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

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 $1° \times 1°$ 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 $6 \times 10^9 \ M_\odot$. 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

---

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:

$$\text{asinh}(\text{ADU}) = \log\left(\alpha \times (\text{ADU} - b) + \sqrt{\alpha^2 \times (\text{ADU} - b)^2 + 1}\right)$$

(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:

$$\mu = -2.5 \log_{10}\left(\frac{\text{ADU} - b}{pixsize^2}\right) + 30.$$

(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
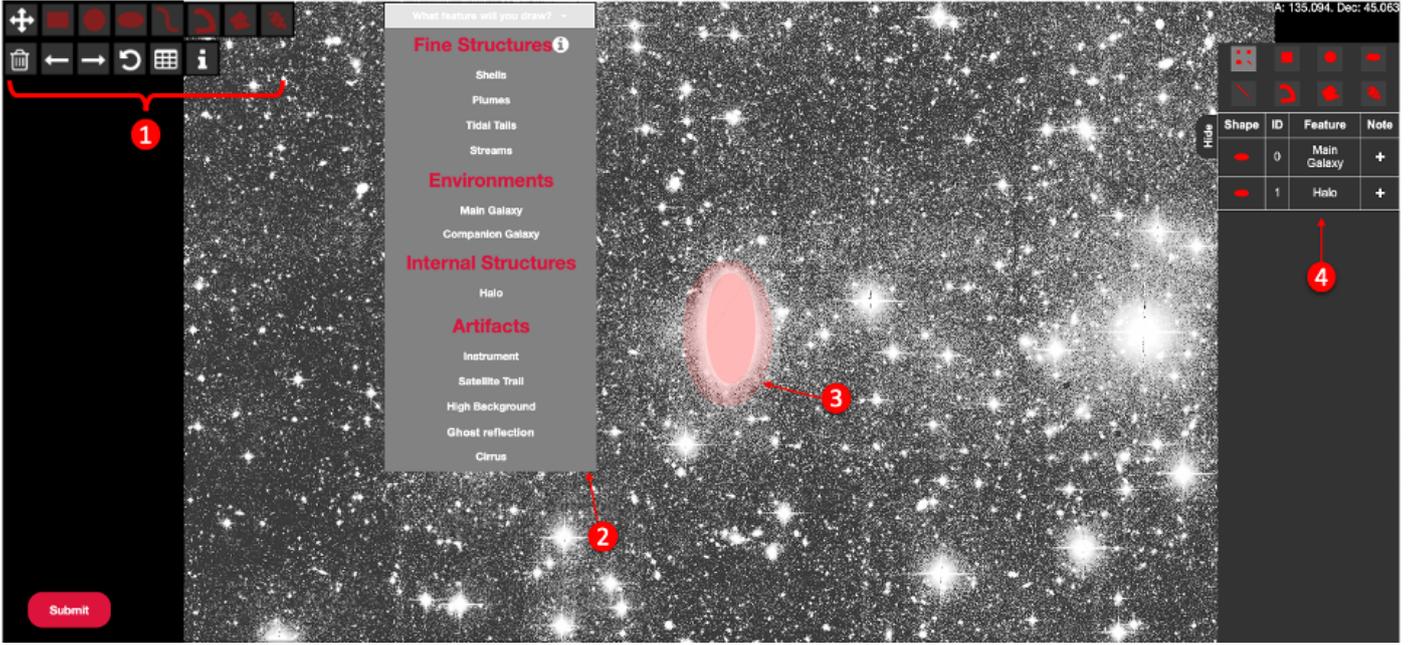

Fig. 1. Annotation interface with its main facilities: drawing buttons (label 1), classification menu (label 2), examples of already drawn annotations (label 3) and summary table (label 4). The background is a CFIS asinh-scaled image.

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:
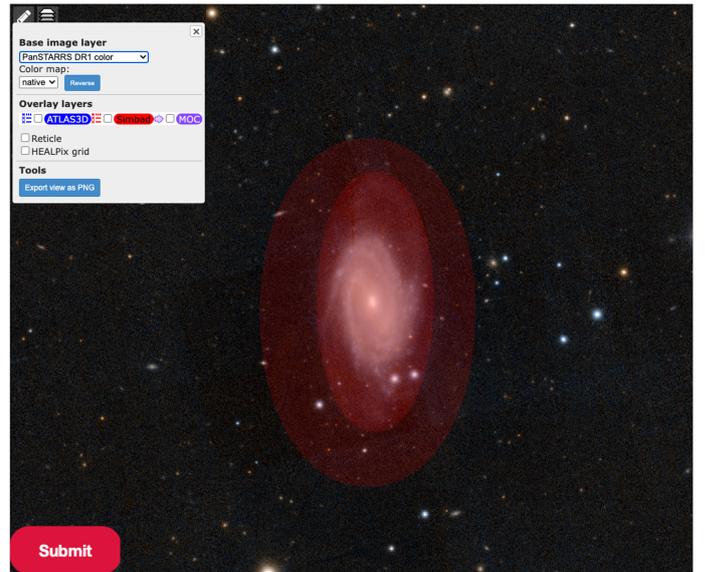

Fig. 2. Zoom on a shallower version of Fig. 1 with the same annotated galaxy. The Pan-STARRS DR1 RGB-color image layer that was used to delineate the "main galaxy" (inner red ellipse) is shown as background.

– 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.
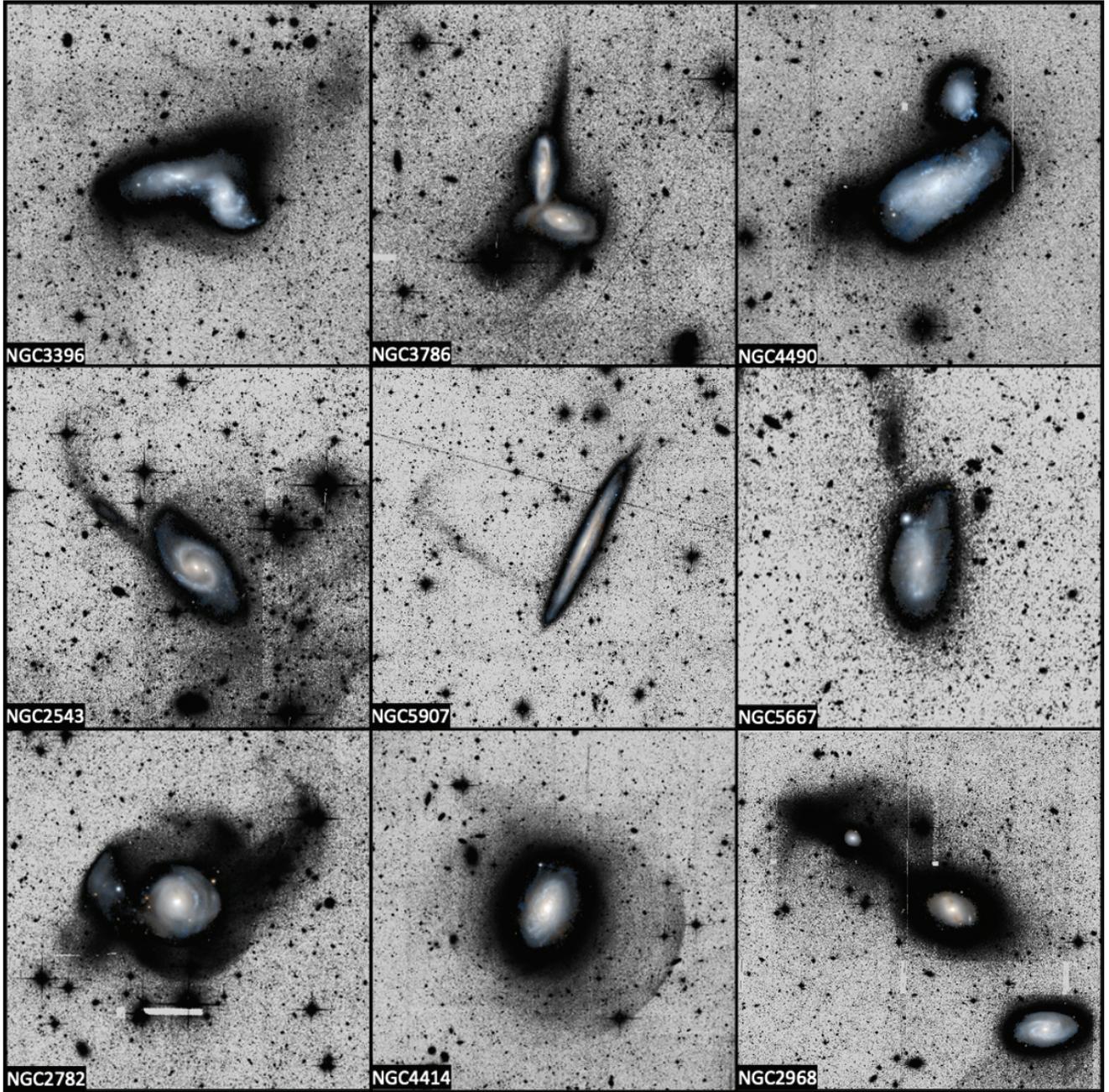
---

