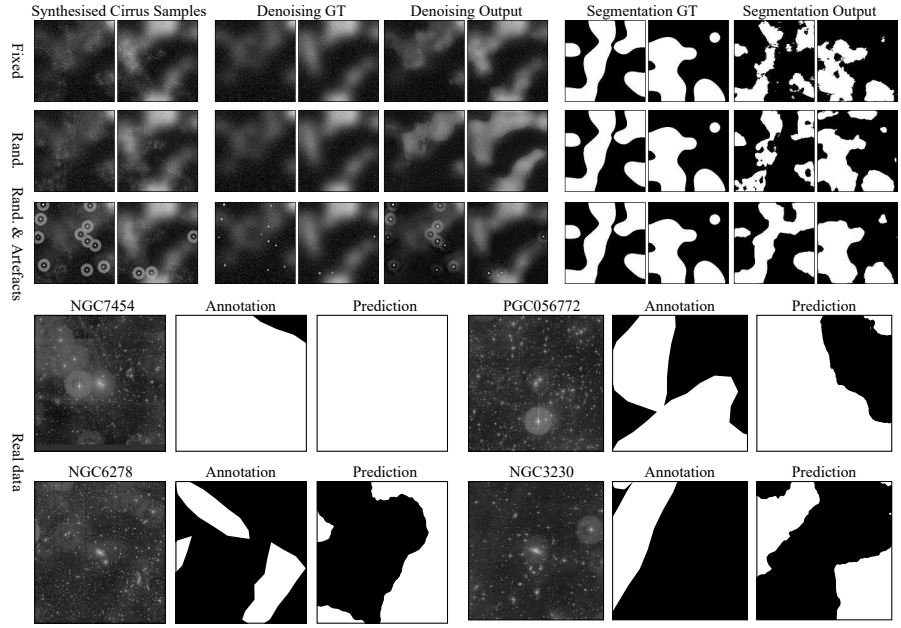

Figure 3: Denoising and segmentation results on real and synthesised samples of galactic cirri generated with fixed rotation; randomised rotation; and randomised rotation with stars and telescope artefacts. These are difficult tasks as the striped textures of cirrus regions are easily confused/obstructed with bright diffuse regions and other objects.

In the segmentation case, given that rotation of cirrus texture does not affect the cloud's envelope, this is a problem where invariance is beneficial. The denoising problem requires equivariance, as isolating the cloud from the detailed background is dependent on the orientation of its streaks. In the first dataset, CNN, $\mathbb{C}$-CNN and GCN performances are close to LGCN's, generating fine segmentations of the cirrus clouds with few missed regions, and similarly good denoising. For the second, with orientation variation, GCN performs marginally best in the segmentation case but its performance drops for denoising. On the other hand, LGCN maintains a quite stable performance for denoising (and also for segmentation) on this dataset. This difference in behaviour may be explained by the use of cyclic convolutions in LGCN that better preserve orientation information due to rotational weight-tying across layers. CNN and $\mathbb{C}$-CNN

Table 3: Segmentation IoU (left), denoising PSNR (middle) on synthesised cirri with fixed and randomised orientation, and with stars and telescope artefacts. Segmentation IoU (right) on real cirrus samples in LSB images. *Gabor convolutions of [7] applied to our base model

	Fixed	Rand.	Rand & artefacts	Fixed	Rand.	Rand & artefacts	Real cirrus
Base model	0.914	0.882	0.806	26.3	25.3	23.5	0.684
$\mathbb{C}$ base model	0.918	0.905	0.839	26.7	25.6	24.5	0.691
GC model*	0.923	**0.920**	0.875	26.7	26.2	24.8	0.685
LGCN	**0.925**	0.918	**0.898**	**27.4**	**27.1**	**25.4**	**0.731**

performances start to fall behind, with relative differences of 4.1% and 1.4% IoU, respectively, for segmentation and 7.1% and 5.9% PSNR for denoising. This separation becomes much larger in the final experiment on the most complex data exhibiting overlapping textured regions and localised objects, with LGCNs outperforming GCNs in both tasks by 2.6% and 2.4%, $\mathbb{C}$-CNN by 7.0% and 3.7%, and plain CNNs by 11.4% and 8.1%. The affect of randomising rotations and even introducing telescope artefacts makes little difference to LGCN's performance for segmentation, demonstrating its strength in generating rotation invariant features that are robust to local disturbances. We see that denoising performance is stable with randomised rotation, indicating equivariant encoding produced by the modulated layers. While performance drops for the third dataset, due to artefacts introducing strong variations locally, LGCN still outperforms other models by a larger margin than without artefacts, showing that feature robustness is exhibited in the equivariant case. The results demonstrate that both the use of complex numbers and the modulation of filters are beneficial. We note that the margin between CNN and $\mathbb{C}$-CNN is significantly the largest on this dataset variation requiring robustness, compared to other dataset variations. LGCN combines both augmentations and cyclic convolutions for further improved results. Visual analysis of the network outputs (see Fig. 3) indicates a possible overfitting for the first two datasets with no telescope artefacts for both the segmentation and denoising tasks, which results in a more difficult generalisation and poorer (visual) quality on test data. This issue may be

due to these two simpler scenarios requiring simpler models and/or fewer training steps, and it will be investigated in future work.

Prediction of cirrus structures in LSB images – We task LGCNs with segmenting cirrus clouds in optical telescope images, demonstrating the effectiveness of our method on a real world problem. This dataset [47] contains 48 expert annotated images of approximate size 5000x5000, with two channels representing different optical wavelength bands. Of the 48 images, we use 32 for training, 8 for validation (for 5-fold validation) and 8 for testing. Across the entire dataset 55% of pixels are labelled as cirrus contaminated. In their original form, objects other than bright stars are indistinguishable from darkness, as the number of photons emitted by the brightest stars is several orders of magnitude larger than other objects such as cirrus clouds. To compensate for this we add an initial layer to the networks prior to projection onto complex space, implementing arcsinh scaling, popular in astronomical image processing, with learned parameters: $Y_0 = \text{arcsinh}\,(aX_0 + b)$, where a, b are real scalars. Following this a sigmoid function is applied and bad pixels are set to zero. Finally scaled images are concatenated onto original input to account for oversaturation of very strong cirrus regions. Initial scaling parameters were determined from a simple gradient descent algorithm, with the target set as auto-scaled versions of images using astronomical image software [48, 49] (GPL v3 license).

Networks are trained over 300 epochs on random crops of size 512x512, which are then downscaled by a factor of two. To mitigate against the limited sample size, we augment data with random flips and 90° rotations, and pretrain networks on an extended version (N=1200) of the synthesised dataset. We also train a standard CNN, a $\mathbb{C}$-CNN, and a GCN [7] for comparison as in the synthesised data experiment, fixing parameter size to be roughly equal. In comparison to the synthesised images, real cirrus regions often exhibit much fainter textures and orientation is more subtle and can vary slightly globally. In addition, training labels may not be fully reliable, due to the difficulty in annotating precisely the borders of cirri – there is an inherent uncertainty associated with each annotation, especially due to the ambiguous nature of the cirrus cloud boundary, so several experts may disagree on the exact location of borders –, and due to the limited number of available expert annotators – in our dataset, 2 to 3 annotators annotated each

image, but for simplicity in this proof-of-concept, we worked only with the annotations of the single most experienced expert, making the simplified assumption that their annotation corresponds to the ground-truth. These factors, in combination with more severe artefacts, background noise, and small training size, make the dataset incredibly challenging. Results are shown in the last column of Table 3: LGCNs achieve an IoU of 73.1%, with an absolute increase of 4.7% over standard CNN, 4.0% over complex CNN and 4.6% over real-valued static Gabor modulation without cyclic convolutions. Notably, GCN barely surpasses the base model and is outperformed by $\mathbb{C}$-CNN, suggesting that only static rotation sensitivity is not sufficient on more challenging datasets, a finding which is supported by results from the previous experiment of synthesised images. LGCNs significantly outperforms compared methods, demonstrating the ability of proposed augmentations to generate robust orientation sensitive features, even on data with extreme contamination. Given that the class balance is 55%, the problem is very difficult, and although this absolute increase represents a significant performance improvement, more progress may be achieved by considering e.g. new architectures to be augmented by our methods, or a multi-scale approach, and a more complete dataset with consensus annotation from several experts.

4.3. Application to boundary detection in natural images

We demonstrate the general applicability of learnable Gabor modulated convolutions on the Berkeley Segmentation Dataset (BSD500) [50, 51]. This task requires the ability to learn equivariant features in the scenario where there is no dominant global orientation, and the network must handle high variations in local feature orientation dependence. The dataset contains natural images of size 321×481 in both portrait and landscape, with 200 training samples, 100 validation samples and 200 testing samples. Each image has associated with it several ground truth labellings produced by different annotators.

