MASARYKOVA
UNIVERZITA

PŘÍRODOVĚDECKÁ FAKULTA

Zkoumání obsahu plynu za skupinou galaxií ve Lvu

Diplomová práce

VLASTIMIL KAPUSTA

Brno, jaro 2026

MASARYKOVA
UNIVERZITA

PŘÍRODOVĚDECKÁ FAKULTA

Zkoumání obsahu plynu za skupinou galaxií ve Lvu

Diplomová práce

VLASTIMIL KAPUSTA

Vedoucí práce: Rhys Taylor, Ph.D.

Ústav teoretické fyziky a astrofyziky

Brno, jaro 2026

MUNI
PŘÍRODOVĚDECKÁ
FAKULTA

Bibliografický záznam

Autor:	Vlastimil Kapusta Přírodovědecká fakulta Masarykova univerzita Ústav teoretické fyziky a astrofyziky

Název práce:	Zkoumání obsahu plynu za skupinou galaxií ve Lvu
Studijní program:	Fyzika
Obor:	Astrofyzika
Vedoucí práce:	Rhys Taylor, Ph.D.
Akademický rok:	2025/2026
Počet stran:	14 + 100
Klíčová slova:	galaxie, neutrální vodík, HI, ultra diffuse galaxie, blind volume, Arecibo, AGES, Leo

Bibliographic record

Author:	Vlastimil Kapusta Faculty of Science Masaryk University Department of Theoretical Physics and Astrophysics

Title of Thesis:	Exploring the gas content beyond the Leo group
Degree Programme:	Physics
Field of Study:	Astrophysics
Supervisor:	Rhys Taylor, Ph.D.
Academic Year:	2025/2026
Number of Pages:	14 + 100
Keywords:	galaxies, neutral hydrogen, HI, ultra diffuse galaxies, blind volume, Arecibo, AGES, Leo

Abstrakt

V této práci se zaměřujeme na zpracování a analýzu dat z extragalaktických rádiových pozorování neutrálního vodíku. Využíváme data z jedné z nejcitlivějších přehlídek HI, která byla provedena Arecibem před jeho zřícením, a to v rámci projektu Arecibo Galaxy Environment Survey (AGES). Zaměřujeme se na pozadí pozorované oblasti, které je bez jakýchkoli předběžných výběrových kritérií. Prohledali jsme oblast v souhvězdí Lva o rozloze 5×4 stupňů, ke katalogizaci a analýze zdrojů HI jsme použili pokročilé vizualizační nástroje. Zahrnuli jsme také optická data z SDSS a dalších katalogů. Poté jsme detekované galaxie zasadili do jejich environmentálního kontextu a analyzovali získaný vzorek jako celek pomocí scaling relations grafů, které většinou odhalily nečekaně rovnoměrné rozložení těchto galaxií. S vysokou mírou jistoty jsme identifikovali 6 ultra diffuse galaxií a dalších 11 dobrých kandidátů. Nakonec jsme analyzovali jeden takový objekt, který je také přítomen v katalogu SMUDGes. Zjistili jsme, že u něj existuje významný posun mezi HI a hvězdnou složkou, ačkoli není jasná příčina tohoto narušení. Kromě toho objekt vykazuje náznak deficitu temné hmoty, ale k potvrzení tohoto zjištění jsou zapotřebí data s vyšším rozlišením.

Abstract

In this thesis, we focus on data processing and analysis of extragalactic radio observation of neutral hydrogen. We use data from one of the most sensitive HI surveys undertaken at Arecibo prior to its collapse, the Arecibo Galaxy Environment Survey (AGES). We focus on the so-called background volume with no preselection criteria. We have searched the 5×4 degree Leo field, using advanced visualization tools to catalogue and analyze the HI sources. We have also included optical data from the SDSS and other catalogues. Then we placed the detected galaxies in their environmental context and analyzed the sample as a whole using scaling relations, which mostly revealed an unexpectedly uniform distribution of those galaxies. We have identified 6 ultra diffuse galaxies with high confidence and an additional 11 good candidates. Finally, we have analyzed one such object also present in the SMUDGes catalog. We find that this has a significant offset between the HI and the stellar component, though there is no clear cause of the disturbance. In addition, the object shows a hint of a deficit in dark matter, but higher resolution data is needed to confirm this.

MUNI
SCI

MASARYKOVA UNIVERZITA
PŘÍRODOVĚDECKÁ FAKULTA
KOTLÁŘSKÁ 2, 611 37 BRNO
IČ: 00216224
DIČ: CZ00216224

ZADÁNÍ DIPLOMOVÉ PRÁCE

Akademický rok: 2025/2026

Ústav:	Ústav teoretické fyziky a astrofyziky
Student:	Ing. Bc. Vlastimil Kapusta
Program:	Fyzika
Specializace:	Astrofyzika

Ředitel *ústavu* PřF MU Vám ve smyslu Studijního a zkušebního řádu MU určuje diplomovou práci s názvem:

Název práce:	Zkoumání obsahu plynu za skupinou galaxií ve Lvu
Název práce anglicky:	Exploring the gas content beyond the Leo group
Jazyk práce:	angličtina

Oficiální zadání:
This project uses data from the Arecibo Galaxy Environment Survey (AGES), one of the most sensitive HI surveys undertaken at Arecibo prior to its collapse. HI, neutral atomic hydrogen, provides important information both on how gas undergoes star formation and how the environment influences galaxy evolution. The survey targeted specific regions to sample the full range of galaxy environments, from isolated galaxies to rich clusters, where galaxies experience different environmental influences. However, given its large bandwidth, the bulk of our detections are actually in the background volume (out to $z < 0.06$, $d = 280$ Mpc), which form the blind "AGES Volume" with no preselection criteria. Individual such fields have contained > 300 detections, including low surface brightness galaxies, galaxies with tails, and early type galaxies rich in HI despite a lack of ongoing star formation. This project will search the 5x4 degree Leo field, using advanced visualisation tools to catalogue and analyse the HI sources (measuring their key parameters of total flux, redshift, velocity width, and presence of asymmetric features indicative of interactions). Comparison samples of non-detections will be obtained from NED, the SDSS, and other surveys, placing the HI detections in their essential environmental context. In particular, other fields – using results from independent optical catalogues – have suggested the presence of significant numbers of Ultra Diffuse Galaxies within AGES data, of which few are currently known with HI. Using newly-developed scripts to identify such galaxies, this project will provide a pilot study to estimate their detection frequency in our HI samples.

Vedoucí práce:	Rhys Taylor, Ph.D.
Datum zadání práce:	21. 12. 2023
V Brně dne:	8. 1. 2026

Zadání bylo schváleno prostřednictvím IS MU.

Ing. Bc. Vlastimil Kapusta, 6. 1. 2026
Rhys Taylor, Ph.D., 6. 1. 2026
Mgr. Dušan Hemzal, Ph.D., 7. 1. 2026

Poděkování

Chtěl bych poděkovat především vedoucímu mé práce, kterým je Rhys Taylor, Ph.D. za ochotu, trpělivost a cenné rady při zpracování této práce.

Prohlášení

Prohlašuji, že jsem svoji diplomovou práci vypracoval samostatně pod vedením vedoucího práce s využitím informačních zdrojů, které jsou v práci citovány.

Vlastimil Kapusta

Contents

Introduction

Since the remarkable detection of astronomical radio emission by Karl Jansky in 1932, which opened a completely new window to the universe, radio astronomy has come a long way and dramatically extended our knowledge about the universe.

Similarly, the detection of the 21 cm line published in 1951 revolutionized studies of the interstellar medium. We were finally able to see the simplest but most abundant atom in space.

This is where the focus of this thesis lies: in neutral hydrogen observation of a blind volume of the nearby extragalactic universe. A part of the sky with no pre-selection criteria, with little to no idea of what to expect. For that, we used an untouched observation of the marvelous but decommissioned Arecibo telescope.

Chapter 1 of this thesis describes the electromagnetic spectrum, then focuses on radio astronomy, in particular, the origin and observation of neutral hydrogen emission. It then describes the Arecibo observatory and the AGES survey.

Since we are dealing with galaxies with unknown history and state of evolution, it is useful to understand galaxy composition, morphology, and processes that can influence their state. Chapter 2 provides an overview.

In chapter 3, we describe various external data sources and tools used for our work.

Chapter 4 is focused on the studied region, our data, their reduction, visualization, and the associated challenges.

In chapter 5 we describe the cataloguing process, which was a large part of the work. We go through the steps of extracting information from our data cube and external databases as well.

Chapter 6 is focused on data analysis. We explore the environment, discuss scaling relations, identify ultra diffuse galaxies in our sample, and analyze one interesting object.

Chapter 7 is a conclusion.

1 Radio Astronomy

Before the development of radio astronomy in the twentieth century, astronomers were limited to a narrow range of visible wavelengths. As described in Marr et al. (2015), the first successful detection of astronomical radio emission by Karl Jansky in 1932 opened a completely new window to the universe. Radio waves are electromagnetic (EM) radiation, just like visible light, but with longer wavelengths λ (lower frequencies ν). For context, figure 1.1 shows the whole EM spectrum with corresponding wavelengths, frequencies, and energies.

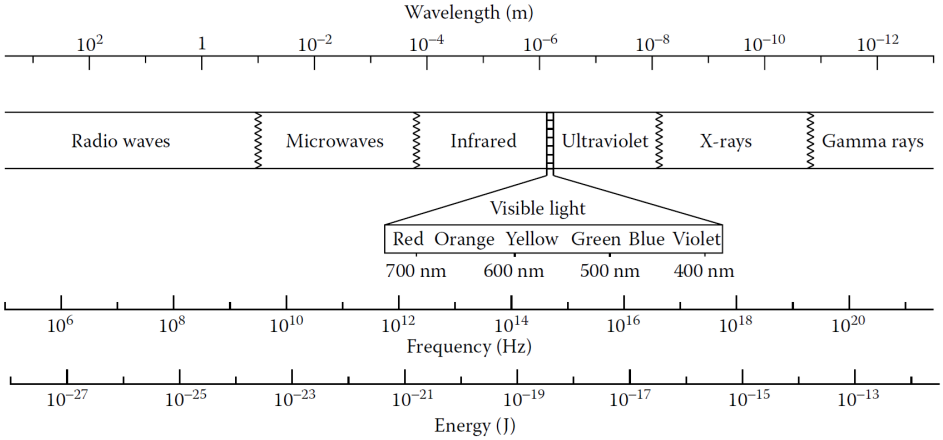


Figure 1.1: Display of all the bands of the entire electromagnetic spectrum, shown in order of energy of the waves, with the lowest energy radiation on the left. Taken from Marr et al. (2015).

When looked through the eyes of a radio telescope, the sky looks very differently than we are used to from optical images. Stars and other objects bright in the visible part of the spectrum are not dominant in radio emission. At these longer wavelengths, different objects and phenomena can be much brighter. For example, the interstellar medium of our Galaxy or other galaxies, the intracluster medium (ICM), the relic radiation from the early universe (the cosmic microwave background), and the synchrotron emission, which is non-thermal and can be produced by sources such as supernova remnants

and quasars, all appear radically different at radio wavelengths to optical.

Since Earth's atmosphere is not uniformly transparent to all wavelengths, and high frequency observations (e.g. UV, X-rays) require space-based facilities. For radio wavelengths, the atmosphere has little effect, except at the very shortest (mm) and longest (a few metres) parts of the band. This means radio telescopes are almost entirely ground-based.

At a basic level, most radio telescopes gather and focus light in a similar way to optical telescopes. The parts of a radio telescope are: a usually parabolic reflector that focuses the waves to a point, where a receiver converts them to a digital signal for processing. There can sometimes be an intermediate stage, where the focused signal travels along a waveguide before digital conversion later along the signal chain. The last part is a computer and a storage device for further analysis.

1.1 Single-dish radio astronomy

As discussed in Lang (2013), since radio waves can be millions of times longer than those of visible light, a radio telescope needs to be typically thousands, and in extreme cases millions, of times larger than an optical telescope to obtain the same resolving power. The limitations of this are determined by engineering constraints.

One way of making high-resolution radio images is baseline interferometry, where signals from multiple smaller telescopes are combined. Examples of such interferometers are the Very Large Array (VLA) or Atacama Large Millimeter/submillimeter Array (ALMA).

Single dishes are still widely used as they can be much faster for surveys and have greater sensitivity (this is the penalty for interferometry). For example, one of the largest fully steerable dishes is the 100m diameter telescope in Effelsberg, West Germany, and the largest one is the 100-by-110-meter Green Bank Telescope in West Virginia, US. One way of constructing a larger reflector is to cover the floor of a valley. This approach was used in the case of Arecibo.

1.2 HI emission

In this section, we provide a brief explanation of the physical mechanism behind the HI emission, which is used in this thesis. More about the connection with galaxies can be found in chapter 2, specifically section 2.3.

Intense sources like young, massive stars, quasars, and other high-energy phenomena can ionize nearby hydrogen in the universe. When we want to know what is in the vast dark places, far away from those high-energy sources, we have to rely on a different mechanism. We know that stars are forming mostly from hydrogen, so there should be plenty of it in the interstellar space. Although typical temperatures of neutral hydrogen are of order 10^4 K, the gas density is extremely low, so neutral hydrogen produces essentially no optical emission.

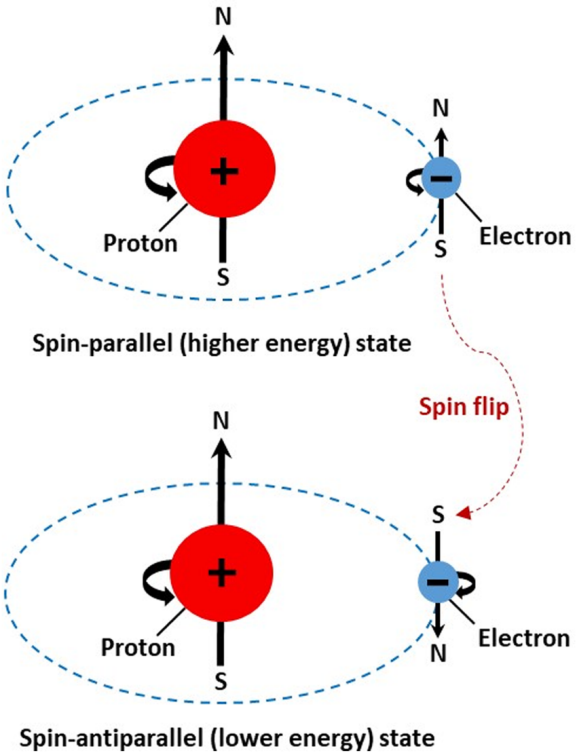


Figure 1.2: Formation of the 21-cm hyperfine transition line for neutral hydrogen. Taken from Allawi et al. (2025).

As explained in Lang (2013), the lowest orbital energy state of atomic hydrogen has a hyperfine splitting arising from the spins of the proton and electron changing from a parallel to an anti-parallel configuration, see figure 1.2. When this transition occurs, the energy equal to the energy difference between these two levels ($\approx 6 \times 10^{-6}$ eV) is released in the form of electromagnetic radiation. This photon has a frequency $\nu = 1.420 \times 10^9$ Hz or wavelength $\lambda = 21.1$ cm which is not in the visible part of the spectrum, but it is a radio wave.

One of the main difficulties in detecting this emission is that the spin-flip does not occur very frequently. The transition is highly forbidden with an extremely small probability $A_{10} = 2.8689 \times 10^{-15}\,\mathrm{s}^{-1}$.

Even though the density of the neutral hydrogen clouds is very low, collisions between atoms can be frequent enough to greatly increase the rate of emission of 21 cm photons by changing their spin orientation from one possible direction to the other one, either leaving them in an unstable configuration or "bumping" them down, causing emission which is much more frequent than their spontaneous return to the lower energy configuration which is not swift at all because the radiative half-life of this hyperfine transition is $\tau_{1/2} = 1/A_{10} \approx 11.1$ million years. An additional factor in our ability to detect the line is simply the fact that there are a very large number of atoms in a typical HI cloud, thus ensuring enough photons are emitted to be detectable.

The detection of the 21 cm line published in the Nature journal (Ewen and Purcell, 1951, Muller and Oort, 1951) revolutionized studies of the interstellar medium. We were finally able to see the simplest but most abundant atom in space.

The observations of the HI line initially relied on pointed observations of individual targets or small samples of galaxies. A major advance happened in the 1990s with the development of multibeam receivers, which enabled efficient, blind HI surveys over large areas of sky. At Arecibo, this culminated in the Arecibo L-band Feed Array (ALFA) and associated surveys such as ALFALFA as described in Giovanelli et al. (2005) and Haynes et al. (2011)

Most recently, the Five-hundred-meter Aperture Spherical Radio Telescope (FAST), equipped with a multibeam receiver, continues this trend through surveys such as FASHI, further improving sensitivity to low-mass and low-column-density HI systems as explained by Nan et al. (2011).

1.3 Arecibo

As described by the official site[1], the Arecibo Telescope in Puerto Rico was a 305-meter wide dish, capable of receiving and emitting signals in the radio part of the electromagnetic spectrum. It studied a wide range of objects from nearby planets, comets, and asteroids to distant pulsars, galaxies, and quasars. Since its completion in 1963, it has been the largest single-aperture telescope in the world for 53 years until 2016, when the Five-hundred-meter Aperture Spherical Telescope (FAST) in China was built.

Figure 1.3: The Arecibo observatory in its full glory.

Its reflective surface was made of almost 40 000 perforated aluminium panels. The shape of the reflector was spherical, not parabolic. 137 meters above the reflector, there was a 900-ton platform of scientific instruments suspended in mid-air on eighteen cables, which were strung from three reinforced concrete towers. The position of the platform could have been adjusted with another six cables with millimeter precision. Inside, there was a set of additional reflectors that focused the incoming radiation to a point and fed it to very sensitive and highly complex radio receivers. To minimize electron noise and to be able to

1. https://www.naic.edu/ao/telescope-description

amplify only the incoming radio signals, which are very weak, these receivers were cooled by liquid helium. The observatory was sensitive to frequencies between 50 MHz up to 10 GHz. That corresponds to a wavelength range from 3 cm to 6 m.

One of the major discoveries made with Arecibo was the discovery of the first ever binary pulsar in 1974. The observed changes in periastron confirmed the predictions of general relativity. It was awarded the 1993 Nobel Prize in Physics. Arecibo also discovered the first millisecond pulsar, suggesting that pulsars can spin up by accreting mass from a companion. In 1992, the first ever exoplanet was discovered by this radio telescope. Interestingly enough, this planet is orbiting a pulsar. In 2016, the first-ever repeating fast radio burst was discovered. Arecibo observations have also significantly constrained the possible time variability of the fine structure constant.

Arecibo played a crucial role in advancing extragalactic HI studies due to its exceptional sensitivity and angular resolution. This is particularly evident when contrasted with HIPASS, the first all-sky HI survey conducted with the Parkes 64 m telescope, which had significantly poorer resolution and sensitivity. As a result, HIPASS catalogued only about 5000 galaxies, whereas even a relatively shallow Arecibo survey such as ALFALFA detected over 30 000 HI sources.

As summarized in Astronomy Magazine (2024), the demise of this marvelous piece of engineering started in 2017 when it was damaged by Hurricane Maria. In 2020, the first cables supporting the platform hanging above the main telescope dish broke and damaged the surface. After that, it was decided to close the site permanently because of financial and safety reasons. Later that year, more cables from one tower broke, causing the instrument platform to crash through the dish.

Figure 1.4: Remains of the Arecibo observatory.

1.4 AGES

The Arecibo Galaxy Environment Survey (AGES) is an HI survey designed to investigate various galactic environments to higher sensitivity than previous neutral hydrogen surveys. The survey itself and its goals are described in detail in Auld et al. (2006).

The installation of the Arecibo L-Band Feed Array (ALFA), which operates between 1.225 and 1.525 GHz made AGES and the other surveys possible. The back-end signal processors are digital Wide-band Arecibo Pulsar Processors (WAPPs). These have been upgraded for spectral line observations and are configured to cover 100 MHz bandwidth with 4096 channels. That means, each channel spans 24.4 kHz $\equiv$ 5.15 km s^{-1} at the rest frequency of HI. At 1.4 GHz, the mean half power beam width is 3.4′, and the mean system temperature is 30 K.

Compared to ALFALFA, which has an effective integration time of $\sim 40 - 48$ seconds per beam, prioritizing survey speed rather than depth, AGES uses a much longer effective integration time per pointing — about ~ 300 seconds (5 minutes) per beam. This deeper integration is what gives AGES its much lower noise and higher sensitivity.

