Characterization of gas-rich galaxies from ALFALFA's survey data

Author: Danae Mañé Fradera
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.[*]

Advisor: Josep Maria Solanes Majúa

Abstract: This research work is the first statistical study of the properties of the galaxies included in the final data release (A-100) of the Arecibo Legacy Fast ALFA survey (ALFALFA). The recently completed ALFALFA survey constitutes a representative sample of the nearby Universe in the 21 cm line of the neutral Hidrogen (HI), with more than 30 000 extragalactic sources out to $z \sim 0.06$. Its substantial overlap with wide-area surveys conducted at other wavelengths has been used to create a consistent multi-wavelength catalog of parameters for galaxies within about 250 Mpc ($\sim$ 800 million light-years) from us. We have characterized the gas-rich galaxies by comparing their content of atomic hydrogen with several properties from the optical window ranging from intrinsic luminosities and colors to estimators of the morphology and structure. Our results corroborate earlier findings that there does not appear to be a significant population of optically dark but HI-rich galaxies and that most HI-rich galaxies are very blue, while we find that their colors and HI masses are more strongly correlated than previously suspected. In addition, we have (not just confirmed but) found a clearer decrement of the bar fraction with increasing HI content, an anti-correlation that is not explained by known trends for more massive (stellar) and redder disc galaxies to host more bars and have lower gas contents.

I. INTRODUCTION

Hydrogen is the most abundant element in the Universe. It can be found in different phases: very hot ionized gas (coronal gas), hot ionized gas (HII regions), hot atomic gas, cold atomic gas (HI clouds) and cold molecular gas (H_2 molecular clouds).

The gaseous content of galaxies is very important for their stellar population. Stars are born from this gas and interact with it in many evolutionary phases. Abundant HI in a galaxy indicates potentially active star formation processes. Gas-rich galaxies can be studied thanks to the spectra of their gaseous components which modulate the shapes of the spectral lines.

This work studies the HI content and stellar properties of galaxies. Neutral hydrogen atoms are abundant in low-density regions of the interstellar medium. The $\lambda = 21$ cm spectral line allows to detect the neutral intergalactic hydrogen in atomic form. This line is due to a hyperfine splitting of the fundamental state of hydrogen atom. Two energy levels result from the magnetic inter-action between the quantized electron and proton spins. When the spins change from parallel (F = 1) to anti-parallel (F = 0), a photon of frequency $\nu = 1420.406$ MHz corresponding to a wavelength of 21 cm is emitted. From the 21 cm line, the total HI mass of a galaxy can be obtained from the following expression,

$$M_H = 2.356 \times 10^5 d^2 \int S_c(v) dv M_\odot,$$

where d is the distance in Mpc and $\int S_c(v)dv$ is the integral of the flux density over the line, with $S_c(v)$ expressed in Jy and v in km/s. It is also possible to obtain dynamic properties such the maximum rotational velocity which is typically measured from the half-width of a galaxy's integrated 21-cm line profile. The HI line is also very useful to determine extragalactic distances with high precision. In addition, it must be taken into account that the total content of HI in galaxies varies strongly with their morphology. Elliptical galaxies typically have less than 10^7 $M_\odot$, S0s can have up to 10^9 $M_\odot$, while Sbs and Scs usually exhibit atomic gas contents of several billion solar masses. Therefore, the relevance of the neutral hydrogen in galaxies increases towards later Hubble types.

For this study we have used the recently published catalog of the 100% complete Arecibo Legacy Fast ALFA (ALFALFA) survey [1], which consists of a total of 31 505 galaxies with properties measured from the 21 cm line. This A.100 catalog consist of 16 columns listing, among other parameters, the ALFALFA identifier, the right ascension and declination of the extragalactic source, its heliocentric velocity, cosmological distance, HI mass, and the observed velocity width of the source line profile at the 50% level of the two line peaks. This survey has used the Arecibo radio-telescope, the most sensitive radio-antenna at L-band (1.4 GHz) in the world. ALFALFA detects galaxies out to $z \sim 0.06$ and HI masses as low as 10^6 $M_\odot$ and as high as $10^{10.8}$ $M_\odot$ with positional accuracies typically better than 20 arcsec. The catalog covers the range of declination between 0 and 36 degrees, the range of right ascension between 0 and 360 degrees and the range of heliocentric velocity between -2000 km/s and 18 000 km/s. We have verified that more than 99% of the sources listed in the A.100 catalog have an extragalactic counterpart. This means that there are no large numbers of dark galaxies—galaxies with no visible stars, but that could be detectable if they contain significant amounts of neutral gas—in the local volume. At the same time,

[*]Electronic address: dmanefra7@alumnes.ub.edu

this also makes it possible to compare ALFALFA's 21-cm data with measurements at other wavelengths in order to determine the most defining characteristics of gas-rich galaxies and reveal correlations between their HI content and other properties that will help us improve our understanding of the way these objects form and evolve.