[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

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

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
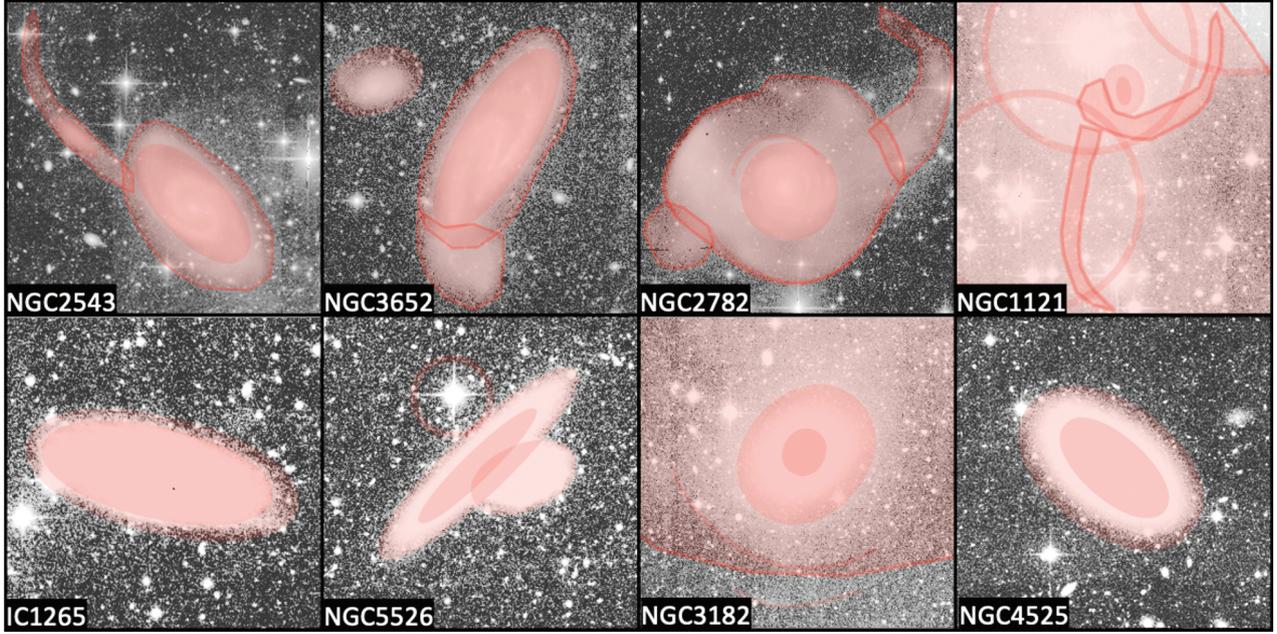
Fig. 4. Examples of annotated galaxies. Some of them present tidal features and pollutants, or a companion galaxy, while others only present the halo and the center of the galaxy.

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.
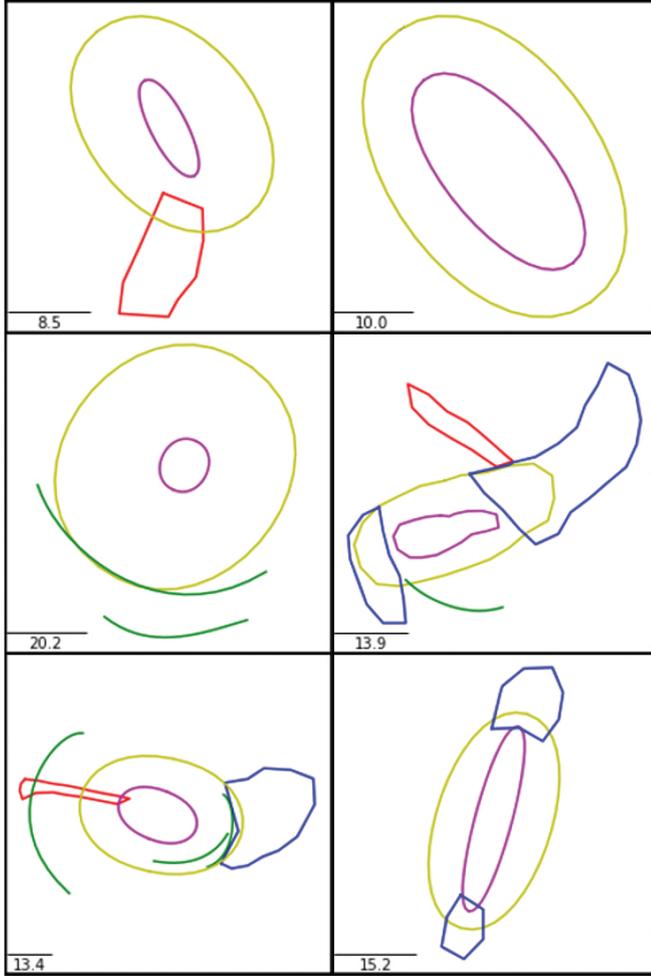
Fig. 5. Examples of annotation thumbnails. The center of the target galaxy is plotted in purple, its halo in yellow, tidal tails in blue, streams in red and shells in green. A scalebar in kiloparsecs is shown at the bottom left of each thumbnail.
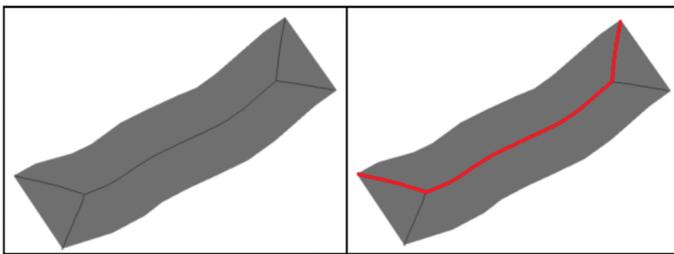

Fig. 6. Example of the longest length in a polygon (in gray). *Left*: medial axis (black line) of the polygon obtained during the skeletonization. *Right*: longest length kept (red line) in the polygon