For a spectral line with a velocity width W, the minimum detectable velocity-integrated flux was approximated as $(\int S(v)\,dv)_{\rm lim} \approx$

$N_\sigma \sigma \sqrt{W/\delta v}\, W$, where $N_\sigma = 5$ is the adopted detection threshold, $\sigma \sim 0.7\,\mathrm{mJy}$ is the rms noise per $10\,\mathrm{km\,s^{-1}}$ channel, δv is the velocity width of an individual spectral channel ($10\,\mathrm{km\,s^{-1}}$), and $W \sim 40\,\mathrm{km\,s^{-1}}$ is the assumed line width. At the distance of the Virgo Cluster ($D = 16\,\mathrm{Mpc}$), this corresponds to a limiting HI mass of $M_{\mathrm{HI}} \sim 5 \times 10^6\,\mathrm{M_\odot}$ according to Giovanelli et al. (2005) and Haynes et al. (2011).

The fact that the Arecibo telescope is fixed poses a limit on observable declinations ($-1^\circ \lesssim \delta \lesssim +38^\circ$) and also on time on source, which is around 2 hours. For AGES, most selected areas comprise a $5^\circ \times 4^\circ$ field with an integration time of 300 s per point in two polarizations.

Environments possible to observe range from apparent voids to galaxy-rich regions associated with galaxy clusters and filamentary structures.

2 Galaxies

A lot of time has passed since the definitive proof by Hubble (1925) that there is more to the universe than the Milky Way and that those faint patches in the sky, then called spiral nebulae, are in fact other small "universes", nowadays called galaxies. Our knowledge in extra-galactic astronomy and cosmology has vastly increased since then, as pointed out in Schneider (2006), mainly thanks to our technical advances. Contemporary large telescopes and sensitive instruments offer a window to the distant universe and, because of the finite speed of light, also a window to the past.

As described in Karttunen et al. (2003), galaxies are gravitationally bound systems composed primarily of stars of various ages, interstellar gas in different phases (neutral, ionized, and molecular), and dust. In addition to these baryonic constituents, current understanding is that galaxies are dominated by dark matter, whose presence is inferred indirectly from dynamical measurements. In particular, spectroscopic observations of the gas reveal rotation velocities that are too high to be explained by the visible matter alone. This implies the existence of a massive, extended dark matter halo. The dominance of dark matter is a key property that distinguishes galaxies from other astrophysical systems.

For a better understanding of galaxies, it is useful to classify them based on their form. There is an important caveat to keep in mind when discussing galaxy morphology: our understanding will always be limited to the galaxies we can detect. Small, faint galaxies are easily missed, but even large, relatively bright galaxies can be missed. If their flux is sufficiently spread out, their signal can be close to the noise level, thus making them difficult to detect. Conversely, if a galaxy is too compact, it can be mistaken for a compact object such as a star. Although our detection limits continue to improve, these changing selection effects must always be taken into account.

2.1 Galaxies are moving away

As written in Mo et al. (2010), the fact that almost all galaxies appear to move away from us, and that their recession velocities increase

in direct proportion to their distances from us, is one of the most important discoveries in modern science. The relation is called the Hubble law, and it allows distance measurement from redshift. The relation is:

$$v_{\mathrm{r}} = H_0 d, \tag{2.1}$$

where v_{r} is the recession velocity, H_0 is the Hubble constant and d is the distance. The recession velocity of a galaxy can be measured from its redshift using the formula $v_{\mathrm{r}} = cz$, where c is the speed of light.

However, the object's observed velocity might be partly caused by its peculiar velocity, v_{pec}. Then the observed velocity is the sum of v_{pec} along the line of sight and the velocity caused by the Hubble expansion ($H_0 d$):

$$v_{\mathrm{r}} = H_0 d + v_{\mathrm{pec}}. \tag{2.2}$$

In a case where $v_{\mathrm{pec}} \ll H_0 d$, the redshift is an accurate distance measurement, otherwise it isn't.

Typical value of peculiar velocity v_{pec} for galaxies is a few hundred kilometers per second, meaning redshifts can be used as a reliable distance approximation for $cz \gg 1000\,km/s$. For rich clusters, this value would be larger. The typical line-of-sight velocity dispersion σ_{los} of cluster member galaxies is of the order of $1000\,km/s$.

Because our sample does satisfy these requirements, we assume pure Hubble flow, meaning we use the redshift as an accurate distance measurement. In our calculations, we use the value of the Hubble constant $H_0 = 71\,\mathrm{km/s/Mpc}$.

2.2 Morphology

As Keel (2023) explains, galaxy morphology is commonly described using the Hubble sequence, which, in its basic form, distinguishes ellipticals, spirals with and without bars, and irregulars. These Hubble types are commonly arranged into the Hubble tuning fork.

There is a historical terminology often used today as well: ellipticals and lenticulars are called early galaxies, and spirals and irregulars

are referred to as late-type galaxies. This is a remnant of a belief that galaxies evolve along the Hubble sequence. Today we know this is not the case, and Hubble himself was careful with that idea: "The nomenclature, it is emphasized, refers to position in the sequence, and temporal connotations are made at one's peril. The entire classification is purely empirical and without prejudice to theories of evolution..." (Hubble, 1927).

Classification of ellipticals is quite straightforward because of their lack of structure. There are eight types recognized (E0-E7). The number represents a projected axial ratio, so E0 appears to be circular. As the number increases, the ellipticals become more elongated. It is difficult to determine their true 3-D shape because we only see the projection.

Spiral galaxies are stellar systems where most of the stars are situated in a relatively thin disk (usually $\sim 1000\times$ wider than thicker). They are divided into ordinary (S) and barred (SB) families. They differ in the size of their bulge and how tightly wound the arms are. Our Galaxy is a representative of such a system; its disk can be seen on the night sky as a dense band of stars.

There have been many modifications applied to the basic Hubble tuning fork. One version by Alan Sandage is the recognition of transitional lenticular (S0) galaxies. Lenticulars are like spirals, just without spiral arms. Another modification was introduced by de Vaucouleurs, who highlighted the continuity of spirals into irregulars. Figure 2.1 shows the Hubble-de Vaucouleurs diagram, which is a widely used extension of the Hubble sequence. Galaxies at the end of the middle line are dwarf spheroidals (dSph). Those are characterized by low luminosity, low gas content, and spherical shape.

Note that there are many other different morphological classification schemes used, which are not discussed here, but for an overview, see Buta (2013).

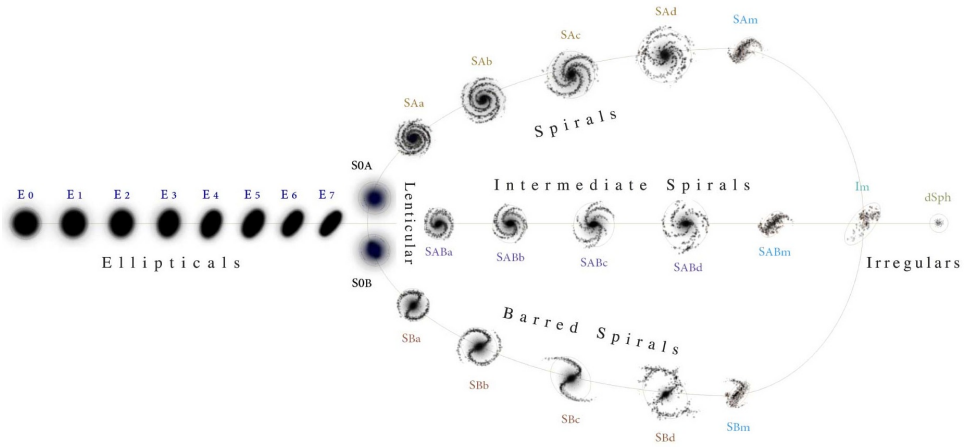


Figure 2.1: The Hubble-de Vaucouleurs diagram. Taken from Ciccolella and Leo (2016).

Beyond their morphological differences, spiral and elliptical galaxies are also distinguished by their star formation activity and stellar populations. Spiral galaxies typically contain significant amounts of cold gas in their disks, which enables ongoing star formation, particularly in their spiral arms. This results in a composite stellar population, with young, massive, blue stars coexisting with an older stellar component. Consequently, spiral galaxies generally appear bluer in optical colors. In contrast, elliptical galaxies are largely devoid of cold gas and dust, and their stellar populations are dominated by old, low-mass stars. Star formation in ellipticals is therefore minimal or absent, leading to uniformly red colors and smooth light distributions. This color difference reflects fundamentally different evolutionary histories, with spirals maintaining prolonged or episodic star formation, while ellipticals experienced most of their star formation at early cosmic times and have since evolved passively.

It is important to note that the currently used morphological distinction is based on galaxy appearance in optical wavelengths, so these classifications are based only on certain components and phenomena occurring in galaxies, as pointed out by Keel (2023). To illustrate this difference, let's look at figure 2.2, where we can see that for M81, with longer wavelengths, the spiral arms become the dominant feature of the galaxy.

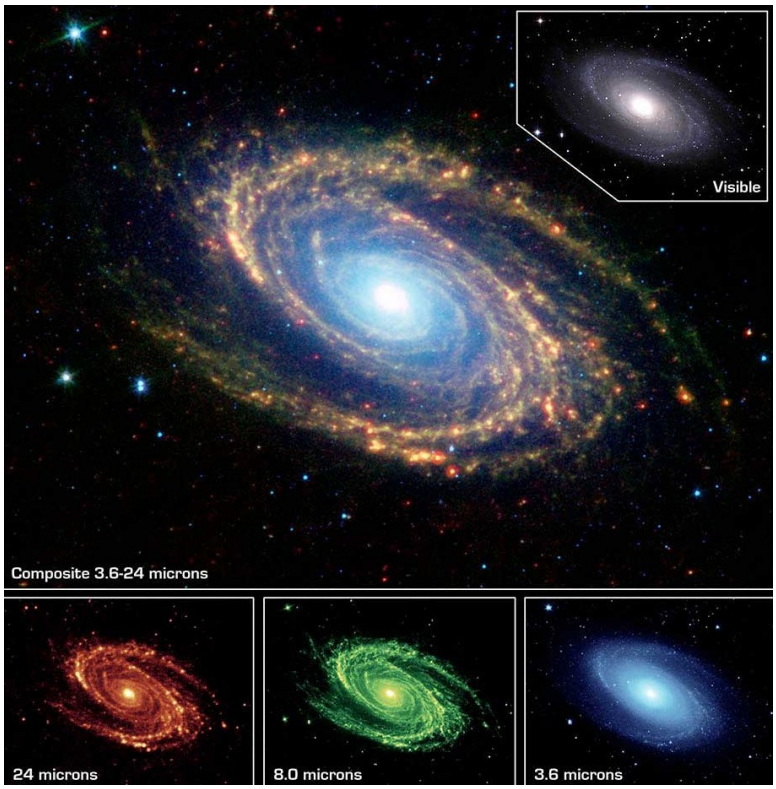


Figure 2.2: Messier Object 81 in different wavelengths. Taken from Hedberg (2025).

In figure 2.3, the same galaxy is shown in X-rays. Here, nothing resembles its spiral arms that are so prominent in the visible part of the spectrum.

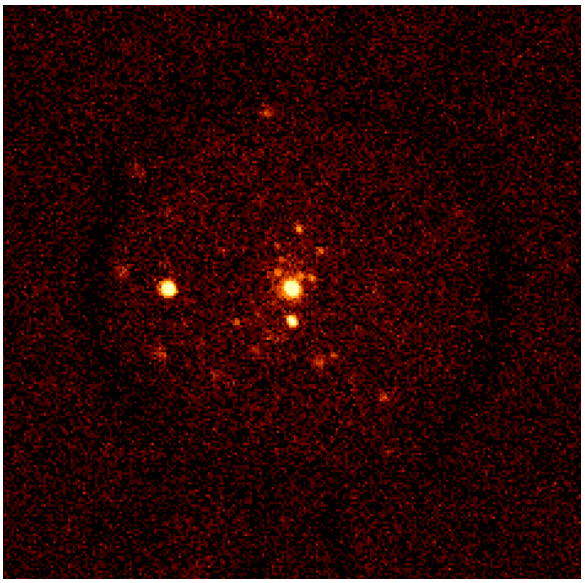


Figure 2.3: M81 in X-rays. Taken from Keel (2023).

2.3 Gas in galaxies

As explained by Keel (2023), the gas content of a galaxy plays a crucial role in its evolution. Within galaxies, gas provides the fuel for star formation, but stars also return gas to the interstellar medium through winds and supernovae. Gas can be affected by both internal and external processes, with the latter being discussed in section 2.6. An excellent tracer of overall gas content in a galaxy is the 21-cm HI line, discussed further in 1.2. In many systems, molecular gas is also important, but because observation of molecular hydrogen (H_2) is severely limited, the most useful tracer of molecular gas has been, for a long time, the molecule CO. Much of a galaxy's molecular gas resides in giant molecular clouds. These clouds are responsible for star formation, so there is a strong link between molecular mass and star-formation indicators.

HI makes up a large fraction of the neutral gas in galaxies and often extends well beyond the stellar disk. Since HI is often found in extended disks, it provides a powerful tracer of disk kinematics and rotation curves, dark matter distribution, warps, tails, and interactions with companions or environment. This is why extended HI maps are widely used to study galaxy dynamics and environment effects.

2.4 Galaxy surface-brightness profiles

Galaxy surface-brightness profiles provide a quantitative description of how stellar light is distributed within galaxies and are essential for structural and evolutionary studies. Historically, disk galaxies were described by different functions than ellipticals. The Sérsic model unifies these descriptions, allowing consistent characterization of galaxies across the full range of morphologies.

2.4.1 Sérsic surface-brightness profile

The Sérsic profile published by Sérsic (1963) is an empirical description of galaxy surface brightness distributions that is able to reproduce the light profiles of both disk and spheroidal systems. The Sérsic index n provides a quantitative measure of the central concentration of stellar light, making the profile a fundamental tool in studies of galaxy structure, morphology, and evolution.

As summarized in Graham and Driver (2005), the standard Sérsic surface-brightness profile is defined as an intensity profile, such that:

$$I(R) = I_e \exp\left\{-b_n\left[\left(\frac{R}{R_e}\right)^{1/n} - 1\right]\right\}. \tag{2.3}$$

Here $I(R)$ is the surface brightness at projected radius R, R_e is the effective radius enclosing half of the total luminosity from the model, and I_e is the intensity at the effective radius R_e. The parameter n is the Sérsic index, which describes the shape of the light profile: $n = 1$ corresponds to an exponential disk profile (99.1% of the flux resides within the inner $4R_e$) typical of spiral galaxies, while $n \approx 4$ reproduces the de Vaucouleurs profile commonly observed in massive elliptical

galaxies, here 84.7% of the flux resides within the inner $4R_e$. The constant b_n depends on n and is defined such that R_e contains half of the total integrated luminosity of the profile. An accurate approximation is $b_n \approx 2n - 1/3$ for $n \gtrsim 1$.

The magnitude form of the profile is:

$$\mu(R) = \mu_e + \frac{2.5\, b_n}{\ln 10} \left[\left(\frac{R}{R_e} \right)^{1/n} - 1 \right], \tag{2.4}$$

where $\mu(R)$ is the surface brightness in mag arcsec^{-2}.

2.5 Galaxies in different environments

In this section, we follow the description of galaxy environments and their properties given in Mo et al. (2010). Galaxies living in relative isolation are often referred to as field galaxies. They are typically gas-rich, star-forming systems, and from a morphological point of view, are therefore mainly blue spirals and irregulars. In the absence of frequent gravitational or hydrodynamical interactions, their evolution is governed primarily by internal processes such as gas accretion, star formation, and feedback.

There appears to be a strong correlation between galaxy properties and their environment. In dense environments such as galaxy groups and clusters, galaxies are surrounded by hot, diffuse intragroup or intracluster medium (IGM/ICM), which enables a variety of environmental processes to act on their gas reservoirs. In addition, the high galaxy number density increases the frequency of gravitational encounters and mergers.

A larger fraction of spirals is found in less dense environments. The fraction of spiral galaxies decreases from approximately 60% in the lowest-density regions to less than 10% in the highest-density environments. The fraction of elliptical galaxies shows the opposite behavior. This is usually interpreted as evidence that galaxies undergo morphological transformations in dense environments.

In clusters, galaxies are on average more massive, redder, more concentrated, less gas-rich, and have lower specific star-formation rates (a rate at which new stars are formed in a galaxy per unit mass) than

their field counterparts. These trends reflect the combined influence of environmental mechanisms such as ram pressure stripping, strangulation, tidal interactions, and galaxy mergers, which progressively remove gas, suppress star formation, and drive the transition from late-type to early-type morphologies. So, field galaxies offer to study intrinsic galaxy evolution, while cluster and group galaxies reflect the impact of the environment.

2.6 Interactions

Having already discussed internal processes, we will now consider external effects acting on galaxies. As described in Boselli et al. (2022), various mechanisms can act upon galaxies in different environments and may consequently alter their evolution. These mechanisms can be divided into two main families, which will be discussed in turn. First are gravitational perturbations with other galaxies or larger structures, which can affect all components of the galaxy. Second are hydrodynamic interactions, which primarily affect the gas.

2.6.1 Gravitational interactions

Gravitational interactions include tidal encounters, harassment, and mergers as explained by R. Taylor (2012). A common characteristic of all such processes is that they can affect both the stars and the gas together.

Tidal encounters

Tidal encounters can generally be divided into low and high-speed ones, where the low-speed encounters usually cause greater disturbance because of the longer duration of gravitational forces, but geometry is also an important factor. Tidal effects can cause, for example, disruption of a stellar disk or promote star formation, or in extreme cases, it can split the galaxy apart as explained in Toomre and Toomre (1972).

Harassment

Harassment describes a case of multiple, high-speed close tidal encounters. This usually happens in clusters. These short, high-speed disturbances have less impact than low-speed ones, but their cumulative effect can have a large impact on galaxy evolution. It is possible that it can cause evolution from spirals to ellipticals.

Mergers

If the relative speed of two galaxies is smaller than the escape velocity, they will merge. We distinguish between two types of mergers: major, where both galaxies are of a comparable mass, and minor, where their masses are very different. Mergers are often raised as an explanation for a wide range of phenomena, from the formation of elliptical galaxies to the formation of bulges.

2.6.2 Hydrodynamic interactions

These interactions contrast with gravitational interactions in that they affect mainly the gas. They include ram pressure stripping, accretion, strangulation, viscous stripping, and thermal evaporation as described in Nulsen (1982) and Cortese et al. (2021).

Ram pressure stripping

Ram pressure stripping (RPS) is a process that occurs because of the intracluster medium, which exerts a force on the interstellar medium and can remove it from the galaxy. This results in one-sided tails of stripped gas, hence the name jellyfish galaxy. Although this process usually doesn't affect stars, it can be a reason for morphological change because of the possibility of a star formation burst in the tail. Some of these tails or stars can be later recaptured by the galaxy.

Accretion

Accretion, on the other hand, is an infall of gas into the galaxy. The gas can be of a primordial origin, or it can be previously stripped material. Accretion may even spark new star formation in ellipticals.

Strangulation

Strangulation, also known as starvation, is an environmental quenching mechanism in groups and clusters. It occurs when the gas reservoir in the galaxy's halo, which would normally replenish the interstellar medium, is removed or prevented from cooling. Without fresh gas inflow, star formation gradually declines on gigayear timescales once the star formation consumes the remaining gas in the disk.

Viscous stripping

Viscous stripping is the removal of gas due to the difference in viscosities between the galaxy's interstellar medium and the surrounding medium. As the hot medium flows past the galaxy, turbulent and laminar viscosity can pull gas off the galaxy's gas disk even when direct ram pressure is insufficient for complete stripping.

Thermal evaporation

Thermal evaporation is caused by heating of the cold galactic gas by the hot ambient medium. This increase in temperature results in the evaporation of galactic gas into the surrounding intracluster or intragroup medium. It does not directly depend on the motion of the galaxy through the medium.

2.7 Ultra diffuse galaxies

Despite huge advances in deep, high-resolution imaging surveys, the low surface brightness objects remain unexplored to a large degree. The term ultra diffuse galaxies (UDGs) was introduced by van Dokkum et al. (2015) and Koda et al. (2015) to describe physically large galaxies with low surface brightness.

As stated in Zaritsky et al. (2023), early interpretations suggested that UDGs in clusters were massive "failed" galaxies quenched by early gas loss, although subsequent work has shown that mainly kinematics make this scenario unlikely for the majority of UDGs. Gas stripping can quench galaxies, but tidal interactions generally make them more compact rather than ultra-diffuse; therefore, tides are highly debatable

to create UDGs, even if the environment affects their evolution. UDGs found in the field imply that a dense environment is not necessary for their formation. Many UDGs are consistent with being dwarf-mass galaxies that formed in dark matter halos with unusually high spin parameters, resulting in being puffed up. The only secure conclusion right now is that there may be multiple pathways to UDG formation, but it is unknown which is the dominant mechanism.