II. CROSS-MATCH WITH EXTERNAL CATALOGS

The combination of ALFALFA with other catalogs opens up amazing possibilities for extragalactic research. The matching of astronomical catalogs, however, is a complex and challenging problem, especially when the match involves surveys with a dense coverage of the sky. Next, we outline the main lines of the adopted cross-match procedure between the ALFALFA catalog and three other catalogs: the Sloan Digital Sky Survey (SDSS), which is a large source of optical properties [3], the Huertas-Company+18 catalog [4], which provides morphological information for the SDSS galaxies in the Legacy Survey, and the NASA SDSS ATLAS (NSA) [5], itself a cross-correlated galaxy dataset, which extends the SDSS measurements to the ultraviolet and the near-infrared.

To begin with, we have restricted the cross-matching to extended objects with spectral information, that have been identified as GALAXY by the SDSS pipeline, and that are located within the northern and southern galactic pole footprints of the ALFALFA survey (limiting z to values < 0.075). We have submitted SQL (Structured Query Language) queries to the SDSS database server (http://cas.sdss.org/dr14/en/tools/search/sql.aspx) to combine different Data Releases in order to put in a single catalog not only a wide range of optical properties but also the different identifiers used for each galaxy by the Sloan survey over the years. Then, we have used the TOPCAT tool [6] of the University of Bristol to match this dataset with the NSA catalog by imposing a maximum tolerance of 3 arcmin in the sky coordinates (right ascension and declination) and of 300 km/s in cz. The outcome has been cross-matched with the Huertas-Company+18 catalog using the DR7 objid identifier. Finally, this last dataset has been cross-matched with the ALFALFA catalog, for each hemisphere separately, using again TOPCAT in Sky + X mode, but now with the "All matches" condition activated because the maximum angular tolerance is set to 120 arcsec to account for the relatively large uncertainty of the ALFALFA pointing, while for cz the uncertainty is limited to 350 km/s. The result is run through a Fortran program, which for each ALFALFA object selects the SDSS galaxy with a redshift closest to that of the ALFALFA galaxy, within the 120-arcsec region defined around it. This program also imposes extra conditions on the apparent magnitudes (e.g. $u < 22$ and $r < 22.2$ mag, tolerance < 0.3 mag, etcetera) and colors to guarantee as much as possible that we are selecting the most probable counterpart. In addition to all this, we have removed from the final datasets galaxies with heliocentric velocity less than 2000 km/s, to exclude those objects with more uncertain both SDSS measurements and peculiar velocity corrections, as well as galaxies with cz in the range 15 000 km/s − 16 000 km/s, where the 21 cm measurements are strongly affected by RFI from the radar at the Arecibo airport. The results of all these cross-matches are a catalog for the northern hemisphere with 14 063 galaxies and another for the southern hemisphere with 864 galaxies.

III. ANALYSIS OF THE DATA

A. Spatial distribution and HI content

To begin the analysis of the data of our catalog resulting from the cross-matches, we represent the distribution of ALFALFA galaxies according to the cosmological distance. We have detected a dip at a distance of ~ 230 Mpc due to a loss in sensitivity of the ALFALFA survey caused by radio frequency interference (RFI) coming from the San Juan airport's radar. We have also detected a peak at about ~ 80 Mpc due to the Pisces-Perseus supercluster that is at this distance. We have also represented

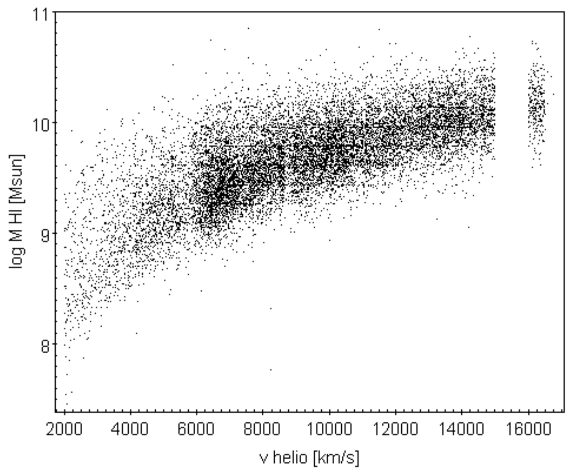

FIG. 1: HI mass vs. heliocentric velocity for the ALFALFA galaxies.