– 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.
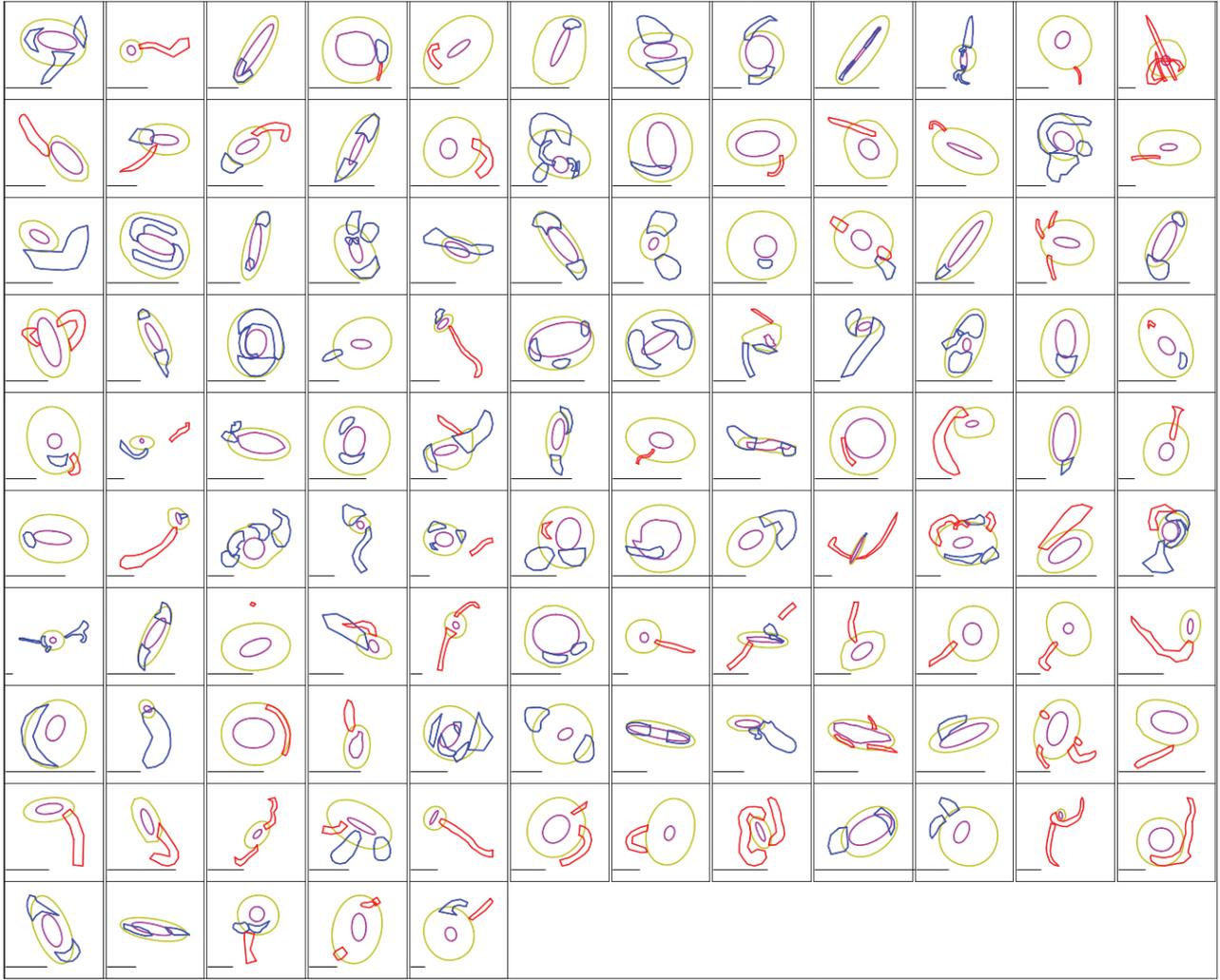
Fig. 7. Thumbnails of all the galaxies having tidal tails or streams. The main galaxy is represented in magenta, the halo in yellow, tidal tails in blue and streams in red. A scalebar representing 30 kpc at the bottom left gives an indication of the size of the structures.

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
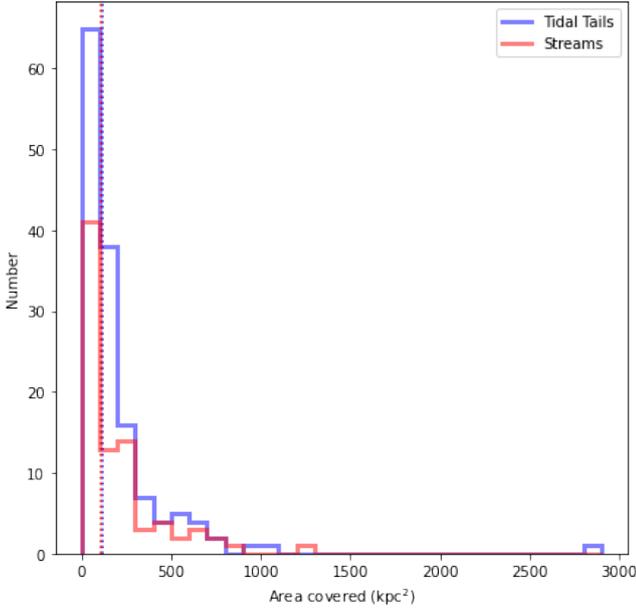
Fig. 8. Histogram of the area (in square kiloparsecs) covered by tidal tails (in blue) and streams (in red), in bins of width 100 kpc². The median of each distribution is represented by the dotted lines.

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 (kpc²)	std	Median area (kpc²)	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 kpc² for tidal tails and 110 kpc² for streams. Most structures cover areas smaller than 300 kpc², and few between 300 and 800 kpc². 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

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.

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.

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
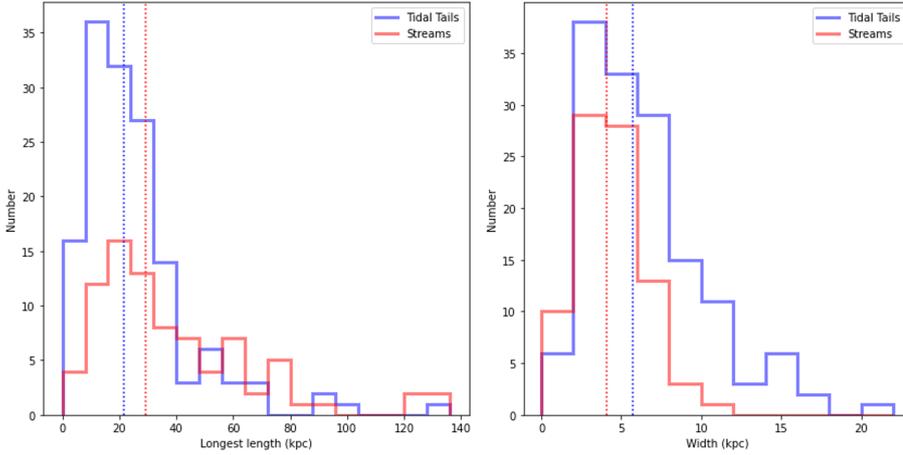
Fig. 9. Histograms of the length in kiloparsecs of tidal tails (in blue) and streams (in red), in bins of width 8 kpc (*left*), and of their width in bins of width 2 kpc (*right*). The median of each distribution is represented by the dotted lines.
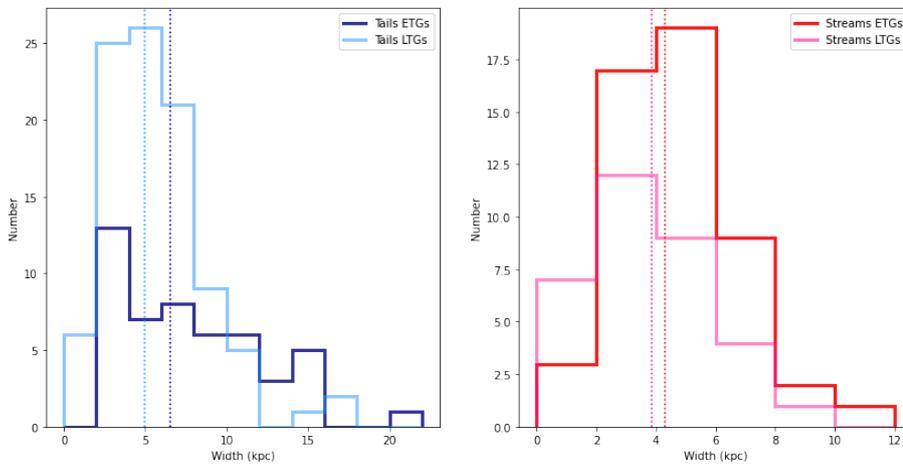

Fig. 10. Histograms of the width in kiloparsecs of tidal tails (*left*) and streams (*right*) in bins of 2 kpc as a function of the morphological type of the galaxies: ETGs are represented by darker shades than LTGs. The median of each distribution is represented by the dotted lines.