One of the main problems when studying ultra diffuse galaxies is the lack of homogeneous UDG data across environments. In Zaritsky et al. (2023), they addressed this issue by searching images provided by the Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Surveys. The provided candidate catalogue contains 7070 objects, from which 585 sources have distance measurements so can be securely assigned as UDGs; the remaining objects could be UDGs, but their distance is unknown. They refer to their catalogue as SMUDGes (Systematically Measuring Ultra-Diffuse Galaxies). There are two quantitative criteria commonly used to classify a galaxy as an UDGs: first is minimal central surface brightness in the g-band $\mu_{0,g} \geq 24\,\mathrm{mag\,arcsec^{-2}}$ and second is minimal effective radius $r_e \geq 1.5\,\mathrm{kpc}$.

3 Data sources

In this thesis, we relied not only on the data from the Arecibo Galaxy Environment Survey (more in section 1.4) but also on other observations, surveys, and catalogues. This chapter describes the major properties of these supplementary data sets.

3.1 SDSS

An important source of optical data for this work was the Sloan Digital Sky Survey (Kollmeier et al., 2025 and York et al., 2000). This is an imaging and spectroscopic survey for millions of sources spread across the entire sky. One of the aims of this project is to build the most detailed Three-Dimensional Imagery of the Universe ever made. Beyond mapping large-scale structure, SDSS was designed to study the formation and evolution of galaxies and quasars, to measure the statistical properties of galaxies in various environments, and to constrain cosmological models. Among its many significant results, the survey has led to the discovery of new faint satellite galaxies of the Milky Way, significantly extending the known population of low-luminosity dwarf systems.

It uses a 2.5 m telescope equipped with a large-format mosaic CCD camera to image the sky in five optical bands (u, g, r, i, z) and two digital spectrographs. There is a comparison of the SDSS optical filters with Johnson-Cousins in figure 3.1. The currently available catalog offers photometry of almost 500 million unique objects from approximately one-third of the sky.

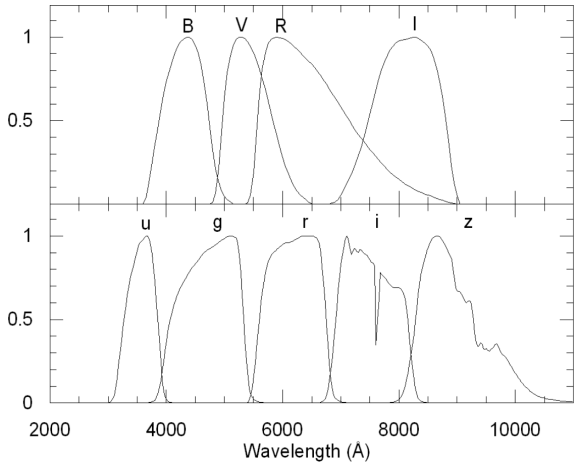


Figure 3.1: Normalized passbands of the Johnson-Cousins BVRI filters (upper panel) and the SDSS ugriz filters (lower panel), taken from Bilir et al. (2008).

We used the Data Release 18 (DR18). It is described in detail in the DR18 data release paper by Almeida et al. (2023). All following attributes were retrieved from the database table `PhotoObjAll`, which contains the full photometric catalog quantities for SDSS imaging:

- right ascension in r-band (J2000)
- declination in r-band (J2000)
- magnitudes in the u, g, r, i, z bands, they are calibrated but not Galactic extinction corrected
- extinction for every band, derived from Schlegel et al. (1998) dust maps and extinction coefficients from Schlafly and Finkbeiner (2011a)
- optical image (from cutout service)
- `expAB` which is an exponential fit b/a in the R-band, from which we can get an inclination angle i because $\cos i = b/a$, where a and b are the major and minor axes, respectively. This assumes the galaxy is a perfectly thin circular disc, as shown by Hubble (1926).

3.1.1 SDSS Visual Tools

We used the SDSS Navigate web interface in conjunction with FRELLED (see section 4.4.1) software, which allows us to open this page at desired coordinates that are selected directly in the data cube. The interface allows navigation in the vicinity of a given location, highlighting photometric objects and objects with spectra that can then be selected to get more detailed information. So it was possible to quickly double-check the optical observation for a given galaxy or galaxy candidate. That way, we could assign `SDSS ID` (unique identifier within the SDSS database) to each of our objects, effectively doing a cross-match.

Moreover, on this Navigate page, there is also an option to go to the Explore Tool, which contains a summary of an object, interactive links to further detailed properties, such as magnitudes, spectrum, redshift measurements, etc.

3.1.2 CasJobs

While the SDSS web interface is suitable for working with individual objects, CasJobs[1] was developed to run large queries (involving even millions of objects) or complex joins across multiple tables. It provides a web-based interface for running customized SQL queries on the SDSS catalog, effectively allowing the survey data to be accessed and analyzed as a relational database.

This tool was especially helpful in pairing all our HI data with optical measurements from SDSS in one operation. More on used data columns can be found in section 3.1 about SDSS.

3.2 Galaxy Zoo

Galaxy Zoo[2] is an incentive to classify a large number of galaxies morphologically. The presumption is that one can tell a lot about a galaxy just from its shape. But because there is way too much astronomical data, like images from the SDSS (see section 3.1) and not enough people to go through them, the founders of this project decided to use

1. https://casjobs.sdss.org/casjobs
2. http://www.galaxyzoo.org

volunteers. Currently, over 700 publications mention "Galaxy Zoo" in the abstract from the NASA Astrophysical Data Service[3].

Unfortunately, for our sample, only a few galaxies with morphology information were present in this database. So we were forced to look elsewhere.

3.3 HyperLeda

This database came into consideration second after Galaxy Zoo didn't help. HyperLeda[4], as written on the official website, is a database and a collection of tools to study the physics of galaxies and cosmology. It is a collection of measurements published in literature and modern surveys, which is combined and homogenized into a comprehensive description of those astronomical objects.

As described in (Paturel et al., 2003 and Makarov et al., 2014), the catalog is restricted to confirmed galaxies only, which means it contains about one million galaxies that are brighter than ≈ 18 B-mag. The database is maintained regularly, and doubtful or erroneous data are flagged.

Attributes in the HyperLeda database are described in detail in Paturel et al. (1997). We used the following fields for our galaxy sample:

- `type`: represents a morphological type, or Hubble type (more in section 2.2). Hyperleda does give more detailed classifications, but they suffer from uncertainty, so we simplify these to the basic morphological types (spiral, elliptical, irregular, or unclassified galaxy).

- `logd25`: this is the decimal logarithm of the length of the projected major axis of a galaxy at the isophotal level 25 mag/arcsec^2 in the B-band. This is a standard system for galactic diameters. This diameter definition is particularly useful for calculating HI deficiencies. Other ways of parameterizing size exist, such as effective radius, which we use for UDG classification (more in sec 2.7). If the galaxy is a disk, then this is its diameter; otherwise, it is a characterization of the outer extent of the galaxy's stellar

3. https://ui.adsabs.harvard.edu
4. http://leda.univ-lyon1.fr

light. According to the convention, the diameters are expressed in log of 0.1 arcmin. For instance a diameter of $10'$ will be given as `logd25` $= 2.00$.

3.4 SMUDGes

Systematically Measuring Ultra-Diffuse Galaxies (SMUDGes) is a catalog of ultra-diffuse galaxy (UDG) candidates presented in Zaritsky et al. (2023). More information on UDGs is given in section 2.7. This catalog contains 7070 candidate objects satisfying the following criteria: central surface brightness in the g-band, $\mu_0 \geq 24\,\mathrm{mag\,arcsec^{-2}}$ and physical half-light radius (the radius from which half of a galaxy's total luminosity is emitted) $r_e \geq 2.5\,\mathrm{kpc}$.

Studying ultra diffuse galaxies is difficult in principle, because firstly, these objects are hard to detect, and there are also difficulties with estimating their properties, like mass and their history. This can lead to conflicting results. Studies of individual UDGs often focus on the larger and brighter ones, resulting in a bias concluding their typical masses (including baryonic and dark matter) are comparable to or even larger than the Large Magellanic Cloud with total mass (dark matter and baryon content) $M_h \sim 1.4 \times 10^{11}\,\mathrm{M}_\odot$ like in Erkal et al. (2019). But statistical studies using larger and more diverse objects usually conclude these objects are less massive ($M_h < 10^{11}\,\mathrm{M}_\odot$), for example, in Beasley and Trujillo (2016) or Amorisco et al. (2018).

Statistical mass estimations are crucial for understanding the reason for their lack of star formation. Because there is a lack of homogeneous UDG data in various environments, the SMUDGes authors decided to search the images provided by the Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Surveys (Legacy Survey), on which more is given in section 3.5. They have selected UDG candidates across a wide variety of environments in sufficient numbers for statistical study. Then they estimated distances and masses for as many of them as possible to determine which candidates satisfy the size criterion.

They found that lower mass UDGs are star-deficient by a factor of ~ 4 relative to the general population of galaxies of the same total mass, and for the most massive ones, the star formation efficiency is

roughly an order of magnitude lower than for the general galaxy population. By star-deficient, the authors mean that, at a given total mass, UDGs contain significantly less stellar mass than typical galaxies of the same mass, indicating that they have converted their available baryons into stars with much lower efficiency than the general galaxy population. So, on average, over the full range of masses, UDGs found are deficient by a factor of 7. In the paper, various effects are considered, like different dark matter distributions in halos, and also the partial responsibility of systematics. However, the conclusion is that whatever process lies behind the formation of ultra diffuse galaxies, it does not simply reorganize the stars to larger radii, but it causes a measurable decrease in integrated star formation rates common for similarly massive galaxies.

3.5 Legacy Survey and DESI

As described in Dey et al. (2019), The Original Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Surveys[5] combine three survey projects (DECaLS, BASS, MzLS) to image $\approx 14\,000\,\mathrm{deg}^2$ of the extragalactic sky visible from the northern hemisphere in optical bands g, r, z, see figure 3.2. There are regular updates, currently the sky covered is over $20\,000\,\mathrm{deg}^2$ including the i band.

5. www.legacysurvey.org

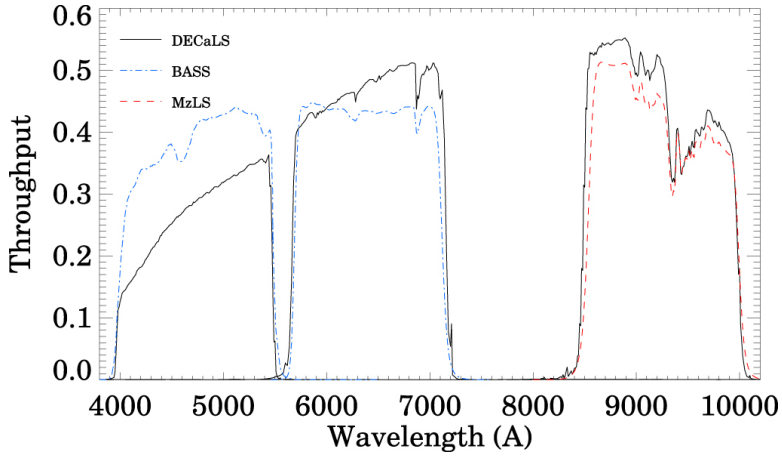


Figure 3.2: Effective bandpasses used for the Legacy Surveys, taken from Dey et al. (2019).

We used the Legacy Survey for fainter objects because its optical images are of higher quality than SDSS, offering significantly greater depth and resolution. Legacy Surveys are 1-2 magnitudes deeper, allowing for the observation of fainter objects and more detailed structures like tidal features that are difficult to see in SDSS images. For comparison, see figure 3.3.

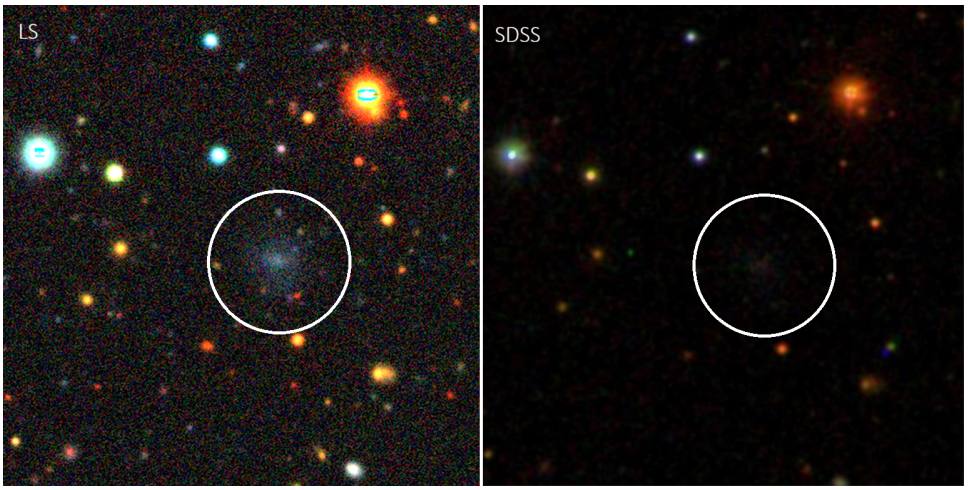


Figure 3.3: Comparison of Legacy Survey (left) and SDSS (right). An ultra diffuse galaxy is highlighted by the circle.

There is an online image viewer available (The Sky Viewer[6]) for interactive data exploration.

In addition to the interactive viewer, the Legacy Survey provides a publicly accessible Cutout API, which enables programmatic retrieval of small image regions centered on user-specified sky coordinates. The API supports requests in multiple bands (g, r, z, and i for newer releases) and allows the user to define the cutout size in pixels or arcseconds. This makes it possible to efficiently extract images for large samples of objects without downloading full survey tiles. This functionality is particularly useful for automated workflows where thousands of objects can be queried and processed in a consistent manner.

3.6 NASA/IPAC Extragalactic Database

The NASA/IPAC Extragalactic Database, or simply NED[7], is a curated database of multi-wavelength data for extragalactic objects, compiling measurements from numerous sky surveys and the published astrophysical literature. It provides observations spanning the electromagnetic spectrum from gamma rays to radio frequencies, with newly published data cross-identified and integrated with existing entries. NED also offers direct links to major NASA mission archives, such as IRSA, HEASARC, and MAST, as well as to the astrophysical literature through ADS and to other external data centers.

NED is integrating data from large survey catalogs such as 2MASS, NVSS, APM, and SDSS, and from several journals like Astrophysical Journal, Astronomy & Astrophysics, Monthly Notices of the Royal Astronomical Society, Nature, and others. Currently, it contains over 1.1 billion distinct objects, and roughly 11 million of them have a redshift measurement.

Since objects can be queried by name, near name or near position (cone), there are two main use-cases we had for NED: first to get redshift measurements for objects we were cataloging and second to find existing already recognized objects in our region that aren't present in

6. https://www.legacysurvey.org/viewer
7. https://ned.ipac.caltech.edu

our data (non-detections), so that we can have more complete idea about the environment.

The key points for us are that NED is very large because it uses such a variety of sources, but also that it's inhomogeneous for the same reason. This makes it an excellent source of supplementary data on a plethora of galaxy parameters, but care must be taken since it isn't necessarily complete, and measurements aren't standardized.

4 Beyond the Leo I group

In this chapter, we will provide information about the studied region. Then we will focus on the data itself, specifically its structure, reduction, and visualization.

4.1 General information

The surveyed region is centered around $(\alpha, \delta) = 10^h\,45^m\,00^s, 11^\circ\,48'\,00''$ and is $5^\circ \times 4^\circ$ in size. The main focus for the observation itself was the Leo I group, also called the M96 group. As described in Auld et al. (2006) the group lies at about $cz \sim 1\,000\,\mathrm{km\,s^{-1}}$. It consists of 4 main galaxies and is well known for its several optically dark HI features. The most prominent one is the Leo Ring, which is described in Stierwalt et al. (2009). With diameter of roughly $\sim 225\,\mathrm{kpc}$, total HI mass of about $1.8 \times 10^9\,\mathrm{M_\odot}$ and minimal star formation it challenges our models of gas accretion and consumption in galaxy groups, and may offer a laboratory for studying dwarf galaxy formation (or lack thereof) in a group environment. The enormous size and substantial mass of the Leo Ring highlight the importance of HI surveys in detecting large structures that are invisible at other wavelengths. This can be even more important for detecting dwarf galaxies, which can easily evade detection in optical surveys.

Even the Leo Ring only occupies a small fraction of the data volume. Our interest lies far beyond the Leo Group, in the unexplored volume which extends to much greater distances…

Thanks to its large bandwidth, most of Arecibo's detections lie in the background volume. In this case out to redshift $z < 0.06$ ($cz \sim 19\,000\,\mathrm{km\,s^{-1}}$) which corresponds to distance $d \sim 280\,\mathrm{Mpc}$. This background part of the observed area has no preselection criteria and is often referred to as the blind AGES volume. according to Auld et al. (2006), the detection limit at this distance is expected to be $M_{\mathrm{HI}} = 1.4 \times 10^9\,M_\odot$.

The data cube used in this work does contain the Leo ring, and it has been partially examined before, but only its foreground section. The results can be found in R. Taylor et al. (2022): five HI clouds with no optical counterparts were discovered. The authors chose to

limit their search within the velocity range $300 - 1\,600\,\mathrm{km\,s^{-1}}$. And that's where our work comes in. We have decided to analyze the blind background volume of this cube, meaning we had no idea what we were going to find. The available sampled velocity range is $1\,600 - 19\,000\,\mathrm{km\,s^{-1}}$ ($z \sim 0.005 - 0.06$). That corresponds to the distance range $d \sim 22 - 280\,\mathrm{Mpc}$.

4.2 The data structure

As is typical for radio astronomy, the observed data came in the form of a data cube as depicted in figure 4.1. It is a 3-dimensional structure containing the brightness of the objects across the sky at different wavelengths (frequencies). In other words, it is a series of images taken at different wavelengths, with each pixel containing spectral information. The file format is FITS, which is commonly used in astronomy.

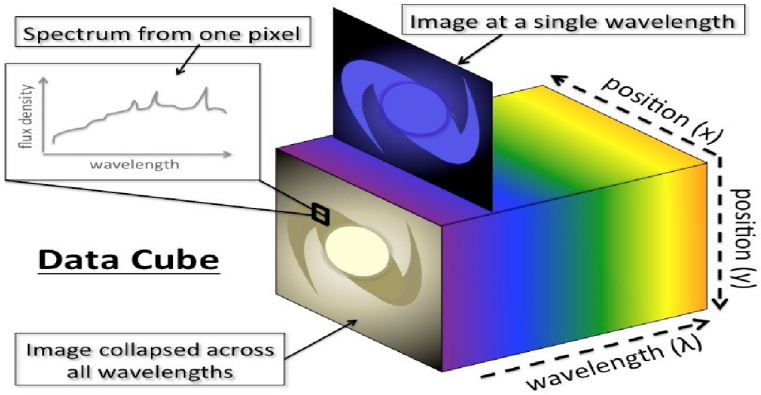


Figure 4.1: Structure of a data cube, taken from Liuzzo (2018).

4.3 Data reduction

The process of converting the raw observational data into a data cube in FITS format is not part of this project. The reduced cube was provided by Robert Minchin, with only a few additional processing steps needed before it could be used for analysis. More technical details on the reduction itself can be found in R. Taylor et al. (2014).

Firstly, we have applied a cleaning script written specifically for AGES data. First, this takes an input cube and applies Hanning smoothing at level 3 along the spectral axis, meaning it sets a three-channel Hann kernel used by the MIRIAD hanning task. More about MIRIAD can be found in section 4.4.1. As explained in Condon and Ransom (2016) in this procedure, each output channel is replaced by a weighted average of itself and its two neighboring channels according to

$$S'_i = \frac{1}{4}S_{i-1} + \frac{1}{2}S_i + \frac{1}{4}S_{i+1}, \tag{4.1}$$

where S_i is the original spectrum and S'_i is the smoothed spectrum. This operation is equivalent to a convolution with the kernel

$$K = \left[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right]. \tag{4.2}$$

Hanning smoothing reduces channel-to-channel noise by averaging adjacent spectral channels, therefore improving the sensitivity for weak signals. However, because neighboring channels are combined, narrow spectral features are broadened, and the velocity resolution is degraded. The smoothing width of three channels represents a compromise between noise reduction and preservation of spectral resolution. In our case, the improvement in sensitivity at the expense of velocity resolution is almost exclusively advantageous because the loss in velocity resolution is negligible, and our challenge is finding weak sources in noise.

The second action the script does is fitting and subtracting a polynomial from the spectral baseline for all spectra. In HI cubes, this is a standard calibration step. In this case, a second-order polynomial was fitted and subtracted from each spectrum to remove instrumental and continuum contributions. As explained in Condon and Ransom (2016), because the line emission or RFI can bias the fit, the iterative sigma-clipping was employed: after an initial fit, channels deviating more than a given threshold are excluded. The polynomial is then fitted to the remaining channels. The iterative nature of this process ensures that line features or RFI do not influence the baseline model and that the resulting baseline is a robust representation of the smooth background.

