

The Complete *Spitzer* Survey of Stellar Structure in Galaxies (CS⁴G)[★]

P. M. Sánchez-Alarcón^{1,2,★★}, H. Salo³, J. H. Knapen^{1,2}, S. Comerón^{2,1}, J. Román^{1,2,4}, A. E. Watkins⁵, R. J. Buta⁶, S. Laine⁷, J. M. Falcón-Ramírez^{2,1}, M. Anetjärvi^{3,8}, E. Athanassoula⁹, A. Bosma⁹, D. A. Gadotti¹⁰, J. L. Hinz¹¹, L. C. Ho^{12,13}, B. W. Holwerda¹⁴, J. Janz^{15,3,16}, T. Kim¹⁷, J. Koda¹⁸, J. Laine^{3,19}, E. Laurikainen³, B. F. Madore^{20,21}, K. Menéndez-Delmestre²², R. F. Peletier²³, M. Querejeta²⁴, A. Ruokanen³, K. Sheth²⁵, and D. Zaritsky²⁶

(Affiliations can be found after the references)

Received 24 July 2024 / Accepted 28 February 2025

ABSTRACT

Context. The *Spitzer* Survey of Stellar Structure in Galaxies (S⁴G), together with its Early Type Galaxy (ETG) extension, stands as the most extensive dataset of deep uniform mid-infrared (mid-IR; 3.6 and 4.5 μm) imaging for a sample of 2817 nearby ($d < 40$ Mpc) galaxies. However, the velocity criterion used to select the original sample results in an additional 422 galaxies without H I detection that should have been included in the S⁴G on the basis of their optical recession velocities.

Aims. In order to create a complete magnitude-, size-, and volume-limited sample of nearby galaxies, we collected 3.6 μm and *i*-band images using archival data from different surveys and complemented it with new observations for the missing galaxies. Since most, but not all, of these galaxies have a Hubble type in Hyperleđa $T_{\text{HL}} > 0$, we denote the sample of these additional galaxies as disc galaxy (DG) extension. We present the Complete *Spitzer* Survey of Stellar Structure in Galaxies (CS⁴G), encompassing a sample of 3239 galaxies (S⁴G+ETG+DG) with consistent imaging, surface brightness profiles, photometric parameters, and revised morphological classification.

Methods. Following the original strategy of the S⁴G survey, we produced masks, surface brightness profiles, and curves of growth using masked 3.6 μm and *i*-band images. From these profiles, we derived the integrated quantities, including total magnitude, stellar mass, concentration parameter, and galaxy size, converting between optical *i*-band and 3.6 μm . We also re-measured these parameters for the S⁴G and ETG to create a homogenous sample. We present new morphologically revised *T*-types, and we showcase mid-IR scaling relations for the stellar mass, galaxy size, concentration index, and morphological type.

Results. Our new masking procedure increases the number of pixels masked out by a factor of five, improving the masking of fainter regions over previous S⁴G data. Our photometric parameters from *i*-band imaging yield measurements consistent with the original sample (S⁴G) and its ETG extension in the 3.6 μm band. The new DG extension consists of galaxies with a wide morphological range ($-5 < T_{\text{HL}} < 10$) and a mass range of $6 < \log(M_*/M_\odot) < 11$. The galaxies in the DG sample have an average mass of $\log(M_*/M_\odot) = 9.21$, an average galaxy isophotal radius at 25.5 mag arcsec⁻² of $R_{25.5} = 7.1$ kpc, and an average concentration index of $C_{82} = 2.92$.

Conclusions. We completed the S⁴G sample by incorporating 422 galaxies into the original dataset. The new galaxies constitute 15% of the total previous sample (S⁴G+ETG), but in the lower-mass range ($M_* < 10^9 M_\odot$), and the disc galaxy extension increases the sample by 36%. The CS⁴G includes at least 99.94% of the complete sample of nearby galaxies, meeting the original selection criteria based on a comparison with the NED database. We make the images and surface brightness profiles available to the community together with the conjunct catalogue of the whole CS⁴G dataset with consistent photometric measurements for 3239 galaxies. The CS⁴G will enable a wide set of investigations into galaxy structure and evolution, and it will complement the optical, near-IR, and mid-IR imaging that will be obtained in the coming years with *Euclid*, *Rubin*, *Roman*, and other research projects.

Key words. catalogs – surveys – galaxies: general – galaxies: photometry – galaxies: stellar content – galaxies: structure

1. Introduction

The local Universe offers an essential sample for studying galaxy formation and evolution. Most local galaxies have a substantial angular size (e.g. Jarrett et al. 2019; Moustakas et al. 2023), which facilitates detailed decompositions and morphological classifications. This also enables a comprehensive exploration

of the roles played in galaxy evolution by structural components such as bulges, bars, rings, or lenses.

The *Spitzer* Survey of Stellar Structure in Galaxies (S⁴G; Sheth et al. 2010) is the most extensive dataset of deep uniform mid-infrared (mid-IR) imaging for a sample restricted by volume, magnitude, and size ($d < 40$ Mpc, $|b| > 30^\circ$, $m_{B,\text{corr}} < 15.5$, and $D_{25} > 1'$). The S⁴G is a survey of 2352 galaxies observed in the mid-IR with the 3.6 μm and 4.5 μm channels of the infrared array camera (IRAC; Fazio et al. 2004) of the *Spitzer* Space Telescope (Werner et al. 2004). IRAC was an infrared camera that simultaneously produced $5'2 \times 5'2$ images in two bands, either in 3.6 and 4.5 μm or 5.8 and 8.0 μm , with a pixel scale of $1'2$. After the cryogenic phase, during the warm mission

* The authors of this paper seek to express deep gratitude for the discussions and expertise from the late Dr. Tom Jarrett, a pioneering infrared astronomer who made invaluable contributions to 2MASS, S⁴G, WISE and other infrared studies.

** Corresponding author: pmsa.astro@gmail.com

of *Spitzer*, only channels 1 (3.6 μm) and 2 (4.5 μm) continued to be operational because they were less affected by the higher operating temperature. The 3.6 μm and 4.5 μm wavelengths are much less affected by dust than optical bands, and thus galaxies imaged in those bands have a nearly constant mass-to-light ratio (M/L) that only varies slightly with age and metallicity. As a result, the 3.6 and 4.5 μm bands are ideal for measuring stellar masses of galaxies (see Meidt et al. 2014; Querejeta et al. 2015). Studying these galaxies has contributed to our comprehension of the impact of gas on bar formation (Comerón et al. 2018; Díaz-García et al. 2019), the mechanisms behind disc breaks and truncations (Laine et al. 2014a, 2016; Comerón et al. 2012), thin and thick discs (Comerón et al. 2011, 2018), the nature of resonance rings (Comerón et al. 2014), morphological features (Laine et al. 2014b; Buta et al. 2015), and the influence of the environment on the gas content of disc galaxies (Laine et al. 2014a; Watkins et al. 2019). The S⁴G plays a crucial role in studying the evolution of galaxies, having led to the publication of numerous research articles over the years (see Watkins et al. 2022, hereafter, W2022, for a review). Recent studies (e.g., Cuomo et al. 2023; Menéndez-Delmestre et al. 2024) continue to emphasise the ongoing relevance of S⁴G and reinforce its enduring impact.

However, the original S⁴G sample suffered from a bias due to the velocity-based selection criterion for distance. The selection was based on HI measurements, which resulted in excluding gas-poor galaxies. Sheth et al. (2013) overcame this bias using optical-band spectroscopic redshifts for early-type galaxies (ETGs) with HyperLeda morphological types $T_{\text{HL}} \leq 0$, incorporating 465 new *Spitzer* observations for these missing galaxies into the survey (published by W2022). The original survey and its extension (S⁴G+ETG) contain 2817 galaxies. Nonetheless, the T -type criterion for the ETGs ($T_{\text{HL}} \leq 0$) does not include 391 late-type, 20 lenticular, and 11 ETGs (based on HyperLeda types) that lack radio-based velocities but meet the distance criterion using optical velocities, as well as the other S⁴G selection criteria. These 422 missed galaxies form the disc galaxy (DG) extension described here.

To complete the survey with these 422 missing galaxies, we collected archival IRAC images and optical i -band images from several surveys (see Sect. 2). We complemented these data with new observations for 18 galaxies lacking archival images. We followed the original strategy of the S⁴G survey by conducting the same photometric analysis applied to the previous extension of 465 ETGs by W2022. We applied the methods described in the S⁴G Pipelines 2 and 3 (P2 and P3; see Muñoz-Mateos et al. 2015, hereafter MM2015). This analysis includes masking images; producing intensity and surface brightness profiles; obtaining position angle and ellipticity profiles; and computing total magnitudes, stellar masses, measures of central concentration, and measures of galaxy size. MM2015 extensively outlined the applications of this structural analysis, all of which are still relevant to this extended sample of gas-poor galaxies that includes mostly disc and dwarf galaxies.

In the current landscape of extensive deep surveys, which are notably focused on exploring vast cosmic volumes (e.g. the Hyper-Suprime Cam Subaru Strategic Program, the *Vera C. Rubin* Observatory’s Legacy Survey of Space and Time, the *Roman* Space Telescope, and the *Euclid* Mission; Aihara et al. 2018; Ivezić et al. 2019; Spergel et al. 2015; Laureijs et al. 2010), a complete census of late-type galaxies at redshift $z = 0$ becomes crucial for comparisons between young and present-day galaxies.

In this work, we present the Complete *Spitzer* Survey of Stellar Structure in Galaxies (CS⁴G) by adding an extension of 422 galaxies that completes the final sample of 3239 galaxies.

The purpose of this complete dataset is to serve as the basis for future analyses, including a detailed analysis of morphology and multi-component decompositions. In Sect. 2, we give an overview of the surveys and instruments used for the acquisition of the infrared and i -band images. In Sect. 3, we describe the completeness of the CS⁴G sample. In Sect. 4, we review the methods used for the photometric analysis, the creation of masks, and the sky background estimation. In Sect. 5, we present the empirical photometric conversion between i -band and 3.6 μm . In Sect. 6, we explain the radial and integrated photometric parameters derived from the images. In Sect. 7, we explore the scaling relationships among the photometric parameters within this new sample. We list the conclusions of our results in Sect. 8. Finally, in Sect. 8, we explain the data products released with this work. Throughout, we use the Λ CDM model with a Hubble-Lemaître constant $H_0 = 75^1 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and a matter density parameter $\Omega_m = 0.3$. The Morphological types used throughout the paper are either HyperLeda morphological types, denoted as T_{HL} , or revised T -types presented in Sect. 6, denoted as T , which are in the Comprehensive de Vaucouleurs revised Hubble-Sandage (CVRHS) system (Buta et al. 2015).

2. S⁴G disc galaxy extension sample, observations, and origin of archival data

The S⁴G DG extension sample, as well as the ETG one (W2022), adhere to the same selection criteria as the main S⁴G project. These criteria include galaxies listed in the HyperLeda² database (Makarov et al. 2014) with radial velocities $v < 3000 \text{ km/s}$, total extinction-corrected blue magnitudes $m_{B,\text{corr}} < 15.5$, blue isophotal diameters $1' < D_{25} < 30'$, and Galactic latitude $|b| > 30^\circ$. Thus, the CS⁴G sample excludes nearby galaxies with angular diameters larger than $30'$ (LMC, SMC, M33, and UMi Dwarf), and the largest galaxy included is NGC 0055 with a diameter of $29'.8$.

The original sample exclusively relied on HI-derived velocities (Sheth et al. 2010), which introduced a bias towards gas-rich late type galaxies. To rectify this bias in the sample, the ETG extension (Sheth et al. 2013; Watkins et al. 2022) incorporates galaxies with morphological types $T_{\text{HL}} \leq 0$ and optical-band spectroscopic velocities. Nevertheless, this limitation to early types still left out many galaxies (see Appendix A for more details): a new search on the HyperLeda database in 2018 led to a list of 422 galaxies with visual-band spectroscopic velocities and meeting the original S⁴G criteria, forming our DG sample (see Appendix B for comments on the exclusion of some galaxies). Of these, 391 galaxies have morphological types $T_{\text{HL}} \geq 0$, 20 galaxies with $-3 \leq T_{\text{HL}} < 0$, and 11 galaxies with $T_{\text{HL}} < -3$. The DG sample consists of these 422 galaxies selected from the HyperLeda database that lacked HI derived velocities.

To keep consistency with the original survey, we collect 3.6 μm and 4.5 μm band imaging from the *Spitzer* Heritage Archive, maintaining the original photometric bands of the survey. When these bands were not available, we used i -band imaging from different surveys. Despite its 3.4 μm band being similar to IRAC 3.6 μm , we decided not to use WISE due to its worse angular resolution. We find 55 mid-IR images from the *Spitzer* Heritage Archive, 11 of which are from S⁴G and ETG footprints, and 367 i -band images. For i -band imaging, we used

¹ We chose this value to be consistent with Sheth et al. (2010).