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
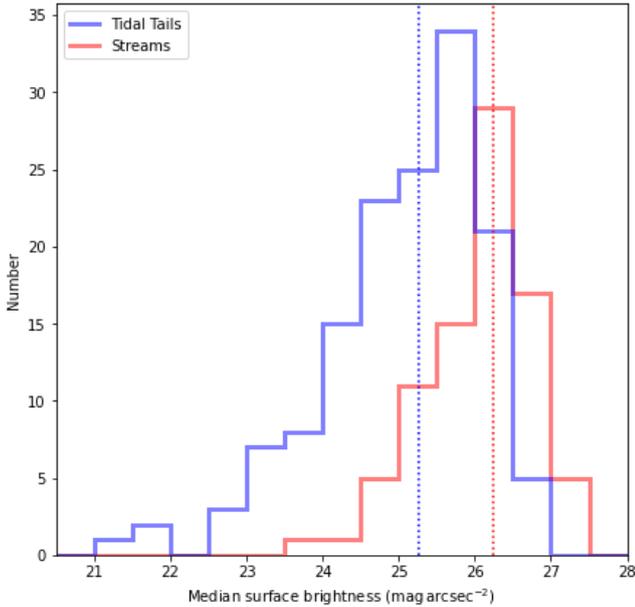
Fig. 11. Histogram of the median surface brightness value in magnitudes per square arcsecond for tidal tails (in blue) and streams (in red), in bins of width 0.5 mag arcsec^{-2}. The median of each distribution is represented by the dotted lines.

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.
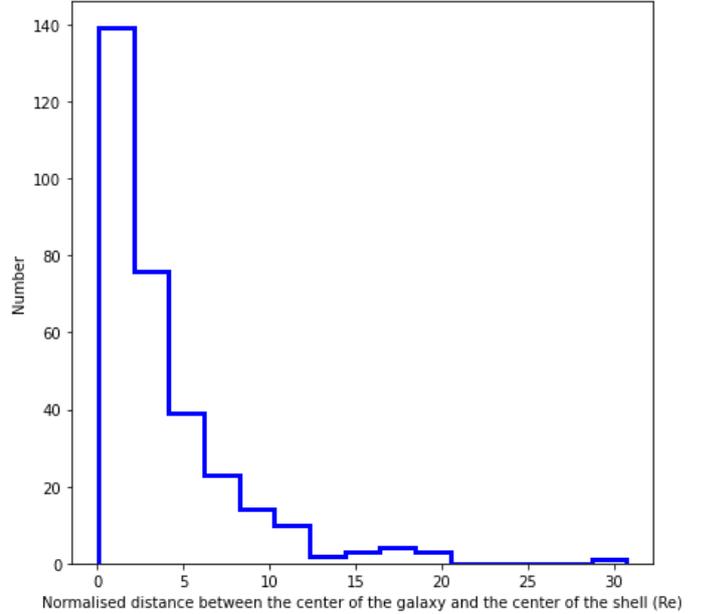
[18] The center of the shell is hence the center of the circle that passes through the circular arc defining the shell.
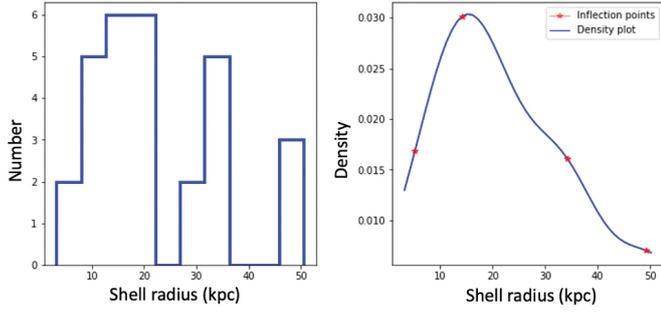
Fig. 13. Example of NGC 0474. *Left*: histogram of the shell radii for this galaxy. *Right*: corresponding density plot with the inflection points of the curve indicated by red stars.

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
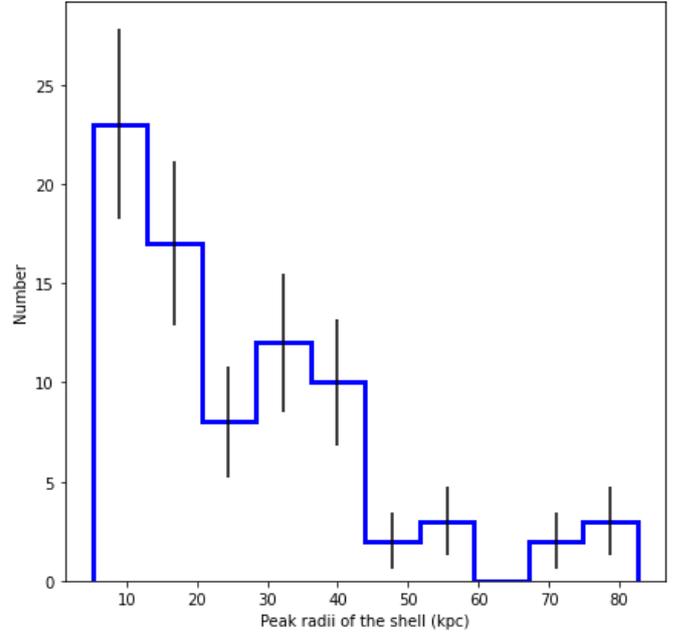
Fig. 14. Histogram of the peak radii of shells (in kiloparsecs) for all galaxies, with Poisson uncertainties plotted as the black error bars in each bin.