the Spänhauer diagram where we can study how the HI mass varies with the heliocentric velocity (Fig 1) or with the cosmological distance. We have found the discontinuity at about 230 Mpc ($\sim 15\,000$ km/s) produced by the RFI, that we had already detected in the distribution of distances. We have also detected a vertical line-up of points at a distance of 17 Mpc (~ 2000 km/s) due to the Virgo cluster and an important accumulation of points at 80 Mpc (~ 6000 km/s) due to the Pisces-Perseus supercluster as we have already commented. The distribution of points in the Spänhauer diagram shows how, as the distance increases, ALFALFA detects brighter galaxies, that is, with larger HI masses. This happens because ALFALFA survey is similar to a magnitude-limited survey in the optical window.

From the Spänhauer diagram and from the distribution of masses, we have found that most of the ALFALFA galaxies are in a HI mass range between 10^9 M$_\odot$ and 10^{10} M$_\odot$.

B. Colors and magnitudes

In this section, we compare the distributions of the A.100 and SDSS-NSA galaxies on the color-magnitude diagram, using (u-r) for the color and the absolute magnitude in r band, M_r. The galaxy color-magnitude diagram shows the relationship between absolute magnitude (a photometric tracer of the luminosity or mass of the galaxies) and their colors (a measure of the star formation rate; SFR). It is comparable to the Hertzsprung-Russell diagram for stars, although unlike them the properties of galaxies are not necessarily completely determined by their location on the diagram. In the case of galaxies, it shows a clear bimodal distribution with two main regions where data accumulate: the 'blue cloud' and the 'red sequence'. Most of the galaxies in the blue cloud are of late type (corresponding to the Spiral and Irregular Hubble types). Late-type objects are characterized by having active star formation (SFR$\gtrsim$ 1 M$_\odot$ yr^{-1}), therefore their light is dominated by young blue giant stars. The red sequence is delineated mostly by early-type galaxies (Elliptical and Lenticular Hubble types), which tend to have depressed or non-existent star formation, so their light comes from red giants and supergiants. Between these two distributions there is an underpopulated space known as the 'green valley', which includes a number of spirals with intermediate characteristics that are likely evolving towards the red sequence.

In Fig 2, we represent the color-magnitude diagram of the ALFALFA galaxies in our sample (blue dots). As expected, since they are mostly gas-rich late-type objects, most of them belong to the blue cloud. We can also see the color-magnitude diagram of the SDSS-NSA galaxies (red contours), where we can differentiate very well the red sequence and the blue cloud. The (u-r) color is a very good indicator of star formation activity that enables a clear separation of the galaxies. According to Strateva et al. (2001) [7], early-type galaxies show preferentially (u-r) colors redder than 2.22. We see that most of the ALFALFA galaxies have (u-r) colors which are below this divider, specifically we find that 92% of them are blue and only 8% are red (at the same depth, the fractions of blue/red objects in the NSA-SDSS galaxies are 70/30, respectively). Baldry et al. (2004) [8] adopted a more elaborated separator governed by the following equation that relates the (u-r) color to the absolute magnitude M_r

$$(u-r) = 2.06 - 0.244 \tanh\left(\frac{M_r + 21.07}{1.09}\right).$$

which we represent by the black curve in Fig 2. Again, we see that ALFALFA galaxies are predominantly located below the black line. According to the Baldry delimiter the blue cloud/red sequence ratio is 82/18. Regarding the absolute magnitude in r band, the bulk of the ALFALFA galaxies range between -21 $< Mr <$ -16.

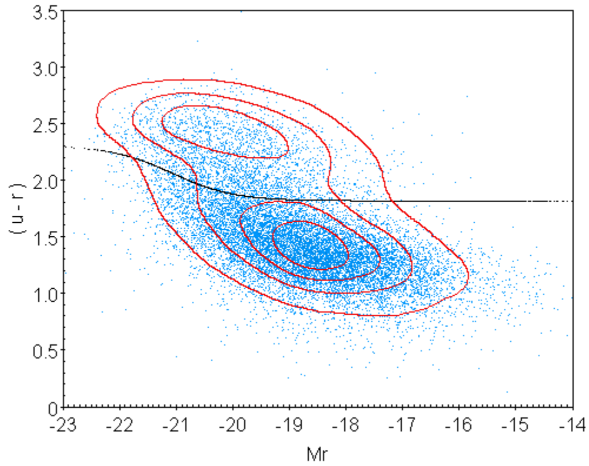

FIG. 2: Color-magnitude diagram: (u-r) vs. absolute magnitude in the r band for ALFALFA galaxies (blue dots) and for SDSS-NSA galaxies (red contours). The graphic incorporates the Baldry et al. delimiter (black line).