² <http://leda.univ-lyon1.fr/>

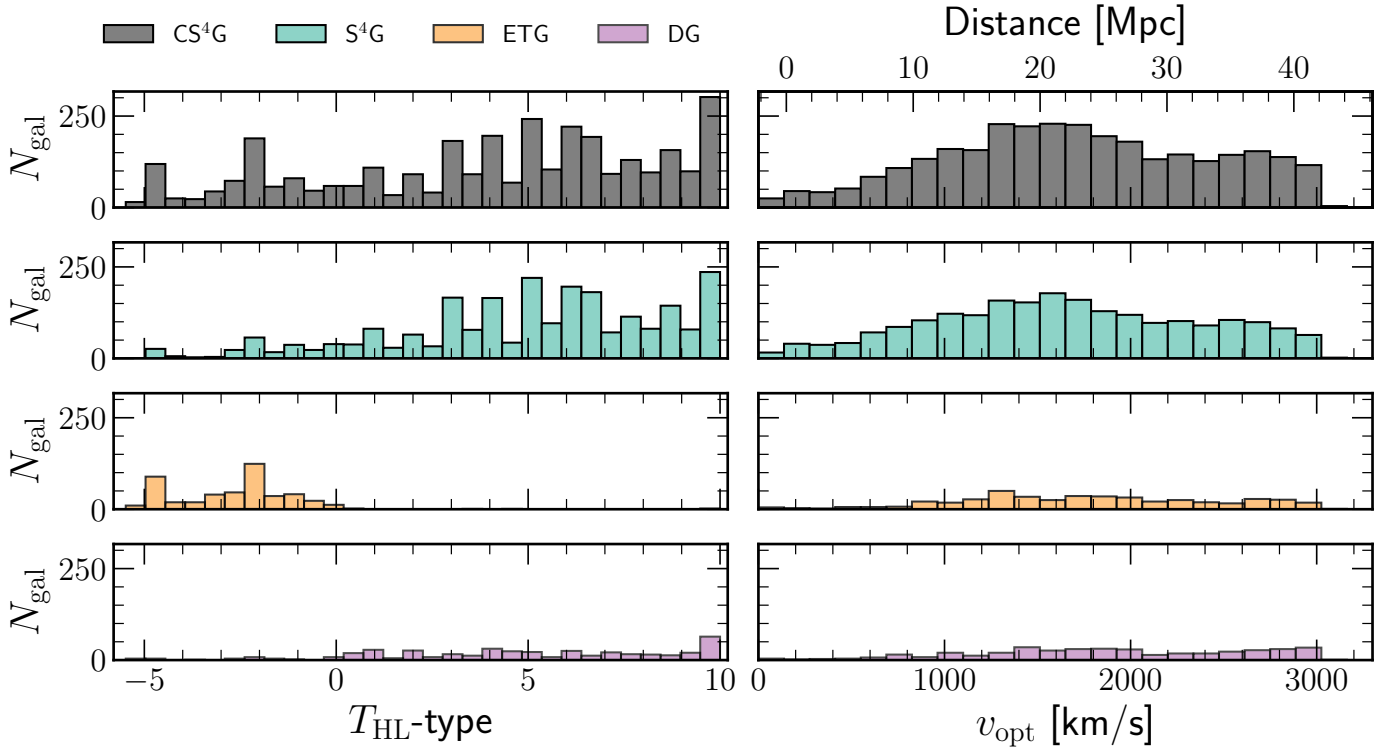


Fig. 1. HyperLeda morphological type (T_{HL} , left) and optical velocity distribution (right) for all the samples. These values are queried from HyperLeda. From top to bottom, each row shows the distribution of the CS⁴G (grey), original S⁴G (blue), ETG (orange), and DG (purple) samples, respectively.

Table 1. Survey and instrument properties.

Survey	Galaxies	Scale	FWHM	Depth
IRAC	44	0'600	2'1	26.8
IRAC ^(a)	11	0'750	2'1	26.8
DES	102	0'263	0'9	28.4
LS	169	0'263 ^(b)	1'0	27.9
SDSS	77	0'396	1'1	26.4
LT	15	0'150	1'4	28.3
NTT	3	0'241	1'2	27.6
HST	1	0'050	0'1	27.5

Notes. The depth column represents the surface brightness limit [3σ , $10'' \times 10''$] in units of mag arcsec^{-2} . ^(a)For this set of galaxies, the IRAC images are from S⁴ G fields. ^(b)For 41 galaxies, the mosaics have a pixel scale of 0'269.

the Dark Energy Survey DR2 (Abbott et al. 2021) for 102 galaxies, the DESI Legacy Imaging Surveys DR10³ (hereafter Legacy Surveys, Dey et al. 2019) for 169 galaxies, the SDSS DR12 (York et al. 2000; Alam et al. 2015) for 77 galaxies and archival data from the *Hubble* Space Telescope for one galaxy. For the remaining 18 galaxies we obtain *i*-band images with the Liverpool Telescope (LT) for 15 northern hemisphere galaxies, and with the New Technology Telescope (NTT) for three galaxies in the southern hemisphere.

We summarise the characteristics of the surveys and instruments in Table 1. We show the number of images of each survey used in this extension, the pixel scale, the average full width at

half maximum (FWHM) spatial resolution, and the average surface brightness depth [3σ , $10'' \times 10''$], measured as explained by Román et al. (2020). In Fig. 1 we show the morphological distribution (left column using HyperLeda types T_{HL}) and the distance distribution (right column) for the whole CS⁴G (grey), the original S⁴G (blue), the ETG (orange), and the DG (purple) samples.

2.1. Spitzer Heritage Archive

We selected sample galaxies observed with the *Spitzer* IRAC camera from the *Spitzer* Heritage Archive, either from the cryogenic or the warm phase of the mission. We found 55 galaxies with 3.6 μm and 4.5 μm IRAC observations. Of these 55 galaxies, 11 were observed in fields of the original S⁴G survey. These images have the same properties as those in the S⁴G survey (Sheth et al. 2010), with 0'75/pixel and similar depth. The remaining 44 had observations from different proposals, or mosaics from the *Spitzer* Enhance Image Products with a pixel scale of 0'60/pixel. The proposal IDs and the Region ID of the frames and mosaics can be found in the headers of the images. Two galaxies, NGC 4516 and NGC 4431, have only 4.5 μm observations available.

2.2. Dark Energy Survey

The Dark Energy Survey (The Dark Energy Survey Collaboration 2005) is a southern hemisphere imaging survey that aims to probe the origin of the accelerating Universe and help uncover the nature of dark energy. The survey is based on optical and mid-infrared imaging with the Dark Energy Camera (DECam; Flaugher et al. 2015) mounted on the 4m

³ <https://legacysurvey.org/dr10>

Blanco telescope at Cerro Tololo Inter-American Observatory in Chile, covering 5000 deg². We identified 102 galaxies of the DG sample with *i*-band images in the second data release of DES (DR2; Abbott et al. 2021). The DECam has a pixel scale of 0′′263, and the survey depth in the *i*-band is 23.8 mag (for a 1′′95 diameter aperture at a S/N of ten) with a mean FWHM of 0′′88. We accessed the images via the DESaccess⁴ web application. We used the image cutout service to download images for 102 galaxies, centred on the galaxy, with size 12 × 12 arcmin².

2.3. DESI Legacy Imaging Surveys

The DESI Legacy Imaging Surveys (Dey et al. 2019) are a combination of three public projects: the Dark Energy Camera Legacy Survey, the Beijing–Arizona Sky Survey, and the Mayall *z*-band Legacy Survey. In its latest data release (DR10), the combined surveys jointly cover ~20 000 deg². This release incorporates additional DECam data from NOIRLab mostly from DES (Abbott et al. 2021), the DELVE Survey⁵ (Drlica-Wagner et al. 2021), and the DECam eROSITA Survey (DeROSITA; Zenteno et al. 2025). There are 169 S⁴G-DG galaxies with *i*-band imaging in the DESI Legacy Imaging Surveys DR10 that are not in the DES DR2 survey. We collected images for these 169 galaxies using the Sky Viewer⁶ URL cutout service, imposing a size of ~11 × 11 arcmin² and at the original pixel scale 0′′263.

2.4. Sloan Digital Sky Survey

The Sloan Digital Sky Survey (York et al. 2000) is probably the most widely used northern hemisphere imaging and spectroscopic survey. We used the Data Release 12 (DR12; Alam et al. 2015) Science Archive Server (SAS) to download individual frames and create mosaics using Swarp (Bertin 2010). Since DR12, no additional imaging has been incorporated into SDSS. We created mosaics centred at the galaxy with a size of 6 × 6 arcmin² for 75 galaxies, and for the galaxies NGC 4964 and UGC 8736, which have a larger angular size, the mosaics are 18 × 18 arcmin² in size.

2.5. Liverpool telescope

Of our S⁴G-DG sample, some galaxies lack mid-IR and *i*-band imaging in public surveys. We observed 15 of these galaxies during the nights of July 16, 2018, December 9, 12, and 14, 2018, January 5 and 12, 2019, March 3 and 16, 2019, and April 11, 2019, with the Optical-Infrared I0:0 camera (Barnsley et al. 2016) at the 2.0 m Liverpool Telescope in the Observatorio del Roque de los Muchachos (Steele et al. 2004). The I0:0 camera has a 4096 × 4112 pixel² CCD, with a pixel scale of 0′′15 and a 10 × 10 arcmin² field-of-view.

2.6. New technology telescope

We observed three further missing galaxies in the southern hemisphere lacking archival mid-IR and *i*-band imaging using the EFOSC2 instrument (Buzzoni et al. 1984) at ESO’s 3.58 m New Technology Telescope (NTT) in Chile during the night of August 11th, 2019. The EFOSC2 instrument is equipped with a 2048 × 2048 pixel CCD, and we used the 2 × 2 binning mode

resulting in a pixel scale of 0′′2408 and a 4.13 × 4.13 arcmin² field-of-view.

2.7. Hubble space telescope

There is existing archival data for one remaining galaxy in the Hubble Legacy Archive (Proposal ID: 9395, PI: Marcella Carollo; Carollo et al. 2007). NGC 2082 was observed with the Advanced Camera for Surveys (ACS, Ford et al. 1998; Sirianni et al. 2005) using the *F814W* (*I*-band) filter with the Wide Field Channel (WFC) camera of the *Hubble* Space Telescope. The pixel scale of the ACS/WFC camera is 0′′050 providing a 3.4 × 3.4 arcmin² field of view, sufficient to capture the entire galaxy.

3. Complete *Spitzer* survey of stellar structure in galaxies

We present CS⁴G, which is formed by the DG sample studied in this work added to the original S⁴G (Sheth et al. 2010) and the ETG extension (W2022). The CS⁴G includes 3239 galaxies selected using the original S⁴G criteria (see Sect. 2).

To assess the completeness of the CS⁴G sample, we selected galaxies mimicking the criterion used in HyperLeda from the NASA/IPAC Extragalactic Database (NED; Chen et al. 2022). We used the Objects with Parameter Constraints tool from the NED website⁷ to select a NED comparison sample. However, this tool does not allow for constraints on the galaxy diameter. From the resulting selection queried by magnitude (*B*-band magnitude brighter than 15.5), volume (redshift less than 0.01), type of object (galaxies) and sky area, we query the diameters. We used an isophotal diameter at 25 mag arcsec⁻² when available, or the median of all diameters available. We find 3005 galaxies that meet the criteria of the CS⁴G according to NED. Of these, 295 are not in the CS⁴G. This list of galaxies includes the LMC, SMC, M33 and UMi Dwarf galaxies that were excluded from the original S⁴G due to their large angular size. We also find four objects misclassified as galaxies in NED. SIMBAD and visual inspection of the images show clearly that they are globular clusters of the Milky Way (UGC 09792, ESO 118-031, and Sextans C) and an emission region from the LMC (ESO 056-019).

We cross-matched the remaining 287 galaxies with the HyperLeda database updated to October 1, 2024. We find that 285 galaxies fail to meet one or more of the S⁴G criteria (*B*-band magnitude, diameter, or velocity) according to HyperLeda. Two galaxies (IC 3418 and ESO 084-12) do not have any velocity measurement in HyperLeda, but the other parameters, magnitude, diameter and galactic latitude, meet the criteria. According to NED, these two galaxies have redshifts well below the limit. They could potentially be included in the CS⁴G. However, to be consistent with the original S⁴G sample selection criteria, these galaxies are not included since they do not have any velocity measurements in HyperLeda.

Figure 2 shows the HyperLeda isophotal diameter at 25.5 mag arcsec⁻² with respect to *B*-band magnitude for the CS⁴G sample (blue points) and for a larger sample of galaxies extracted from HyperLeda (grey dots). Grey points inside the region of the CS⁴G are galaxies that fail other criteria while blue points outside the region are galaxies from the original S⁴G that the newer measurements in HyperLeda made them outside the

⁴ <https://des.ncsa.illinois.edu/desaccess/>

⁵ <https://datalab.noirlab.edu/delve>

⁶ <https://www.legacysurvey.org/viewer>

⁷ <https://ned.ipac.caltech.edu/byparams>

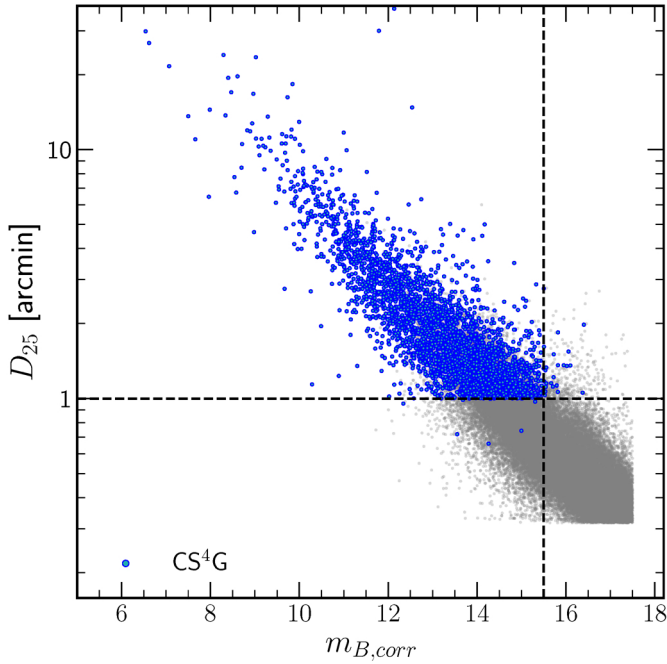


Fig. 2. HyperLeda isophotal diameter at 25.5 mag arcsec⁻² (from logd25 parameter) in arcminutes with respect to the *B*-band magnitude (btc parameter). Blue points represent the CS⁴G sample, and grey dots represent a larger sample from HyperLeda. The vertical and horizontal dashed lines represent the thresholds used to build the CS⁴G. There are some cases where some galaxies in the CS⁴G quadrant are not real galaxies meeting the criteria (as explained in Appendix B).

region. For the sake of completeness, we included them in the CS⁴G. The CS⁴G describes the largest and brightest galaxies of the local Universe followed by a smooth transition to fainter galaxies as seen with the representation of a larger sample in grey dots.

In conclusion, a comparison to NED identifications indicates that there are no new galaxies that meet the original sample selection criteria of the S⁴G. Thus, the completeness of the CS⁴G is 100%. However, there are two potential galaxies (0.06% of the sample) that could meet the criteria in the future if the velocity from NED is included and validated in HyperLeda. In this sense, the CS⁴G contains at least 99.94% of galaxies of the local Universe falling within the CS⁴G selection criteria.

4. Data preparation

The original S⁴G sample and the ETG extension were observed with the *Spitzer* IRAC camera and thus have similar properties, such as the pixel scale (0'.75), units of MJy sr⁻¹ and depth ($\mu_{AB,lim} \sim 27$ mag arcsec⁻²), as described in MM2015 and W2022. Since the DG extension was observed with a variety of instruments, it does not share these technical details. We adapted Pipeline 1 and 2 (P1 and P2, see MM2015) to produce image mosaics for new observations and mask images for the 422 galaxy images in the disc galaxy sample. We follow the procedures used for the original S⁴G (MM2015, Salo et al. 2015) and the ETG extension (W2022) to have consistent measurements throughout the surveys. We explain in detail the differences in the methods used in the following sections while summarising the similarities to the previous analyses.