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
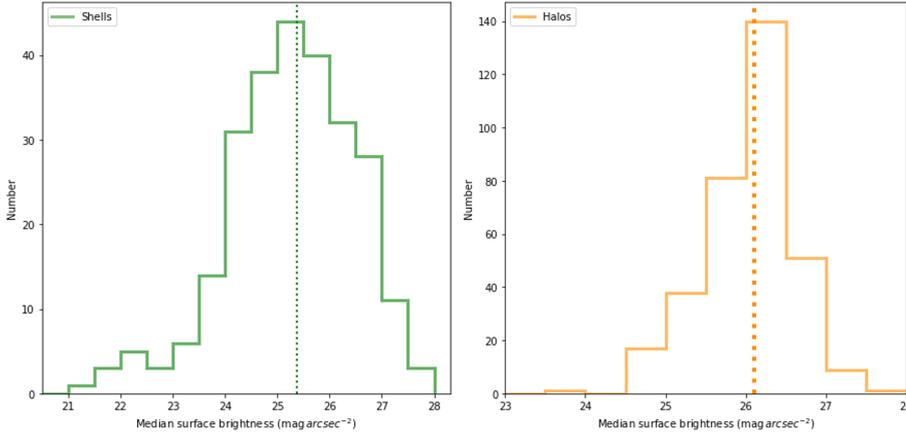
Fig. 15. Histograms of the median surface brightness value in magnitudes per square arcsecond for shells (*left*) and for halos (*right*), in bins of width 0.5 mag arcsec^{-2}. The medians are represented by the dotted lines.
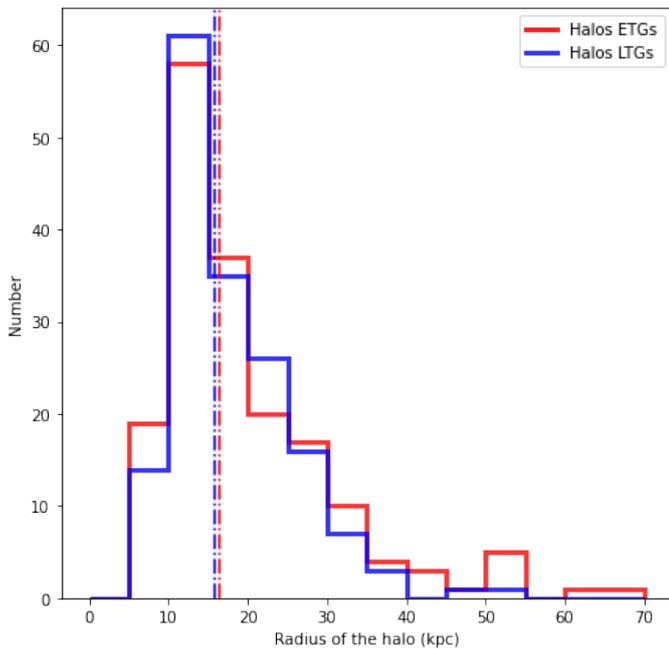

Fig. 16. Histogram of the radius of the halos in kiloparsecs as a function of the morphological type, in bins of width 5 kpc. The median of each distribution is represented by the dotted line.

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
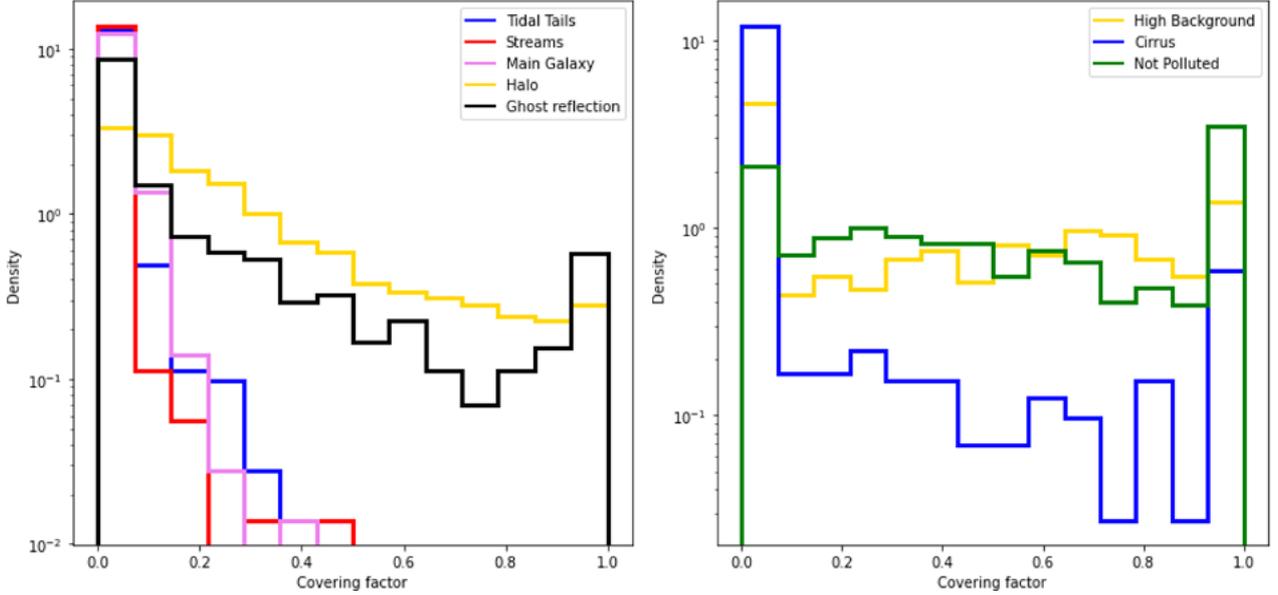
Fig. 17. Histograms of the covering factor. *Left:* covering factor for stellar structures and ghost reflections in a selected box of side length $20\,R_e$ centered on the primary galaxy. *Right:* covering factor for high background, cirrus and clean ("Not Polluted") regions in a field of view of $30.95 \times 30.95'$ 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.
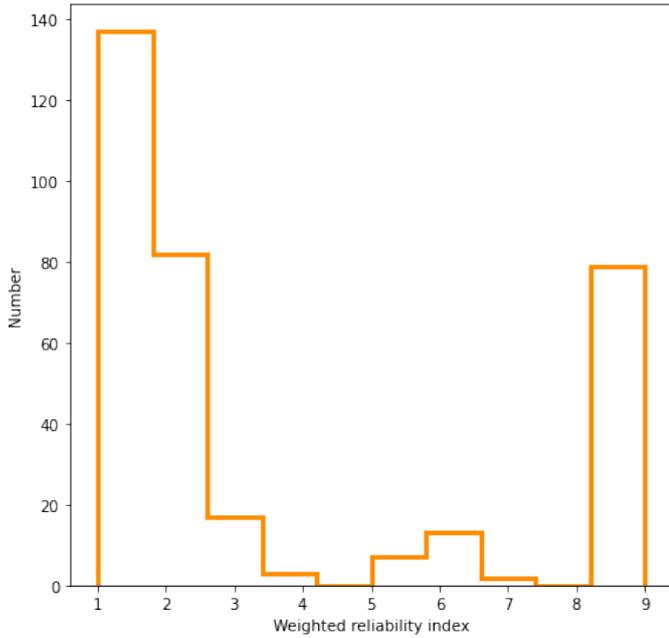
Fig. 18. Histogram of the weighted reliability index for the galaxies. A higher index corresponds to a cleaner image around the halo.