Because the whole data cube was quite large (1.14 GB), any operations on this data set would be slow and make cataloguing difficult. We therefore decided to cut it by velocity into 4 smaller cubes, which we then analyzed separately. This was done manually using MIRIAD (more in section 4.4.1).

The final but very useful step for data visualization was creating a Signal-to-noise cube. S/N cubes represent spectral data normalized by the local noise level, providing a uniform statistical measure of significance across a data cube (Condon & Ransom, 2016). These are widely used in manual and automated source finding. HI surveys search for faint emission in noisy data. Using S/N units allows a uniform detection threshold independent of local noise variations. This is especially important when noise varies across the cube. In our case, the edges of the cube were much noisier, and this transformation helped to mitigate this issue. We use the S/N cube only for source finding. For measuring the parameters of the detections, we use the original flux cube after the cleaning steps described earlier.

4.4 Visualization

In this section, we will look at the tools used for data visualization and the obstacles we have faced during this process.

4.4.1 FRELLED

As a primary tool for visualizing, cataloguing, and processing the FITS datacube, we have used the FRELLED software (FITS Realtime Explorer of Low Latency in Every Dimension). As described in R. Taylor (2025a), it is a Python-based extension to a popular 3D modeling tool, Blender, which offers a powerful platform for data inspection and manipulation. FRELLED utilises the capabilities of Blender to examine and analyze standard radio astronomy 3D data sets, mainly HI observations. Its 3D mode allows rapid visual cataloguing of more prominent sources, while the 2D mode allows looking for the faintest objects and features.

FRELLED was developed as part of AGES (see section 1.4). The standard technique of data inspection is to first use the volumetric 3D

view, mask bright sources using regions (cuboids marking the position of the identified object), and then search for fainter objects in 2D. The 3D mode allows one to see the structure in its entirety, but because of the rendering in Blender, the weaker sources are often masked by noise. So in 2D mode, where we see individual slices, even faint features can be spotted without obscuration, but defining boundaries of the source is a little more laborious, as we have to go through the individual slices.

FRELLED is meant for manual visual extraction of the sources. Although there are tools and efforts to automate this process, currently, visual source extraction outperforms most algorithms in terms of reliability and often completeness. Manual extraction is subjective and prone to human errors, so it's important to employ techniques minimizing these risks, like doing multiple searches by different people. While an automated approach is reliable for bright sources, usually the most interesting sources are at low S/N (signal to noise) levels, where automated tools can't be trusted yet.

An important feature for source extraction of FRELLED is the capability to display renzograms. As described in R. Taylor (2015), they are contour plots of the flux in each channel of a source. Each channel gets assigned a different color. An example of a renzogram is in the image 5.6. Especially for faint sources, these contours can help to decide whether the emission is real or spurious.

The direct SDSS image overlay feature displays SDSS images directly in FRELLED. This makes it very easy to compare HI emission with optical observations and determine if the galaxy has an obvious optical counterpart. Most optical counterparts can be identified immediately, with the majority of HI sources showing only a single galaxy near the HI coordinates. Cases where the galaxy is not clearly visible, or which have multiple counterparts, require more careful checking, as described in section 5. It is also possible to include a renzogram. So in one view, we can have an HI source with contours and an SDSS image. For example, see image 5.6.

Another useful feature for cataloging is the ability to query external databases, specifically NED and the SDSS (see sections 3.6 and 3.1). Because FRELLED supports standard astronomical world coordinates, we can quickly see if the possible source has a counterpart in different

wavelengths or if it has an independent redshift measurement. For example, see image 5.6.

The MIRIAD radio interferometry data-reduction package (Sault et al., 1995) is also integrated into FRELLED. This tool is commonly used in radio astronomy for data reduction and spectral line analysis. The main procedure we use is the *mbspect* spectral line analysis task, which is a standard way to determine the precise coordinates of a source by fitting a 2D Gaussian, measuring its line width, recessional velocity, rms (root-mean-square noise), S/N (signal to noise ratio), and total flux.

The *mbspect* requires that the user manually inputs coordinates and initial velocity estimates, but this is not necessary in our use-case because FRELLED handles all this automatically with user-defined regions, which we use for galaxy masking. For more complicated cases where the automatic profile range and other parameters used for fitting are not good enough, FRELLED offers an interface to set their values. For example, the user can set the Hanning smoothing, order of the polynomial used when fitting the baseline, range, and mask for the profile, and other parameters.

4.4.2 Radio Frequency Interference

As described in Giovanelli and Haynes (2015), radio frequency interference (RFI) is the worst enemy of a spectro-photometric, extragalactic HI survey. For a better idea of what the situation is at Arecibo, see figure 4.2 (taken from Giovanelli and Haynes (2015)), which shows the fraction of RFI disturbance as a function of frequency averaged over a year.

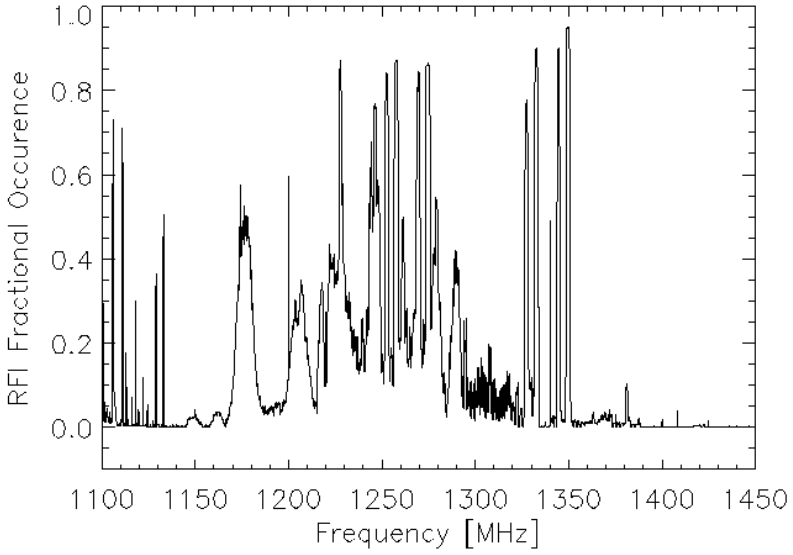


Figure 4.2: Fraction of the time that observations are severely affected by RFI at any given frequency in L-Band at the Arecibo Observatory. The 21 cm line rest frequency is 1420.4 MHz. Image taken from Giovanelli and Haynes (2015).

We can calculate the corresponding observed frequency range using the classical Doppler approximation $\nu \approx \nu_0 \left(1 - \frac{v}{c}\right)$, where $\nu_0 = 1420\,\text{MHz}$ is neutral hydrogen's 21-cm line rest frequency, v is the receding velocity and c is the speed of light. The relevant frequency range for our observation is between $1300 - 1400\,\text{MHz}$. The main disturber is the Global Positioning System (GPS) polluting near 1381 MHz. Another source of RFI is the San Juan airport radar transmitting between 1325 and 1350 MHz. Those are not the only ones by any means; several other sources are present but tend to be much weaker.

For further illustration, figure 4.3 is a position-velocity diagram of part of our data. There are 2 sources highlighted by a rectangle. Both are real, but the left one is almost completely disguised by the radio frequency interference.

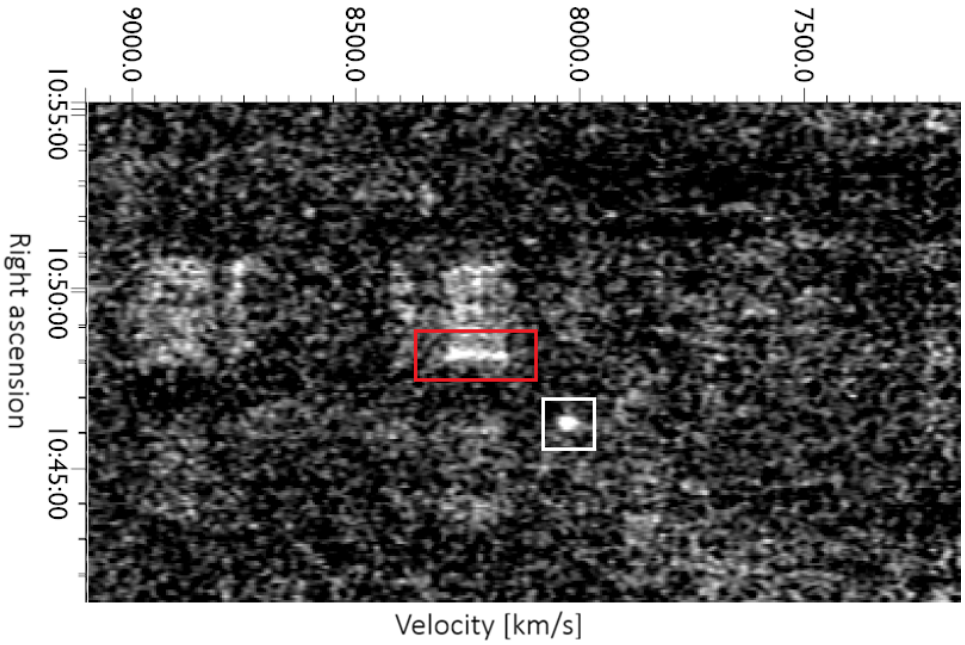


Figure 4.3: Galaxy marked by red rectangle is hidden by RFI.

Even worse case of RFI is in image 4.4. In the left part, around the velocity 18 000 km/s, there is a normal noise level. To the right, around 17 000 km/s, there is heavier pollution, but the worst part is between 16 000 km/s and 15 000 km/s, where the radio frequency interference dominates. In these RFI dominated belts, all efforts to look for galaxies are pointless.

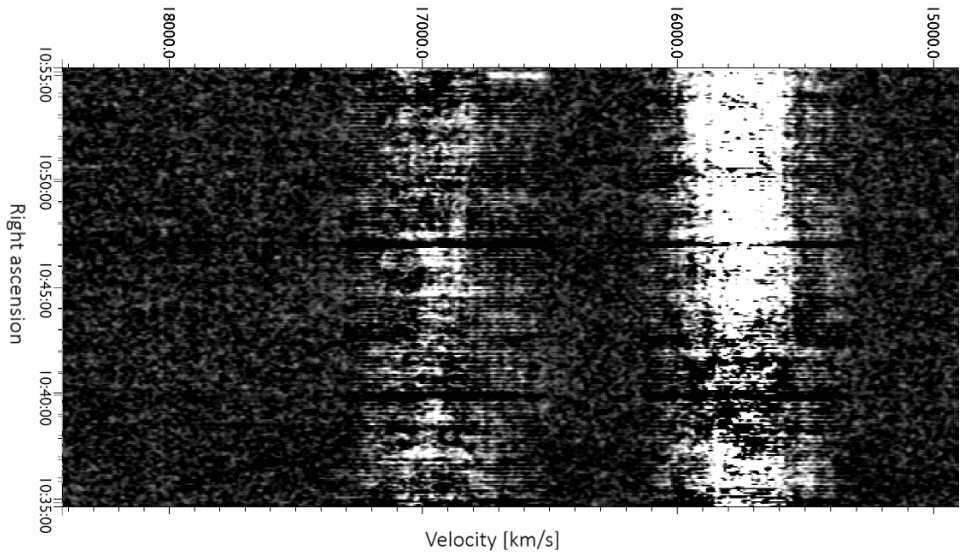


Figure 4.4: Area heavily polluted by RFI spanning the whole width of the cube.

4.4.3 Noise

In addition to radio frequency interference, the data are affected by random fluctuations that originate from the receiver electronics and natural background emission from the sky, as explained in Thompson et al. (2017). In radio astronomy, the noise is primarily caused by thermal noise, produced by receiver electronics, then sky background, e.g., cosmic microwave radiation or galactic synchrotron emission, atmosphere, and unresolved astrophysical emission. Unlike RFI, which is structured and external, this noise is intrinsically random; it manifests as Gaussian-distributed variations around the mean signal, and sets the fundamental sensitivity limit of the observation, limiting detection of weak sources.

The noise's strength is quantified by the root-mean-square. The RMS noise represents the standard deviation of intensity fluctuations in line-free spectral channels and provides a quantitative measure of observational sensitivity.

The signal-to-noise (S/N) is defined as the ratio of measured intensity to the RMS noise level and expresses the statistical significance of a detected signal.

5 Cataloguing

This chapter describes the cataloging itself: What techniques we have employed, and what the outputs of this step were. We also describe getting data from other databases.

5.1 The process

We have decided to do the whole cataloguing step manually, by eye, and as a preliminary phase to learn the process, I have gone through a trial cataloguing using a similar data cube from AGES (see section 1.4) previously analyzed by several experts. During this phase, I have learned how to use the FRELLED (see section 4.4.1) and all the available tools. This made it possible to compare my findings with a standard catalogue of known completeness and reliability. This was necessary to make sure my own catalogue would give results in agreement with previous searches (a more detailed discussion on the biases of visual source extraction is given in R. Taylor (2025b) section 5.4.4).

Then we started the cataloguing of the Leo data itself. The adopted strategy was a standard procedure when using FRELLED: first identify bright sources in 3D mode, then go through individual 2D slices and look for the faintest signals or sources masked by the RFI (see sections 4.4.3 and 4.4.2). Everything even remotely resembling a signal was marked for further analysis. To be sure nothing is overlooked, I have gone through the whole cube several times in different directions with different display settings in FRELLED, mainly the data levels, opacity, and clipping, which control how the data will be imaged. All of this required some experimentation to find a good range. After this initial step, we had 235 candidate objects.

5.1.1 Extracting HI data

The next thing to do was to obtain information from the HI data itself for each candidate source. That included finding the center coordinates, velocity, line width, peak flux, total flux (robust moment 0 which is the integrated HI intensity), rms, AASN (as defined by Saintonge (2007), it is a signal-to-noise metric optimized for the ALFALFA

survey's HI source detection pipeline, where values larger than 6.5 tend to indicate real sources) and HI mass using standard HI mass estimation equation from Roberts (1962) in logarithmic form:

$$\log_{10}(M_{\mathrm{HI}}) = \log_{10}(2.36 \times 10^5 \times (v/H_0)^2 \times T_{\mathrm{flux}}), \qquad (5.1)$$

where v is the recession velocity, $H_0 = 71\,\mathrm{km/s/Mpc}$ is the Hubble parameter, T_{flux} is the total flux. The result is in logarithmic solar masses.

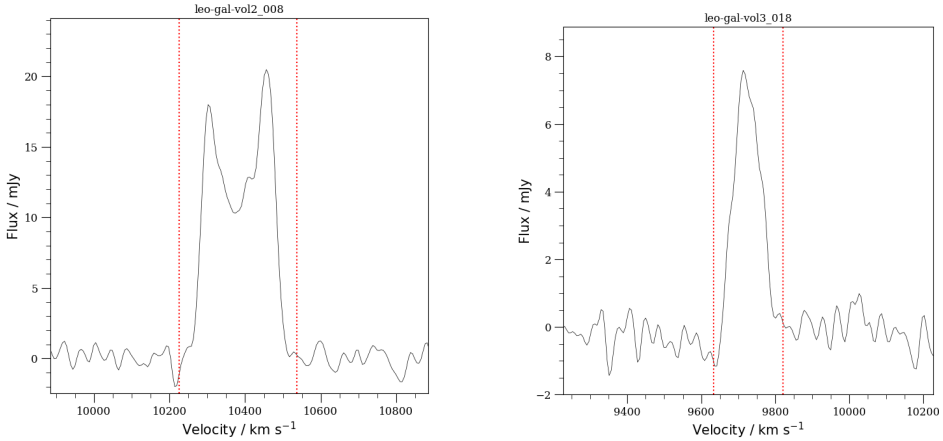
To estimate the key parameters of HI flux (i.e. mass, line width, coordinates, and systemic velocity), we have used a custom script used as standard for processing AGES data, which utilises Miriad and its mbspect task.

After setting initial constraints via FRELLED, like an estimate of the position, velocity, and line width; for bright sources, mbspect can usually refine these very accurately as long as those estimates are reasonable. For fainter sources, I had to provide more careful initial estimates, otherwise mbspect would often fail to provide a solution.

5.1.2 Spectrum

An important attribute of an HI source is its spectrum. Besides obtaining useful information like redshift or rotational velocity, its appearance can help to decide if the source is real or just noise.

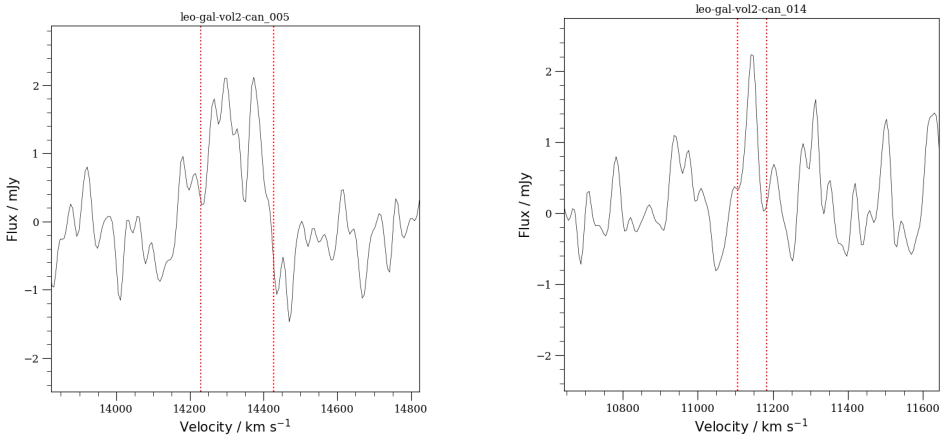
The classic HI galaxy spectrum is known as a double horn profile. It arises from the rotation of a galactic disk, where the two peaks represent emission from the approaching and receding sides at maximum rotational velocity. In a typical rotation curve, the speed at the center is zero, then rises rapidly to some value which remains constant out to large distances. This means that the amount of gas moving at high speeds is much greater than that moving at lower speeds. This means that less HI contributes to the flux at the zero relative motion (systemic velocity), causing the central dip. The example in the figure 5.1a shows almost an ideal double-horn with a high S/N ratio. Objects close to face-on tend to show as a single peak in the spectrum, like in the figure 5.1b. Thanks to steep edges in those examples, it is easy to select channels containing the emission from those with just noise, which is marked by the red vertical dotted lines.



(a) Double horn. **(b)** Single peak.

Figure 5.1: High S/N, nicely defined spectra.

These obvious cases were not rare in our sample, but the situation was often more complicated, as in figure 5.2. These galaxies have a much lower S/N ratio (the absolute values of flux are much lower than in the previous example). The profiles are shallow overall, making it difficult to determine the edges of the emission and making it subjective to some degree.



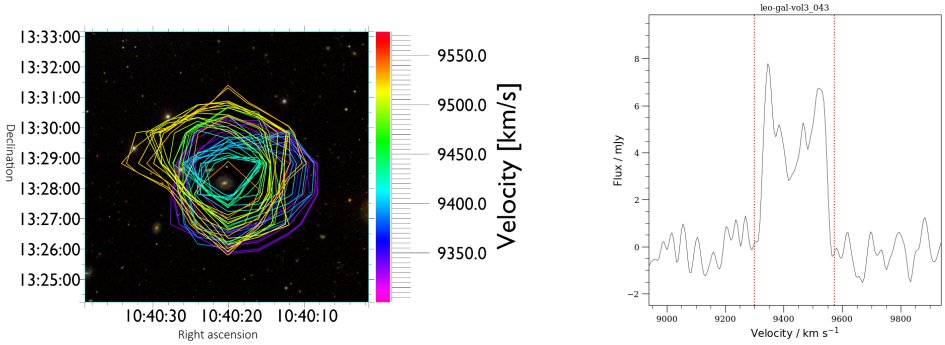
(a) Multiple peaks. **(b)** Single peak.

Figure 5.2: Low S/N spectra with unclear edges.

Many strange profiles can arise when galaxies are disturbed by the outside environment or by mutual interactions. But this is outside of the scope of this thesis.

5.1.3 The decision process

The next challenge was to separate the real sources from the spurious ones. For bright galaxies, this was straightforward because it was clear it wasn't noise or RFI. Additionally, almost exclusively, there was a clear optical counterpart in the SDSS, often with a matching redshift measurement. An example of an unambiguous bright source is in picture 5.3 where we can see the renzogram of the HI signal, optical counterpart from the SDSS right in the middle, and the spectrum, which is a nice double horn profile with a high S/N ratio.



(a) Renzogram of HI emission and SDSS overlay.

(b) Spectrum.