We replicate the pipeline of one of the highest performing methods, RCF [52], and replace convolutional operators with learnable Gabor modulated convolutions. This methodology uses a pretrained VGG16 [53] based architecture, taking 'side' outputs from each convolutional block that represent coarser scale edges as network depth increases. These side outputs are then fused together with a 1×1 convolutional layer. The

Table 4: Boundary detection results on the BSD500 [50] dataset. *Our parameter restricted implementation. †ImageNet pretrained.

	Kivinen et al. [54]	DexiNed [55]	RCF* [52]	H-Net [2]	LGCN-RCF	RCF† [52]
ODS	0.702	0.728	0.707	0.726	0.727	0.806
OIS	0.715	0.745	0.720	0.742	0.747	0.823
# params	-	4.41M	1.80M	0.12M	1.88M	14.84M

final prediction is then computed as the average between all side outputs and the fused output. We denote our modified implementation as LGCN-RCF: an additional Gabor convolutional layer is used to create an orientation channel; all convolutional layers apart from the fusion layer are replaced with cyclic Gabor convolutions; orientation features are pooled prior to side output. We train for 250 epochs with 3-fold validation. Results are shown in Table 4, using the optimal dataset scale (ODS) and optimal image scale (OIS) metrics defined in [50]. LGCN-RCF achieves 0.727 ODS and 0.747 OIS, which is a strong result considering in each epoch we train on one random augmentation per image, as opposed to other methods which use the entire range of augmentations per image (due to lack of compute and time, thus care is to be taken when comparing results). Our method also significantly outperforms our implementation of RCF [52] with parameters restricted to match LGCN-RCF, demonstrating the benefit of modulating with complex Gabor filters on tasks with natural images.

5. Conclusion

We presented a framework for incorporating adaptive modulation into complex-valued CNNs. This framework was used to design an orientation robust network with convolutional layers using Gabor modulated weights, where complex convolutional filters and Gabor parameters are learned simultaneously. A cyclic convolutional layer was proposed to retain rotational information throughout layers and encourage equivariance. Our architecture is able to generate unconstrained representations dependent on exact orientations, without interpolation artefacts. We validated this empirically for three use cases, with experiments designed to test properties of both invariance and equivariance

to orientation. We first verified that LGCNs are able to effectively produce rotation invariant features on the rotated MNIST dataset. An ablation study was performed to assess in turn and in combination the effect of two proposed augmentations to GCNs [7], namely using complex-valued weights and learning parameters of modulating Gabor filters. Secondly, we carried out experiments on a purpose designed dataset of varying difficulty. The architecture's modulated layers were able to create fine segmentations in synthetic and real images despite local disturbances. The presented LGCN architecture achieved strong denoising scores in comparison to standard CNNs, even on contaminating cirrus cloud structures with randomised orientation. Clear performance improvements were observed for both use cases, demonstrating the effectiveness of the augmentations. Thirdly, we applied an LGCN architecture to boundary detection in natural images and achieved strong metrics in comparison to other non-pretrained methods. The successful augmentation of three different architectures also demonstrates the general applicability of our method, and it may be applied to more complex DNNs and application scenarios in the future.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] T. S. Cohen, M. Welling, Steerable cnns, in: Proceedings of the 5th International Conference on Learning Representations, 2017.

[2] D. E. Worrall, S. J. Garbin, D. Turmukhambetov, G. J. Brostow, Harmonic Networks: Deep Translation and Rotation Equivariance, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5028–5037.

[3] M. Weiler, F. A. Hamprecht, M. Storath, Learning Steerable Filters for Rotation Equivariant CNNs, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 849–858.

[4] D. Marcos, M. Volpi, N. Komodakis, D. Tuia, Rotation equivariant vector field networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 5048–5057.

[5] D. Marcos, M. Volpi, D. Tuia, Learning rotation invariant convolutional filters for texture classification, in: 2016 23rd International Conference on Pattern Recognition (ICPR), 2016, pp. 2012–2017.

[6] D. Laptev, N. Savinov, J. M. Buhmann, M. Pollefeys, TI-POOLING: transformation-invariant pooling for feature learning in convolutional neural networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 289–297.

[7] S. Luan, C. Chen, B. Zhang, J. Han, J. Liu, Gabor convolutional networks, IEEE Transactions on Image Processing 27 (2018) 4357–4366.

[8] Y. Zhou, Q. Ye, Q. Qiu, J. Jiao, Oriented response networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 519–528.

[9] X. Cheng, Q. Qiu, R. Calderbank, G. Sapiro, RotDCF: decomposition of convolutional filters for rotation-equivariant deep networks, in: Proceedings of the 7th International Conference on Learning Representations, 2018.

[10] T. S. Cohen, M. Welling, Group Equivariant Convolutional Networks, in: International conference on Machine Learning, 2016, pp. 2990–2999. doi:10.1017/S1464793101005656.

[11] S. Dieleman, J. De Fauw, K. Kavukcuoglu, Exploiting Cyclic Symmetry in Convolutional Neural Networks, International Conference on Machine Learning 48 (2016) 1889–1898. doi:10.1525/aa.1976.78.4.02a00030.

[12] E. J. Bekkers, M. W. Lafarge, M. Veta, K. A. J. Eppenhof, J. P. W. Pluim, R. Duits, Roto-translation covariant convolutional networks for medical image analysis, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, 2018, pp. 440–448.

[13] D. G. Lowe, Object recognition from local scale-invariant features, in: Proceedings of the seventh IEEE international conference on computer vision, volume 2, 1999, pp. 1150–1157.

[14] G. M. Haley, B. S. Manjunath, Rotation-invariant texture classification using a complete space-frequency model, IEEE transactions on Image Processing 8 (1999) 255–269.

[15] S. Arivazhagan, L. Ganesan, S. P. Priyal, Texture classification using Gabor wavelets based rotation invariant features, Pattern Recognition Letters 27 (2006) 1976–1982. doi:10.1016/J.PATREC.2006.05.008.

[16] P. Y. Simard, D. Steinkraus, J. C. Platt, others, Best practices for convolutional neural networks applied to visual document analysis., in: Icdar, volume 3, 2003.

[17] A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, Advances in neural information processing systems (2012) 1097–1105.

[18] D. Ciregan, U. Meier, J. Schmidhuber, Multi-column deep neural networks for image classification, in: 2012 IEEE conference on computer vision and pattern recognition, 2012, pp. 3642–3649.

[19] J. Wu, Y. Yu, C. Huang, K. Yu, Deep multiple instance learning for image classification and auto-annotation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3460–3469.

[20] A. Kanazawa, A. Sharma, D. W. Jacobs, Locally Scale-Invariant Convolutional Neural Networks, Deep Learning and Representation Learning Workshop: Neural Information Processing System (2014).

[21] K. Sohn, H. Lee, Learning Invariant Representations with Local Transformations, in: Proceedings of the 29th International Coference on International Conference on Machine Learning, ICML'12, Omnipress, Madison, WI, USA, 2012, pp. 1339–1346.

[22] S. Dieleman, K. W. Willett, J. Dambre, Rotation-invariant convolutional neural networks for galaxy morphology prediction, Monthly Notices of the Royal Astronomical Society 450 (2014) 1441–1459. doi:10.1186/1748-5908-4-32.

[23] R. Gens, P. M. Domingos, Deep symmetry networks, in: Advances in neural information processing systems, 2014, pp. 2537–2545.

[24] J. F. Henriques, A. Vedaldi, Warped convolutions: Efficient invariance to spatial transformations, in: Proceedings of the 34th International Conference on Machine Learning-Volume 70, 2017, pp. 1461–1469.

[25] M. Jaderberg, K. Simonyan, A. Zisserman, others, Spatial transformer networks, in: Advances in neural information processing systems, 2015, pp. 2017–2025.

[26] C.-H. Lin, S. Lucey, Inverse compositional spatial transformer networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2568–2576.

[27] X. Jia, B. De Brabandere, T. Tuytelaars, L. V. Gool, Dynamic filter networks, in: Advances in Neural Information Processing Systems, 2016, pp. 667–675.

[28] T. S. Cohen, M. Geiger, J. Köhler, M. Welling, Spherical cnns, in: Proceedings of the 6th International Conference on Learning Representations, 2018.

[29] J.-H. Jacobsen, B. De Brabandere, A. W. M. Smeulders, Dynamic steerable blocks in deep residual networks, in: Proceedings of the British Machine Vision Conference, 2017.

[30] M. Finzi, S. Stanton, P. Izmailov, A. G. Wilson, Generalizing convolutional neural networks for equivariance to lie groups on arbitrary continuous data, in: International Conference on Machine Learning, 2020, pp. 3165–3176.

[31] M. Weiler, G. Cesa, General E (2)-Equivariant Steerable CNNs, in: Advances in Neural Information Processing Systems, 2019, pp. 14334–14345.

[32] L. Sifre, S. Mallat, Rotation, scaling and deformation invariant scattering for texture discrimination, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2013, pp. 1233–1240.