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 ≥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).
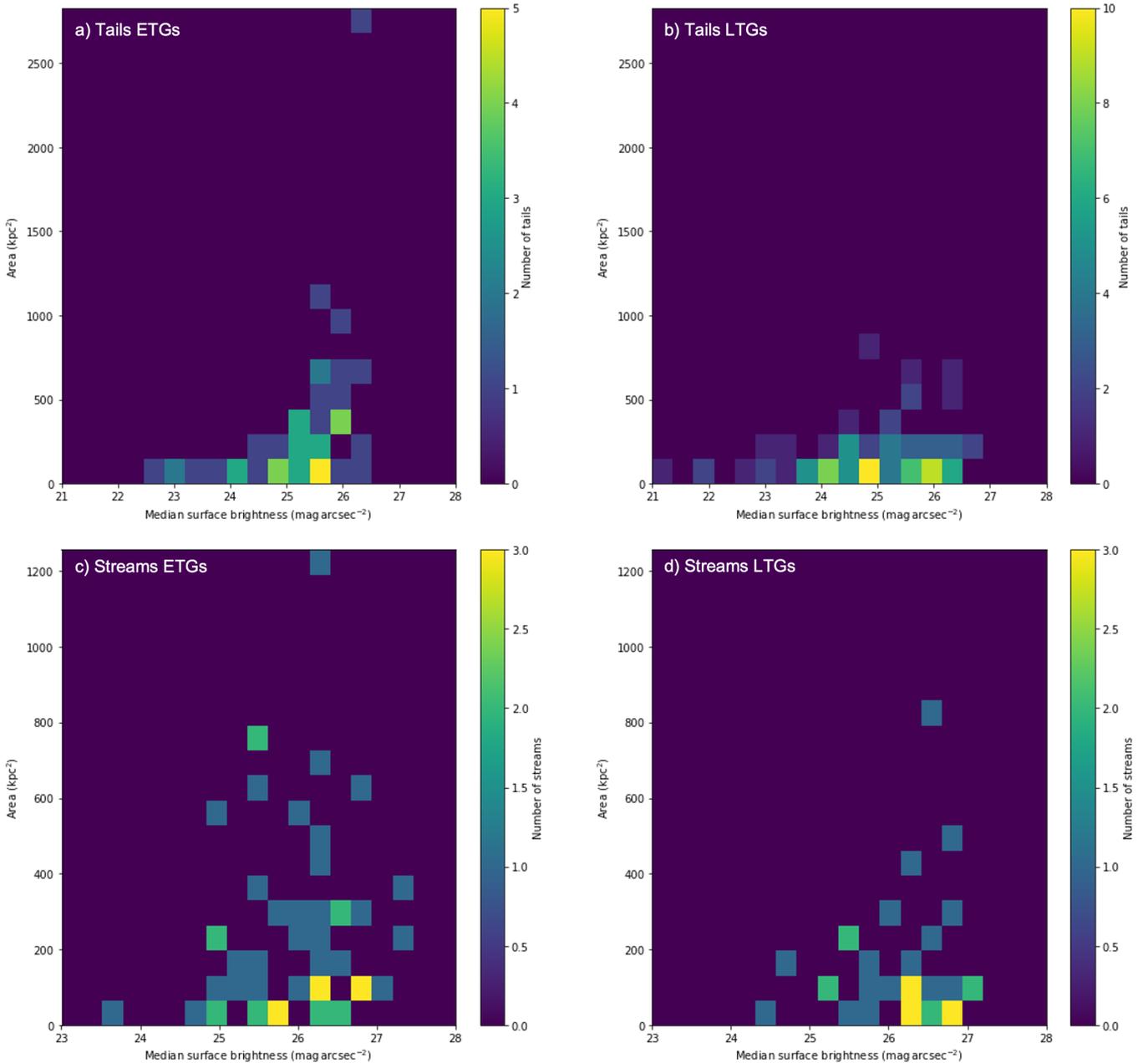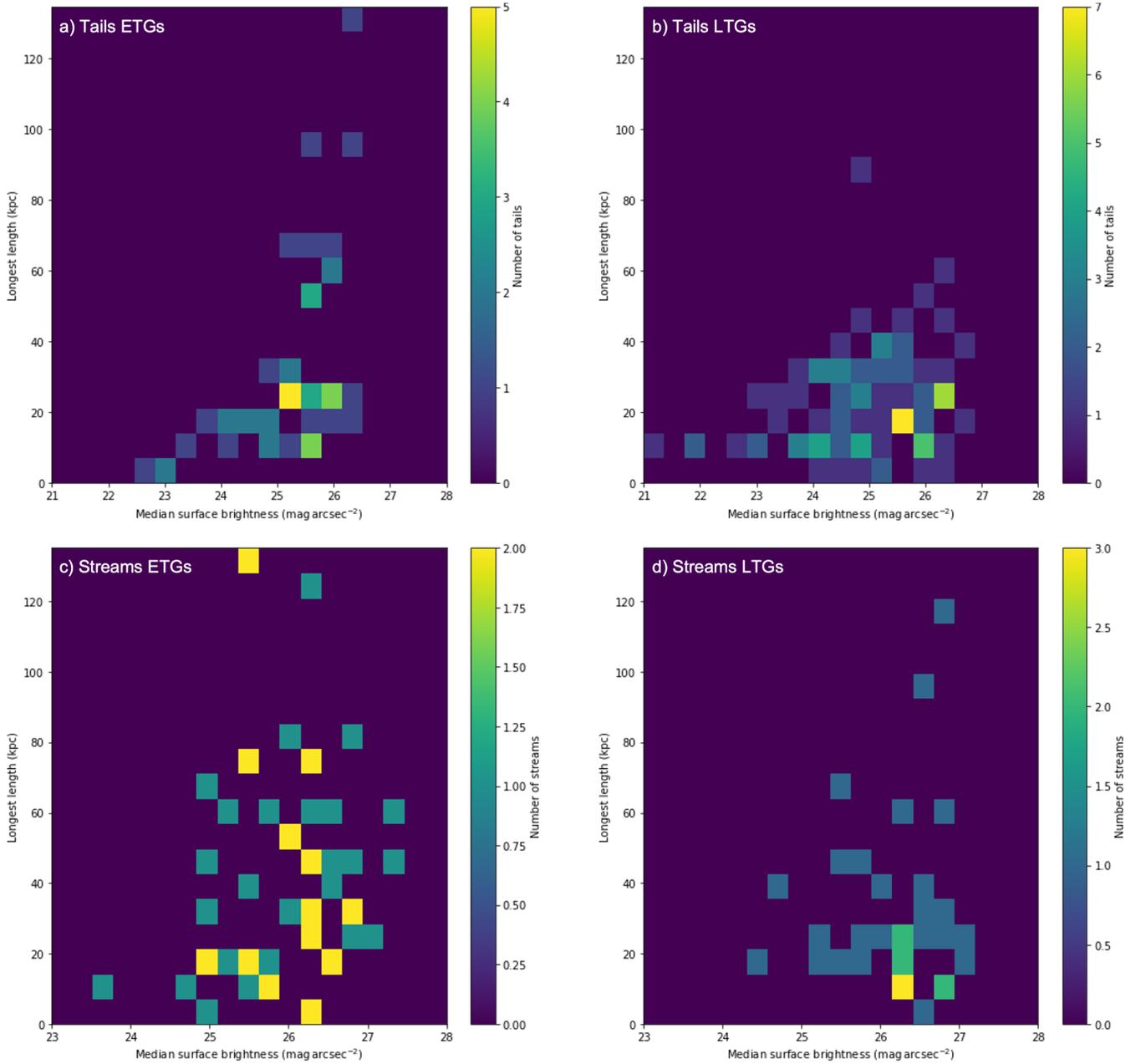
Fig. 19. 2D histograms of the area (in square kiloparsecs) of tidal tails and streams as a function of the median surface brightness (in mag arcsec^{-2}). This 2D histogram for tidal tails around ETGs is presented in *panel a* and for LTGs in *panel b*. For streams, the histogram for ETGs is visible 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.

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 $1.1 \times 10^{11} M_\odot$ 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 $1.1 \times 10^{11} M_\odot$.

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 $1.8 \times 10^{11} M_\odot$. 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 $1.5 \times 10^{11} M_\odot$.