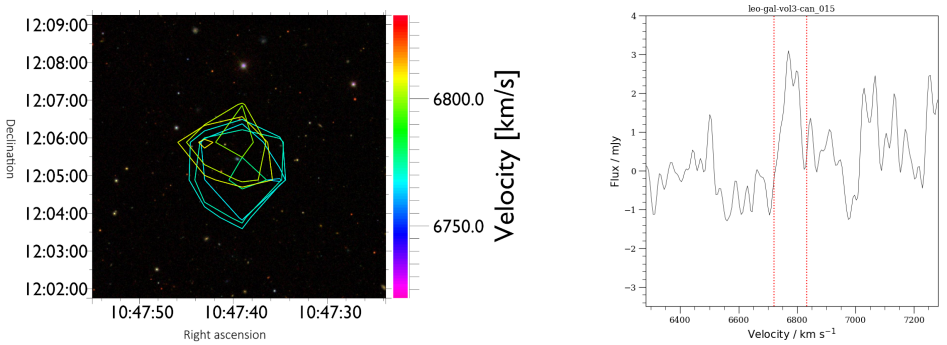
Figure 5.3: Clear, accepted source.

With the decreasing strength of the detections, the decision whether the signal is real or spurious became harder. We checked all the available information we could get from our HI data: appearance of the signal and its environment (noise level and RFI presence), its width (how many channels does it span), the spectrum, renzogram, and other properties, mainly the AASN parameter.

For the sources close to the noise level, it usually wasn't enough to look only at the HI data to make a decision. For that, we tried to find

an optical counterpart in the SDSS or Legacy Survey, then we looked for redshift measurements in the NED to see if it matches our value.

In figure 5.4, we present a candidate source that was accepted, but the decision was not straightforward. The HI is smooth, as can be seen in the renzogram, but the signal is quite weak. The spectrum is close to noise, with quite low S/N (AASN being 7, which is right on the edge of being a secure detection). With just this information, it would be hard to make a decision. But there is a faint optical counterpart right in the middle with a matching redshift measurement from NED, which confirmed that the signal is real.

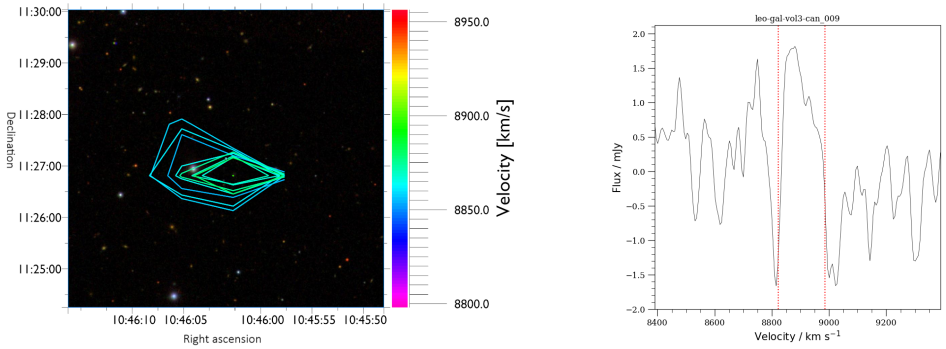


(a) Renzogram and SDSS overlay.

(b) Spectrum.

Figure 5.4: Accepted source with faint optical counterpart.

In Figure 5.5 is an example of a possible source. In the left panel (5.5a) is the renzogram and image from the SDSS. The HI signal is very weak, spanning only a few channels, and it is very close to the noise level. There is no optical counterpart and no redshift measurement in NED. The spectrum in 5.5b is also noisy, with low S/N and an unclear boundary. We have decided to reject this signal.



(a) Renzogram and SDSS overlay. **(b)** Spectrum.

Figure 5.5: Rejected source.

To make the decision-making process more systematic, I have created a table of attributes describing the source. As their summary, I used a confidence level value, which was assigned based on a combination of the source characteristics and parameters mentioned above. The debatable ones were subjected to more thorough analysis and sometimes discussion with the supervisor.

5.1.4 Interesting features

During the cataloguing, I have extended the table with additional qualitative flags to categorize the sources. Those characteristics were not necessary for accepting or rejecting the signal, but for classifying whether the source has some interesting feature. They are summarized in the table 5.1. Those flags serve as an indicator that an object might be worth exploring separately in future work.

Table 5.1: Interesting features assigned to some sources.

Feature	Meaning
Extended	The HI halo is much larger than optical counterpart
Offset	The HI center is outside the visual part of galaxy
Hairy	The HI emission is visibly messy
Merger	There are 2 galaxies merging, also optically
Interacting	There is an interaction of galaxies' HI halos
Interesting	Other interesting feature, described in notes

The following images show objects representing those categories. There is a renzogram of the HI signal and the SDSS photo in the background. Figure 5.6 is an example of a hairy signal (left) and an extended HI halo(right).

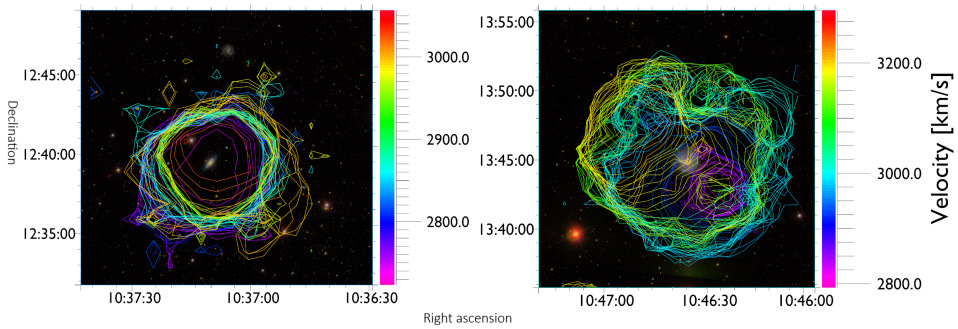


Figure 5.6: Example of hairy signal (left) and extended emission (right).

In figure 5.7, we can see a merger of two galaxies (left); they have a halo that looks like it belongs to a single galaxy. Their optical parts are in contact with each other. In the right panel, two galaxies are far away in optical, but their HI halos are already interacting.

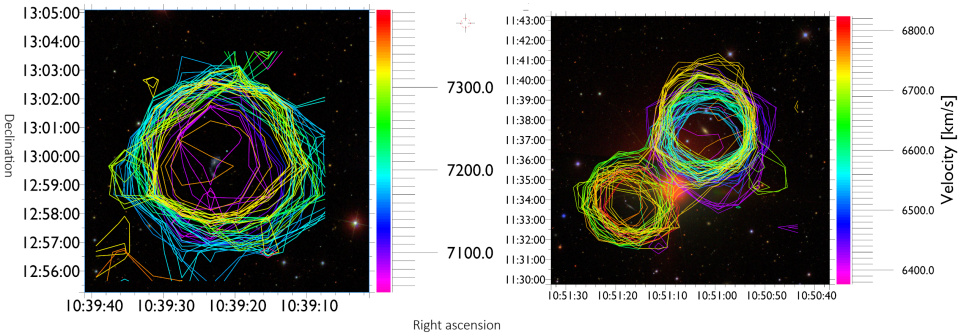


Figure 5.7: Example of a merger (left) and two interacting halos (right).

The last example is a galaxy that I have classified as being interesting because no obvious optical counterpart was visible in the SDSS. When we examined the much deeper DESI Legacy Survey data, we found that there is a very faint cloud that also has a large offset to the HI center. More on this particular source can be found in section 6.5.

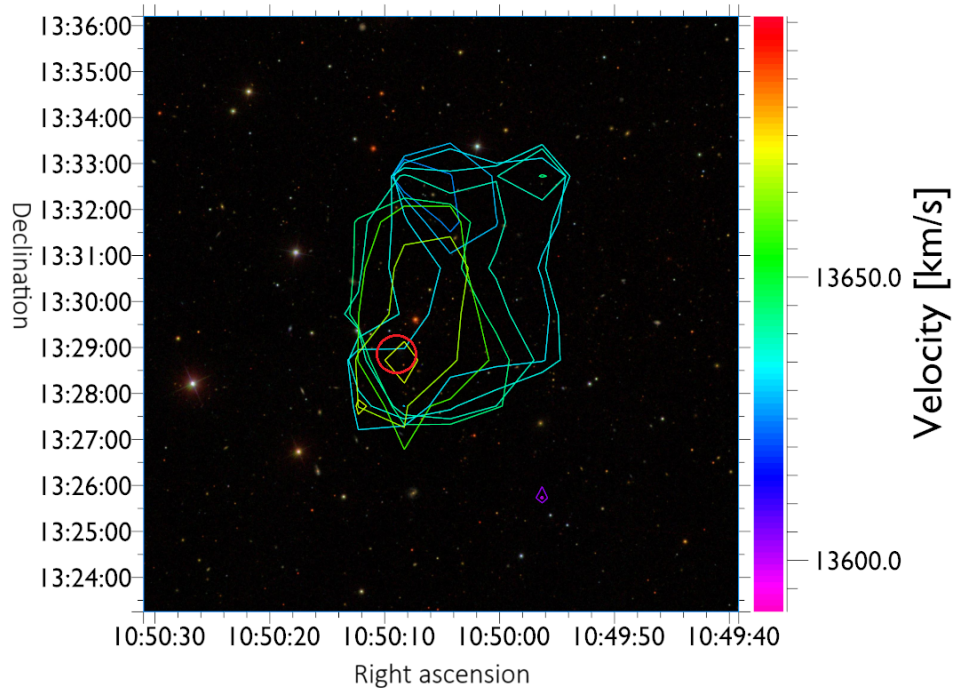


Figure 5.8: Example of an interesting object. There is no obvious optical counterpart, and a one-sided HI tail of an isolated object.

5.1.5 Result

After this quite long and tedious process, we had a list of all candidate sources with interesting features described, and most importantly, a decision whether we accept or reject the detection was made. We have also excluded objects with multiple possible optical counterparts. The further analysis was performed only with those accepted detections. Their number was 163.

5.2 Getting data from other surveys and catalogues

In this section, we will describe the process of extracting photometric measurements for our detections from the SDSS database. Then we will explain how we used those values to estimate stellar masses of the

galaxies. The last part talks about data extraction from the HyperLeda database.

5.2.1 SDSS cross-match

After having all the sources identified and having exhausted the options of getting information from the HI data, we wanted to get optical measurements from the SDSS for each of our galaxies. There is a way of automating this process using SQL queries, which allows finding the closest object, but because of the offsets of HI emission to the optical center of galaxies and the fact that SDSS has really a lot of photometric objects that sometimes represent only one physical object, the accuracy of the automatic assignment was low, and a wrong ID was often assigned. Because of that, we have decided to do the cross-matching with SDSS manually. Though quite tedious, the result was an accurate cross-match.

Once the SDSS identification numbers for each object were obtained, downloading the data was straightforward. It was a matter of constructing an SQL (Structured Query Language) command to query the SDSS database and submitting it to CasJobs (see section 3.1.2). The result was a table combining our HI data with SDSS magnitudes, extinctions, inclination angle, photometric coordinates, and optical image.

5.2.2 Using SDSS photometry

The next step was to exploit the SDSS photometry data. From the apparent magnitudes in the u, g, r, i, z bands together with the distance, which we have from the HI velocities, it is possible to calculate the absolute magnitudes using the well-known distance modulus. To improve the accuracy of the calculation, we have also included sky coordinates to account for Galactic extinction. The Galactic extinction corrections use the models from Schlafly and Finkbeiner (2011b) that are the standard corrections used in NED. The distance modulus then becomes $m - M = 5\log_{10}(d) - 5 + A$, where m is the apparent magnitude, M is the absolute magnitude, d is the distance in parsecs, and A is the extinction in magnitudes.

From the absolute magnitudes, we could calculate the stellar mass of the galaxies. There are numerous recipes for calculating the stellar masses, and there is no consensus on which one is the most correct. Because of this, the masses serve more as an indicator to discover general trends and to find outliers in the sample. Since we are looking for ultra diffuse galaxies or objects that don't fit the usual scaling relations, we do not need precise stellar masses – for our purposes, consistency is more important than deciding on the exact stellar mass recipe.

We have decided for widely used prescription from E. N. Taylor et al. (2011), where the calculation is based on the g − i color index:

$$\log_{10}\left(\frac{M_*}{L_i}\right) = -0.68 + 0.70\,(g-i), \tag{5.2}$$

where M_* is the stellar mass and L_i is the i-band luminosity. The luminosity is derived from the absolute magnitude M_i via

$$L_i = 10^{-0.4\,(M_i - M_{i,\odot})}, \tag{5.3}$$

where $M_{i,\odot} = 4.58$ is the solar absolute magnitude in the i-band.

This is the procedure used in (among others) R. Taylor et al. (2022) for estimating the baryonic Tully-Fisher relation, thus allowing us to compare our findings with the earlier results. An important caveat is that this method is designed for normal, star-forming galaxies. It may break down for extreme objects.

The photometric calculations were adapted from existing implementations by R. Taylor (2023). This code is in the form of a web app, where all values have to be input manually. I have modified this code in Python to work with tabular data.

5.2.3 HyperLeda

The next step was to match our data with the HyperLeda catalogue (more in sec 3.3) to obtain the morphological type and the length of the major axis for our galaxies. This was quite straightforward; it was a matter of writing a script for downloading the database and doing a cross-match, this time automatic, by finding the nearest HyperLeda object for each of our galaxies' photometric coordinates. As a result, we got the desired information for 75% of our sample.

6 Data analysis

This section describes the process of plotting and analyzing all the information we have obtained from the HI itself, and also from the external databases.

6.1 Environment

We wanted to put detected galaxies into an environmental context to examine if gas loss is occurring in different environments, and if so, whether this has any effect on the galaxies themselves. We needed to compare gas-detected and non-detected samples to examine the effect of environment.

6.1.1 Wedge diagram/polar plot

A polar plot, also called a wedge diagram, is shown in figure 6.1. This displays each HI detection in a 2D slice with right ascension or declination on one axis and its velocity, which directly corresponds with the distance on the second axis. For plotting, we have used a modified version of a code by Boris Deshev used in Deshev et al. (2022). For a more complete picture of the environment, we have included non-detections from NED into the diagram as well.

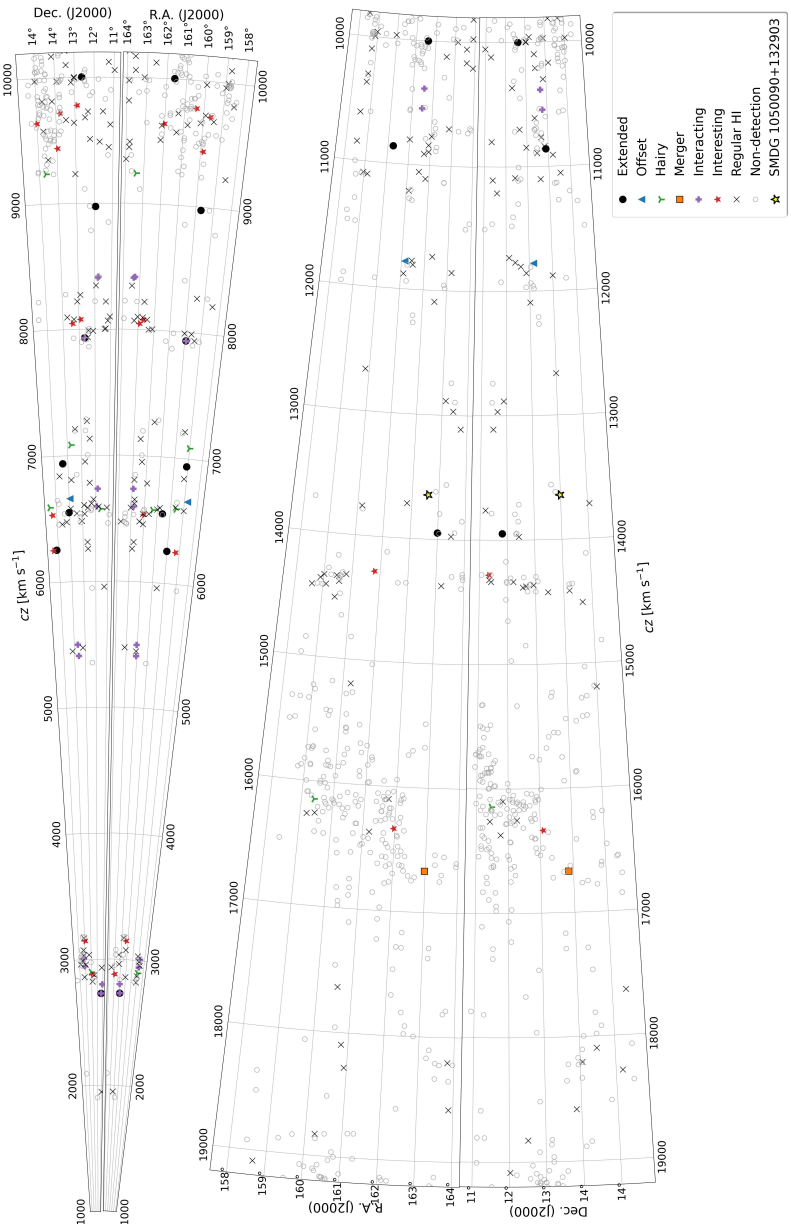


Figure 6.1: Polar plot showing detected HI galaxies and optical galaxies from NED (hollow circles) in RA-cz and Dec-cz planes. For clarity, the plot is split in 2 halves by velocity. The plot is meant to be viewed on a larger screen. SMDG 1050090+132903 (more in section 6.5) marked by a yellow star.

The wedge diagram reveals several overdensities in redshift space at $cz \approx 3000, 6500, 10000, 16000$ km s^{-1}, indicating the presence of large-scale structures.

A cross-match with NED shows that these peaks align with known galaxy groups and clusters. This is visible in figure 6.2, which is a velocity histogram with known clusters marked by red lines. At $cz \approx 3000$ and 6500 km s^{-1} there are only galaxy groups known. The ratio of HI galaxies vs non-detections is reduced in the densest peaks, confirming that HI-rich galaxies are less common in high-density environments, which is consistent with environmental processes that remove neutral gas in clusters. The HI does preferentially trace gas-rich galaxies located in the outskirts of clusters rather than their dense cores.

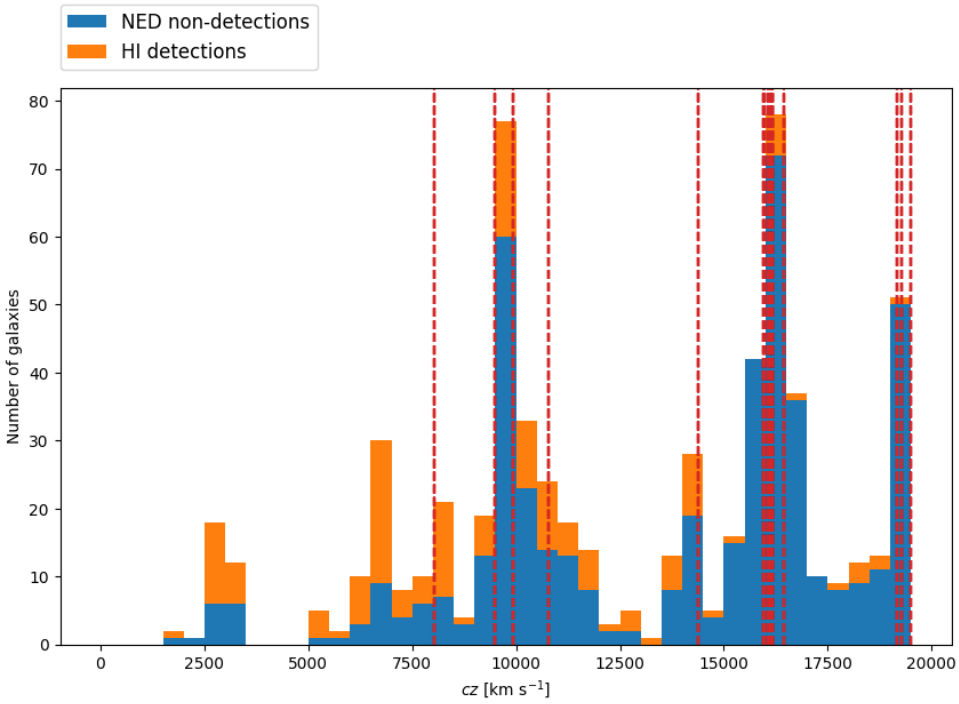


Figure 6.2: Velocity histogram. Catalogued clusters from NED are marked by red vertical lines.

The identified galaxy clusters were detected by automated algorithms and they are likely rather loose structures than massive clusters

because no prominent Finger-of-God structures are evident, even with included non-detections as shown in figure 6.1.

6.1.2 HI deficiency

The HI deficiency parameter is used to quantify how much atomic hydrogen a galaxy is missing compared to similar isolated galaxies. It is a dimensionless quantity defined as

$$\mathrm{DEF_{HI}} = \log_{10}\left(\frac{M_{\mathrm{HI}}^{\mathrm{expected}}}{M_\odot}\right) - \log_{10}\left(\frac{M_{\mathrm{HI}}^{\mathrm{observed}}}{M_\odot}\right), \qquad (6.1)$$

where $M_{\mathrm{HI}}^{\mathrm{expected}}$ is the expected HI mass of a typical field galaxy. The $M_{\mathrm{HI}}^{\mathrm{observed}}$ is the measured neutral hydrogen mass.