4.1. Pipeline 1. Mosaics

For the new LT and NTT observations we follow the procedures in Pipeline 1 as described in MM2015 to achieve consistency with the whole survey. The pipeline initially reduces all the frames, subtracting the dark and bias and correcting the flat field. Then, it aligns the background levels of individual exposures by utilising overlapping regions. Subsequently, it combines all frames following standard dither/drizzle procedures (Fruchter & Hook 2002). The final mid-IR images are delivered in units of MJy sr⁻¹ and for the optical images in nanomaggies (Abazajian et al. 2003) per pixel (the pixel size can be found in Table 1) with a zero point of 22.5 mag arcsec⁻².

Figure 3 shows examples of images of different galaxies from the DG sample, the original S⁴G and the ETG extension. We used images of the different surveys and instruments in the DG sample to show their differences in angular resolution and depth (see Table 1). We sort the galaxies according to their stellar mass in the *x*-axis and to their morphological type in the *y*-axis (see Sect. 6). The SDSS examples are the images with the noisiest backgrounds, while the DES and LS imaging surveys have more depth and a better angular resolution. When selecting galaxies for each panel, we first look for galaxies within the DG sample, and when they are not available, we select them from the original S⁴G. For panels with multiple options, the galaxy shown is chosen randomly. Blank spaces are regions where there are no galaxies in the CS⁴G. Despite their differences, the set of images used in the DG samples traces the stellar content of their galaxies similarly to those used in the original S⁴G.

4.2. Pipeline 2. Masking

The Pipeline 2 of the S⁴G and the ETG extension produced masks of contaminating sources, including foreground stars, background galaxies, and scattered light artefacts, using SExtractor (Bertin & Arnouts 1996) with three different thresholds. Originally, three different masks on the 3.6 μ m mosaic images, with high, medium, and low detection thresholds, were produced to identify sources both far from the target galaxy and overlapping in projection with the galaxy. We adopt a similar strategy to produce masks for the DG sample using a combination of SExtractor(v.2.25.0) masks with different thresholds and configurations to improve the detection of the lower S/N regions (i.e. low surface brightness features).

We first smooth the image using the Bayesian noise reduction technique FABADA (Sánchez-Alarcón & Ascasibar 2023) with a standard deviation of the noise measure on the image with hard sigma clipping (rejecting pixels above 2.5 σ). We then combine three different runs of SExtractor in the smoothed image. For the first two, we vary the background size (BACK_SIZE), and detection threshold (DETECT_THRESH). In the first iteration, we run it with an intermediate threshold (DETECT_THRESH = 1.2) and large background size (BACK_SIZE > 50 \times FWHM) which produces good masking of extended regions in the image at intermediate S/N regions. We then run SExtractor again, now optimised for point source detection, using a lower threshold (DETECT_THRESH = 0.9) and smaller background size (BACK_SIZE < 10 \times FWHM). We vary the background size according to the image resolution (measured as FWHM) to increase the detection of smaller point sources in higher-resolution images.

In the last step, we aim to increase the detection of faint point sources and deblend the sources inside the galaxy. We run

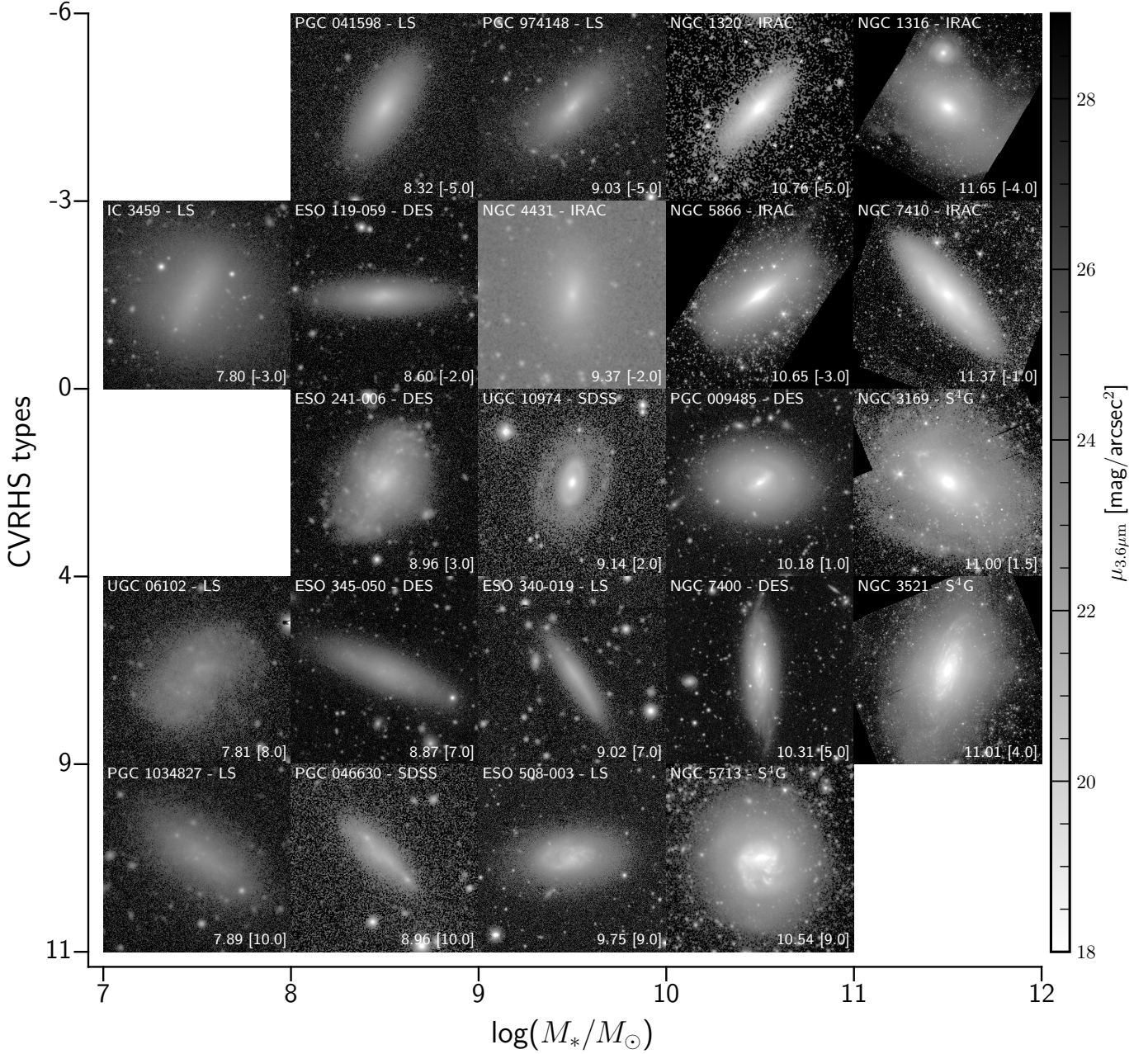


Fig. 3. Examples of images of galaxies from the CS⁴G sample. Galaxies are sorted according to their stellar mass (x -axis) and morphological type (y -axis) measured as explained in Sect. 6. The galaxy name and the survey or instrument of the image are shown in the upper part of each image. The stellar mass and the revised morphological type (in brackets) are shown in the bottom part of each image. Each image is centred on the galaxy and the field of view is set to be 1.4 times the isophotal radius at 25.5 mag arcsec⁻² ($R_{25.5}$). The radius, $R_{25.5}$, is measured in the 3.6 μ m band as described in Sect. 6.

SExtractor on the residual of the image enhanced with **FABADA** minus a Gaussian smoothed version of the image (with a kernel width of the FWHM). This residual highlights the regions that are brighter than the average of the surrounding area, i.e., the most likely peaks of sources. Then we masked regions compatible with being stars, masking rounder objects weighted by the inverse of the distance to the centre of the galaxy (the closer to the centre of the galaxy, the rounder the region must be to be masked). As a final step, to improve the detection of the fainter regions of the sources, which are more affected by the wings of the point spread function (PSF), we dilate the final mask images using a Gaussian kernel of size 3×3 pixel² and $\sigma = 1$ pixel. We

visually inspect all the masks produced to confirm that all the external sources are masked. When needed, we edit masks manually to overcome some problematic results, and we make use of colour images when available to disentangle bright regions of the galaxy from other sources, usually higher-redshift sources with redder colours.

Using the same method described here we compute a new mask image for NGC 0936 from the original S⁴G. We show its impact on the surface brightness profiles and background characterisation in Fig. 4 where we compare the mask produced by the original technique of the S⁴G and our new method on the galaxy NGC 0936. The left panel shows the original 3.6 μ m band image

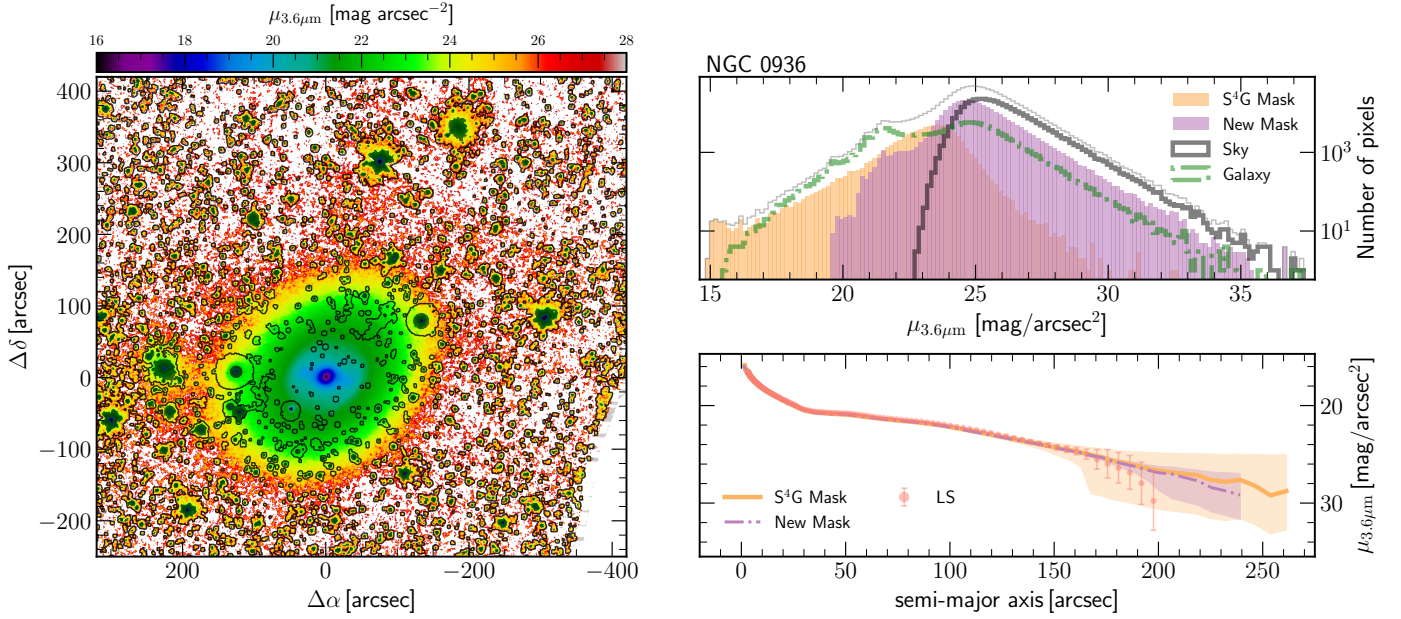


Fig. 4. Difference between the masking of the original S⁴G and our new method. The left panel shows the surface brightness map in the 3.6 μm band of the galaxy NGC 0936 from the original S⁴G image. The shaded region represents the original mask of the S⁴G, while the black contour lines show the regions masked using the method explained in Sect. 4.2. The top right panel shows the histogram of the image (thin grey line) in surface brightness and the different contributions of the pixels in the original S⁴G mask (orange), the pixels added to the mask with the new method (purple), the pixels within the galaxy (green dotted line) and the sky pixels (black line), which are the pixels that are neither masked nor used to create the profile of the galaxy. The bottom right panel shows the surface brightness profiles in mag arcsec^{-2} using the original mask (orange line) and the new mask (purple dashed line). We also show the converted 3.6 μm profile (using the recipe from Sect. 5) using the Legacy Surveys (LS) imaging. The filled region represents the uncertainties due to the sky background subtraction.

of the galaxy. On top of the image, we show the original S⁴G mask with darker shaded regions and the new mask computed with our method using contours (black line). The top right panel shows the histogram (thin grey line) of the image but with the pixels converted to surface brightness and with the sky background subtracted. We show the contribution to the histogram of the pixels masked by the original mask image (orange), the pixels added by our new mask (purple), and the pixels of the galaxy (green dashed line). We determined the pixels of the galaxy as those used to derive the radial surface brightness profile of the galaxy (see Sect. 6) up to a distance of 250 arcsec where the profile reaches the noise level. The remaining pixels that were not masked and not used to construct the profile of the galaxy constitute the background sky (thick black line). In the bottom right panel, we show the radial surface brightness profile measured using the original S⁴G mask and the new mask. We also show the radial profile measured using the deepest and widest imaging available, the *i*-band image of the Legacy Survey. We convert the *i*-band profile to the 3.6 μm band using our recipe described in Sect. 5.

Compared with the original S⁴G, we increase the number of pixels masked by almost a factor of five, from 5% on the original mask to 24% in the new mask. The original S⁴G fails to mask out the faint outskirts of the sources, with an increase of unmasked pixels above $\sim 24 \text{ mag arcsec}^{-2}$, where both distributions cross (Fig. 4). This results in 89% of pixels with a surface brightness fainter than $22 \text{ mag arcsec}^{-2}$ which are not masked out by the original S⁴G. When measuring the background value, we obtain a difference of 0.4% between the old and new masks. However, we decrease the error on the background value by 52% when using the new mask. This is also seen in the profiles panel (Fig. 4; bottom right), where the error on the profile

(the shaded region) is larger for the original S⁴G profile (orange) than for the new one (purple). The integrated magnitude of the galaxy we measure is 9.54 ± 0.04 and 9.55 ± 0.01 mag using the original S⁴G and new mask, respectively. Both values are compatible, with a small difference of 0.01 mag but with an error four times larger when using the original S⁴G mask. This effect occurs because the boxes used to measure the background value are contaminated with low surface brightness (LSB) pixels from sources not properly masked when using the original S⁴G mask, which increases the variance of the pixels, increasing the error on the background value. However, not properly masking LSB regions does not significantly affect the measured value of the sky background. Our new method is more effective in masking the outskirts of sources, i.e. detecting pixels down to a lower S/N (LSB regions) than the original S⁴G method.