[33] J. Bruna, S. Mallat, Invariant scattering convolution networks, IEEE transactions on pattern analysis and machine intelligence 35 (2013) 1872–1886.

[34] R. Ejbali, M. Zaied, A dyadic multi-resolution deep convolutional neural wavelet network for image classification, Multimedia Tools and Applications 77 (2018) 6149–6163.

[35] S. Fujieda, K. Takayama, T. Hachisuka, Wavelet convolutional neural networks, arXiv preprint arXiv:1805.08620 (2018).

[36] X. Wang, B. Zhang, C. Li, R. Ji, J. Han, X. Cao, J. Liu, Modulated convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 840–848.

[37] H. Khan, B. Yener, Learning filter widths of spectral decompositions with wavelets, in: Advances in Neural Information Processing Systems, 2018, pp. 4601–4612.

[38] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, C. J. Pal, Deep Complex Networks, in: Proceedings of the 6th International Conference on Learning Representations, 2018.

[39] M. Arjovsky, A. Shah, Y. Bengio, Unitary evolution recurrent neural networks, in: International Conference on Machine Learning, 2016, pp. 1120–1128.

[40] W. T. Freeman, E. H. Adelson, others, The design and use of steerable filters, IEEE Transactions on Pattern analysis and machine intelligence 13 (1991) 891–906.

[41] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, in: Proceedings of the IEEE international conference on computer vision, 2015, pp. 1026–1034.

[42] S. Scardapane, S. Van Vaerenbergh, A. Hussain, A. Uncini, Complex-valued neural networks with nonparametric activation functions, IEEE Transactions on Emerging Topics in Computational Intelligence (2018).

[43] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE 86 (1998) 2278–2324.

[44] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, in: Proceedings of the 3rd International Conference on Learning Representations, 2014.

[45] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, 2015, pp. 234–241.

[46] T. M. Quan, D. G. C. Hildebrand, W.-K. Jeong, FusionNet: A Deep Fully Residual Convolutional Neural Network for Image Segmentation in Connectomics, Frontiers in Computer Science 3 (2021) 34. URL: https://www.frontiersin.org/article/10.3389/fcomp.2021.613981. doi:10.3389/fcomp.2021.613981.

[47] E. Sola, P.-A. Duc, F. Richards, A. Paiement, M. Urbano, J. Klehammer, M. Bílek, J.-C. Cuillandre, S. Gwyn, Characterization of low surface brightness structures in annotated deep images, A&A (in prep).

[48] T. Boch, P. Fernique, Aladin Lite: Embed your Sky in the browser, Astronomical data analysis software and systems XXIII 485 (2014) 277.

[49] F. Bonnarel, P. Fernique, O. Bienaymé, D. Egret, F. Genova, M. Louys, F. Ochsenbein, M. Wenger, J. G. Bartlett, The ALADIN interactive sky atlas-A reference tool for identification of astronomical sources, Astronomy and Astrophysics Supplement Series 143 (2000) 33–40.

[50] P. Arbelaez, M. Maire, C. Fowlkes, J. Malik, Contour Detection and Hierarchical Image Segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 898–916. URL: http://dx.doi.org/10.1109/TPAMI.2010.161. doi:10.1109/TPAMI.2010.161.

[51] D. Martin, C. Fowlkes, D. Tal, J. Malik, A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, in: Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, volume 2, 2001, pp. 416–423. doi:10.1109/ICCV.2001.937655.

[52] Y. Liu, M.-M. Cheng, X. Hu, K. Wang, X. Bai, Richer convolutional features for edge detection, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 3000–3009.

[53] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, in: International Conference on Learning Representations, 2015.

[54] J. Kivinen, C. Williams, N. Heess, Visual Boundary Prediction: A Deep Neural Prediction Network and Quality Dissection, in: S. Kaski, J. Corander (Eds.), Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics, volume 33 of *Proceedings of Machine Learning Research*, PMLR, Reykjavik, Iceland, 2014, pp. 512–521. URL: http://proceedings.mlr.press/v33/kivinen14.html.

[55] X. S. Poma, E. Riba, A. Sappa, Dense extreme inception network: Towards a robust cnn model for edge detection, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2020, pp. 1923–1932.

Bibliography

Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2003, AJ, 126, 2081

Abbott, T. M. C., Abdalla, F. B., Allam, S., et al. 2018, ApJS, 239, 18

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

Abraham, R. G., van den Bergh, S., Glazebrook, K., et al. 1996, ApJS, 107, 1

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

Abraham, R. G. & van Dokkum, P. G. 2014, PASP, 126, 55

Ackermann, S., Schawinski, K., Zhang, C., Weigel, A. K., & Turp, M. D. 2018, MNRAS, 479, 415

Adams, S. M., Zaritsky, D., Sand, D. J., et al. 2012, AJ, 144, 128

Afruni, A., Fraternali, F., & Pezzulli, G. 2019, A&A, 625, A11

Agertz, O., Teyssier, R., & Moore, B. 2009, MNRAS, 397, L64

Aihara, H., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, S8

Alabi, A. B., Romanowsky, A. J., Forbes, D. A., Brodie, J. P., & Okabe, N. 2020, MNRAS, 496, 3182

Ambartsumian, V. A. 1961, AJ, 66, 536

Amorisco, N. C. 2015, MNRAS, 450, 575

Amorisco, N. C. 2016, Monthly Notices of the Royal Astronomical Society, 464, 2882

Ann, H. B., Seo, M., & Ha, D. K. 2015, ApJS, 217, 27

Annibali, F., Beccari, G., Bellazzini, M., et al. 2020, MNRAS, 491, 5101

Antoja, T., Ramos, P., Mateu, C., et al. 2020, A&A, 635, L3

Arp, H. 1965, ApJ, 142, 402

Arp, H. 1966, ApJS, 14, 1

Arp, H. C. 1972, in External Galaxies and Quasi-Stellar Objects, ed. D. S. Evans, D. Wills, & B. J. Wills, Vol. 44, 380

Atkinson, A. M., Abraham, R. G., & Ferguson, A. M. N. 2013, ApJ, 765, 28

Balcells, M. & Quinn, P. J. 1990, ApJ, 361, 381

Barchi, P. H., de Carvalho, R. R., Rosa, R. R., et al. 2020, Astronomy and Computing, 30, 100334

Barnes, J. E. 1988, ApJ, 331, 699

Barnes, J. E. 1992, ApJ, 393, 484

Barnes, J. E. & Hernquist, L. 1992, Nature, 360, 715

Baugh, C. M., Benson, A. J., Cole, S., Frenk, C. S., & Lacey, C. 2003, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini, 91

Bekki, K. 1998, ApJ, 499, 635

Bekki, K. 1999, ApJL, 510, L15

Bekki, K. 2009, MNRAS, 399, 2221

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289

Bell, E. F., Naab, T., McIntosh, D. H., et al. 2006, ApJ, 640, 241

Belokurov, V. 2013, NewAR, 57, 100

Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, ApJL, 642, L137

Benson, A. J. & Bower, R. 2010, arXiv e-prints, arXiv:1004.1162

Bernardi, M., Shankar, F., Hyde, J. B., et al. 2010, MNRAS, 404, 2087

Berrier, J. C., Stewart, K. R., Bullock, J. S., et al. 2008, The Astrophysical Journal, 690, 1292

Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393

Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 228

Bertin, E., Pillay, R., & Marmo, C. 2015, Astronomy and Computing, 10, 43

Bertola, F. & Capaccioli, M. 1975, ApJ, 200, 439

Bezanson, R., van Dokkum, P., & Franx, M. 2012, ApJ, 760, 62

Bianconi, M., Smith, G. P., Haines, C. P., et al. 2018, MNRAS, 473, L79

Bickley, R. W., Bottrell, C., Hani, M. H., et al. 2021, MNRAS, 504, 372

Bigiel, F., Leroy, A. K., Walter, F., et al. 2011, ApJL, 730, L13

Bílek, M., Duc, P.-A., Cuillandre, J.-C., et al. 2020, MNRAS, 498, 2138

Bílek, M., Duc, P.-A., & Sola, E. 2022, arXiv e-prints, arXiv:2210.02478

Bílek, M., Fensch, J., Ebrová, I., et al. 2021, arXiv e-prints, arXiv:2111.14886

Bílek, M., Thies, I., Kroupa, P., & Famaey, B. 2018, A&A, 614, A59

Bílek, M., Thies, I., Kroupa, P., & Famaey, B. 2019, arXiv e-prints, arXiv:1908.07537

Binggeli, B., Sandage, A., & Tammann, G. A. 1985, AJ, 90, 1681

Bland-Hawthorn, J., Sutherland, R., Agertz, O., & Moore, B. 2007, ApJL, 670, L109

Bluck, A. F. L., Conselice, C. J., Buitrago, F., et al. 2012, The Astrophysical Journal, 747, 34