The interpretation of the HI deficiency is following: $\mathrm{DEF_{HI}} = 0$: normal HI content, $\mathrm{DEF_{HI}} > 0$: HI deficient. Values around $0.3 - 0.5$ mean moderate deficiency. For example $\mathrm{DEF_{HI}} = 1$ indicates severe stripping. It means the galaxy has 10 times less HI than a typical isolated galaxy should have.

For the expected HI mass, we have adopted a formula from Haynes and Giovanelli (1984) which links a galaxy's total mass to the optical size of the disk:

$$\log M_{\mathrm{HI}}^{\mathrm{expected}} = a + b \log D. \qquad (6.2)$$

D is the optical diameter of the galaxy, it is the isophotal diameter at 25 mag arcsec^{-2}. We've taken it from HyperLeda parameter `logd25` which is described in section 3.3. Empirical coefficients a and b are weakly depending on the morphological type. They are determined from samples of isolated galaxies, which are assumed to represent the normal HI content. We have used coefficients independent of morphology. We have used $a = 7.3$ and $b = 1.72$, which are recent values from Jones et al. (2018) calibrated on 844 isolated galaxies. This paper has the largest and most carefully selected galaxy sample to date.

We have restricted our analysis to 122 galaxies with available HyperLeda morphology classification and optical diameter. The resulting HI deficiency plot is in figure 6.3. As we can see, it is centered around

0, indicating that most galaxies have HI masses close to the expected value. There are no severe deficiencies or over-abundances.

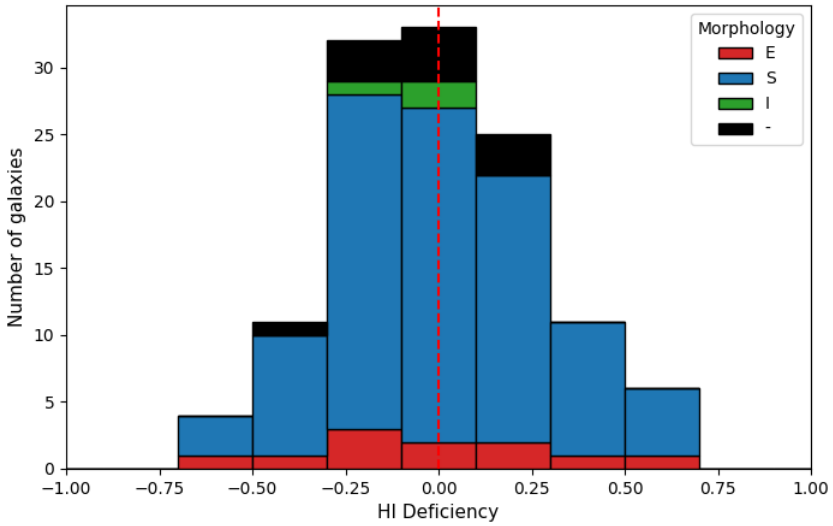


Figure 6.3: Distribution of HI deficiency stacked by morphology (elliptical, spiral, irregular, unknown. No deficiency marked by a dashed line.

The mean HI deficiency of the sample is $\langle \mathrm{DEF_{HI}} \rangle = 0.01 \pm 0.02$, with only 12% of galaxies having $|\mathrm{DEF_{HI}}| > 0.3$, which exceeds the intrinsic scatter of the calibration. No galaxies exhibit strong deficiency ($|\mathrm{DEF_{HI}}| > 0.7$).

This distribution is characteristic of galaxies residing in the field or low-density environments, where environmental effects such as ram-pressure stripping are weak or absent. The lack of strongly HI-deficient galaxies further supports the conclusion that the sample does not probe the dense cores of massive clusters, where such effects are typically most pronounced. This is, of course, partly caused by the obvious fact that the more HI-deficient the galaxy is, the more difficult it is to detect it with an HI survey. In other words, regions of strong stripping might be present, but hard to detect in our survey as they would render the galaxies undetectable in HI.

6.2 Offset plot

During the cataloging, we have noticed there seems to be a systematic offset between the HI center of each galaxy and the optical center from the SDSS. As a small side-quest, we've decided to investigate a little further by plotting it. The result can be seen in the figure 6.4 with the mean marked by an orange cross.

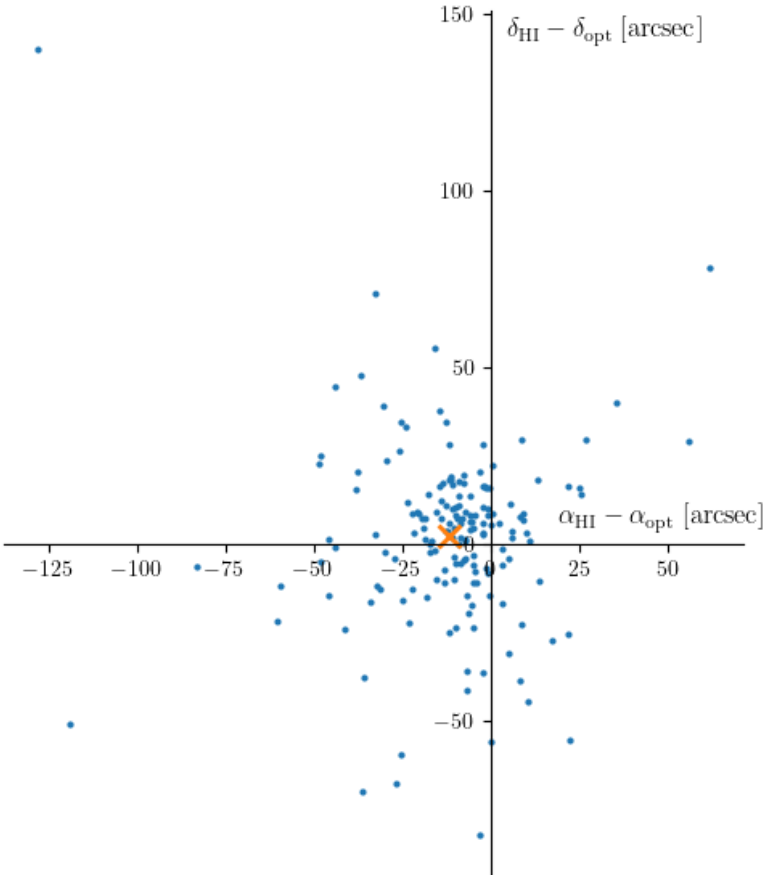


Figure 6.4: Offset plot with marked mean.

There is a visible offset in right ascension and possibly in declination. We have quantified those positional offsets between the HI detections and their optical counterparts. The RA offsets show a mean offset of $\langle \Delta\alpha \rangle = -12 \pm 2$ arcsec, indicating a systematic shift. But the magnitude of the offset is affected by a few outliers with low S/N,

where the center is determined with large uncertainty. After excluding those, the offset becomes $\langle \Delta\alpha \rangle = -10 \pm 2$ arcsec. The Dec offsets show a mean shift of $\langle \Delta\delta \rangle = 2 \pm 2$ arcsec, which is not statistically significant.

The pointing error of the telescope was about 10 arcsec, but this shouldn't cause a systematic shift. The RA offset is a known problem due to the reference coordinate in the FITS header being recorded incorrectly. We expect an average of about 8 arcsec.

We conclude that the found offset of $\langle \Delta\alpha \rangle = -10 \pm 2$ arcsec falls into the expected range. More importantly, it is not enough to affect any science, given the 3.5 arcmin beam.

6.3 Scaling relations

We have decided to create this set of plots, commonly referred to as the scaling relations, to serve as diagnostic tools for identifying trends and outliers, e.g., galaxies with high gas fraction but low luminosity, or very red but HI-rich galaxies. In these plots, we distinguish galaxies by their morphology and also by their interesting features (as described in section 5.1.4).

6.3.1 Mass to Light

The HI mass-to-light ratio M_{HI}/L_g scaling relation provides a measure of the gas content relative to the stellar component. When plotted against absolute magnitude M_g, it reveals systematic trends in galaxy evolution, with lower-luminosity systems typically exhibiting higher gas fractions.

The M_{HI} is the HI mass, M_g is the absolute magnitude in the g-band corrected for extinction. The g-band luminosity L_g is calculated from the absolute magnitude as

$$L_g = 10^{-0.4\left(M_g - M_{g,\odot}\right)}, \tag{6.3}$$

where $M_{g,\odot} = 5.11$ is the solar absolute magnitude in the same band from Willmer (2018). This relation yields luminosities in units of the solar luminosity.

The ratio M_{HI}/L_g thus represents the HI mass-to-light ratio in solar units. It provides a measure of the gas richness of galaxies by comparing the mass of neutral hydrogen to the stellar light traced by the g-band luminosity.

Figure 6.5 shows the mass-to-light scaling relation. It reveals an expected trend: a clear increase of M_{HI}/L_g toward fainter magnitudes. Ellipticals mixed with spirals, especially at low mass-to-light ratios, are likely due to their uncertain classification. Visual inspection of the SDSS images revealed that some galaxies are misclassified by the HyperLeda catalog. There is a mixture of effects, e.g., some objects classified as ellipticals are actually edge-on spirals, some disks are bulge-dominated and misclassified because of that as ellipticals. In most of those uncertain cases, the available resolution of the SDSS is not good enough for a reliable determination. Not all ellipticals are misclassified though.

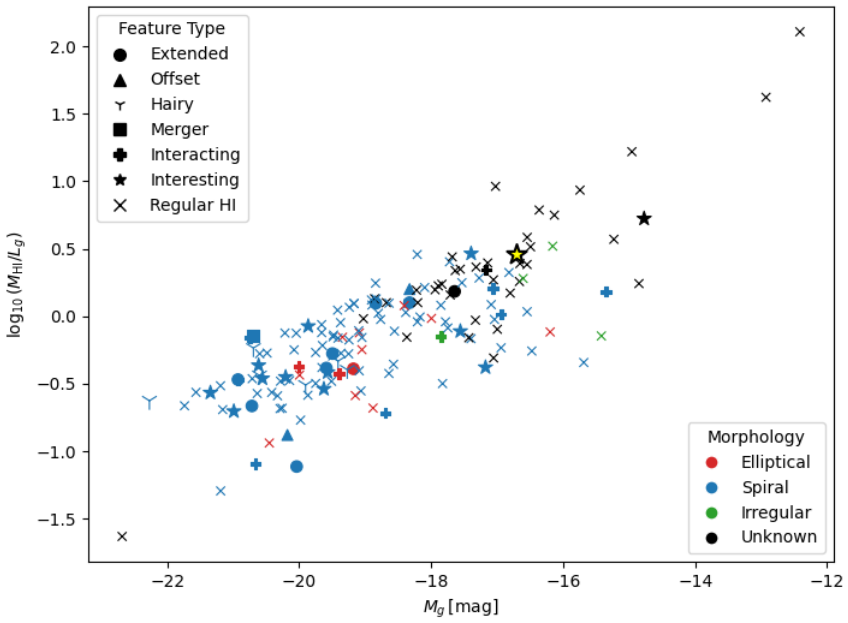


Figure 6.5: Mass to light scaling relation. SMDG 1050090+132903 marked by a yellow star.

6.3.2 HI mass fraction

The HI mass fraction is defined as the ratio of neutral hydrogen mass to stellar mass, $M_{\mathrm{HI}}/M_{\star}$. This quantity provides a direct measure of the gas content relative to the stellar component and represents the amount of fuel available for future star formation.

When plotted as a function of stellar mass, the typical trend is that low-mass galaxies have high gas fraction and high-mass galaxies have low gas fraction because massive galaxies have already converted more gas into stars.

In the figure 6.6 is our version of this plot. It shows the expected trend very strongly with very little scatter. Some ellipticals appear at moderate gas fractions (~ 0 to -0.5), but this is again because of their uncertain classification.

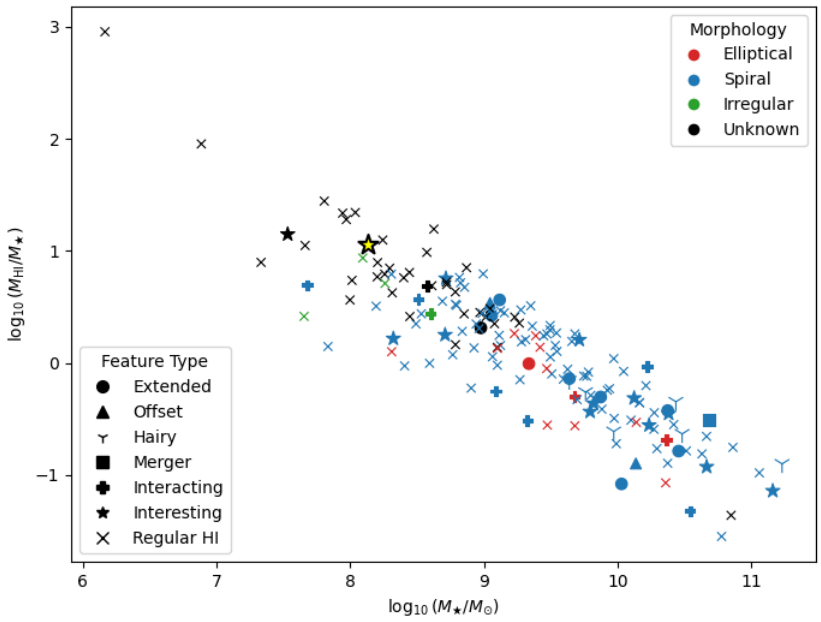


Figure 6.6: HI mass fraction scaling relation. SMDG 1050090+132903 marked by a yellow star.

6.3.3 Color magnitude diagram

The color–magnitude diagram relates the optical color of galaxies, in our case the $g - i$ color index, to their luminosity. Galaxies are typically distributed into two main populations: the red sequence, consisting of quiescent systems dominated by old stellar populations, and the blue cloud, composed of star-forming galaxies with younger stars. An intermediate region, often referred to as the green valley, is thought to host galaxies transitioning between these two states. The position of a galaxy in this diagram reflects its star formation history and evolutionary stage.

Figure 6.7 shows the color magnitude diagram of our sample. It shows no clear separation of populations, likely because our HI-selected sample probes very little of the red sequence. This has the most scatter of all the scaling relations, but this is expected and still reasonable, as the sample includes a heterogeneous population of galaxies and is biased toward HI-rich systems. Ellipticals are mostly redder ($\sim 0.7 - 1.0$) as expected. Some ellipticals might look too blue ($\sim 0.5 - 0.6$), but again, this is due to misclassification.

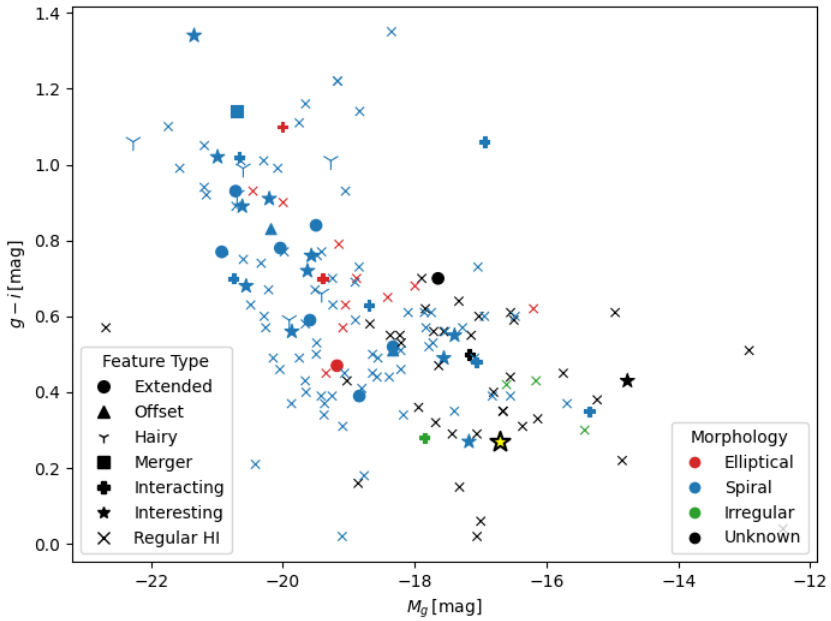


Figure 6.7: Color magnitude scaling relation. SMDG 1050090+132903 marked by a yellow star.

6.3.4 Tully-Fisher relationship

The Tully–Fisher relation (TFR) links the rotational velocity of a galaxy to its luminosity or baryonic mass. It reflects the connection between the total mass of a galaxy and its gravitational potential. In its baryonic form (BTFR), the relation connects the total baryonic mass (stars and gas) to the rotational velocity, providing a probe of galaxy dynamics and dark matter content. The Tully–Fisher relation is widely used as a distance indicator and as a tool for studying galaxy formation, evolution, and testing ΛCDM and other theories like modified Newtonian dynamics (MOND).

For HI observations, the rotational velocity is estimated from the observed line width of the 21-cm profile. Corrections for galaxy inclination are required to recover the intrinsic rotational velocity, as

projection effects can significantly broaden or narrow the observed line profile.

To ensure reliable measurements of rotational velocities and baryonic masses of galaxies included in the Tully–Fisher relationship, we apply several additional quality cuts to the sample. First, galaxies with low inclination angles ($i < 45^\circ$) are excluded, since the correction of observed line widths by a factor of $\sin i$ becomes highly uncertain for nearly face-on systems. Galaxies with high HI deficiency (HIdef > 0.6) are also removed, as environmental processes such as gas stripping can significantly alter the HI distribution and distort the line profile, leading to biased velocity width measurements. Finally, we require a sufficiently high signal-to-noise ratio S/N > 10 to ensure accurate measurement of the HI line width. We have adopted the same cuts for all our TFRs.

I have explicitly included the object SMDG 1050090+132903 from the SMUDGes database, although it doesn't satisfy those stricter restrictions, for more information see section 6.5.

For both baryonic and optical Tully-Fisher relationships, we have created two versions of the plot. First is the basic one, where we simply used our data after doing the quality cuts and applying an inclination correction. But for the plots to be scientifically significant and comparable to the literature, numerous corrections to convert the observables to baryonic mass and velocity have to be applied. This is the reason why when we plotted fits from literature on our sample, the relation isn't a great fit to the data, but it was not too far off, as most of the corrections are small.

Significant outliers can identify objects which are dynamically different from typical galaxies, such as the gas clouds found in the main Leo region. Lack of outliers confirms that our distance assignments from a simple Hubble flow model are probably reliable, which is also a valuable check.

Baryonic Tully-Fisher relationship

As a standard BTFR, we have used the calibration from McGaugh (2012), which was calibrated using a carefully selected sample of gas-rich galaxies. The relation itself is given as:

$$M_{\text{baryon}} = A \times v_c^4 \tag{6.4}$$

Where v_c is the corrected velocity and $A = 47 \pm 6\,\mathrm{M}_\odot\mathrm{km}^{-4}\mathrm{s}^4$. The one-sigma scatter is given as 0.24 dex. See the appendix to the AGES Leo paper by R. Taylor et al. (2022) for the details of how all the corrections are done. They include corrections to both line width and velocity according to stellar mass and turbulence. Stellar is mass corrected for internal extinction and baryonic mass is corrected for helium.

The resulting basic plot before applying the advanced corrections is shown in the figure 6.8. There is also the McGaugh's fit and ± 1 and 2σ intervals. Our sample does follow the same general trend, but most of the galaxies are centered around the bottom 2σ line.

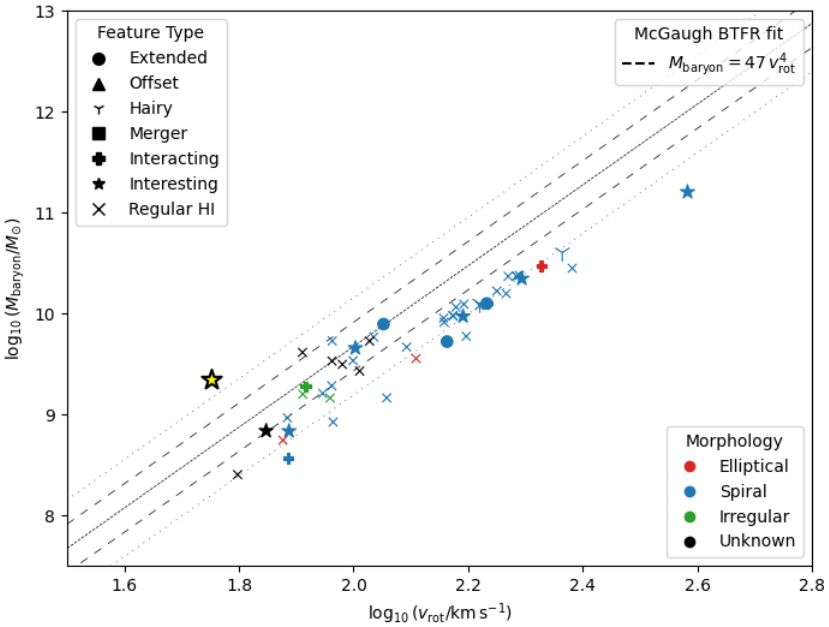


Figure 6.8: Baryonic Tully-Fisher relationship without corrections. SMDG 1050090+132903 marked by a yellow star.