The three profiles shown in the bottom right panel match up to a surface brightness of $26 \text{ mag arcsec}^{-2}$ when the profiles start to diverge. The profiles from IRAC imaging (orange and purple) reach the outskirts of the galaxy and extend down to $\sim 29\text{--}30 \text{ mag arcsec}^{-2}$, which is expected according to its surface brightness limit. However, they differ due to the mask used to measure the profiles. Using the original S⁴G mask, we incorporate signal from contaminating sources which results in a brighter tail on the profile. Despite its deeper surface brightness limit (see Table 1), the Legacy Surveys fails to preserve the outskirts of the galaxy. This is an effect of the Legacy Surveys pipeline where the sky background is overestimated and extended objects usually show signatures of destroyed LSB regions. This oversubtraction problem has also been noticed in other works (see Liu et al. 2023). However, this is limited to LSB regions fainter than $\sim 27 \text{ mag arcsec}^{-2}$ and therefore does not affect our results.

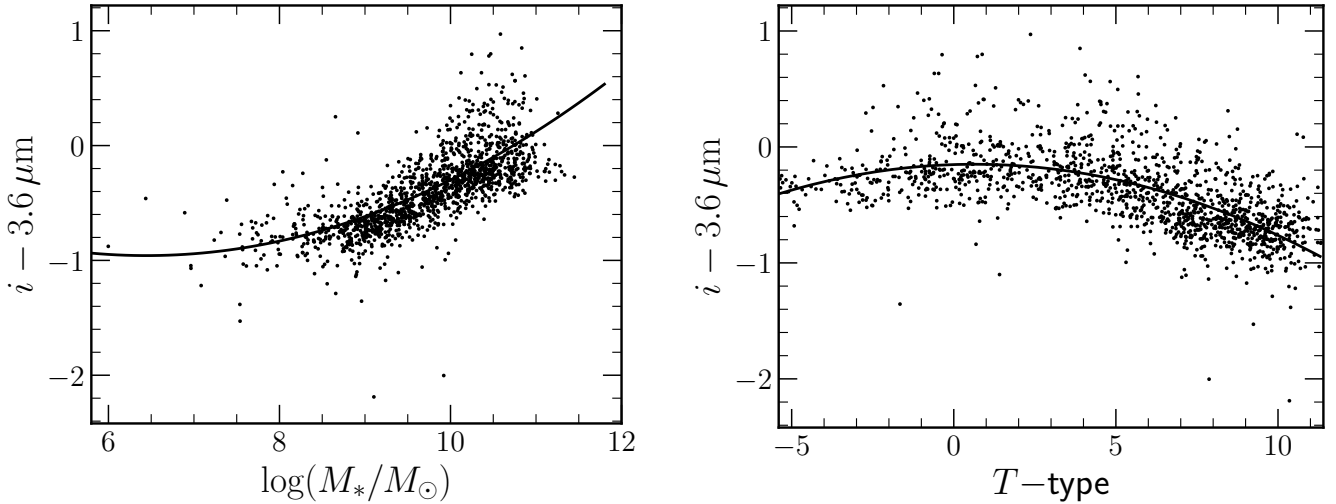


Fig. 5. Optical to infrared conversion using galaxies in the original S⁴G sample. The left panel shows the $i - 3.6 \mu\text{m}$ colour versus the stellar mass relationship. The right panel shows the $i - 3.6 \mu\text{m}$ colour plotted against the CVRHS revised T -types. To have a more homogeneous distribution in the right-hand panel, we added a random offset according to the uncertainties of the morphological revised T -types as reported by Buta et al. (2015), from a normal distribution with a width of their uncertainties, $\mathcal{N}(0, 0.52)$.

4.3. Pipeline 3. Sky background

We follow a similar procedure as that in Salo et al. (2015), using a semi-automatic IDL routine to manually select regions that are used to estimate the sky background, typically 20–30 locations outside the visible galaxy. We avoid placing the boxes on the image edges and in crowded areas. The local sky values in these chosen locations are determined by calculating medians of non-masked pixels within 30 pixel by 30 pixel boxes. Subsequently, the global sky background (sky) and its uncertainty (DSKY) are derived from the mean and standard deviation of these local values, respectively.

5. Optical to infrared correction

Since the goal of the DG extension is to provide images that can be used alongside those of the original S⁴G and the ETG extension, we need to convert from i -band to $3.6 \mu\text{m}$ AB magnitudes. As a benchmark for the calibration we used the 1388 S⁴G galaxies that have both an SDSS i -band image compiled in Knapen et al. (2014) and a well-defined $\mu_{3.6\mu\text{m}} = 26.5 \text{ mag arcsec}^{-2}$ isophote in MM2015. For each galaxy we derive i -band and $3.6 \mu\text{m}$ asymptotic magnitudes with profiles obtained out to the $\mu_{3.6\mu\text{m}} = 26.5 \text{ mag arcsec}^{-2}$ isophote while keeping the orientation parameters fixed to the values presented in MM2015. For the $3.6 \mu\text{m}$ images, we used the aperture correction described in Eq. (1) of W2022. Pipeline 2 masks were accounted for while producing the profiles.

Figure 5 shows the derived $i - 3.6 \mu\text{m}$ colours as a function of stellar mass (left panel) and morphological T -type from Buta et al. (2015) (right panel). We fitted the data with order-two polynomials. The resulting relation between the colour and the stellar mass is

$$i - 3.6 \mu\text{m} = (0.052 \pm 0.007)(\log(M_*/M_\odot))^2 - (0.67 \pm 0.13)\log(M_*/M_\odot) + (1.20 \pm 0.61), \quad (1)$$

and between the colour and the T -type

$$i - 3.6 \mu\text{m} = -(0.0070 \pm 0.0004)T^2 + (0.009 \pm 0.004)T - (0.152 \pm 0.011). \quad (2)$$

The two expressions are represented as continuous black curves in their respective figures. The first expression has a root-mean-square deviation of 0.17 mag and the second one of 0.19 mag.

We convert our i -band apparent magnitudes to $3.6 \mu\text{m}$ apparent magnitudes using the mass relationship (Eq. (1)), which has the lowest sum of the squares residual. Since we need the stellar mass of a galaxy to estimate the $i - 3.6 \mu\text{m}$ colour and the stellar mass is one of the parameters we want to derive from our images, we apply an iterative procedure to estimate the colours. We first assume that the colour $i - 3.6 \mu\text{m} = 0$ and then estimate the stellar mass of the galaxy (see Sect. 6.3). With this first mass estimate, we calculate the $i - 3.6 \mu\text{m}$ colour using Eq. (1), we transform the i -band magnitudes to $3.6 \mu\text{m}$ and calculate the mass with the $3.6 \mu\text{m}$ magnitude again. We repeat this process until we reach convergence in the $i - 3.6 \mu\text{m}$ colour. We reach the convergence after 5–9 iterations to the fifth decimal value. This uncertainty is well below our uncertainty in the integrated magnitudes.

6. Derived quantities

We follow the original S⁴G procedures (MM2015, Salo et al. 2015) along with those in the ETG extension (W2022) to derive photometric quantities. We measure radial profiles of flux and surface brightness, position angle and ellipticity profiles for the DG sample. For the whole set of galaxies in the CS⁴G, we measure asymptotic magnitudes, concentration indices, and sizes. The parameters from the three samples are homogenised using the same methods.

All the parameters derived from this analysis can be found in an electronic table accessible from the CDS and IRSA (see Sect. 8).

6.1. Radial profiles

To derive radial profiles, we used the implementation of the iterative ellipse-fitting method described by Jedrzejewski (1987) in the Photutils package (Bradley et al. 2022, v1.5.0) of Astropy (Astropy Collaboration 2013, 2018, 2022, v5.1). This method is efficient in tracing the structure of a galaxy, following

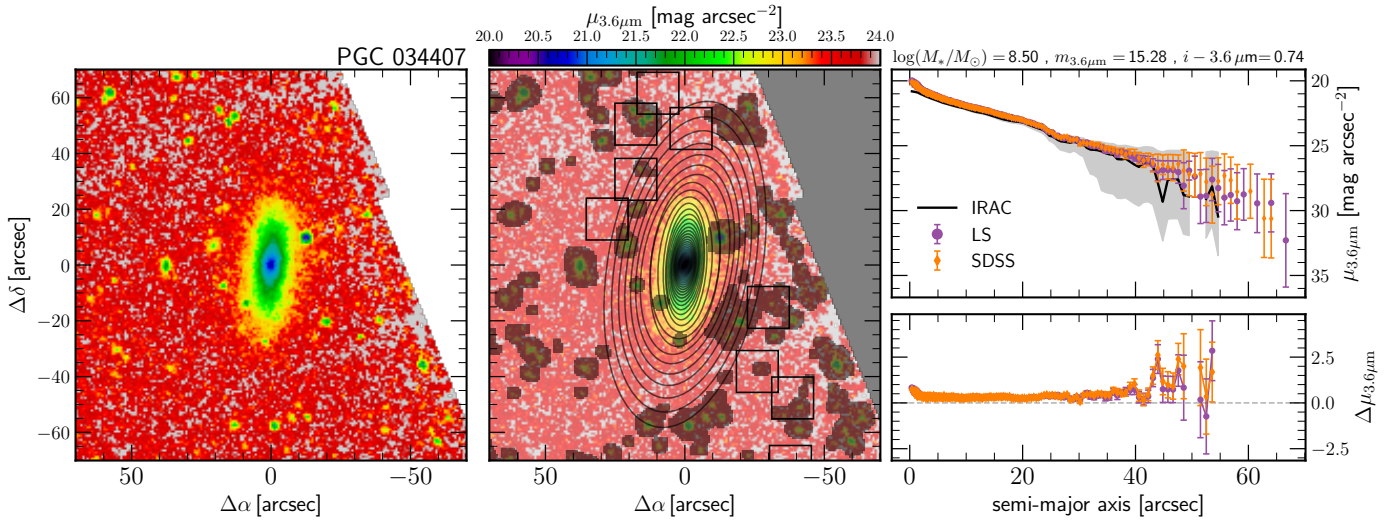


Fig. 6. Surface brightness maps and profiles. The left panel shows the surface brightness map of PGC 34407 in the 3.6 μm band. The middle panel shows the surface brightness map overlaid with the mask image with shaded regions, and the elliptical apertures used to measure the radial profile. The top right panel shows the radial surface brightness profile derived using the 3.6 μm band (black curve), the i -band from the Legacy Imaging Surveys (purple circles), and i -band from the SDSS (orange diamonds). The radial profiles of both the i -band images have been converted to 3.6 μm using the recipe described in Sect. 5. The lower right panel shows the difference between the mid-IR profile and the converted optical profiles.

bars and other features found in the isophotes (e.g. Salo et al. 2015, W2022, Sánchez-Alarcón et al. 2023). This is the same method used in the original sample (MM2015), but implemented in the Python programming language.

We first estimate the galaxy orientation using the image moments on the masked image to improve the success of the fitting procedure, which requires a first ellipse to initialize. Previously, the implementation in Salo et al. (2015) of the same method iterated until the fitting was done through the entire galaxy but sometimes it stopped prematurely, missing the outskirts or failing to fit the inner regions, and forcing a restart of the fitting procedure. To accelerate the process and make it more efficient, we estimate the maximum radius of the profile as the radius where we reach the value that was on the order of the sky level. This way, the fitting is more robust and successful for almost the whole sample ($\geq 90\%$). After a visual inspection of the profile, we decide the maximum radius of the profile. We keep the centre position fixed through the whole fitting procedure and measure the position angle (PA) and the ellipticity (ϵ) for each radial bin. We determine the centre using the same method as in Salo et al. (2015), in which a first guess is done manually and the centre is refined as the location where the brightness gradient is zero. We increase the radius logarithmically by 2% for each radial bin. This adapted step size allows us to have an optimal radial resolution in the inner regions, with wider ellipses in the outskirts reaching higher S/N. We estimate the uncertainties as the quadratic sum of the error on the fit procedure, and the uncertainty of the sky value measured in P3.

We show an example of the surface brightness profile measured of PGC 34407 in the 3.6 μm band in Fig. 6. The left panel shows the surface brightness map of the 3.6 μm image and the middle panel shows the same map with the mask regions shaded and the elliptical apertures used to measure the profiles, along with the boxes used to measure the sky background value and its uncertainty. The top right panel shows the 3.6 μm surface brightness radial profiles of the IRAC imaging, and the profiles from the optical LS and SDSS surveys, converted to 3.6 μm using our recipe described in Sect. 5. In the lower right panel, we show

the difference between the converted optical profiles and the IRAC profile. We used the same elliptical apertures to measure the three profiles. The elliptical isophotes successfully follow the orientation of the inner and the outer regions of the galaxy and the profiles reach the faint outskirts. We observe that the three profiles exhibit consistent behaviour. We see the expected differences, in the inner region caused by the PSF and in the outer part due to the depth of the different images. We reach surface brightnesses of ≥ 25.1 and 26.2 mag arcsec $^{-2}$ [3σ , $10'' \times 10''$] in the mid-IR and i -band, respectively, for 95% of the galaxies. Some galaxies have contaminating light from brighter neighbouring sources (i.e. bright stars or galaxies) that would require a model subtraction to describe the faintest regions.

We do not measure radial profiles for the S 4 G and ETG galaxies, instead, we used the ones derived from MM2015 and W2022. However, when converting profiles to physical units, we used updated distances.

6.2. Asymptotic magnitudes

We derive asymptotic magnitudes from the curve of growth (hereafter, c.o.g.) of the 3.6 μm AB magnitude following the same procedure as in MM2015 and W2022. We measure the cumulative sum over the flux within elliptical apertures with constant ellipticity (ϵ), and position angle (PA), and logarithmically grown by 2% in semi-major axis. We set the orientation values, PA and ϵ , to match the outskirts of the galaxy which were measured by P3. Given the depth of the DG extension, the c.o.g. flattens in the outskirts of the galaxy and the relationship between the local gradient and the magnitude enclosed within each elliptical aperture becomes roughly linear. We can measure the asymptotic magnitudes as the y -intercept of the linear fit of the last points on the magnitude-local gradient plane. For galaxies with i -band images, we convert the i -band magnitudes to 3.6 μm magnitudes.

W2022 show in their Fig. 5 a representation of the same method as what was used in this work to measure the asymptotic magnitudes for two galaxies in the sample with different

inclinations, luminosity, and radial extent. They also discuss the different effects on the uncertainty of this method introduced in the magnitude derived. This method is accurate enough that the uncertainty introduced is smaller than the systematic uncertainty induced by the sky subtraction.

6.3. Absolute magnitudes and stellar masses

To transform apparent magnitudes to absolute magnitudes, we query distances from the NED database (Chen et al. 2022). We used redshift-independent measurements when available (for 97 galaxies, ~23%) and redshifts when not (for 325 galaxies, ~77%). There are four galaxies with redshift-independent measurements with very high and unexpected values, of the order of ~100 Mpc. For these galaxies, we used another distance indicator such as the redshift when available or optical velocities from HyperLeda (see Appendix D for a discussion). We investigated the use of Cosmic-Flow corrected distances (Carrick et al. 2015) and found that, for nearby galaxies, redshift-independent distances or simple redshift-derived estimates provide a more reliable determination of distance.