Blum, R. D., Burleigh, K., Dey, A., et al. 2016, in American Astronomical Society Meeting Abstracts, Vol. 228, American Astronomical Society Meeting Abstracts #228, 317.01

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

Bode, P., Ostriker, J. P., & Turok, N. 2001, ApJ, 556, 93

Bois, M., Emsellem, E., Bournaud, F., et al. 2011, Monthly Notices of the Royal Astronomical Society, 416, 1654

Borlaff, A., Trujillo, I., Román, J., et al. 2019, A&A, 621, A133

Boselli, A., Fossati, M., Ferrarese, L., et al. 2018, A&A, 614, A56

Boselli, A., Fossati, M., & Sun, M. 2022, A&A Rv, 30, 3

Bothun, G. D., Impey, C. D., Malin, D. F., & Mould, J. R. 1987, AJ, 94, 23

Boulade, O., Charlot, X., Abbon, P., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, ed. M. Iye & A. F. M. Moorwood, 72–81

Bridge, C. R., Carlberg, R. G., & Sullivan, M. 2010, ApJ, 709, 1067

Brough, S., Collins, C., Demarco, R., et al. 2020, arXiv e-prints, arXiv:2001.11067

Brough, S., van de Sande, J., Owers, M. S., et al. 2017, ApJ, 844, 59

Buitrago, F., Trujillo, I., Conselice, C. J., et al. 2008, ApJL, 687, L61

Bullock, J. S. & Johnston, K. V. 2005, ApJ, 635, 931

Burbidge, E. M. & Burbidge, G. R. 1963, ApJ, 138, 1306

Butcher, H. & Oemler, A., J. 1984, ApJ, 285, 426

Byrd, G. & Valtonen, M. 1990, ApJ, 350, 89

Byun, W., Sheen, Y.-K., Ho, L. C., et al. 2018, AJ, 156, 249

Camps, P. & Baes, M. 2014, arXiv e-prints, arXiv:1410.1629

Canalizo, G., Bennert, N., Jungwiert, B., et al. 2007, ApJ, 669, 801

Cappellari, M. 2008, MNRAS, 390, 71

Cappellari, M. 2016, ARA&A, 54, 597

Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126

Cappellari, M., Emsellem, E., Bacon, R., et al. 2007, MNRAS, 379, 418

Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011a, MNRAS, 413, 813

Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011b, MNRAS, 416, 1680

Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2012, Nature, 484, 485

Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2013a, MNRAS, 432, 1862

Cappellari, M., Scott, N., Alatalo, K., et al. 2013b, MNRAS, 432, 1709

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, in Interstellar Dust, ed. L. J. Allamandola & A. G. G. M. Tielens, Vol. 135, 5–10

Carlsten, S. G., Greco, J. P., Beaton, R. L., & Greene, J. E. 2020, ApJ, 891, 144

Casteels, K. R. V., Bamford, S. P., Skibba, R. A., et al. 2013, MNRAS, 429, 1051

Cheng, T.-Y., Huertas-Company, M., Conselice, C. J., et al. 2021, MNRAS, 503, 4446

Cheselka, M. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 172, Astronomical Data Analysis Software and Systems VIII, ed. D. M. Mehringer, R. L. Plante, & D. A. Roberts, 349

Chiboucas, K., Karachentsev, I. D., & Tully, R. B. 2009, AJ, 137, 3009

Chung, A., van Gorkom, J. H., Kenney, J. D. P., Crowl, H., & Vollmer, B. 2009, AJ, 138, 1741

Cimatti, A., Cassata, P., Pozzetti, L., et al. 2008, A&A, 482, 21

Colbert, J. W., Mulchaey, J. S., & Zabludoff, A. I. 2001, AJ, 121, 808

Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168

Combes, F. 2008, in Astronomical Society of the Pacific Conference Series, Vol. 396, Formation and Evolution of Galaxy Disks, ed. J. G. Funes & E. M. Corsini, 325

Combes, F. & Tiret, O. 2010, in American Institute of Physics Conference Series, Vol. 1241, Invisible Universe, ed. J.-M. Alimi & A. Fuözfa, 154–161

Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2007, ApJ, 668, 826

Conselice, C. J. 2009, MNRAS, 399, L16

Conselice, C. J., Bershady, M. A., Dickinson, M., & Papovich, C. 2003, AJ, 126, 1183

Conselice, C. J., Rajgor, S., & Myers, R. 2008, MNRAS, 386, 909

Cooper, A. P., Cole, S., Frenk, C. S., et al. 2010, MNRAS, 406, 744

Cortese, L., Bendo, G. J., Isaak, K. G., Davies, J. I., & Kent, B. R. 2010, MNRAS, 403, L26

Cox, T. J., Dutta, S. N., Di Matteo, T., et al. 2006, ApJ, 650, 791

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11

Cuillandre, J.-C. J., Withington, K., Hudelot, P., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8448, Observatory Operations: Strategies, Processes, and Systems IV, ed. A. B. Peck, R. L. Seaman, & F. Comeron, 84480M

Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686

Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156

Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, ApJ, 626, 680

Dalcanton, J. J., Spergel, D. N., & Summers, F. J. 1997, ApJ, 482, 659

Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, ApJ, 695, 101

Davies, J. I., Phillipps, S., & Disney, M. J. 1988, MNRAS, 231, 69P

Davies, J. I., Phillipps, S., & Disney, M. J. 1989, MNRAS, 239, 703

Davies, J. I., Phillipps, S., & Disney, M. J. 1990, MNRAS, 244, 385

Davies, J. I., Wilson, C. D., Auld, R., et al. 2010, MNRAS, 409, 102

Davies, R. L. & Birkinshaw, M. 1986, ApJL, 303, L45

de Vaucouleurs, G. 1959, Handbuch der Physik, 53, 275

Deason, A. J., Belokurov, V., Evans, N. W., & Johnston, K. V. 2013, ApJ, 763, 113

Dehnen, W., Odenkirchen, M., Grebel, E. K., & Rix, H.-W. 2004, AJ, 127, 2753

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

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

Dekel, A. & Silk, J. 1986, ApJ, 303, 39

Di Matteo, P., Pipino, A., Lehnert, M. D., Combes, F., & Semelin, B. 2009, A&A, 499, 427

Dieleman, S., Willett, K. W., & Dambre, J. 2015, MNRAS, 450, 1441

Disney, M. J. 1976, Nature, 263, 573

Doliva-Dolinsky, A., Martin, N. F., Thomas, G. F., et al. 2022, ApJ, 933, 135

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

Domínguez Sánchez, H., Martin, G., Damjanov, I., et al. 2023, arXiv e-prints, arXiv:2303.03407

Dressler, A. 1980, ApJ, 236, 351

Dressler, A., Smail, I., Poggianti, B. M., et al. 1999, ApJS, 122, 51

Dubinski, J., Mihos, J. C., & Hernquist, L. 1996, ApJ, 462, 576

Duc, P.-A. 2012, in Astrophysics and Space Science Proceedings, Vol. 28, Dwarf Galaxies: Keys to Galaxy Formation and Evolution, 305

Duc, P.-A., Cuillandre, J.-C., Karabal, E., et al. 2015, MNRAS, 446, 120

Duc, P. A. & Mirabel, I. F. 1994, A&A, 289, 83

Duc, P.-A. & Renaud, F. 2013, in Tides in Astronomy and Astrophysics (Springer Berlin Heidelberg), 327–369

Duncan, J. C. 1923, ApJ, 57, 137

Dupraz, C. & Combes, F. 1986, A&A, 166, 53

Durret, F., Chiche, S., Lobo, C., & Jauzac, M. 2021, A&A, 648, A63

Ebrova, I. 2013, arXiv e-prints, arXiv:1312.1643

Ebrová, I., Bílek, M., & Jungwiert, B. 2019, arXiv e-prints, arXiv:1909.07393

Ebrová, I., Bílek, M., Vudragović, A., Yıldız, M. K., & Duc, P.-A. 2021, A&A, 650, A50

Ebrova, I. & collaborators. 2023, in prep

Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, ApJ, 136, 748

Elmegreen, B. G., Kaufman, M., & Thomasson, M. 1993, ApJ, 412, 90

Emsellem, E., Cappellari, M., Krajnović, D., et al. 2011, MNRAS, 414, 888

Emsellem, E., Cappellari, M., Krajnović, D., et al. 2007, MNRAS, 379, 401

Eneev, T. M., Kozlov, N. N., & Sunyaev, R. A. 1973, A&A, 22, 41

Euclid Collaboration, Borlaff, A. S., Gómez-Alvarez, P., et al. 2022, A&A, 657, A92