To show the impact for these corrections, compare figure 6.8 (uncorrected) with figure 6.9 (after corrections). Our AGES data are shown as black filled circles, and McGaugh's as open black circles. Again, it also shows McGaugh's relation with ± 1 and 2σ. Here, our sample is in excellent alignment with the calibration.

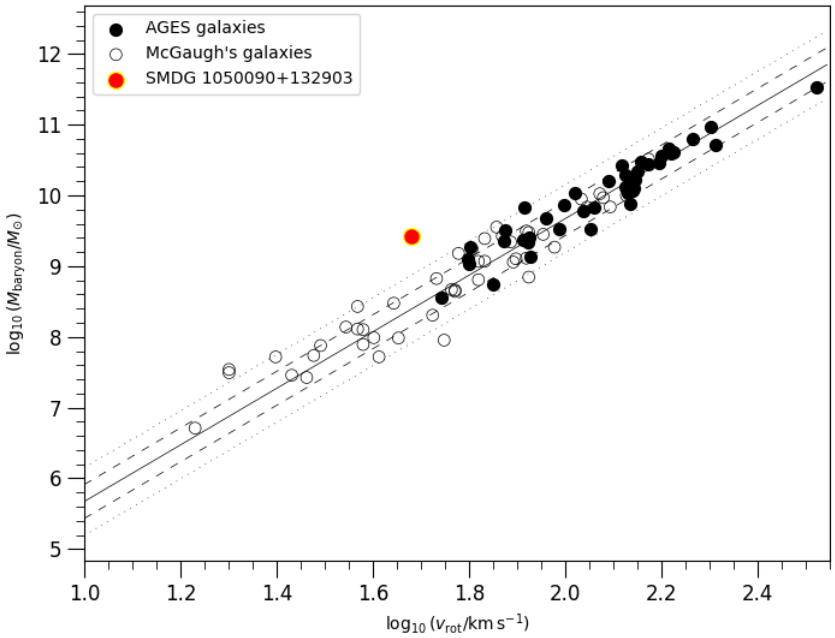


Figure 6.9: Baryonic Tully-Fisher relationship with corrections.

Optical Tully-Fisher relationship

For the optical Tully–Fisher relation, we have adopted the calibration from Kourkchi et al. (2020), which is based on a sample of approximately 600 spiral galaxies, and the fit can be written as

$$M_g = -8.04(\log_{10}(W_{\rm corr}) - 2.5) - 20.18, \tag{6.5}$$

where $W_{\rm corr}$ is the corrected HI line width. The intrinsic scatter is 0.48 mag (1σ). Their analysis combines high-quality HI line width

measurements with multi-band optical photometry from SDSS and infrared data from WISE.

In figure 6.10 is the basic version, where the linewidth is corrected only for inclination. The Kourkchi relation is shown together with its 1σ and 2σ confidence intervals. Similarly to uncorrected BTFR, the sample is shifted towards the bottom 2σ line.

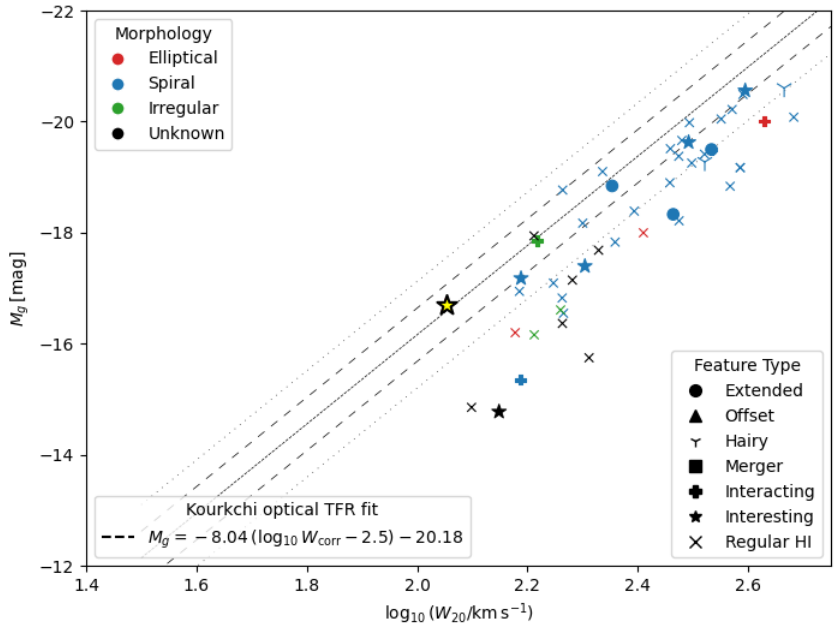


Figure 6.10: Optical Tully-Fisher relationship without corrections. SMDG 1050090+132903 marked by a yellow star.

Figure 6.11 presents the version with the optical magnitudes corrected for both Galactic and internal extinction using a color-based prescription following Kourkchi et al. (2020), resulting in a pseudo-magnitude that minimizes inclination-dependent effects and reduces scatter. In addition, the velocity widths are corrected using a multi-step procedure including instrumental, redshift, and turbulence corrections. This version of optical TFR thus provides a robust, independent check of the baryonic relation, avoiding uncertainties associated with

stellar mass estimates. Although visually this appears slightly more scattered than in the BTFR, our data are still in good alignment with the Kourkchi relation.

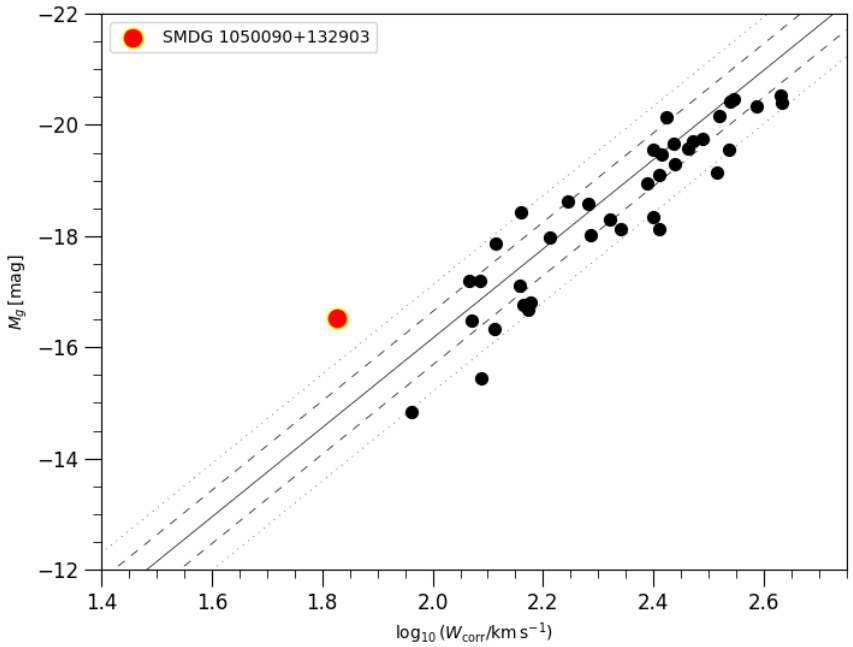


Figure 6.11: Optical Tully-Fisher relationship with corrections.

6.3.5 Scaling relations summary

In the color-magnitude diagram, ellipticals are on the redder side of the relation, meaning they are redder than a typical galaxy from the sample, suggesting there is at least some morphological difference: they are not typical spirals.

Gas fraction of ellipticals in the HI mass fraction relation is normal; they don't seem to be particularly gas-rich or gas-depleted compared to a typical galaxy. None of them are unusually gas-rich; they are all within the general scatter of the relation. None of them is an outlier. They are not gas-rich relative to the whole population: they are just relatively gas-rich for a red object, based on a comparison of their gas fractions with typical values for galaxies of similar color.

Some spirals in the color magnitude diagram appear redder than normal spirals, but they have a typical gas fraction. Their red color implies a reduced star formation activity, which raises the question as to what suppresses star formation in these relatively gas-rich objects. It is interesting to consider whether this relates to the gas-rich "ellipticals", which, though likely not true elliptical galaxies, nevertheless share their red colors. Gas-loss driven evolution of galaxies from the red sequence to the blue cloud has been suggested, for example, by Hughes and Cortese (2009), and it may be that we are witnessing a similar but weaker effect without a full morphological change. However, this requires a more detailed analysis of the gas fraction of the objects and a more exact quantification of their morphology, which is beyond the scope of the current work.

Another interesting set of objects are the so-called "hairy" galaxies. They were previously found in the Virgo cluster, and we do not know how they came to be. They don't have systematically enhanced gas fractions, which would be expected if they were predominantly undergoing gas-rich mergers. Conversely, they are not consistently gas-poor or red, which would indicate strong environmental stripping. Moreover, finding them in different environments points to the creation mechanism not being an environmental effect like ram pressure stripping. There might be multiple mechanisms causing these, but it seems we do not need RPS. There might be something interesting going on. The fact wee see high S/N hairy objects makes it much more likely that the cause is astrophysical – something is really disturbing the gas – but we can't completely eliminate the possibility of a data reduction artifact. We would need higher-resolution data to make further conclusions.

There might be something interesting going on, but it might still be some data reduction artifact not noticed before. They are not all faint objects, which would suggest the effects of low S/N.

None of those relations revealed any major outliers, which would have to be glaringly obvious. Our relations are rather nice, tidy, mostly following expected trends, with galaxies uniformly scattered.

6.4 Identifying UDGs

One of the goals of this work was to determine the number of Ultra diffuse galaxies (more in section 2.7) in the observed background volume. To do that, we have decided to do our own photometry of our sources using the Sérsic profile fitting script called *Photomass* from Ebrová et al. (2025). The photometry fitting is performed on images downloaded from Legacy Survey (more in section 3.5) using GALFIT (Peng et al., 2002). The script automatically downloads Legacy Survey images of the desired objects, does the masking, performs the GALFIT photometry, and calculates the stellar masses.

However, *Photomass* does not perform multi-band fitting or stacking, which is a drawback for UDG detection. That's why we used an extension of the code that stacks optical bands (g, r, z) before fitting to improve the detection of ultra-diffuse galaxies, which are intrinsically faint and often close to the noise limit in individual images.

From the fitted absolute parameters, we derived the central surface brightness μ_0 and physical effective radius. The script also outputs $\langle \mu_e \rangle_g$. This is the mean surface brightness in g-band within the effective radius R_e. While most people use μ_0 for determining UDGs, van der Burg et al. (2016) use the mean, though in r-band rather than g-band.

In the SMUDGes database, the authors fit the structural parameters (Sérsic index, R_e) on the stacked grz image, then use these to measure the central surface brightness in each band. The extensive modifications required to the *Photomass* script mean that we don't yet do this – instead, we measure everything in each band independently. This is likely the reason our central surface brightnesses are significantly brighter than the mean values: the lower sensitivity of the individual images means the fit will be worse. The mean value is therefore the more reliable of the parameters here, making this the better guide to UDG-status.

We have decided to be generous with the selection criteria for UDGs: anything with central or average surface brightness > 23.5 mag arcsec^{-2} and $R_e > 1.0$ kpc will be a candidate, anything with central or average surface brightness > 24.0 and $R_e > 1.5$ kpc will be secure detection.

The final step included checking the script output. Candidates with incorrectly estimated radius were removed. The resulting UDG

detections and candidates are in the table 6.1. There are 6 clear detections (including SMDG 1050090+132903, which is the first one in the table with Name 1237664131018719712) and 11 candidate objects. From the SMUDGes and NED search, five of the clear detections are completely new, and none of them were previously detected in HI.

As described in R. Taylor et al. (2026b), there are currently fewer than 300 UDGs detected in HI, so this is a good extension of the list. This can also help us estimate how many other AGES detections we expect to be UDGs or near-UDGs. We have $6 + 11$ good plus near UDGs from 163 objects, about 10%. For the whole published AGES sample to date (1231 objects), that means we potentially have ~ 120 objects already detected, just waiting to be classified. The full sample size will likely at least double this, suggesting that AGES is a very powerful resource for data mining for gas-rich UDGs.

Table 6.1: Properties of Ultra Diffuse Galaxies detections in the first 6 rows and UDG candidates in the rest of the table. $\mu_{0,g}$ and $\langle\mu_e\rangle_g$ are in units [mag arcsec^{-2}].

Name	$\mu_{0,g}$	$\langle\mu_e\rangle_g$	RA [deg]	Dec [deg]	D [Mpc]	R_e [kpc]
1237664131018719712	24.12	25.36	162.53718	13.48371	192.18	5.31
1237661948634661064	23.28	24.32	163.42932	10.76962	141.65	3.41
1237664131017474329	22.84	24.11	159.69394	12.98948	101.55	3.19
1237661950781751346	21.74	24.07	162.32637	12.37884	97.90	5.22
1237661068716409201	22.98	24.07	161.67341	13.00711	92.11	2.48
1237660670357471474	21.70	24.05	161.27921	11.48305	230.62	4.36
1237661949171663030	22.66	23.99	163.57304	11.22477	184.68	5.21
1237661949707550974	22.33	23.95	161.38297	11.29498	137.03	5.16
1237661950780440972	22.61	23.93	159.26640	12.07387	156.21	2.51
1237661950781424016	22.43	23.93	161.54677	12.26428	166.24	2.65
1237661950782013650	23.09	23.86	162.86814	12.43939	40.56	1.72
1237664131554083010	22.73	23.73	159.05329	13.44954	41.69	3.02
1237664132091085042	22.95	23.62	159.23674	13.93426	42.10	2.44
1237660670893883545	21.77	23.60	160.10536	11.99503	137.00	5.69
1237661950244815055	21.16	23.59	162.16246	11.90995	133.93	4.58
1237661950782013668	22.21	23.57	162.95283	12.24240	92.73	4.04
1237661069789233351	22.92	23.52	159.47066	13.71978	42.72	3.59

6.5 SMDG 1050090+132903

From all the potential UDGs we found, one galaxy is also identified in the SMUDGes catalog (see section 3.4) as SMDG 1050090+132903. In our catalog, this object corresponds to leo-gal-vol2-can_033, and its SDSS object ID is 1237664131018719712. We adopt the SMUDGes designation throughout this work.

Optical image from Legacy Survey and SDSS is in figure 3.3. Its renzogram and SDSS overlay from FRELLED is in figure 6.12 with the galaxy itself marked by the red circle. It is barely visible even on a screen. A clearer image from the Legacy Survey is in the figure. 6.13.

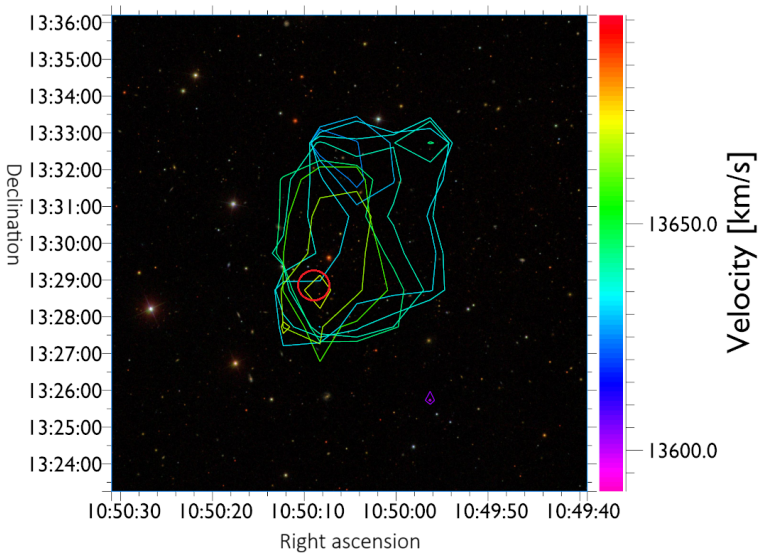


Figure 6.12: SMDG 1050090+132903: object matched with the SMUDGes database.

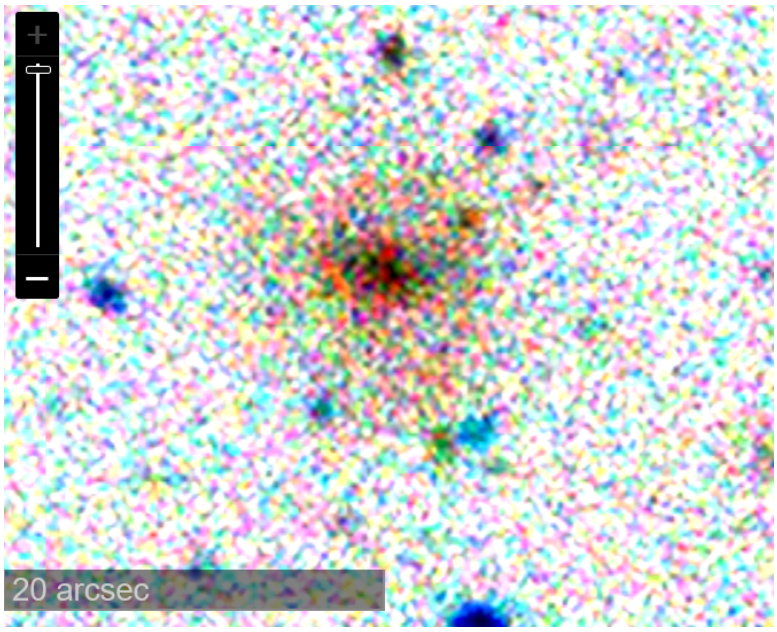


Figure 6.13: SMDG 1050090+132903: matched object with SMUDGes database, image from Legacy Survey.

This galaxy doesn't stand out as an outlier in mass-to-light, HI mass fraction, or color magnitude scaling relations, see figures 6.5, 6.6, and 6.7, where it is marked by a yellow star with a black edge. But it does have a higher M_{HI}/L_g or $M_{\mathrm{HI}}/M_\star$ than most other galaxies of comparable mass. This may argue towards UDGs being extreme examples of typical galaxies rather than having formed by some fundamentally different mechanism. They appear to be gas-rich, but not extraordinarily so.

It does stand out in the baryonic Tully-Fisher relations in figures 6.8, 6.9, and optical Tully-Fisher relations in figures 6.10 and 6.11. Various other UDGs have been found with anomalously low line widths. The implied lack of dark matter in these systems is interesting because this marks them out as radically different from conventional galaxies. This is why we have decided to investigate further.

The main problem here is the inclination that is in the SDSS estimated to be 36°, which is close to being face-on. This small inclination would be enough for it to be excluded from the TFR plot by itself because of the quality cuts. Moreover, even at first look this value is

hard to believe with any confidence. We have tried to fit an ellipse using DS9 to do the inclination estimate ourselves, and the possible values ranged from 0° (face-on) to approximately 40°. The SDSS estimate thus seems to be on the upper end. Since the inclination directly influences the rotational velocity estimate, even slight changes have a dramatic effect on the position in the TFR. When we estimate the inclination to 25°, it already puts this UDG right into the expected 2σ region of the BTFR. Setting the $i = 22°$ puts the galaxy inside the 1σ confidence level of the relation. A better observation would be needed to determine the inclination more precisely.

What is also interesting is the extended tail of the HI signal. While the HI emission is relatively weak, the coherent structure of the contours suggests this feature is likely real and not an artifact of noise. If so, this would be one of very few UDGs detected which are currently experiencing gas removal, as discussed more fully in R. Taylor et al. (2026a).

This galaxy is not currently a member of any cluster or group, but it could still have experienced an environmental effect in its past, e.g. an interaction responsible for the optical-HI offset. An ongoing interaction with another system is also possible since several other galaxies are visible nearby in projection, but they don't have a reliable redshift measurement, so we can't be sure.

As a result, we conclude that, for a lack of better data, we cannot say there is any significant deviation from the Baryonic Tully-Fisher relationship. A more detailed analysis and dedicated observation would be needed to make a verdict, but that is outside the scope of this project.

7 Conclusions

In this thesis, we learned about radio astronomy, in particular, about the intricacies of the neutral hydrogen observations using the Arecibo radio telescope. We delved into the world of galaxies, their morphology, and interactions, and we specifically described ultra diffuse galaxies. We then explored possibilities of utilizing relevant external data sources and also discussed challenges connected to processing and visualizing radio data.

One of the purposes of this thesis was to explore the unprocessed observation of a Leo field with no pre-selection criteria, use advanced visualization techniques, create a catalog of all the galaxies present and analyze the sample. For that, we have learned how to process the radio data and do the cataloguing with a FRELLED software.