We also updated the distances of the galaxies in the S⁴G and ETG samples. We find 1920 (82%) and 247 (53%) galaxies with redshift-independent measurements for the S⁴G and ETG, respectively. This results in 2264 (70%) galaxies with redshift-independent measurements for the entire CS⁴G.

Then, we transform the 3.6 μm absolute magnitudes to stellar masses. As reported by Querejeta et al. (2015), in the 3.6 μm band we can assume a constant mass-to-light ratio $\Upsilon_{3.6\mu\text{m}}$ which varies by 10%–30% on integrated galaxy scales. This uncertainty does not affect our general results in terms of statistical behaviour in trends. We assume a Chabrier initial mass function (Chabrier 2003) and use a constant mass-to-light ratio of $\Upsilon_{3.6\mu\text{m}} \sim 0.6 (M_{\odot}/L_{\odot})_{3.6\mu\text{m}}$ (see Meidt et al. 2014; Querejeta et al. 2015; Comerón et al. 2018) to estimate the stellar mass for all the galaxies.

6.4. Concentration indices

The spatial distribution of light contains valuable information about the morphological classification of galaxies and their evolution. Concentration indices can be directly measured from the light distribution of a galaxy. Again following MM2015 and W2022, we measure the concentration parameters C_{31} (de Vaucouleurs 1977) and C_{82} (Kent 1985), defined as

$$C_{31} = R_{75}/R_{25}, \text{ and} \quad (3)$$

$$C_{82} = 5 \log (R_{80}/R_{20}), \quad (4)$$

where R_x is the radius containing $x\%$ of the total luminosity of the galaxy. As in MM2015 and W2022, to avoid assumptions about the shapes of the light profiles, we measure R_x from the c.o.g. We extrapolate the total luminosities of galaxies out to infinity rather than measuring them within set apertures.

We interpolate the c.o.g. to improve the resolution of the profile and increase the accuracy of the measurement. We estimate the uncertainty of the concentration indices for each galaxy by performing a Monte Carlo sampling on the value of the sky, modelling 1000 random sky values from a normal distribution $\mathcal{N}(\text{SKY}; \text{DSKY})$, measuring the concentration indices and estimating the error on the value as the standard deviation from each value.

6.5. Galaxy sizes

We measure the effective (or half-light) radius, R_e , along with the two isophotal radii, $R_{25.5}$ and $R_{26.5}$, denoting the radii at which the surface brightness profile of the galaxy reaches 25.5 and 26.5 mag arcsec⁻², respectively, in the 3.6 μm and i -bands. We also apply an inclination correction to measure the isophotal radius at 25.5 and 26.5 mag arcsec⁻² ($R_{25.5,\text{corr}}$ and $R_{26.5,\text{corr}}$) in both bands. We follow the correction used in W2022 and multiply the intensity profiles by the axial ratio. These radii are measured from the major axis of the elliptical apertures. We measure R_e using the c.o.g. and employing the asymptotic magnitudes to set the total flux of the galaxy. We estimate the uncertainties of the values using the same Monte Carlo strategy as in the concentration indices. We measure the galaxy sizes in the 3.6 μm and 4.5 μm channels when available, and for those galaxies with only i -band we measure it from the i -band profile and the transformed 3.6 μm profile.

6.6. Average surface brightness levels

Using the fitted radial surface brightness profile, we measure the mean surface brightness within the effective radii ($\bar{\mu}_e$) and within 1 kpc ($\bar{\mu}_{1\text{kpc}}$). We interpolate the profile to have a higher resolution and extract more precise values. The uncertainties are again estimated using the Monte Carlo strategy. We measure these values for the DG but also for the entire S⁴G and the ETG extension.

6.7. Morphological revision

The morphological classifications of the sample galaxies have been made in a comprehensive version of the de Vaucouleurs (1959) revised Hubble-Sandage (CVRHS) system (Buta et al. 2015). This includes the recognition of many features of interest to extragalactic observers that were considered mere details in the past. The hallmark of the system is the prominence of galactic rings and lenses and the ease with which these can be added to the system. Rings and lenses are considered primary tracers of galactic secular evolution (Knapen 2012). The CVRHS classification also includes barlenses (Laurikainen et al. 2011; Laurikainen & Salo 2017) denoting the inner lens-like components, actually forming part of the bar, embedded in thin bars in massive galaxies. The galaxies were classified using the 3.6 μm and i -band images presented in this work, together with g -band images from the SDSS and LS when available. The g -band is closest in wavelength to the B -band, the historical band originally used for galaxy morphological study.

A noteworthy finding of our study is that the DG galaxy sample contains a significant number of extreme late-type spirals (i.e. types Scd and later). Typically, such galaxies tend to be fairly rich in HI (Buta et al. 1994). Similarly, there appears to be an unusual number of spindles (edge-ons) in the sample.

Images in the g -band were not available for all of the galaxies, and also some images had poor seeing. Since the classifications are based on only a single examination (i.e. without a second classification of the same objects using the same images), we assume that the uncertainty is $\sigma(T) = 0.7$ stage intervals (Buta et al. 2015, 2019).

7. Discussion

We present a collection of scaling relations using the derived photometric parameters for the DG extension, in comparison

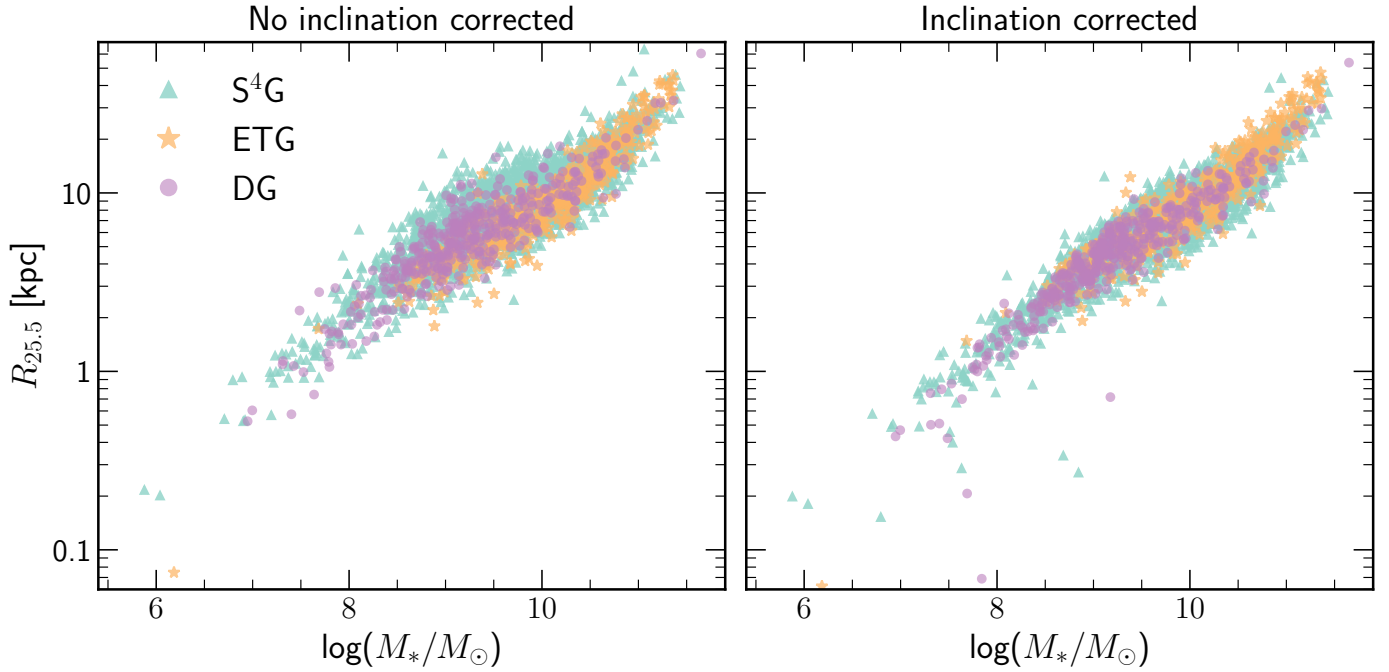


Fig. 7. Stellar mass–size relation showing the correlation of size, traced by the isophotal radius ($R_{25.5}$) with the stellar mass. The DG, ETG, and original S⁴G samples are represented with purple circles, orange stars, and blue triangles respectively. The isophotal radius axis is on a logarithmic scale. The left panel shows the size of galaxies with no correction, and the left panel shows the size of galaxies with inclination correction.

with the original sample (MM2015) and the ETG extension (W2022). We begin with the mass–size relation, then we explore the parameter space defined by the photometric parameters and the regions where the different morphological types of galaxies reside. We finish by exploring the difference in H I content for the different samples.

7.1. Mass–size relation

In the left side panel of Fig. 7 we plot the isophotal radius at 25.5 mag arcsec^{−2} ($R_{25.5}$) against the stellar mass of the galaxy for the DG extension (purple circles), the ETG extension (orange stars) and the original sample (blue triangles). MM2015 and W2022 showed this relation (in their Figs. 14 and 11, respectively), demonstrating the expected monotonic trend, where galaxies with a higher stellar mass are also larger. Our photometric measurements, properly converted to 3.6 μ m, reproduce this trend, with excellent agreement between our new galaxies and the original S⁴G and the ETG samples. The DG sample populates, in general, the intermediate-low mass region ($7 \lesssim \log(M_*/M_\odot) \lesssim 11$) with a few massive galaxies with stellar mass above $10^{11} M_\odot$.

The right panel of Fig. 7 shows the mass–size relation using inclination-corrected isophotal sizes (see Sect. 6.5). This correction reduces the root mean square deviation of the sizes from the average trends in mass bins by 0.014 dex, leading to a tighter relation with lower dispersion. This behaviour was also discussed in W2022 (their Fig 11), but the corrected values were not published. However, for a few edge-on galaxies, the correction introduces discrepancies, increasing their deviation from the expected trend.

Intriguingly, the DG sample contains six large and massive galaxies: NGC 1316, NGC 1404, NGC 4125, NGC 4552, NGC 7172, and NGC 7410. They can be identified in the top-right corner of Fig. 7, with size $R_{25.5} > 15$ kpc and stellar

mass $\log(M_*/M_\odot) > 10.9$. The largest galaxy is NGC 7410 with $R_{25.5} = 51$ kpc and $\log(M_*/M_\odot) = 11.36$ while the most massive is NGC 1316 with $R_{25.5} = 43$ kpc and $\log(M_*/M_\odot) = 11.65$. The galaxies NGC 1316 and NGC 1404 are part of the Fornax Cluster (see e.g. Ferguson & Sandage 1988, and follow-up papers). These galaxies follow the same trend as their counterparts in the S⁴G and ETG samples.

We also find three dwarf galaxies with mass below $\log(M_*/M_\odot) < 7$ and smaller than $R_{25.5} < 0.6$ kpc. These galaxies are PGC 3097691, UGC 04879, and UGC 08308.

These galaxies underscore the significance of incorporating the DG extension to improve the completeness throughout the mass range of the sample.

7.2. General trends

In Fig. 8, we present a corner plot of the different photometric parameters derived in this work to explore the different relationships between the parameters in a compact way. We defined two sub-samples from the S⁴G, the late-type (S⁴G-LTG with $T > 0$) and the early-type (S⁴G-ETG with $T < 0$) to compare with the DG and ETG extensions. We represent the different parameter spaces defined by the stellar mass, $\log(M_*/M_\odot)$, the size of the galaxy measured as the isophotal radius at 25.5 mag arcsec^{−2}, $\log(R_{25.5})$, the concentration index, C_{82} , and the morphological type, T , for the different samples: the original S⁴G-LTGs (blue triangles) and S⁴G-ETGs (down-pointing red triangles), the ETG extension (orange stars), and the DG extension (purple stars). The curves represent the average values per mass bin in the middle two rows and per morphological-type bins in the bottom row.

Along the diagonal of the figure, the panels (from top-left to bottom-right) show the distribution of stellar mass, isophotal radius, concentration index, and morphological type, respectively, for the four different samples. The sample size difference