Eyre, A. & Binney, J. 2011, MNRAS, 413, 1852

Falcón-Barroso, J. & Knapen, J. H. 2013, Secular Evolution of Galaxies

Famaey, B. & McGaugh, S. S. 2012, Living Reviews in Relativity, 15, 10

Feldmann, R., Mayer, L., & Carollo, C. M. 2008, ApJ, 684, 1062

Fenner, Y. & Gibson, B. K. 2003, PASA, 20, 189

Ferguson, A. M. N., Irwin, M. J., Ibata, R. A., Lewis, G. F., & Tanvir, N. R. 2002, AJ, 124, 1452

Fernique, P., Allen, M. G., Boch, T., et al. 2015, A&A, 578, A114

Fernique, P., Boch, T., Donaldson, T., et al. 2014, MOC - HEALPix Multi-Order Coverage map Version 1.0, IVOA Recommendation 02 June 2014

Ferrarese, L., Côté, P., Cuillandre, J.-C., et al. 2012, ApJS, 200, 4

Fliri, J. & Trujillo, I. 2016, MNRAS, 456, 1359

Forbes, D. A., Brodie, J. P., & Grillmair, C. J. 1997, AJ, 113, 1652

Forrest, B., Marsan, Z. C., Annunziatella, M., et al. 2020, ApJ, 903, 47

Förster Schreiber, N. M., Genzel, R., Bouché, N., et al. 2009, ApJ, 706, 1364

Fortson, L., Masters, K., Nichol, R., et al. 2012, in Advances in Machine Learning and Data Mining for Astronomy, ed. M. J. Way, J. D. Scargle, K. M. Ali, & A. N. Srivastava, 213–236

Franx, M., Illingworth, G., & Heckman, T. 1989, ApJ, 344, 613

Freeman, K. C. 1970, ApJ, 160, 811

Fujita, Y. 2004, PASJ, 56, 29

Garstang, R. H. 1989, PASP, 101, 306

Gavazzi, G., Boselli, A., Scodeggio, M., Pierini, D., & Belsole, E. 1999, MNRAS, 304, 595

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

Genel, S., Bouché, N., Naab, T., Sternberg, A., & Genzel, R. 2010, ApJ, 719, 229

Genzel, R., Tacconi, L. J., Gracia-Carpio, J., et al. 2010, MNRAS, 407, 2091

Ghosh, A., Urry, C. M., Wang, Z., et al. 2020, ApJ, 895, 112

Gibbons, S. L. J., Belokurov, V., & Evans, N. W. 2014, MNRAS, 445, 3788

Gilbert, K. M., Guhathakurta, P., Beaton, R. L., et al. 2012, ApJ, 760, 76

Gilhuly, C., Hendel, D., Merritt, A., et al. 2020, ApJ, 897, 108

Gilhuly, C., Merritt, A., Abraham, R., et al. 2022, ApJ, 932, 44

Girardi, M., Biviano, A., Giuricin, G., Mardirossian, F., & Mezzetti, M. 1993, ApJ, 404, 38

Giri, G., Barway, S., & Raychaudhury, S. 2023, MNRAS, 520, 5870

Gnedin, O. Y. 2003, ApJ, 582, 141

Goto, T., Okamura, S., Sekiguchi, M., et al. 2003a, PASJ, 55, 757

Goto, T., Yamauchi, C., Fujita, Y., et al. 2003b, MNRAS, 346, 601

Graham, M. T., Cappellari, M., Bershady, M. A., & Drory, N. 2019, arXiv e-prints, arXiv:1910.05136

Greco, J. P., Greene, J. E., Price-Whelan, A. M., et al. 2018, PASJ, 70, S19

Greene, J. E., Leauthaud, A., Emsellem, E., et al. 2017, ApJL, 851, L33

Greener, M. J., Merrifield, M., Aragón-Salamanca, A., et al. 2021, MNRAS, 502, L95

Gunn, J. E. & Gott, J. Richard, I. 1972, ApJ, 176, 1

Habas, R., Marleau, F. R., Duc, P.-A., et al. 2020, MNRAS, 491, 1901

Hatch, N. A., Cooke, E. A., Muldrew, S. I., et al. 2017, MNRAS, 464, 876

Hausen, R. & Robertson, B. 2019, Morpheus: Library to generate morphological semantic segmentation maps of astronomical images, Astrophysics Source Code Library, record ascl:1906.012

Helmi, A. 2004, ApJL, 610, L97

Hendel, D. & Johnston, K. V. 2015, MNRAS, 454, 2472

Hennig, C. 2007, Comput. Statist. Data Analysis, 52, 258

Hernquist, L. 1989, Nature, 340, 687

Hernquist, L. & Quinn, P. J. 1987, ApJ, 312, 1

Hester, J. A. 2006, The Astrophysical Journal, 647, 910

Hibbard, J. E. & Mihos, J. C. 1995, AJ, 110, 140

Hill, A. R., Muzzin, A., Franx, M., et al. 2017a, ApJ, 837, 147

Hill, A. R., Muzzin, A., Franx, M., & Marchesini, D. 2017b, ApJL, 849, L26

Hilz, M., Naab, T., & Ostriker, J. P. 2013, MNRAS, 429, 2924

Hilz, M., Naab, T., Ostriker, J. P., et al. 2012, MNRAS, 425, 3119

Hodges, J. L. 1958, Arkiv for Matematik, 3, 469

Hoffman, L., Cox, T. J., Dutta, S., & Hernquist, L. 2010, ApJ, 723, 818

Holmberg, E. 1941, ApJ, 94, 385

Hood, C. E., Kannappan, S. J., Stark, D. V., et al. 2018, ApJ, 857, 144

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

Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 477, 1578

Howard, S., Keel, W. C., Byrd, G., & Burkey, J. 1993, ApJ, 417, 502

Hoyle, B., Masters, K. L., Nichol, R. C., Jimenez, R., & Bamford, S. P. 2012, MNRAS, 423, 3478

Huang, Q. & Fan, L. 2022, ApJS, 262, 39

Hubble, E. 1929a, Proceedings of the National Academy of Science, 15, 168

Hubble, E. P. 1925, Popular Astronomy, 33, 252

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

Hubble, E. P. 1929b, ApJ, 69, 103

Huertas-Company, M., Gravet, R., Cabrera-Vives, G., et al. 2015, ApJS, 221, 8

Ibata, R., Irwin, M., Lewis, G., Ferguson, A. M. N., & Tanvir, N. 2001a, Nature, 412, 49

Ibata, R., Lewis, G. F., Irwin, M., Totten, E., & Quinn, T. 2001b, ApJ, 551, 294

Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, Nature, 370, 194

Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017, ApJ, 848, 128

Illingworth, G. 1977, ApJL, 218, L43

Impey, C. & Bothun, G. 1997, ARA&A, 35, 267

Impey, C., Bothun, G., & Malin, D. 1988, ApJ, 330, 634

Impey, C. D., Sprayberry, D., Irwin, M. J., & Bothun, G. D. 1996, ApJS, 105, 209

Iodice, E., Capaccioli, M., Grado, A., et al. 2016, ApJ, 820, 42

Iodice, E., Spavone, M., Capaccioli, M., et al. 2017, ApJ, 839, 21

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111

Jackson, T. M., Pasquali, A., La Barbera, F., More, S., & Grebel, E. K. 2021, arXiv e-prints, arXiv:2102.02241

Janowiecki, S., Mihos, J. C., Harding, P., et al. 2010, ApJ, 715, 972

Jarrett, T. H., Polletta, M., Fournon, I. P., et al. 2006, AJ, 131, 261

Javanmardi, B., Martinez-Delgado, D., Kroupa, P., et al. 2016, A&A, 588, A89

Jedrzejewski, R. & Schechter, P. L. 1989, AJ, 98, 147

Ji, I., Hasan, I., Schmidt, S. J., & Tyson, J. A. 2018, PASP, 130, 084504

Ji, I., Peirani, S., & Yi, S. K. 2014, A&A, 566, A97

Johnston, K. V. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 136, Galactic Halos, ed. D. Zaritsky, 365

Johnston, K. V., Bullock, J. S., Sharma, S., et al. 2008, ApJ, 689, 936

Johnston, K. V., Hernquist, L., & Bolte, M. 1996, ApJ, 465, 278

Joung, M. R., Bryan, G. L., & Putman, M. E. 2012, ApJ, 745, 148

Joye, W. A. & Mandel, E. 2003, in Astronomical Society of the Pacific Conference Series, Vol. 295, Astronomical Data Analysis Software and Systems XII, ed. H. E. Payne, R. I. Jedrzejewski, & R. N. Hook, 489

Kado-Fong, E., Greene, J. E., Hendel, D., et al. 2018, ApJ, 866, 103

Karabal, E., Duc, P. A., Kuntschner, H., et al. 2017, A&A, 601, A86

Karachentsev, I. D. & Nasonova, O. G. 2010, MNRAS, 405, 1075

Karademir, G. S., Remus, R.-S., Burkert, A., et al. 2019, MNRAS, 487, 318

Katz, N., Keres, D., Dave, R., & Weinberg, D. H. 2003, in Astrophysics and Space Science Library, Vol. 281, The IGM/Galaxy Connection. The Distribution of Baryons at z=0, ed. J. L. Rosenberg & M. E. Putman, 185