We have constructed a catalogue of the galaxies present in our data cube. We also used the SDSS, HyperLeda, SMUDGes, Legacy Survey, and NED to extend our HI data with optical quantities.

With all that information, we investigated the observed galaxies and their environment. For that purpose, we have constructed a polar plot, HI deficiency diagram, and scaling relations (color magnitude diagram, HI mass fraction, mass to light ratio, baryonic and optical Tully-Fisher relationships). This analysis revealed unexpected uniformity of those scaling relations with no prominent trends or obvious outliers. Our relations are rather well-correlated and with low scatter, mostly following expected trends, with galaxies showing little evidence of any systematic trends in gas fraction or colour based on their morphology.

By performing our own photometry, we determined the number of ultra diffuse galaxies in our background volume, which resulted in the detection of six ultra diffuse galaxies, five of which are completely new, and none of them were previously detected in HI. Another 11 objects were marked as good UDG candidates. This is an important result because there are currently fewer than 300 UDGs known in HI.

Finally, there was one UDG with a large spatial offset between the HI and the optical because of an extended tail. The number of UDGs confirmed to be observed actually in the process of gas loss is very small, so this is important. We don't know that much about how they

evolve, and the moment of gas loss is crucial for understanding this. There is also a relatively large offset from the Tully-Fisher relationships, suggesting a lack of dark matter, potentially marking it out as radically different from other galaxies. Further analysis did not confirm this deviation being conclusive, a better observation would be needed to make a verdict. Recent discoveries of the UDGs emphasize the need for further investigations, in particular to see if other UDGs also show evidence of disturbed gas or if this one is exceptional.

Bibliography

Allawi, Y., Moneer, E., Shbat, M., Dyab, W., Jaser, M., & Alwadai, N. (2025). Design and optimization of a low-cost 5-m radio telescope at princess nourah university (pnu), saudi arabia. *Radio Science*, *60*. https://doi.org/10.1029/2024RS008170

Almeida, A., Anderson, S. F., Argudo-Fernández, M., Badenes, C., Barger, K., Barrera-Ballesteros, J. K., Bender, C. F., Benitez, E., Besser, F., Bird, J. C., Bizyaev, D., Blanton, M. R., Bochanski, J., Bovy, J., Brandt, W. N., Brownstein, J. R., Buchner, J., Bulbul, E., Burchett, J. N., … Zasowski, G. (2023). The eighteenth data release of the sloan digital sky surveys: Targeting and first spectra from sdss-v. *The Astrophysical Journal Supplement Series*, *267*(2), 44. https://doi.org/10.3847/1538-4365/acda98

Amorisco, N. C., Monachesi, A., Agnello, A., & White, S. D. M. (2018). The globular cluster systems of 54 coma ultra-diffuse galaxies: Statistical constraints from hst data. *Monthly Notices of the Royal Astronomical Society*, *475*(3), 4235–4251. https://doi.org/10.1093/mnras/sty116

Astronomy Magazine. (2024, October). *Arecibo telescope was doomed by hurricane damage and human failures, says report* [Accessed: 2025-12-16]. https://www.astronomy.com/science/arecibo-telescope-was-doomed-by-hurricane-damage-and-human-failures-says-report/

Auld, R., Minchin, R. F., Davies, J. I., Catinella, B., Van Driel, W., Henning, P. A., Linder, S., Momjian, E., Muller, E., O'Neil, K., Sabatini, S., Schneider, S., Bothun, G., Cortese, L., Disney, M., Hoffman, G. L., Putman, M., Rosenberg, J. L., Baes, M., … Spekkens, K. (2006). The arecibo galaxy environment survey: Precursor observations of the ngc 628 group. *Monthly Notices of the Royal Astronomical Society*, *371*(4), 1617–1640. https://doi.org/10.1111/j.1365-2966.2006.10761.x

Beasley, M. A., & Trujillo, I. (2016). Globular clusters indicate that ultra-diffuse galaxies are dwarfs. *The Astrophysical Journal*, *830*(1), 23. https://doi.org/10.3847/0004-637X/830/1/23

Bilir, S., Ak, S., Karaali, S., Cabrera-Lavers, A., Chonis, T. S., & Gaskell, C. M. (2008). Transformations between 2mass, sdss and bvri

photometric systems: Bridging the near-infrared and optical. *Monthly Notices of the Royal Astronomical Society*, *384*(3), 1178–1188. https://doi.org/10.1111/j.1365-2966.2007.12783.x

Boselli, A., Fossati, M., & Sun, M. (2022). Ram pressure stripping in high-density environments. *The Astronomy and Astrophysics Review*, *30*(1). https://doi.org/10.1007/s00159-022-00140-3

Buta, R. J. (2013). Galaxy morphology. In T. D. Oswalt & W. C. Keel (Eds.), *Planets, stars and stellar systems* (pp. 1–89, Vol. 6). Springer. https://doi.org/10.1007/978-94-007-5609-0_1

Ciccolella, A., & Leo, M. D. (2016). Hubble – de vaucouleurs galaxy morphology diagram [Wikimedia Commons, CC BY 3.0. Accessed: 2025-11-15]. https://en.wikipedia.org/wiki/File:Hubble_-_de_Vaucouleurs_Galaxy_Morphology_Diagram.png

Condon, J. J., & Ransom, S. M. (2016). *Essential Radio Astronomy*. Princeton University Press.

Cortese, L., Catinella, B., & Smith, R. (2021). The dawes review 9: The role of cold gas stripping on the star formation quenching of satellite galaxies. *Publications of the Astronomical Society of Australia*, *38*. https://doi.org/10.1017/pasa.2021.18

Deshev, B., Taylor, R., Minchin, R., Scott, T. C., & Brinks, E. (2022). The arecibo galaxy environment survey (ages): Xi. the expanded abell 1367 field: Data catalogue and hi census over the surveyed volume. *Astronomy and Astrophysics*, *665*, A155. https://doi.org/10.1051/0004-6361/202243103

Dey, A., Schlegel, D. J., Lang, D., Blum, R., Burleigh, K., Fan, X., Findlay, J. R., Finkbeiner, D., Herrera, D., Juneau, S., Landriau, M., Levi, M., McGreer, I., Meisner, A., Myers, A. D., Moustakas, J., Nugent, P., Patej, A., Schlafly, E. F., … Zhou, Z. (2019). Overview of the desi legacy imaging surveys. *The Astronomical Journal*, *157*(5), 168. https://doi.org/10.3847/1538-3881/ab089d

Ebrová, I., Bílek, M., & Eliášek, J. (2025). Photometric stellar masses for galaxies in DESI Legacy Imaging Surveys. A&A, *704*, Article A232, A232. https://doi.org/10.1051/0004-6361/202453448

Erkal, D., Belokurov, V., Laporte, C. F. P., Koposov, S. E., Li, T. S., Grillmair, C. J., Kallivayalil, N., Price-Whelan, A. M., Evans, N. W., Hawkins, K., Hendel, D., Mateu, C., Navarro, J. F., del Pino, A., Slater, C. T., Sohn, S. T., & Collaboration), (O. O. A. T.

(2019). The total mass of the large magellanic cloud from its perturbation on the orphan stream. *Monthly Notices of the Royal Astronomical Society*, *487*(2), 2685–2700. https://doi.org/10.1093/mnras/stz1371

Ewen, H. I., & Purcell, E. M. (1951). Observation of a line in the galactic radio spectrum: Radiation from galactic hydrogen at 1,420 mc./sec. *Nature*, *168*(4270), 356–356.

Giovanelli, R., & Haynes, M. P. (2015). Extragalactic hi surveys. *The Astronomy and Astrophysics Review*, *24*(1). https://doi.org/10.1007/s00159-015-0085-3

Giovanelli, R., Haynes, M. P., Kent, B. R., Perillat, P., Saintonge, A., Brosch, N., Catinella, B., Hoffman, G. L., Stierwalt, S., Spekkens, K., Lerner, M. S., Masters, K. L., Momjian, E., Rosenberg, J. L., Springob, C. M., Boselli, A., Charmandaris, V., Darling, J. K., Davies, J., … van Zee, L. (2005). The Arecibo Legacy Fast ALFA Survey. I. Science Goals, Survey Design, and Strategy. AJ, *130*(6), 2598–2612. https://doi.org/10.1086/497431

Graham, A. W., & Driver, S. P. (2005). A concise reference to (projected) sérsic r1/n quantities, including concentration, profile slopes, petrosian indices, and kron magnitudes. *Publications of the Astronomical Society of Australia*, *22*(2), 118–127. https://doi.org/10.1071/as05001

Haynes, M. P., & Giovanelli, R. (1984). Neutral hydrogen in isolated galaxies. IV. Results for the Arecibo sample. AJ, *89*, 758–800. https://doi.org/10.1086/113573

Haynes, M. P., Giovanelli, R., Martin, A. M., Hess, K. M., Saintonge, A., Adams, E. A. K., Hallenbeck, G., Hoffman, G. L., Huang, S., Kent, B. R., Koopmann, R. A., Papastergis, E., Stierwalt, S., Balonek, T. J., Craig, D. W., Higdon, S. J. U., Kornreich, D. A., Miller, J. R., O'Donoghue, A. A., … Wilcots, E. M. (2011). The Arecibo Legacy Fast ALFA Survey: The α.40 H I Source Catalog, Its Characteristics and Their Impact on the Derivation of the H I Mass Function. AJ, *142*(5), Article 170, 170. https://doi.org/10.1088/0004-6256/142/5/170

Hedberg, J. (2025). Galaxies — slide 18 [PHYS 45400 Introduction to Astrophysics, The City College of New York].

Hubble, E. P. (1926). Extragalactic nebulae. ApJ, *64*, 321–369. https://doi.org/10.1086/143018

Hubble, E. P. (1927). The classification of spiral nebulae. *The Observatory*, *50*, 276–281.

Hubble, E. P. (1925). Cepheids in spiral nebulae. *Pop. Astr.; Vol. 33; Page 252-255*, *33*.

Hughes, T. M., & Cortese, L. (2009). The migration of nearby spirals from the blue to red sequence: AGN feedback or environmental effects? MNRAS, *396*(1), L41–L45. https://doi.org/10.1111/j.1745-3933.2009.00658.x

Jones, M. G., Espada, D., Verdes-Montenegro, L., Huchtmeier, W. K., Lisenfeld, U., Leon, S., Sulentic, J., Sabater, J., Jones, D. E., Sanchez, S., & Garrido, J. (2018). The AMIGA sample of isolated galaxies. XIII. The HI content of an almost "nurture free" sample. A&A, *609*, Article A17, A17. https://doi.org/10.1051/0004-6361/201731448

Karttunen, H., Kröger, P., Oja, H., Poutanen, M., & Donner, K. (2003). *Fundamental astronomy*. Springer Berlin Heidelberg.

Keel, W. C. (2023). Galaxies [Course lecture notes, University of Alabama. Accessed: 2025-03-18].

Koda, J., Yagi, M., Yamanoi, H., & Komiyama, Y. (2015). Approximately a thousand ultra diffuse galaxies in the coma cluster. https://arxiv.org/abs/1506.01712

Kollmeier, J. A., Rix, H.-W., Aerts, C., Aird, J., Alfaro, P. V., Almeida, A., Anderson, S. F., Arranz, Ó. J., Arseneau, S. M., Assef, R., Aviram, S., Aydar, C., Badenes, C., Bandyopadhyay, A., Barger, K., Barkhouser, R. H., Bauer, F. E., Bender, C., Besser, F., … de J. Zermeño, R. (2025). Sloan digital sky survey-v: Pioneering panoptic spectroscopy. https://arxiv.org/abs/2507.06989

Kourkchi, E., Tully, R. B., Anand, G. S., Courtois, H. M., Dupuy, A., Neill, J. D., Rizzi, L., & Seibert, M. (2020). Cosmicflows-4: The calibration of optical and infrared tully–fisher relations. *The Astrophysical Journal*, *896*(1), 3. https://doi.org/10.3847/1538-4357/ab901c

Lang, K. (2013). *Essential astrophysics*. Springer Berlin Heidelberg. https://books.google.cz/books?id=0txukQEACAAJ

Liuzzo, E. (2018, April). Alma capabilities [Accessed 30-May-2023]. https://www.alma.inaf.it/images/Alma_bettina_UniTO2018.pdf

Makarov, D., Prugniel, P., Terekhova, N., Courtois, H., & Vauglin, I. (2014). HyperLEDA. III. The catalogue of extragalactic distances. A&A, *570*, Article A13, A13. https://doi.org/10.1051/0004-6361/201423496

Marr, J., Snell, R., & Kurtz, S. (2015). *Fundamentals of radio astronomy: Observational methods*. Taylor & Francis.

McGaugh, S. S. (2012). The Baryonic Tully-Fisher Relation of Gas-rich Galaxies as a Test of ΛCDM and MOND. AJ, *143*(2), Article 40, 40. https://doi.org/10.1088/0004-6256/143/2/40

Mo, H., van den Bosch, F., & White, S. (2010). *Galaxy formation and evolution*. Cambridge University Press.

Muller, C., & Oort, J. (1951). Observation of a line in the galactic radio spectrum: The interstellar hydrogen line at 1,420 mc./sec., and an estimate of galactic rotation. *Nature*, *168*(4270), 357–358.

Nan, R., LI, D., JIN, C., WANG, Q., ZHU, L., ZHU, W., ZHANG, H., YUE, Y., & QIAN, L. (2011). The five-hundred-meter aperture spherical radio telescope (fast) project. *International Journal of Modern Physics D*, *20*(06), 989–1024. https://doi.org/10.1142/s0218271811019335

Nulsen, P. E. J. (1982). Transport processes and the stripping of cluster galaxies. MNRAS, *198*, 1007–1016. https://doi.org/10.1093/mnras/198.4.1007

Paturel, G., Andernach, H., Di Nella, H., Durand, N., Garnier, R., Gouguenheim, L., Lanoix, P., Marthinet, M. C., Petit, C., Rousseau, J., Theureau, G., & Vauglin, I. (1997). Extragalactic database: Vii. reduction of astrophysical parameters. *Astronomy and Astrophysics Supplement Series*, *124*(1), 109–122. https://doi.org/10.1051/aas:1997354

Paturel, G., Petit, C., Prugniel, P., Theureau, G., Rousseau, J., Brouty, M., Dubois, P., & Cambrésy, L. (2003). HYPERLEDA. I. Identification and designation of galaxies. A&A, *412*, 45–55. https://doi.org/10.1051/0004-6361:20031411

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. (2002). Detailed structural decomposition of galaxy images. *The Astronomical Journal*, *124*(1), 266–293. https://doi.org/10.1086/340952

Roberts, M. S. (1962). The neutral hydrogen content of late-type spiral galaxies. AJ, *67*, 437–446. https://doi.org/10.1086/108752

Saintonge, A. (2007). The arecibo legacy fast alfa survey. iv. strategies for signal identification and survey catalog reliability. *The Astronomical Journal*, *133*, 2087–2096. https://doi.org/10.1086/513515

Sault, R. J., Teuben, P. J., & Wright, M. C. H. (1995, January). A Retrospective View of MIRIAD. In R. A. Shaw, H. E. Payne, & J. J. E. Hayes (Eds.), *Astronomical data analysis software and systems iv* (p. 433, Vol. 77). https://doi.org/10.48550/arXiv.astro-ph/0612759

Sérsic, J. L. (1963). Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy. *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina*, *6*, 41–43.

Schlafly, E. F., & Finkbeiner, D. P. (2011a). Measuring Reddening with Sloan Digital Sky Survey Stellar Spectra and Recalibrating SFD. ApJ, *737*(2), Article 103, 103. https://doi.org/10.1088/0004-637X/737/2/103

Schlafly, E. F., & Finkbeiner, D. P. (2011b). Measuring Reddening with Sloan Digital Sky Survey Stellar Spectra and Recalibrating SFD. ApJ, *737*(2), Article 103, 103. https://doi.org/10.1088/0004-637X/737/2/103

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. (1998). Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds. ApJ, *500*(2), 525–553. https://doi.org/10.1086/305772

Schneider, P. (2006). *Extragalactic astronomy and cosmology*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-33175-9

Stierwalt, S., Haynes, M. P., Giovanelli, R., Kent, B. R., Martin, A. M., Saintonge, A., Karachentsev, I. D., & Karachentseva, V. E. (2009). The arecibo legacy fast alfa survey. ix. the leo region h i catalog, group membership, and the h i mass function for the leo i group. *The Astronomical Journal*, *138*(2), 338–361. https://doi.org/10.1088/0004-6256/138/2/338

Taylor, E. N., Hopkins, A. M., Baldry, I. K., Brown, M. J. I., Driver, S. P., Kelvin, L. S., Hill, D. T., Robotham, A. S. G., Bland-Hawthorn, J., Jones, D. H., Sharp, R. G., Thomas, D., Liske, J., Loveday, J., Norberg, P., Peacock, J. A., Bamford, S. P., Brough, S., Colless, M., … Wijesinghe, D. (2011). Galaxy And Mass Assembly

(GAMA): stellar mass estimates. MNRAS, *418*(3), 1587–1620. https://doi.org/10.1111/j.1365-2966.2011.19536.x

Taylor, R., Minchin, R. F., Herbst, H., Davies, J. I., Rodriguez, R., & Vazquez, C. (2014). The Arecibo Galaxy Environment Survey - VII. A dense filament with extremely long H I streams. MNRAS, *443*(3), 2634–2649. https://doi.org/10.1093/mnras/stu1305

Taylor, R., Partík, V., & Minchin, R. (2026a). Ultra-diffuse galaxies in clusters: The peculiar gas loss of vcc 1964. *Astronomy & Astrophysics*, *707*, A107. https://doi.org/10.1051/0004-6361/202557543

Taylor, R. (2015). Frelled : A realtime volumetric data viewer for astronomers. https://arxiv.org/abs/1510.03589

Taylor, R. (2025a). Frelled reloaded: Multiple techniques for astronomical data visualisation in blender. https://arxiv.org/abs/2501.02919

Taylor, R. (2023). Photcalc.py [Accessed: 2026-03-20].

Taylor, R. (2025b). Quantifying the completeness and reliability of visual source extraction: An examination of eight thousand data cubes by eye. A&A, *696*, Article A113, A113. https://doi.org/10.1051/0004-6361/202451606

Taylor, R. (2012). The Virgo Cluster Through The AGES. *American Astronomical Society Meeting Abstracts #220*, *220*, Article 514.06, 514.06.

Taylor, R., Köppen, J., Jáchym, P., Minchin, R., Palouš, J., Rosenberg, J. L., Schneider, S., Wünsch, R., & Deshev, B. (2022). The arecibo galaxy environment survey. xii. optically dark h i clouds in the leo i group. *The Astronomical Journal*, *164*(6), 233. https://doi.org/10.3847/1538-3881/ac96e8

Taylor, R., Partík, V., & Minchin, R. (2026b). Ultra diffuse galaxies in clusters : The peculiar gas loss of vcc 1964. https://arxiv.org/abs/2601.05948

Thompson, A. R., Moran, J. M., & Swenson, G. W. (2017). *Interferometry and synthesis in radio astronomy* (3rd). Springer.

Toomre, A., & Toomre, J. (1972). Galactic Bridges and Tails. ApJ, *178*, 623–666. https://doi.org/10.1086/151823

van Dokkum, P. G., Abraham, R., Merritt, A., Zhang, J., Geha, M., & Conroy, C. (2015). Forty-seven milky way-sized, extremely diffuse galaxies in the coma cluster. *The Astrophysical Journal*

Letters, *798*(2), L45. https://doi.org/10.1088/2041-8205/798/2/L45

van der Burg, R. F. J., Muzzin, A., & Hoekstra, H. (2016). The abundance and spatial distribution of ultra-diffuse galaxies in nearby galaxy clusters. A&A, *590*, Article A20, A20. https://doi.org/10.1051/0004-6361/201628222

Willmer, C. N. A. (2018). The absolute magnitude of the sun in several filters. *The Astrophysical Journal Supplement Series*, *236*(2), 47. https://doi.org/10.3847/1538-4365/aabfdf

York, D. G., Adelman, J., Anderson, J. E., Jr., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouser, R., Bastian, S., Berman, E., Boroski, W. N., Bracker, S., Briegel, C., Briggs, J. W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Carr, M. A., … Yasuda, N. (2000). The sloan digital sky survey: Technical summary. *The Astronomical Journal*, *120*(3), 1579. https://doi.org/10.1086/301513

Zaritsky, D., Donnerstein, R., Dey, A., Karunakaran, A., Kadowaki, J., Khim, D. J., & Spekkens. (2023). Systematically measuring ultra-diffuse galaxies (smudges). v. the complete smudges catalog and the nature of ultradiffuse galaxies. *The Astrophysical Journal Supplement Series*, *267*(2), 27. https://doi.org/10.3847/1538-4365/acdd71