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.
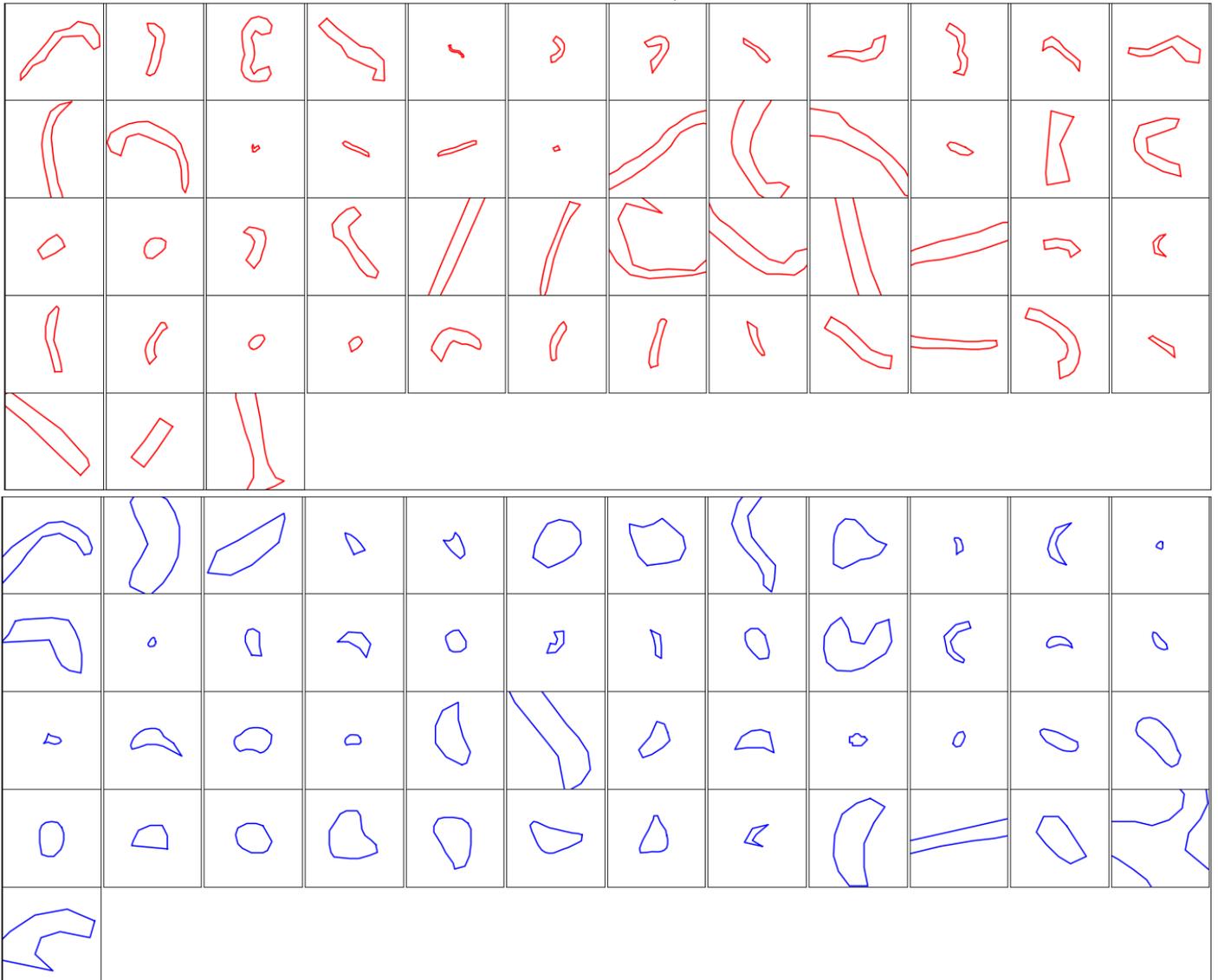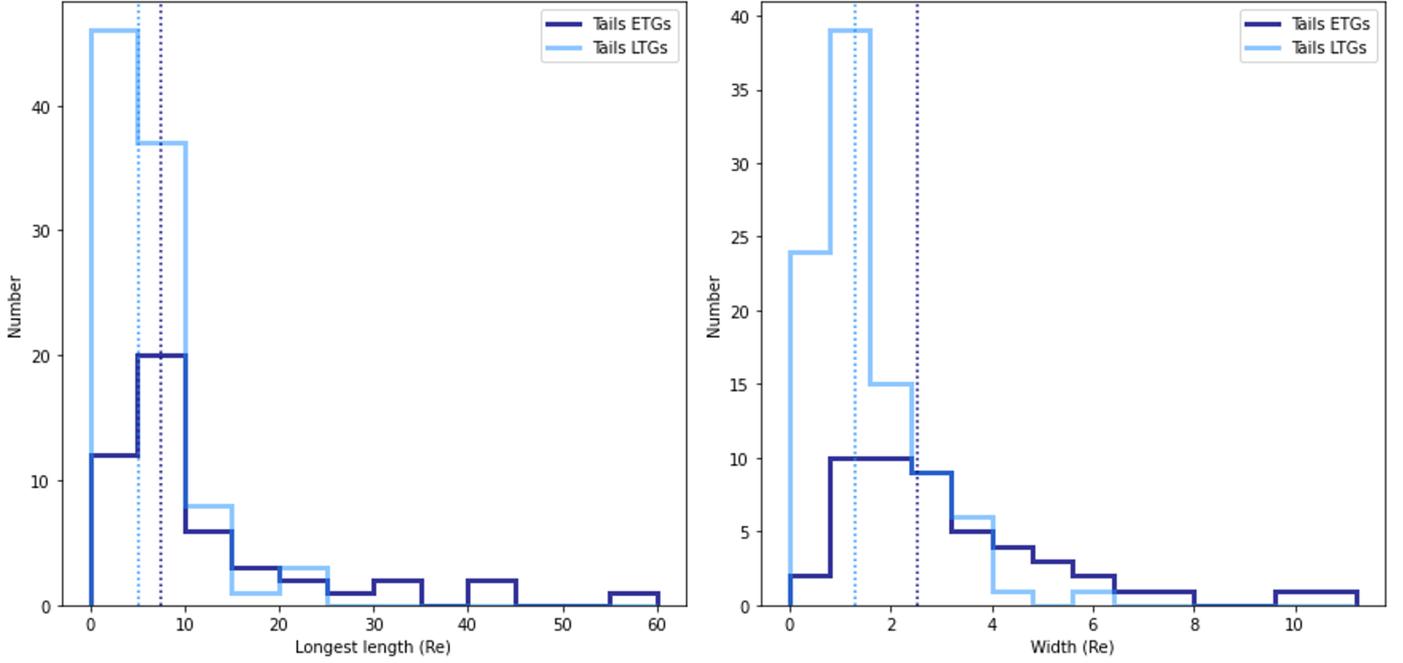
Fig. A.2. Thumbnails of the streams (*top, in red*) and tidal tails (*bottom, in blue*) identified around the ETGs, plotted in boxes of 50×50 kpc. They are sorted by increasing mass of the host galaxy, starting from the top-left for the lightest ETG to the bottom-right for the most massive ETG.