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 54

Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, MNRAS, 264, 201

Kaviraj, S. 2010, MNRAS, 406, 382

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

Kawata, D. & Mulchaey, J. S. 2008, ApJL, 672, L103

Kawinwanichakij, L., Papovich, C., Quadri, R. F., et al. 2014, The Astrophysical Journal, 792, 103

Keenan, D. W. & Innanen, K. A. 1975, AJ, 80, 290

Keenan, P. C. 1935, ApJ, 81, 355

Kelvin, L. S., Hasan, I., & Tyson, J. A. 2023, MNRAS, 520, 2484

Kennicutt, R. C. & Evans, N. J. 2012, ARA&A, 50, 531

Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160

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

Khalid, A. & collaborators. 2023, in prep

Kim, T., Sheth, K., Hinz, J. L., et al. 2012, ApJ, 753, 43

Kluge, M., Neureiter, B., Riffeser, A., et al. 2020, ApJS, 247, 43

Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82

Knapen, J. H. & Trujillo, I. 2017, in Astrophysics and Space Science Library, Vol. 434, Outskirts of Galaxies, ed. J. H. Knapen, J. C. Lee, & A. Gil de Paz, 255

Koch, E. & Rosolowsky, E. 2015, MNRAS, 452, 3435

Koposov, S., Belokurov, V., Evans, N. W., et al. 2008, ApJ, 686, 279

Koposov, S. E., Erkal, D., Li, T. S., et al. 2023, MNRAS, 521, 4936

Koposov, S. E., Yoo, J., Rix, H.-W., et al. 2009, ApJ, 696, 2179

Kormendy, J. 1979, ApJ, 227, 714

Kormendy, J. & Bahcall, J. N. 1974, AJ, 79, 671

Kormendy, J. & Kennicutt, Robert C., J. 2004, ARA&A, 42, 603

Krajnović, D., Emsellem, E., Cappellari, M., et al. 2011, MNRAS, 414, 2923

Kravtsov, A. 2010, Advances in Astronomy, 2010, 1

Kroupa, P. 2015, Canadian Journal of Physics, 93, 169

Lackner, C. N., Cen, R., Ostriker, J. P., & Joung, M. R. 2012, MNRAS, 425, 641

Lagos, C. d. P., Schaye, J., Bahé, Y., et al. 2018, MNRAS, 476, 4327

Lagos, C. d. P., Theuns, T., Stevens, A. R. H., et al. 2017, MNRAS, 464, 3850

Larson, R. B. 1972, Nature, 236, 21

Larson, R. B. 1975, MNRAS, 173, 671

Larson, R. B. & Tinsley, B. M. 1978, ApJ, 219, 46

Larson, R. B., Tinsley, B. M., & Caldwell, C. N. 1980, ApJ, 237, 692

Lauberts, A. 1974, A&A, 33, 231

Leavitt, H. S. & Pickering, E. C. 1912, Harvard College Observatory Circular, 173, 1

Lee, J. H., Kim, M., Ree, C. H., et al. 2013, ApJL, 762, L4

L'Huillier, B., Combes, F., & Semelin, B. 2012, A&A, 544, A68

Lim, S. & collaborators. 2023, submitted

Lintott, C. J., Schawinski, K., Slosar, A., et al. 2008, MNRAS, 389, 1179

López-Sanjuan, C., Balcells, M., Pérez-González, P. G., et al. 2009, A&A, 501, 505

Lotz, J. M., Jonsson, P., Cox, T. J., & Primack, J. R. 2008, MNRAS, 391, 1137

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

LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv e-prints, arXiv:0912.0201

Luppino, G. A., Mezger, M. R., & Miyazaki, S. 1995, in New Developments in Array Technology and Applications, ed. A. G. D. Philip, K. Janes, & A. R. Upgren, Vol. 167, 297

Lux, H., Read, J. I., Lake, G., & Johnston, K. V. 2013, MNRAS, 436, 2386

Magnier, E. A. & Cuillandre, J. C. 2004, PASP, 116, 449

Malhan, K. & Ibata, R. A. 2019, MNRAS, 486, 2995

Malhan, K., Ibata, R. A., & Martin, N. F. 2018, MNRAS, 481, 3442

Malin, D. F. 1977, AAS Photo Bulletin, 16, 10

Malin, D. F. 1978, Nature, 276, 591

Malin, D. F. & Carter, D. 1983, ApJ, 274, 534

Mancillas, B., Duc, P.-A., Combes, F., et al. 2019, A&A, 632, A122

Mantha, K. B., McIntosh, D. H., Ciaschi, C. P., et al. 2019, MNRAS, 486, 2643

Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2017, arXiv e-prints, arXiv:1707.03396

Martin, G., Bazkiaei, A. E., Iodice, M. S. E., et al. 2022, MNRAS

Martin, G., Kaviraj, S., Hocking, A., Read, S. C., & Geach, J. E. 2020, MNRAS, 491, 1408

Martin, N. F., Ibata, R. A., McConnachie, A. W., et al. 2013, ApJ, 776, 80

Martin, N. F., Ibata, R. A., Rich, R. M., et al. 2014, ApJ, 787, 19

Martínez-Delgado, D., Cooper, A. P., Román, J., et al. 2023, A&A, 671, A141

Martínez-Delgado, D., Gabany, R. J., Crawford, K., et al. 2010, AJ, 140, 962

Masters, K. L., Lintott, C. J., Hart, R. E., et al. 2019, MNRAS, 487, 1808

McConnachie, A. W., Irwin, M. J., Ibata, R. A., et al. 2009, Nature, 461, 66

McDermid, R. M., Alatalo, K., Blitz, L., et al. 2015, MNRAS, 448, 3484

McGaugh, S. S., Bothun, G. D., & Schombert, J. M. 1995, AJ, 110, 573

McIntosh, D. H., Guo, Y., Hertzberg, J., et al. 2008, MNRAS, 388, 1537

Merritt, A., Pillepich, A., van Dokkum, P., et al. 2020, MNRAS, 495, 4570

Merritt, A., van Dokkum, P., Abraham, R., & Zhang, J. 2016, ApJ, 830, 62

Meza, A., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003, ApJ, 590, 619

Mihos, C. 2003, arXiv e-prints, astro

Mihos, J. C. 1995, ApJL, 438, L75

Mihos, J. C. 2016, in The General Assembly of Galaxy Halos: Structure, Origin and Evolution, ed. A. Bragaglia, M. Arnaboldi, M. Rejkuba, & D. Romano, Vol. 317, 27–34

Mihos, J. C. 2019, arXiv e-prints, arXiv:1909.09456

Mihos, J. C., Durrell, P. R., Ferrarese, L., et al. 2015, ApJL, 809, L21

Mihos, J. C., Harding, P., Feldmeier, J., & Morrison, H. 2005, ApJL, 631, L41

Milgrom, M. 1983, ApJ, 270, 371

Mirabel, I. F., Dottori, H., & Lutz, D. 1992, A&A, 256, L19

Miskolczi, A., Bomans, D. J., & Dettmar, R. J. 2011, A&A, 536, A66

Miville-Deschênes, M. A., Duc, P. A., Marleau, F., et al. 2016, A&A, 593, A4

Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution

Montes, M. 2019, arXiv e-prints, arXiv:1912.01616

Montes, M., Brough, S., Owers, M. S., & Santucci, G. 2021, ApJ, 910, 45

Mood, A. 1950, Introduction to the Theory of Statistics, McGraw-Hill, 394

Moore, B., Ghigna, S., Governato, F., et al. 1999, ApJL, 524, L19

Moore, B., Katz, N., Lake, G., Dressler, A., & Oemler, A. 1996, Nature, 379, 613

Moore, B., Lake, G., & Katz, N. 1998, ApJ, 495, 139

Morales, G., Martínez-Delgado, D., Grebel, E. K., et al. 2018, A&A, 614, A143

Mosenkov, A. V., Reshetnikov, V. P., Skryabina, M. N., & Shakespear, Z. 2022, Research in Astronomy and Astrophysics, 22, 115003

Muñoz, R. P., Eigenthaler, P., Puzia, T. H., et al. 2015, ApJL, 813, L15

Müller, O. & Jerjen, H. 2020, A&A, 644, A91

Müller, O., Jerjen, H., & Binggeli, B. 2017, A&A, 597, A7

Müller, O. & Schnider, E. 2021, The Open Journal of Astrophysics, 4, 3

Murali, C., Katz, N., Hernquist, L., Weinberg, D. H., & Davé, R. 2002, The Astrophysical Journal, 571, 1

Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, ApJL, 699, L178

Naab, T., Johansson, P. H., Ostriker, J. P., & Efstathiou, G. 2007, ApJ, 658, 710