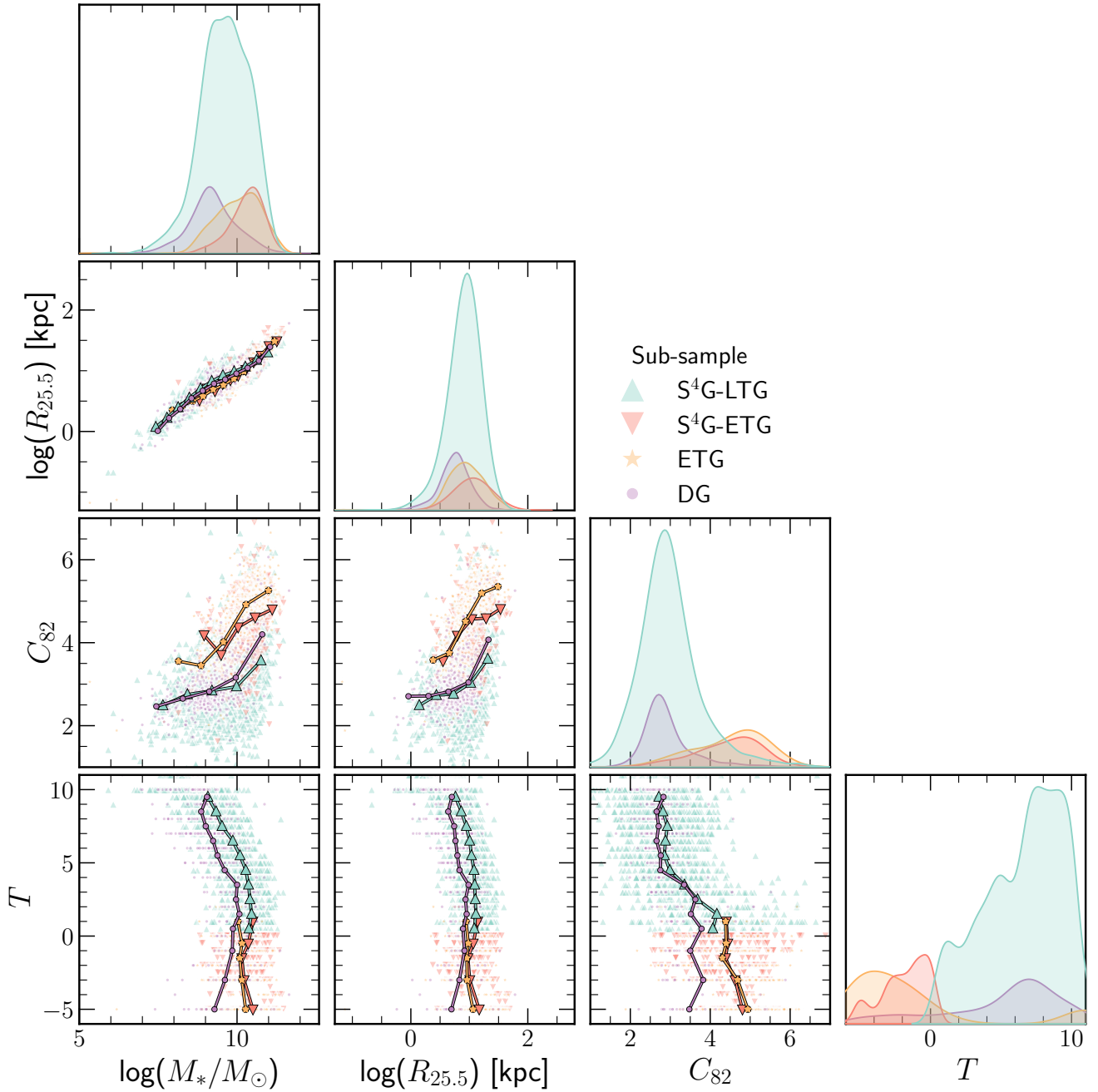


Fig. 8. Corner plot of the parameter space described by some of the photometric parameters derived for the DG sample together with the original sample and ETG extension. From top-left to bottom-right, the diagonal of the figure shows the absolute distribution of the stellar mass, $\log(M_*/M_\odot)$, the galaxy size measured by the isophotal radii at 25.5 mag arcsec⁻², $\log(R_{25.5})$, the concentration index, C_{82} , and the morphological types, T , for the different samples, the original S⁴G-LTGs (blue triangles) and S⁴G-ETGs (down-pointing red triangles), the ETG extension (orange stars), and the DG extension (purple circles). These distributions are heavily smoothed. Each column and row represents one of the parameters. The first column (left) shows, from top to bottom, galaxy size, concentration index and morphological type versus stellar mass. The second column shows the concentration index and morphological type versus galaxy size. The last column shows morphological types versus concentration index. Curves represent the average values within mass bins for the middle two rows and within morphological types bins for the last row.

immediately stands out. While the ETG extension (465) is larger than the number of ETG galaxies in the original S⁴G (282), the DG sample (422) is much smaller than the LTG part of the original S⁴G (2070), yet it is essential for the completeness of the S⁴G.

The stellar masses show different trends in the histogram, with distribution means and standard deviations of 9.23 ± 0.78 , 9.77 ± 0.83 , 10.11 ± 0.82 , and 10.46 ± 0.64 $\log(M_*/M_\odot)$ for

the DG, S⁴G-LTG, ETG, and S⁴G-ETG, respectively. While the S⁴G-ETG and ETG have broader distributions, the DG and S⁴G-LTG distributions are narrower and more skewed. The DG and the ETG sub-samples are 0.54 dex and 0.35 dex, respectively, less massive on average than the LTG and ETG sub-samples of the original S⁴G.

The stellar mass panels (first column) indicate the stellar mass–size relation (second row, same as Fig. 7), the

concentration indices (third row), and the morphological types (fourth row) with respect to the stellar mass. As discussed previously, the parameters derived in this analysis for the DG sample, and properly corrected to the 3.6 μm band, follow similar trends as the S⁴G and ETG. Lower-mass galaxies show smaller radii, lower concentration indices, and higher values of morphological types, while massive galaxies have larger sizes and broader distributions in concentration index and morphological types. These trends are easily seen with the average curves.

The sizes of galaxies show narrow and aligned distributions for all the sub-samples with some difference in the logarithmic means and standard deviations, $\log(R_{25.5}/\text{kpc}) = 0.74 \pm 0.35$, 0.97 ± 0.27 , 0.98 ± 0.31 , and 1.11 ± 0.27 for the DG, S⁴G-LTG, ETG, and S⁴G-ETG sub-samples, respectively. In units of kiloparsecs, this is 7 ± 5 , 11 ± 6 , 11 ± 9 , and 14 ± 8 kpc, respectively. The galaxies in the DG and ETG extension are, on average, 35% smaller than the LTG of the original S⁴G, while the galaxies in the ETG extension are 20% larger.

The size-related panels (second column) show the concentration indices (third row), and the morphological types (fourth row) with respect to the isophotal radii. Again, as galaxy size is well correlated with stellar mass, we find expected trends, as previously reported by MM2015 and W2022. Smaller galaxies (i.e. less massive) have lower values of C_{82} and later types. However, these relations are less strongly correlated than the ones with respect to stellar mass.

The concentration indices C_{82} (third row) show different distributions (in the third column) for the DG and ETG sub-samples but very similar ones for the extensions and the original S⁴G. The mean values for each sub-sample are 2.93 ± 0.74 , 3.12 ± 0.74 , 4.31 ± 0.99 , and 4.19 ± 0.97 for the DG, S⁴G-LTG, ETG, and S⁴G-ETG, respectively. These differences are less significant than 10%. However, in general, these figures and average lines show two clear populations that can be identified as LTGs and ETGs, with lower concentration and higher concentration, respectively. As previously reported in many works (see, e.g. MM2015, W2022, and references therein), the light concentration of a galaxy correlates with its morphological type, and it is one of the parameters typically used for galaxy classification for higher-redshift and unresolved objects. However, among the disc galaxies, there is a notable scatter due to S0 galaxies having both high and intermediate values of concentration indices.

The morphological type distributions (fourth row) show the variety of galaxies in each sub-sample. All samples combined show a homogeneous and representative of the population in the local Universe. The mean value and standard deviations of T -types for each sub-sample are 4.44 ± 4.3 , 6.54 ± 2.77 , -1.25 ± 5.17 , and -1.8 ± 1.58 for the DG, S⁴G-LTG, ETG, and S⁴G-ETG, respectively. These average values are expected and a consequence of the definition of each sub-sample. We find that for similar morphological types, the galaxies in the DG and ETG extensions are less massive, smaller, and less concentrated than their counterparts in the original S⁴G, as shown in the average values. The plot of the morphological types with respect to the concentration index (third column) also shows the expected clear transition between less-concentrated LTGs and highly concentrated ETGs.

Our recipes for the conversion between optical i -band and 3.6 μm band yield results consistent with the original S⁴G sub-sample and the ETG extension observed with *Spitzer*, as shown in all the panels of Fig 8. By adding this new extension, we increase the sample by 422 new galaxies, which represents 18% of the original S⁴G, and 36% of the sub-sample with mass below $\log(M_*/M_\odot) < 9$. Together with the ETG extension, the

survey increased the number of galaxies by 38%, significantly improving its completeness.

7.3. Gas content

Since the DG sub-sample originated from the lack of radio-derived velocities, we may expect that galaxies in this sub-sample exhibit a low H I content. This particularity raises another interesting reason to study this sub-sample and include it in the CS⁴G.

We show the distribution of the H I mass fraction, $\log(M_{\text{HI}}/M_*)$, with respect to the stellar mass, $\log(M_*/M_\odot)$, for all the sub-samples in the left panel of Fig. 9. We used the magnitude of the 21 cm line from HyperLeda to estimate the mass of H I. There are 188 (44%), 2011 (97%), 65 (14%), and 234 (83%) available measurements for the DG, S⁴G-LTG, ETG, and S⁴G-ETG sub-samples, respectively. We follow a similar procedure as Zwaan et al. (1997); Namumba et al. (2023) to obtain H I masses. We measure H I masses as follows:

$$M_{\text{HI}} = 2.36 \times 10^5 d^2 F_{21} \times \frac{1}{1+z}, \quad (5)$$

where d is the distance to the galaxy in Mpc, F_{21} is the 21 cm line integrated flux in Janskys reported by HyperLeda, and z is the redshift of the galaxy. The DG, S⁴G-LTG, ETG, and S⁴G-ETG sub-samples are shown in purple circles, blue triangles, orange stars, and down-pointing red triangles, respectively. The lines represent the average values within the same bins of stellar mass. The right column shows a violin plot with the H I mass $\log(M_{\text{HI}}/M_\odot)$ distribution for all the sub-samples.

Despite the lack of observational H I data in much of the DG sub-sample, the average values of the DG galaxies are below the averages of their counterparts in the original S⁴G-LTG. The disc galaxies in our new DG sub-sample, with the same mass as their LTG counterparts in the original S⁴G, have a lower H I mass fraction. The ETG and S⁴G-ETG sub-samples show no clear difference in their average values. However, only 14% of the galaxies in the ETG sample have H I measurements. If these are the ETG galaxies richest in gas, they might represent the upper limit of H I mass, while the remaining ones might have a lower content of gas. These ETG galaxies most probably are borderline galaxies between ellipticals and spirals. In the right panel, we see a similar trend where the distribution of the DG sub-sample is centred at lower values of H I mass than the S⁴G-LTG, with a longer tail towards the low-mass end while the ETGs and S⁴G-ETG show similar distributions.

Both the ETG and DG sub-samples have lower stellar masses and, as discussed in the previous subsection, lower fractions of H I content than the S⁴G counterparts with a larger difference for the DG sub-sample. The exclusion of many of these galaxies from the original S⁴G may be attributed to their lower mass, lower luminosity, and reduced H I fraction. Their lack of radio-derived velocities stems from either insufficiently deep data or a lack of H I observations at the time the original S⁴G project was defined.

8. Conclusions

We have presented CS⁴G, a survey of 3239 galaxies with consistent homogenised photometric parameters. We joined the original sample of S⁴G with both the ETG (W2022) and DG (this paper) extensions to deliver a single catalogue to the community. Additionally, we homogenised all measurements in the

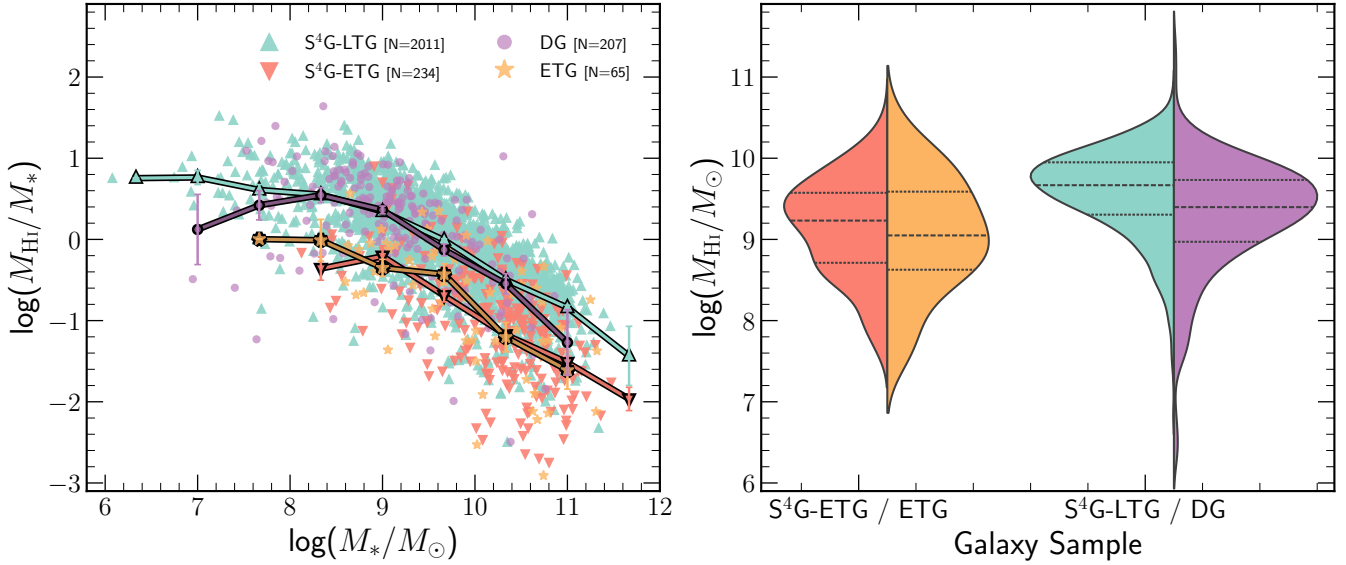


Fig. 9. HI mass content distribution. The left panel shows the logarithmic gas fraction, $\log(M_{\text{HI}}/M_*)$, versus the stellar mass $\log(M_*/M_\odot)$. The right column shows a violin plot of the HI mass relative distributions for all the sub-samples. The x-axis shows the relative distribution of all sub-samples. The lines show the median values (dashed) and the 25% and 75% percentiles (dotted). The S⁴G-LTG, S⁴G-ETG, ETG, and DG sub-samples are shown with blue triangles, down-pointing red triangles, orange stars, and purple circles, respectively. The numbers shown in the legend of the left panel indicate the number of galaxies with HI measurements in each sub-sample.

catalogue and included new measurements of effective radius and mean surface brightness levels within the effective radius and within 1 kpc for the whole CS⁴G.

We incorporated 401 DGs and 21 elliptical galaxies into the sample. We release archival and new *i*-band imaging for 367 galaxies of the DG extension and 55 3.6 μm and 4.5 μm band images from the *Spitzer* Heritage Archive. We also include a number of images (respectively 102, 169, 77, and one) from DES, LS, SDSS, and HST. We observed 18 galaxies in the *i*-band using the LT and the NTT. We analysed all the images using the original S⁴G methods (MM2015, Salo et al. 2015, W2022) and derived radial surface brightness profiles, curves of growth, asymptotic magnitudes, stellar masses, effective radii, isophotal radii, and concentration parameters. We derived a recipe to transform optical *i*-band to 3.6 μm magnitudes using images from the original S⁴G and optical *i*-band images. Using this recipe, we converted *i*-band magnitudes to 3.6 μm to obtain absolute parameters.

The CS⁴G parameters allowed us to study different scaling relations and find the following results:

- We improved the completeness of the survey by adding 422 galaxies to the original sample. This represents 15% of the total previous sample (S⁴G+ETG). However, for low-mass galaxies ($M_* < 10^9 M_\odot$), this sample represents an increment of 36%;
- We measured all the parameters described for the three samples (S⁴G, ETG, and DG) using the same methods, creating a consistent and homogenised CS⁴G sample;
- Our recipe for the conversion between the optical *i*-band to the infrared 3.6 μm band yields measurements consistent with those from the original survey;
- We improved the mask images of the DG by increasing the number of pixels masked by a factor of five in comparison with the original S⁴G masks. We masked regions 2 mag arcsec⁻² fainter. This does not affect the galaxy integrated magnitudes, but it results in a 52% lower error on the background characterization;

- The DG extension consists of galaxies with masses $7 \leq \log(M_*/M_\odot) \leq 11$. It contains six massive galaxies with $\log(M_*) > 11$ and three galaxies in a tail of lower-mass dwarf galaxies $\log(M_*) < 7$;
- The DG galaxies are, on average, 0.23 dex less massive and 34% smaller in size than the LTGs of the original S⁴G. They have similar concentration indices (within 5%) and later morphological types;
- The DG sample galaxies show a lower HI gas fraction than the LTGs in the original S⁴G. However, we lack HI measurements for 86% and 56% of the ETG and DG samples, respectively. Further radio measurements are needed to confirm this finding;
- The CS⁴G encompasses at least 99.94% of the complete sample of nearby galaxies meeting the selection criteria of the S⁴G in the local Universe.