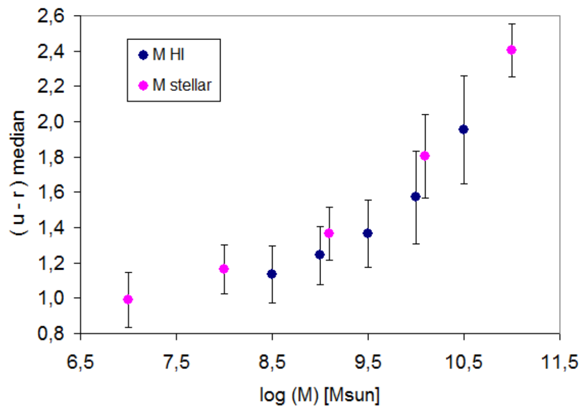

FIG. 3: (u-r) color median vs. HI mass (blue dots) and stellar mass (pink dots) for ALFALFA galaxies of our sample.

We have also investigated the dependence of the (u-r) color with HI mass and compared it with the much better known correlation between color and stellar mass. To our surprise we find that both masses vary in much the same way, as we can see in Fig 3. Nevertheless, this does not prevent the ratio between the masses of neutral gas and stars from being strongly and very tightly anti-correlated with color in agreement with earlier work [2]. Why u-band light from recent star formation correlates so well with atomic gas, when stars form in molecular gas, is a question that deserves a thorough examination.

C. Morphology

In this section we investigate the relation between morphology and neutral hydrogen content using different indicators. In the first place, we use as a proxy for morphology the inverse light-concentration index in the r-band, $C_{59,r}$ = R50/R90 that represents the ratio of the radii containing 50% and 90% of the light using Petrosian (i.e. distance-independent) fluxes. The value $C_{59,r} = 0.38$ represents a morphological separator in r-band [7]. Galaxies with $C_{59,r} < 0.38$ are early-type (E, S0, Sa) and galaxies with $C_{59,r} > 0.38$ are late-type (Sb, Sc, Irr). In Fig 4, we plot the bivariate distribution of inverse light-concentration index, $C_{59,r}$, and the (u-r) color for the ALFALFA and for the SDSS-NSA galaxies up to $z = 0.06$ (the reader must be aware that for this plot we are using the raw distributions for the nearly flux-limited ALFALFA and SDSS samples). It shows that the vast majority of the ALFALFA galaxies have a value of $C_{59,r}$ above 0.38, consistent with them being mostly late-type blue objects.

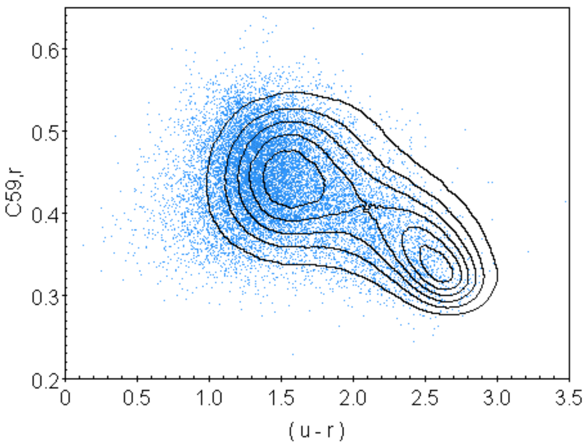

FIG. 4: Inverse light-concentration index in the r-band, $C_{59,r}$ vs. (u-r) color for ALFALFA galaxies (blue dots) and for SDSS-NSA galaxies (black contours).

Actually, it is possible to directly analyze the morphology-HI content relationship by using the automated morphological classification of the ∼ 700 000 SDSS galaxies belonging to the Legacy Survey recently published in H. Domínguez Sánchez et al. (2018) [4]. As shown in Table 1, almost all ALFALFA galaxies have a late-type classification, with 92% of them being Sa or later, and 8% being S0 [11] , while elliptical galaxies do not reach 0.1%.

[11] The fact that the fraction of S0 is larger than that of Sa, 6%, may be due to the difficulty of distinguishing in the SDSS images these two similar Hubble classes.