Naab, T., Oser, L., Emsellem, E., et al. 2014, MNRAS, 444, 3357

Nadler, E. O., Gluscevic, V., Boddy, K. K., & Wechsler, R. H. 2019, ApJL, 878, L32

Naiman, J. P., Pillepich, A., Springel, V., et al. 2017, arXiv e-prints, arXiv:1707.03401

Nair, P. B. & Abraham, R. G. 2010, ApJS, 186, 427

Nelson, D., Pillepich, A., Springel, V., et al. 2019, arXiv e-prints, arXiv:1902.05554

Nelson, D., Pillepich, A., Springel, V., et al. 2017, arXiv e-prints, arXiv:1707.03395

Nevin, R., Blecha, L., Comerford, J., et al. 2023, arXiv e-prints, arXiv:2303.06249

Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, ApJ, 746, 162

Nibauer, J., Bonaca, A., & Johnston, K. V. 2023, arXiv e-prints, arXiv:2303.17406

Nipoti, C., Londrillo, P., & Ciotti, L. 2007, MNRAS, 381, L104

Ocvirk, P., Pichon, C., & Teyssier, R. 2008, Monthly Notices of the Royal Astronomical Society, 390, 1326

Oemler, Augustus, J. 1974, ApJ, 194, 1

Oh, S., Kim, K., Lee, J. H., et al. 2018, ApJS, 237, 14

Okazaki, T. & Taniguchi, Y. 2000, ApJ, 543, 149

Olson, K. M. & Kwan, J. 1990, ApJ, 361, 426

Oogi, T. & Habe, A. 2013, in The Intriguing Life of Massive Galaxies, ed. D. Thomas, A. Pasquali, & I. Ferreras, Vol. 295, 183–183

Oser, L., Ostriker, J. P., Naab, T., Johansson, P. H., & Burkert, A. 2010, ApJ, 725, 2312

Ostriker, J. P. & Steinhardt, P. J. 1995, Nature, 377, 600

Pawlik, M. M., Wild, V., Walcher, C. J., et al. 2016, MNRAS, 456, 3032

Peñarrubia, J., Benson, A. J., Walker, M. G., et al. 2010, MNRAS, 406, 1290

Pearson, K. 1895, Proceedings of the Royal Society of London, 58, 240–242

Pearson, W. J., Wang, L., Trayford, J. W., Petrillo, C. E., & van der Tak, F. F. S. 2019, A&A, 626, A49

Peirani, S., Crockett, R. M., Geen, S., et al. 2010, MNRAS, 405, 2327

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, AJ, 139, 2097

Penoyre, Z., Moster, B. P., Sijacki, D., & Genel, S. 2017, MNRAS, 468, 3883

Peters, S. P. C., van der Kruit, P. C., Knapen, J. H., et al. 2017, Monthly Notices of the Royal Astronomical Society, 470, 427

Phillipps, S., Disney, M. J., Kibblewhite, E. J., & Cawson, M. G. M. 1987, MNRAS, 229, 505

Pillepich, A., Nelson, D., Hernquist, L., et al. 2017, arXiv e-prints, arXiv:1707.03406

Pillepich, A., Nelson, D., Springel, V., et al. 2019, arXiv e-prints, arXiv:1902.05553

Pillepich, A., Vogelsberger, M., Deason, A., et al. 2014, MNRAS, 444, 237

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13

Pogge, R. W. & Martini, P. 2002, ApJ, 569, 624

Poggianti, B. M., Fasano, G., Omizzolo, A., et al. 2016, AJ, 151, 78

Poggianti, B. M., Smail, I., Dressler, A., et al. 1999, ApJ, 518, 576

Pop, A.-R., Pillepich, A., Amorisco, N. C., & Hernquist, L. 2018, MNRAS, 480, 1715

Poulain, M., Marleau, F. R., Habas, R., et al. 2021, MNRAS, 506, 5494

Prieur, J. L. 1990, Status of shell galaxies., ed. R. Wielen, 72–83

Prunet, S., Fouque, S., & Gwyn, S. 2015

Putman, M. E. 2017, in Astrophysics and Space Science Library, Vol. 430, Gas Accretion onto Galaxies, ed. A. Fox & R. Davé, 1

Quinn, P. J. 1984, ApJ, 279, 596

Rees, M. J. & Ostriker, J. P. 1977, MNRAS, 179, 541

Regnault, N., Conley, A., Guy, J., et al. 2009, A&A, 506, 999

Rich, R. M., Mosenkov, A., Lee-Saunders, H., et al. 2019, MNRAS, 490, 1539

Richards, F. 2023, PhD thesis, Swansea University

Richards, F., Paiement, A., Xie, X., Sola, E., & Duc, P.-A. 2020, arXiv e-prints, arXiv:2011.11734

Robotham, A. S. G., Driver, S. P., Davies, L. J. M., et al. 2014, MNRAS, 444, 3986

Rodriguez-Gomez, V., Genel, S., Vogelsberger, M., et al. 2015, MNRAS, 449, 49

Rodriguez-Gomez, V., Pillepich, A., Sales, L. V., et al. 2016, MNRAS, 458, 2371

Rudick, C. S., Mihos, J. C., Frey, L. H., & McBride, C. K. 2009, ApJ, 699, 1518

Rudick, C. S., Mihos, J. C., & McBride, C. 2006, ApJ, 648, 936

Russell, B., Torralba, A., Murphy, K., & Freeman, W. 2008, International Journal of Computer Vision, 77, 157–173

Samland, M. & Gerhard, O. E. 2003, A&A, 399, 961

Sancisi, R., Fraternali, F., Oosterloo, T., & van der Hulst, T. 2008, A&A Rv, 15, 189

Sandage, A. 1961, The Hubble Atlas of Galaxies

Sandage, A. 1976, AJ, 81, 954

Sandin, C. 2014, A&A, 567, A97

Schade, D., Lilly, S. J., Crampton, D., et al. 1995, ApJL, 451, L1

Schulze, F., Remus, R.-S., & Dolag, K. 2017, Galaxies, 5

Schwartzenberg, J. M., Phillipps, S., Smith, R. M., Couch, W. J., & Boyle, B. J. 1995, MNRAS, 275, 121

Schweizer, F. 1982, ApJ, 252, 455

Schweizer, F. 1998, in Saas-Fee Advanced Course 26: Galaxies: Interactions and Induced Star Formation, ed. J. Kennicutt, R. C., F. Schweizer, J. E. Barnes, D. Friedli, L. Martinet, & D. Pfenniger, 105

Schweizer, F. 2006, arXiv e-prints, astro

Schweizer, F. & Seitzer, P. 1988, ApJ, 328, 88

Sekachev, B., Manovich, N., Zhiltsov, M., et al. 2020, opencv/cvat: v1.1.0

Shapley, H. & Curtis, H. D. 1921, Bulletin of the National Research Council, 2, 171

Sharina, M. E., Karachentsev, I. D., Dolphin, A. E., et al. 2008, MNRAS, 384, 1544

Sheen, Y.-K., Yi, S. K., Ree, C. H., & Lee, J. 2012, ApJS, 202, 8

Silk, J. 1977, ApJ, 211, 638

Silk, J. 2003, Ap&SS, 284, 663

Silk, J. & Rees, M. J. 1998, A&A, 331, L1

Slater, C. T., Harding, P., & Mihos, J. C. 2009, PASP, 121, 1267

Smirnov, A. A., Savchenko, S. S., Poliakov, D. M., et al. 2023, MNRAS, 519, 4735

Sola, E., Duc, P.-A., Richards, F., et al. 2022, A&A, 662, A124

Sola, E., Duc, P.-A., Urbano, M., & collaborators. 2023, in prep

Somerville, R. S. & Primack, J. R. 1998, arXiv e-prints, astro

Sorce, J. G., Gottlöber, S., Hoffman, Y., & Yepes, G. 2016, MNRAS, 460, 2015

Spearman, C. 1904, American Journal of Psychology, 15, 72–101

Spindler, A., Geach, J. E., & Smith, M. J. 2021, MNRAS, 502, 985

Spitler, L. R., Longbottom, F. D., Alvarado-Montes, J. A., et al. 2019, arXiv e-prints, arXiv:1911.11579

Springel, V. & Hernquist, L. 2005, ApJL, 622, L9

Springel, V., Pakmor, R., Pillepich, A., et al. 2017, arXiv e-prints, arXiv:1707.03397

Stewart, K. R., Bullock, J. S., Wechsler, R. H., Maller, A. H., & Zentner, A. R. 2008, ApJ, 683, 597

Storey-Fisher, K., Huertas-Company, M., Ramachandra, N., et al. 2020, arXiv e-prints, arXiv:2012.08082

Stringer, M. J. & Benson, A. J. 2007, MNRAS, 382, 641

Student. 1908, Biometrika, 1–25

Szomoru, D., Franx, M., & van Dokkum, P. G. 2012, ApJ, 749, 121

Sánchez-Alarcón, P. M., Román, J., Knapen, J. H., et al. 2023, The AMIGA sample of isolated galaxies. XIV. Disc breaks and interactions through ultra-deep optical imaging