In summary, our measurements and our study of the scaling relations show good agreement with previous studies yet also highlight specific details that are worthy of further investigation. All images, profiles, and derived parameters are made available via the NASA/IPAC Infrared Science Archive and the CDS. The CS⁴G will serve as a local benchmark for comparing higher redshift studies from upcoming surveys such as *Euclid*, *Rubin*, *Roman*, and others.

Data availability

The CS⁴G catalogue and datasets, including Table E.2 are available at the CDS via anonymous ftp to cdsarc.cds.unistra.fr (130.79.128.5) or via <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/697/A38>. The data will soon be available at IRSA, <https://irsa.ipac.caltech.edu/data/SPITZER/S4G/overview.html>. The catalogue contains all the information presented in previous samples homogenised with the methods described together with the additional bands studied in this work, the *i*-band (columns with suffix 3, e.g. mag3) and the converted 3.6 μm (inserted in

columns with suffix 1, e.g. mag1) magnitudes. We include some extra parameters (R_e , $\bar{\mu}_e$, $\bar{\mu}_{1\text{kpc}}$), and inclination-corrected isophotal radii ($R_{25.5,\text{corr}}$ and $R_{26.5,\text{corr}}$) derived for the original S⁴G and ETG that were not previously published. They are explained in Sect. 6. We include all the parameters queried from the HyperLeda database. We include the CVRHS morphological classification for the DG extension and for the S⁴G (Buta et al. 2015) and ETG samples (W2022).

Acknowledgements. This work is based in part on observations made with the Spitzer Space Telescope, which was operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. Based on observations collected at the European Organisation for Astronomical Research in the southern hemisphere under ESO programme 0103.B-0586(A). Based on observations made with the Liverpool Telescope operated on the island of La Palma by Liverpool John Moores University in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias with financial support from the UK Science and Technology Facilities Council. We acknowledge support from the Agencia Estatal de Investigación del Ministerio de Ciencia, Innovación y Universidades (MCIU/AEI) under the grants “The structure and evolution of galaxies and their outer regions” and the European Regional Development Fund (ERDF) with references PID2019-105602GB-I00/10.13039/501100011033 and PID2022-136505NB-I00/10.13039/501100011033. Co-funded by the European Union (MSCA Doctoral Network EDUCADO, GA 101119830 and Widening Participation, ExGal-Twin, GA 101158446). SC acknowledges funding from the State Research Agency (AEI) of the Spanish Ministry of Science, Innovation, and Universities under the grant “The relic galaxy NGC 1277 as a key to understanding massive galaxies at cosmic noon” with reference PID2023-149139NB-I00. J.R. acknowledges financial support from the Spanish Ministry of Science and Innovation through the project PID2022-138896NB-C55. AEW acknowledges support from the STFC [grant number ST/X001318/1]. This work was authored by an employee of Caltech/IPAC under contract No. 80GSFC21R0032 with the National Aeronautics and Space Administration. LCH was supported by the National Science Foundation of China (11991052, 12233001), the National Key R&D Program of China (2022YFF0503401), and the China Manned Space Project (CMS-CSST-2021-A04, CMS-CSST-2021-A06). TK acknowledges support from the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. RS-2023-00240212) and the NRF grant funded by the Korean government (MSIT) (No. 2022R1A4A3031306). EA and AB gratefully acknowledge financial support from the CNES (Centre National d’Études Spatiales, France). DAG was supported by STFC grants ST/T000244/1 and ST/X001075/1. JK acknowledges support from NSF through grant AST-2006600. We acknowledge the usage of the HyperLeda database (<http://leda.univ-lyon1.fr>). This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France. Funding for the Sloan Digital Sky Survey V has been provided by the Alfred P. Sloan Foundation, the Heising-Simons Foundation, the National Science Foundation, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. SDSS telescopes are located at Apache Point Observatory, funded by the Astrophysical Research Consortium and operated by New Mexico State University, and at Las Campanas Observatory, operated by the Carnegie Institution for Science. The SDSS web site is www.sdss.org. SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration, including Caltech, The Carnegie Institution for Science, Chilean National Time Allocation Committee (CNTAC) ratified researchers, The Flatiron Institute, the Gotham Participation Group, Harvard University, Heidelberg University, The Johns Hopkins University, L’Ecole polytechnique fédérale de Lausanne (EPFL), Leibniz-Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Extraterrestrische Physik (MPE), Nanjing University, National Astronomical Observatories of China (NAOC), New Mexico State University, The Ohio State University, Pennsylvania State University, Smithsonian Astrophysical Observatory, Space Telescope Science Institute (STScI), the Stellar Astrophysics Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Illinois at Urbana-Champaign, University of Toronto, University of Utah, University of Virginia, Yale University, and Yunnan University. The DESI Legacy Imaging Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS), the Beijing-Arizona Sky Survey (BASS), and the Mayall z-band Legacy Survey (MzLS). DECaLS, BASS and

MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF’s NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory (LBNL). Legacy Surveys also uses data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. Legacy Surveys was supported by: the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility; the U.S. National Science Foundation, Division of Astronomical Sciences; the National Astronomical Observatories of China, the Chinese Academy of Sciences and the Chinese National Natural Science Foundation. LBNL is managed by the Regents of the University of California under contract to the U.S. Department of Energy. The complete acknowledgements can be found at <https://www.legacysurvey.org/acknowledgment/>. **Software:** This work made use of Astropy (<http://www.astropy.org>) a community-developed core Python package and an ecosystem of tools and resources for astronomy (Astropy Collaboration 2013, 2018, 2022); Photutils, an Astropy package for detection and photometry of astronomical sources (Bradley et al. 2022); Matplotlib (Hunter 2007); NumPy (Harris et al. 2020); SciPy (Virtanen et al. 2020); Pandas (The pandas development team 2020); TOPCAT (Taylor 2005); SExtractor (Bertin & Arnouts 1996); Swarp (Bertin 2010); and SAO Image DS9 (Joye & Mandel 2003).

References

- Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2003, *AJ*, **126**, 2081
- Abbott, T. M. C., Adamów, M., Aguena, M., et al. 2021, *ApJS*, **255**, 20
- Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, *PASJ*, **70**, S4
- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, **219**, 12
- Astropy Collaboration (Robitaille, T. P., et al.) 2013, *A&A*, **558**, A33
- Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, **156**, 123
- Astropy Collaboration (Price-Whelan, A. M., et al.) 2022, *ApJ*, **935**, 167
- Barnes, K. L., van Zee, L., Dale, D. A., et al. 2014, *ApJ*, **789**, 126
- Barnsley, R. M., Jermak, H. E., Steele, I. A., et al. 2016, *J. Astron. Telesc. Instrum. Syst.*, **2**, 015002
- Bertin, E. 2010, Swarp: Resampling and Co-adding FITS Images Together, Astrophysics Source Code Library [[record ascl:1010.0681](https://doi.org/10.1010.0681)]
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
- Bradley, L., Sipőcz, B., Robitaille, T., et al. 2022, <https://doi.org/10.5281/zenodo.6825092>
- Buta, R., Mitra, S., de Vaucouleurs, G., & Corwin, H. G., Jr. 1994, *AJ*, **107**, 118
- Buta, R. J., Sheth, K., Athanassoula, E., et al. 2015, *ApJS*, **217**, 32
- Buta, R. J., Verdes-Montenegro, L., Damas-Segovia, A., et al. 2019, *MNRAS*, **488**, 2175
- Buzzoni, B., Delabre, B., Dekker, H., et al. 1984, *The Messenger*, **38**, 9
- Cantiello, M., Blakeslee, J. P., Ferrarese, L., et al. 2024, *ApJ*, **966**, 145
- Carollo, C. M., Scarlata, C., Stiavelli, M., Wyse, R. F. G., & Mayer, L. 2007, *ApJ*, **658**, 960
- Carrick, J., Turnbull, S. J., Lavaux, G., & Hudson, M. J. 2015, *MNRAS*, **450**, 317
- Chabrier, G. 2003, *PASP*, **115**, 763
- Chen, T. X., Schmitz, M., Mazzarella, J. M., et al. 2022, *ApJS*, **260**, 5
- Comerón, S., Elmegreen, B. G., Knapen, J. H., et al. 2011, *ApJ*, **741**, 28
- Comerón, S., Elmegreen, B. G., Salo, H., et al. 2012, *ApJ*, **759**, 98
- Comerón, S., Salo, H., Laurikainen, E., et al. 2014, *A&A*, **562**, A121
- Comerón, S., Salo, H., & Knapen, J. H. 2018, *A&A*, **610**, A5
- Cuomo, V., Debattista, V. P., Racz, S., et al. 2023, *MNRAS*, **518**, 2300
- de Vaucouleurs, G. 1959, *Handbuch Phys.*, **53**, 275
- de Vaucouleurs, G. 1977, in *Evolution of Galaxies and Stellar Populations*, eds. B. M. Tinsley, D. C. Larson, & R. B. Gehret, 43
- Dey, A., Schlegel, D. J., Lang, D., et al. 2019, *AJ*, **157**, 168
- Díaz-García, S., Salo, H., Knapen, J. H., & Herrera-Endoqui, M. 2019, *A&A*, **631**, A94
- Drlica-Wagner, A., Carlin, J. L., Nidever, D. L., et al. 2021, *ApJS*, **256**, 2
- Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, *ApJS*, **154**, 10
- Ferguson, H. C., & Sandage, A. 1988, *AJ*, **96**, 1520
- Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, *AJ*, **150**, 150
- Ford, H. C., Bartko, F., Bely, P. Y., et al. 1998, *SPIE Conf. Ser.*, **3356**, 234
- Fruchter, A. S., & Hook, R. N. 2002, *PASP*, **114**, 144

- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, **585**, 357
- Hunter, J. D. 2007, *Comput. Sci. Eng.*, **9**, 90
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, **873**, 111
- Jarrett, T. H., Cluver, M. E., Brown, M. J. I., et al. 2019, *ApJS*, **245**, 25
- Jedrzejewski, R. I. 1987, *MNRAS*, **226**, 747
- Joye, W. A., & Mandel, E. 2003, in *Astronomical Society of the Pacific Conference Series*, 295, *Astronomical Data Analysis Software and Systems XII*, eds. H. E. Payne, R. I. Jedrzejewski, & R. N. Hook, 489
- Karachentsev, I. D., Karachentseva, V. E., Dolphin, A. E., et al. 2000, *A&A*, **363**, 117
- Kent, S. M. 1985, *ApJS*, **59**, 115
- Knapen, J. H. 2012, *Proc. Int. Astron. Union*, **10**, 326
- Knapen, J. H., Erroz-Ferrer, S., Roa, J., et al. 2014, *A&A*, **569**, A91
- Laine, J., Laurikainen, E., Salo, H., et al. 2014a, *MNRAS*, **441**, 1992
- Laine, S., Knapen, J. H., Muñoz-Mateos, J.-C., et al. 2014b, *MNRAS*, **444**, 3015
- Laine, J., Laurikainen, E., & Salo, H. 2016, *A&A*, **596**, A25
- Laureijs, R. J., Duvet, L., Escudero Sanz, I., et al. 2010, *SPIE Conf. Ser.*, **7731**, 77311H
- Laurikainen, E., & Salo, H. 2017, *A&A*, **598**, A10
- Laurikainen, E., Salo, H., Buta, R., & Knapen, J. H. 2011, *MNRAS*, **418**, 1452
- Liu, Q., Abraham, R., Martin, P. G., et al. 2023, *ApJ*, **953**, 7
- Makarov, D., Prugniel, P., Terekhova, N., Courtois, H., & Vauglin, I. 2014, *A&A*, **570**, A13
- Meidt, S. E., Schinnerer, E., van de Ven, G., et al. 2014, *ApJ*, **788**, 144
- Menéndez-Delmestre, K., Gonçalves, T. S., Sheth, K., et al. 2024, *MNRAS*, **527**, 11777
- Moustakas, J., Lang, D., Dey, A., et al. 2023, *ApJS*, **269**, 3
- Muñoz-Mateos, J. C., Sheth, K., Regan, M., et al. 2015, *ApJS*, **219**, 3
- Namumba, B., Román, J., Falcón-Barroso, J., et al. 2023, *MNRAS*, **521**, 5177
- Querejeta, M., Meidt, S. E., Schinnerer, E., et al. 2015, *ApJS*, **219**, 5
- Román, J., Trujillo, I., & Montes, M. 2020, *A&A*, **644**, A42
- Salo, H., Laurikainen, E., Laine, J., et al. 2015, *ApJS*, **219**, 4
- Sánchez-Alarcón, P. M., & Ascasibar, Y. 2023, *RAS Tech. Instrum.*, **2**, 129
- Sánchez-Alarcón, P. M., Román, J., Knapen, J. H., et al. 2023, *A&A*, **677**, A117
- Sheth, K., Regan, M., Hinz, J. L., et al. 2010, *PASP*, **122**, 1397
- Sheth, K., Armus, L., Athanassoula, E., et al. 2013, *Spitzer Space Telescope – General Observer Proposal #10043: Not Dead Yet! Completing Spitzer’s Legacy with Early Type Galaxies*, Spitzer Proposal ID: 10043, principal Investigator: Kartik Sheth, NRAO Technical Contact: Kartik Sheth, NRAO.
- Sirianni, M., Jee, M. J., Benítez, N., et al. 2005, *PASP*, **117**, 1049
- Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints [arXiv:1503.03757]
- Steele, I. A., Smith, R. J., Rees, P. C., et al. 2004, *SPIE Conf. Ser.*, **5489**, 679
- Taylor, M. B. 2005, in *Astronomical Society of the Pacific Conference Series*, 347, *Astronomical Data Analysis Software and Systems XIV*, eds. P. Shopbell, M. Britton, & R. Ebert, 29
- The Dark Energy Survey Collaboration 2005, arXiv e-prints [arXiv:astro-ph/0510346]
- The pandas development team 2020, <https://doi.org/10.5281/zenodo.3509134>
- Venhola, A., Peletier, R., Laurikainen, E., et al. 2017, *A&A*, **608**, A142
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nat. Methods*, **17**, 261
- Watkins, A. E., Laine, J., Comerón, S., Janz, J., & Salo, H. 2019, *A&A*, **625**, A36
- Watkins, A. E., Salo, H., Laurikainen, E., et al. 2022, *A&A*, **660**, A69
- Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, *ApJS*, **154**, 1
- York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, *AJ*, **120**, 1579
- Zenteno, A., Kluge, M., Kharkrang, R., et al. 2025, *A&A*, submitted [arXiv:2503.21066]
- Zwaan, M. A., Briggs, F. H., Sprayberry, D., & Sorar, E. 1997, *ApJ*, **490**, 173
-
- ¹ Instituto de Astrofísica de Canarias, c/ Vía Láctea s/n, 38205 La Laguna, Tenerife, Spain
- ² Departamento de Astrofísica, Universidad de La Laguna, 38206 La Laguna, Tenerife, Spain
- ³ Space Physics and Astronomy Research Unit, University of Oulu, Pentti Kaiteran katu 1, 90014 Oulu, Finland
- ⁴ Departamento de Física de la Tierra y Astrofísica, Universidad Complutense de Madrid, 28040 Madrid, Spain
- ⁵ Centre of Astrophysics Research, School of Physics, Astronomy and Mathematics, University of Hertfordshire, Hatfield, UK
- ⁶ Department of Physics and Astronomy, University of Alabama, Box 870324, Tuscaloosa, AL 35487, USA
- ⁷ IPAC, Mail Code 314-6, Caltech, 1200 E. California Blvd., Pasadena, CA 91125, USA
- ⁸ Max-Planck-Institut für Extraterrestrische Physik, Giessenbach-Str. 1, 85748 Garching, Germany
- ⁹ Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
- ¹⁰ Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK
- ¹¹ MMT Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ 85721, USA
- ¹² Kavli Institute for Astronomy and Astrophysics, Peking University, 12 Beijing 100871, PR China
- ¹³ Department of Astronomy, School of Physics, Peking University, Beijing 100871, PR China
- ¹⁴ University of Louisville, Department of Physics and Astronomy, 102 Natural Science Building, Louisville, KY 40292, USA
- ¹⁵ Finnish Centre of Astronomy with ESO (FINCA), Vesilinnantie 5, 20014 University of Turku, Finland
- ¹⁶ Specim, Spectral Imaging Ltd., Elekroniikkatie 13, 90590 Oulu, Finland
- ¹⁷ Department of Astronomy and Atmospheric Sciences, Kyungpook National University, Daegu 702-701, Republic of Korea
- ¹⁸ Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794-3800, USA
- ¹⁹ Normet Oy, Elekroniikkatie 8, 90590 Oulu, Finland
- ²⁰ The Observatories, Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA
- ²¹ Department of Astronomy & Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
- ²² Observatório do Valongo, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
- ²³ Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
- ²⁴ Observatorio Astronómico Nacional (IGN), C/Alfonso XII, 3, 28014 Madrid, Spain
- ²⁵ NASA Headquarters Mary W. Jackson Building, 300 E Street SW, Washington, DC 20546, USA
- ²⁶ Steward Observatory and Department of Astronomy, University of Arizona, 933 N. Cherry Ave., Tucson, AZ 85721, USA