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.
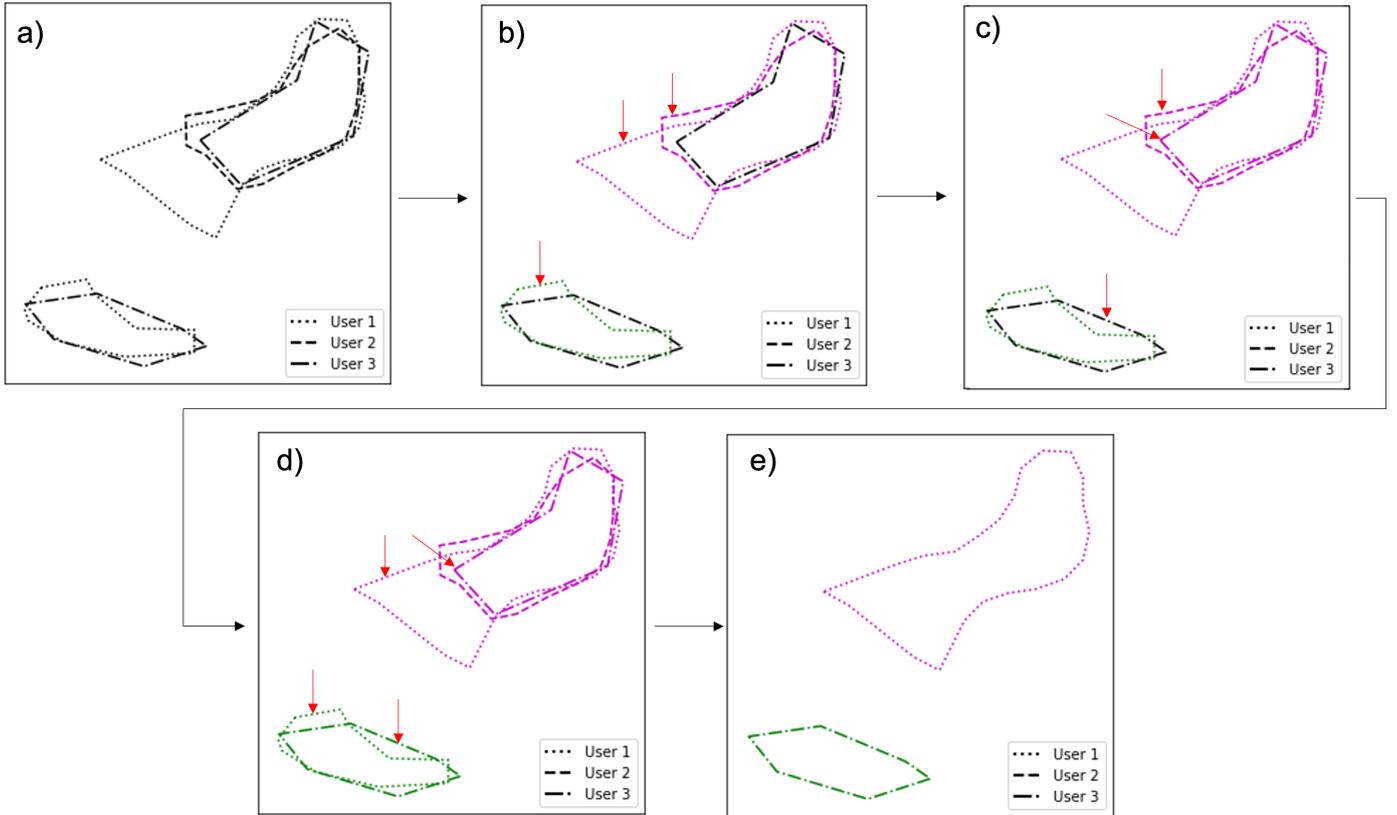

Fig. C.1. Illustration of the selection process for tidal tails and streams for one galaxy. Black annotations indicate that the unique identifier has not been attributed yet, while each color corresponds to one unique identifier. The linestyles correspond to the different users. The red arrows indicate 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.
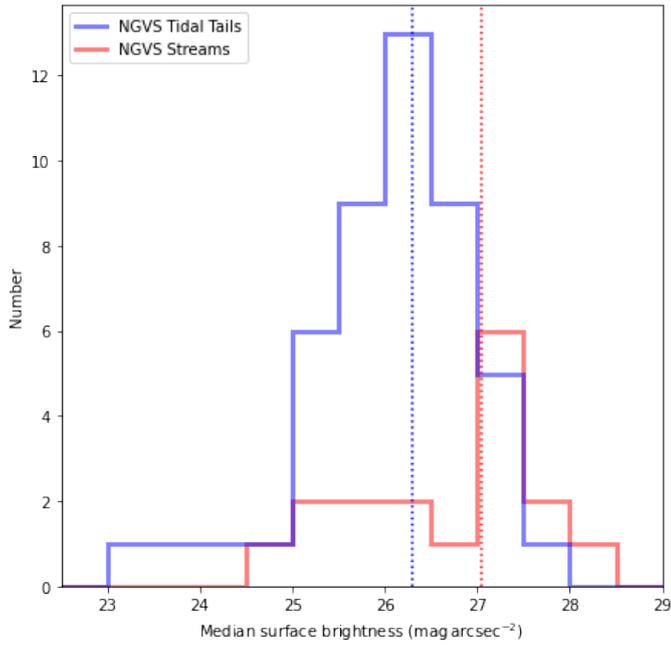
Fig. D.1. For NGVS only: histogram of the median surface brightness values in magnitudes per square arcsecond for tidal tails (in blue) and streams (in red), in bins of 0.5 mag arcsec^{-2}. The median of each distribution is represented by the dotted lines.
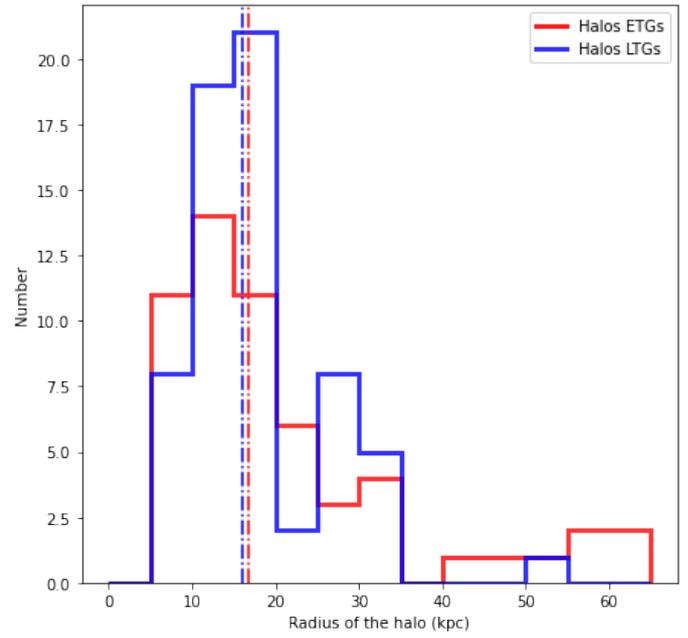

Fig. D.3. For NGVS only: histogram of the radius of the halos, in bins of 5 kpc, as a function of the morphological type. The median of each distribution is represented by the dotted line.
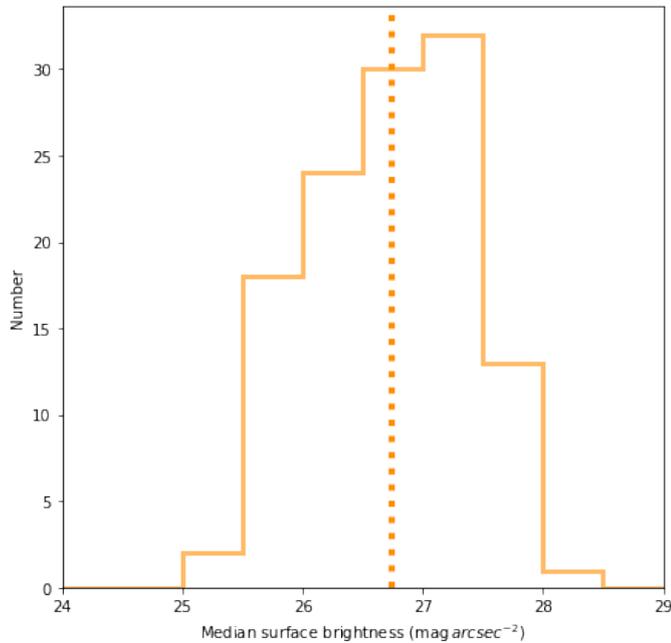

Fig. D.2. For NGVS only: histogram of the outer median surface brightness values in magnitudes per square arcsecond for halos, in bins of 0.5 mag arcsec^{-2}. The median is represented by the dotted line.