Tacchella, S., Conroy, C., Faber, S. M., et al. 2022, ApJ, 926, 134

Tal, T., van Dokkum, P. G., Nelan, J., & Bezanson, R. 2009, AJ, 138, 1417

Tanoglidis, D., ÄiprijanoviÄ, A., & Drlica-Wagner, A. 2020, DeepShadows: Finding low-surface-brightness galaxies in survey images, Astrophysics Source Code Library, record ascl:2011.026

Thomas, D., Greggio, L., & Bender, R. 1999, MNRAS, 302, 537

Thomas, G. F., Famaey, B., Ibata, R., Lüghausen, F., & Kroupa, P. 2017, A&A, 603, A65

Tinsley, B. M. 1981, ApJ, 250, 758

Tiret, O. & Combes, F. 2007, arXiv e-prints, arXiv:0712.1459

Tody, D. 1986, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 627, Instrumentation in astronomy VI, ed. D. L. Crawford, 733

Toft, S., van Dokkum, P., Franx, M., et al. 2007, ApJ, 671, 285

Tohill, C., Ferreira, L., Conselice, C. J., Bamford, S. P., & Ferrari, F. 2021, ApJ, 916, 4

Tolman, R. C. 1930, Proceedings of the National Academy of Science, 16, 511

Tolman, R. C. 1934, Relativity, Thermodynamics, and Cosmology

Toomre, A. 1977, in Evolution of Galaxies and Stellar Populations, ed. B. M. Tinsley & D. C. Larson, Richard B. Gehret, 401

Toomre, A. & Toomre, J. 1972, ApJ, 178, 623

Trujillo, I., D'Onofrio, M., Zaritsky, D., et al. 2021, A&A, 654, A40

Trujillo, I., Ferreras, I., & de La Rosa, I. G. 2011, MNRAS, 415, 3903

Trujillo, I., Feulner, G., Goranova, Y., et al. 2006, MNRAS, 373, L36

Trujillo, I. & Fliri, J. 2016, ApJ, 823, 123

Tully, R. B. & Shaya, E. J. 1984, ApJ, 281, 31

Turner, J. A., Phillipps, S., Davies, J. I., & Disney, M. J. 1993, MNRAS, 261, 39

Valenzuela, L. M. & Remus, R.-S. 2022, arXiv e-prints, arXiv:2208.08443

Valls-Gabaud, D. & MESSIER Collaboration. 2017, in Formation and Evolution of Galaxy Outskirts, ed. A. Gil de Paz, J. H. Knapen, & J. C. Lee, Vol. 321, 199–201

van de Sande, J., Bland-Hawthorn, J., Brough, S., et al. 2017, MNRAS, 472, 1272

van den Bergh, S. 1976, ApJ, 206, 883

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., et al. 2014, arXiv e-prints, arXiv:1407.6245

van Dokkum, P. G. 2005, AJ, 130, 2647

van Dokkum, P. G., Abraham, R., & Merritt, A. 2014, ApJL, 782, L24

van Dokkum, P. G. & Conroy, C. 2010, Nature, 468, 940

van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, The Astrophysical Journal, 677, L5

van Dokkum, P. G., Whitaker, K. E., Brammer, G., et al. 2010, ApJ, 709, 1018

Vandame, B. 2001, in Mining the Sky, ed. A. J. Banday, S. Zaroubi, & M. Bartelmann, 595

Vázquez-Mata, J. A., Hernández-Toledo, H. M., Avila-Reese, V., et al. 2022, MNRAS, 512, 2222

Veale, M., Ma, C.-P., Thomas, J., et al. 2017, MNRAS, 464, 356

Vega-Ferrero, J., Domínguez Sánchez, H., Bernardi, M., et al. 2021, MNRAS, 506, 1927

Venhola, A., Peletier, R., Laurikainen, E., et al. 2017, A&A, 608, A142

Venkatapathy, Y., Bravo-Alfaro, H., Mayya, Y. D., et al. 2017, AJ, 154, 227

Vorontsov-Velyaminov, B. 1962, in Problems of Extra-Galactic Research, ed. G. C. McVittie, Vol. 15, 194

Vorontsov-Vel'Yaminov, B. A. 1965, Soviet Astronomy, 8, 649

Vulcani, B., Marchesini, D., De Lucia, G., et al. 2016, ApJ, 816, 86

Walmsley, M., Ferguson, A. M. N., Mann, R. G., & Lintott, C. J. 2019, MNRAS, 483, 2968

Walmsley, M., Lintott, C., Géron, T., et al. 2022, MNRAS, 509, 3966

Wang, B., Cappellari, M., Peng, Y., & Graham, M. 2020, MNRAS, 495, 1958

Watkins, A. E., Mihos, J. C., & Harding, P. 2015, ApJL, 800, L3

Webb, T., Noble, A., DeGroot, A., et al. 2015, ApJ, 809, 173

Wen, Z. Z., Zheng, X. Z., & An, F. X. 2014, ApJ, 787, 130

White, S. D. M. & Frenk, C. S. 1991, ApJ, 379, 52

White, S. D. M. & Rees, M. J. 1978, MNRAS, 183, 341

Wilkinson, A., Sparks, W. B., Carter, D., & Malin, D. A. 1987, in Structure and Dynamics of Elliptical Galaxies, ed. P. T. de Zeeuw, Vol. 127, 465

Willmer, C. N. A. 2018, ApJS, 236, 47

Yoon, Y., Ko, J., & Kim, J.-W. 2023, ApJ, 946, 41

Yoon, Y. & Lim, G. 2020, ApJ, 905, 154

York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579

Yorke, H. W., Tenorio-Tagle, G., Bodenheimer, P., & Rozyczka, M. 1989, A&A, 216, 207

Zasov, A. V. 1968, Astrophysics, 4, 173

Zhang, Z., Zou, Z., Li, N., & Chen, Y. 2022, Research in Astronomy and Astrophysics, 22, 055002

Zheng, Z. J., Peng, Q. Y., Vienne, A., Lin, F. R., & Guo, B. F. 2022, A&A, 661, A75

Zirm, A. W., van der Wel, A., Franx, M., et al. 2007, ApJ, 656, 66

Zwicky, F. 1953, Physics Today, 6, 7

Zwicky, F. 1956, Ergebnisse der exakten Naturwissenschaften, 29, 344

Zwicky, F. 1962, PASP, 74, 70

Zwicky, F. 1963, Leaflet of the Astronomical Society of the Pacific, 9, 17

UNIVERSITÉ DE STRASBOURG

Elisabeth SOLA

Observatoire astronomique de Strasbourg

Sonder la croissance des galaxies avec les données de faible brillance de surface (LSB) de relevés d'imagerie profonde

Résumé

Les modèles actuels prédisent la formation des galaxies par fusions successives et accrétion de gaz. Ces interactions gravitationnelles arrachent des étoiles à leurs galaxies, qui forment alors de grandes structures étendues appelées débris de marée, dont les propriétés dépendent du type de collision. Leur étude est cruciale car elle permet de remonter l'histoire passée de la galaxie. Cependant, leur faible brillance de surface (LSB) rend leur détection compliquée.

De larges analyses quantitatives de ces débris sont nécessaires afin de contraindre les modèles d'évolution galactique. Pour cela, durant ma thèse j'ai utilisé et contribué à améliorer un nouvel outil d'annotation spécialement dédié à l'annotation des structures LSB superposées sur des images profondes du télescope CFHT. J'ai caractérisé quantitativement cet échantillon statistiquement significatif de débris de marée autour de 475 galaxies proches et massives. Je me suis ensuite concentrée sur l'influence de la masse et de l'environnement sur la croissance des galaxies à travers les débris LSB. Les résultats obtenus sont cohérents avec la littérature et les modèles actuels.

Mots-clés : astrophysique, archéologie galactique, structures de faible brillance de surface, débris de marée

Abstract

Current models predict that galaxies form through successive mergers and accretion of gas. These gravitational interactions strip galaxies of their stars, which then form extended and diffuse structures, known as tidal features, whose properties depend on the type of merger that occurred. Their study is crucial as it enables to probe the past assembly history of galaxies. However, their detection is made complicated by their Low Surface Brightness (LSB).

Large quantitative analyses of these tidal features are needed to constrain models of galactic evolution. To that end, during my PhD I used and contributed to develop a new annotation tool dedicated to the annotation of LSB structures superimposed on deep images from the CFHT telescope. I quantitatively characterized this statistically significant sample of tidal features around 475 nearby massive galaxies. Then, I focused on the influence of the mass and environment on galactic growth through LSB features. The results are consistent with the literature and with the current models.

Keywords : astrophysics, galactic archaeology, low surface brightness structures, tidal debris