Appendix A: Origin of the bias in the S⁴G sample

The S⁴G aimed to construct a volume-, magnitude-, and size-limited survey of galaxies in the local Universe and study the spatial distribution of light through mid-infrared imaging. Originally, when the sample of the S⁴G was selected (Sheth et al. 2010), the query of the galaxies in the HyperLeda database returned 2331 galaxies. For this selection, radio HI velocities as a query for distance were used to select galaxies closer than 40 Mpc. However, the use of radio-derived HI velocities biased the original sample towards gas-rich late-type galaxies and missed a high fraction (over 50%) of ETGs, which typically have lower gas content.

Sheth et al. (2013) used HyperLeda values based on optical spectroscopy to determine the distance to galaxies to search for ETGs ($T_{\text{HL}} \leq 0$) without HI velocities. This introduced 465 new ETGs (W2022), lacking radio HI measurements in the HyperLeda database, which made them miss the criteria of the original S⁴G but meeting the volume-limited criteria of 40 Mpc using optical-derived distances. Nevertheless, the specific criterion to search for ETGs ($T_{\text{HL}} \leq 0$) introduced a new bias against late-type galaxies.

Later on, when a study of the Fornax cluster with S⁴G authors involved was being carried out (Venhola et al. 2017), they found out that many disc galaxies were not in the S⁴G despite meeting all requisites. This finding made the authors revise the original sample and select disc galaxies using optical velocities, giving rise to the DG extension. As explained in Sect. 3, the DG extension mimics the criterion of the ETG extension but searches for late-type galaxies ($T_{\text{HL}} > 0$). Intriguingly, several galaxies with $T_{\text{HL}} \leq 0$ were not included in the early-type extension but are now included in the disc extension. We are not sure why these galaxies were missed in the original or ETG sample, but one possible explanation is that new measurements could have been added in HyperLeda for these galaxies between the time each of the samples was done (2013 and 2018⁸). Another explanation, at least for galaxies with large apparent size (such as NGC 1316), is that the original S⁴G excluded them on purpose not to expend many pointings of *Spitzer* on these galaxies. Now, with this publication, we fulfil the original aim of constructing a

⁸ Between 2018 and 2024, there have been no significant changes in the parameters of CS⁴G galaxies in HyperLeda

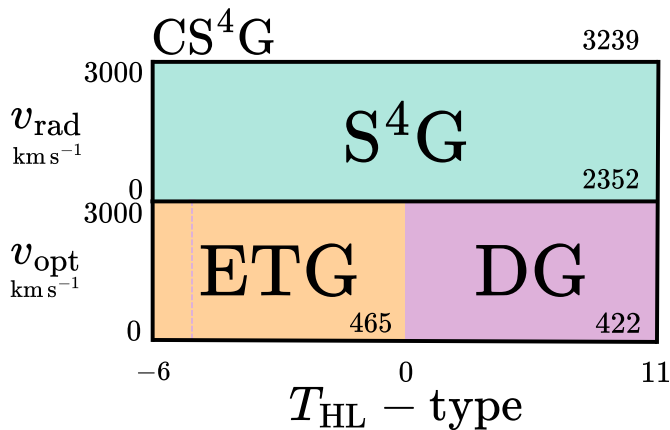


Fig. A.1: Illustration of the CS⁴G sample. The CS⁴G contains the S⁴G, ETG and DG sample. It mimics the criteria used to define the S⁴G, but including also optical velocities to define the ETG and DG samples.

volume-, magnitude-, and size-limited survey by adding 422 disc galaxies to the sample.

We present the CS⁴G, a combination of the S⁴G, ETG and DG sub-samples, with a final total size of 3239 galaxies in the local Universe. Figure A.1 shows an illustration of the CS⁴G and the differences between the ETG and DG samples. The x -axis represents the morphological T -types, while the y -axis represents the recessional velocity of the galaxies. The ETG and DG samples differ from the S⁴G by the wavelength used to measure the recessional velocity, optical and radio HI velocities, respectively. They also differ by the morphological types, while ETG includes $T_{\text{HL}} \leq 0$, the DG mostly includes $T_{\text{HL}} > 0$ with 33 $T_{\text{HL}} \leq 0$ galaxies that missed the ETG (represented with the purple dashed line).

Appendix B: Galaxies not forming part of the DG sample

The exact query used in HyperLeda, mimicking the selection criteria of the CS⁴G is the following:

```
SELECT pgc, objname, al2000, de2000, logd25,
b2, t, bt, btc, vmaxg, vmaxs, v, vrot, modz, mod0,
modbest, mabs, vrad, vopt, objtype
```

```
WHERE (objtype='G' OR objtype='?')
AND btc<15.5 AND logd25 >1.
AND logd25<2.48 AND abs(b2)>30.
AND (vrad <3000 OR vopt <3000)
```

This query executed on October 2024⁹ resulted in 430 new galaxies not included in the S⁴G and ETG. However, of these galaxies, eight were rejected because the galaxy appeared to be misclassified, it was a stellar stream related to another galaxy, a system of interacting galaxies, or it was a background galaxy. The galaxies rejected were: PGC 015573, PGC 166170, PGC 3097827, PGC 014117, PGC 014118, PGC 014121, PGC 068481, and ESO 056-019.

PGC 015573 is misclassified as a galaxy. We deduce it must be a star system embedded in a dust cloud that produces an extended emission which could be interpreted as a galaxy in shallower data. PGC 166170, also known as KK 208, is a stellar stream of M83 with an estimated stellar mass of $\sim 1 \times 10^8 M_{\odot}$ (Barnes et al. 2014). The galaxy PGC 3097827, also known as F8D1, is an Ultra Diffuse Galaxy (UDG) associated with M81 (Karachentsev et al. 2000) with a surface brightness peak of 25.81 mag arcsec⁻² and an integrated magnitude fainter than 15.5 mag. The galaxies PGC 014117, PGC 014118, and PGC 014121 appear to be systems of interacting galaxies, with individual galaxies having diameters below the limit. PGC 068481 is a bright knot in the galaxy ESO 602-003, which is included in the S⁴G. Finally, ESO 056-019 is a region of the LMC. Rejecting these 8 entries results in a total sample of 422 galaxies in the DG extension.

Appendix C: Notes on the analysis of UGC 07636 and PGC 3097691

The galaxy UGC 07636 is a companion of M49 embedded in the galaxy's outskirts. The galaxy M49 is in the original S⁴G, and its morphological decomposition was done by Salo et al. (2015). We have used the residual S⁴G image, after subtracting

⁹ This is the date of the last execution

the model fitted by Salo et al. (2015), to measure the radial profile of UGC 07636 properly. The dwarf galaxy PGC 3097691 is also included in the sample. At a distance of 0.8 Mpc the Legacy Surveys imaging can resolve the stars, and despite being analysed, the surface brightness technique is not optimal for studying the structure of the object.

Appendix D: Distance selection

When selecting the distance to the galaxies to transform apparent to absolute magnitudes, we aim to be consistent with the original S⁴G, just as we have been with the sample analysis. In the S⁴G, while the sample selection was based on HyperLeda velocities, the distances were selected from the NED database. A redshift-independent measurement was used in the original S⁴G, and if that was unavailable, a distance derived from the redshift was used assuming $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$. We used the same methodology. However, there are a few cases where the redshift-independent distance is incompatible with its redshift and others where the NED distances are incompatible with the HyperLeda optical velocities. Figure D.1 shows the distances obtained from the HyperLeda radial velocities in comparison with the NED distances.

The NED distances are selected as explained above and are a combination of redshift-independent measurements and redshift-derived distances. We also show the lines representing the volume limit criterion of the S⁴G of 40 Mpc. These lines divide the figure into four quadrants. The first quadrant, in the bottom left, shows galaxies with distances from the two different sources (NED and HyperLeda) within the criteria of the S⁴G (below 40 Mpc). The third quadrant, in the upper right, as expected, is empty since it represents galaxies not meeting the criterion in either of the sources. The second and fourth quadrants show galaxies meeting the criterion in one of the

sources but not in the other. Since the sample selection criterion of the S⁴G was only based on the HyperLeda database, we would expect not to see any galaxies in the fourth quadrant. In spite of this, we find two galaxies (ESO 288-049 and ESO 603-006) with optical velocities corresponding to distances larger than 40 Mpc in the HyperLeda database. These two galaxies do have HI radio velocity measurements below 3000 km s^{-1} . Moreover, there are 25 galaxies with NED distances over the 40 Mpc criterion, but with optical velocities meeting the selection criterion (seen in the second quadrant). For the particular case of the galaxy PGC 040869, we used the redshift-independent distance from Cantiello et al. (2024). NED and HyperLeda both register a redshift estimated distance of 0.43 Mpc. However, the galaxy appears to be one of the many dwarfs around the Virgo Cluster. Cantiello et al. (2024) obtained a distance estimation of 13.86 Mpc using surface brightness fluctuations, which links the galaxy to the Virgo Cluster as expected.

In summary, our criterion for assigning distances is to ensure they meet the requirement of being less than 40 Mpc, regardless of the source. We prioritised our selections in the following order: first, we used NED redshift-independent distances. Second were NED redshift distances, and third were optical velocities from HyperLeda. If a measurement from a higher-priority source does not meet the $d < 40 \text{ Mpc}$ criterion, we move to the next priority source.

Appendix E: CVRHS classifications

The morphological classification and their numerical values (described in Sect. 6.7) can be found in the electronic table (see Sect. 8) in the columns CVRHS and CVRHS_τ, respectively. Table E.1 shows the transformation from the morphological classification to numerical values. The description of all symbols can be found in Buta et al. (2015). The symbol K stands for galaxies showing a knotted outer ring.

Table E.1: Numerical T -type transformation.

Stage	Numerical index (T -type)
cE	-6
E	-5
E ⁺	-4
S0 ⁻	-3
S0	-2
S0 ⁺	-1
S0/a	0
Sa	1
Sab	2
Sb	3
Sbc	4
Sc	5
Scd	6
Sd	7
Sdm	8
Sm	9
Im	10
dE, dS0, dSph	11

References. de Vaucouleurs (1977); Buta et al. (2015); W2022

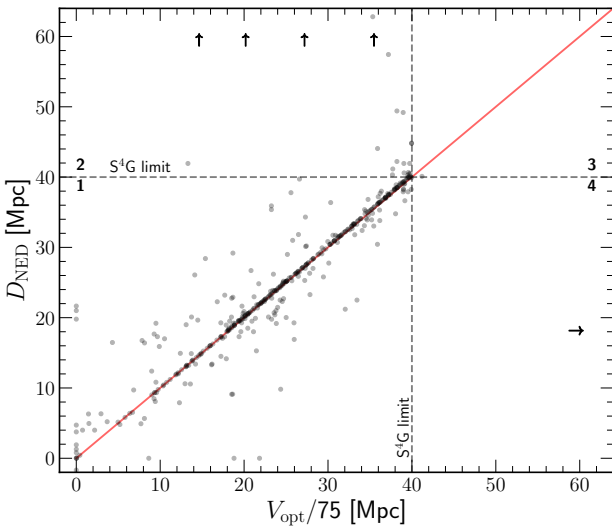


Fig. D.1: Comparison of distances. The x -axis shows the distances obtained using the optical velocities from the HyperLeda database. The y -axis shows the distances obtained from the NED database, a combination of redshift-independent measurements and redshift distances. The red line shows the one-to-one relation. The arrow shows galaxies with distances above the limits of the figure but located at the position of the other axis values. The dashed grey lines show the volume limit criterion of the S⁴G of 40 Mpc.