Hubble Type	E	S0	Sa	Sb	Sc	Sd
ALFALFA catalog	0.09	7.63	6.29	32.86	51.80	1.33
Domínguez et al. catalog	1.25	54.19	20.88	17.21	6.36	0.10

TABLE I: Percentages of galaxies for each Hubble type.

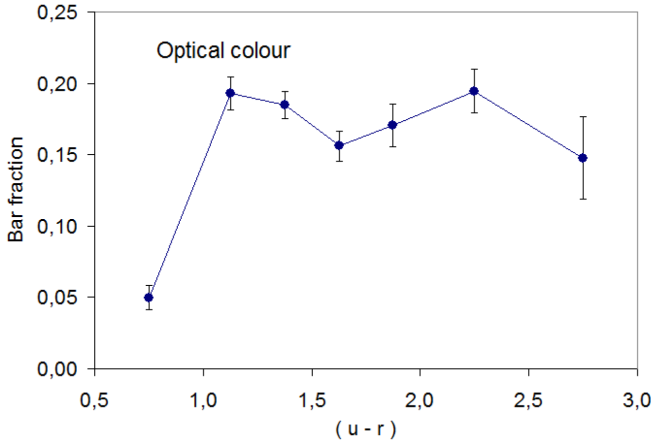

FIG. 5: Bar fraction vs. optical color.

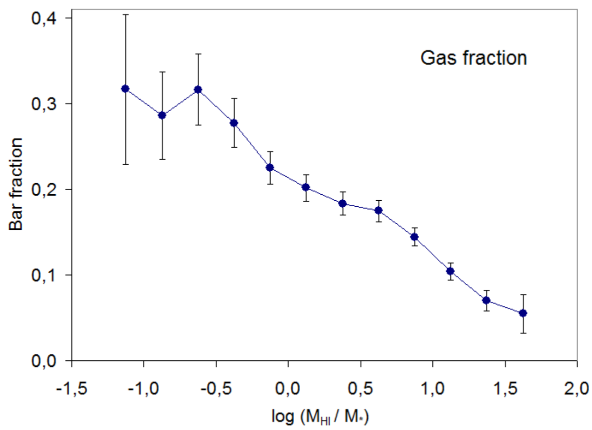

FIG. 6: Bar fraction vs. gas fraction.

We have also taken advantage of the fact the the H. Domínguez Sánchez+18 catalog gives two different estimates of the probability that a galaxy has a (strong) bar, one based on the Nair & Abraham (2010) [9] bar estimator and other on the guided visual classification of bars carried out by the Galaxy Zoo 2, to study the relationship between bar fraction and gas content. In the present work, however, we only discuss the results based on the Nair & Abraham's methodology since their are somewhat less noisy. We consider a galaxy is barred when its probability of having a bar is greater than 0.5. Consistently, with similar trends for barred galaxies seen recently both locally and at higher redshift [e.g. Karen L. Masters et al. (2012) [10]], we see a clear increase in the bar fraction with redder colors, decreased luminosity and in galaxies with more prominent bulges (see e.g. Fig 5). In contrast, we find that the bar fraction in HI-rich galaxies is nearly independent on color, though the result that calls the most attention is that this fraction is significantly lower among gas-rich disc galaxies than gas-poor ones (Fig. 6). This is not explained by known trends for more massive (stellar) and redder disc galaxies to host more bars and have lower gas fractions. Possible causal explanations of this finding are that bars cause the atomic gas of disc galaxies to be used up more quickly, or that increasing the atomic gas content in a disc galaxy may inhibit bar formation and, by a knock-on effect, reduce or halt star formation, reddening the global color.

IV. CONCLUSIONS AND FUTURE WORK

- By combining the recently published A.100 catalog of the ALFALFA survey with data from other wide area blind extragalactic surveys, we have compiled an extensive multi-wavelength (from the UV to the NIR) photometric (in narrow- and broad-band filters) and spectroscopic dataset for 14 460 HI-rich galaxies of the local universe ($z \lesssim 0.06$). In its present state, the catalog lists up a total of 51 entries per galaxy, including luminosities in several bands, sizes, inclinations, cosmic distances, total gaseous and stellar masses, morphologies, and bar fractions, plus a series of ID numbers that should facilitate the cross-match with other catalogs and the future additions of more parameters.
- We have verified, using different indicators, that the ALFALFA galaxies have overwhelmingly blue colors and that, accordingly, they are mostly located on the blue cloud region of the color-magnitude diagram.
- We have analyzed the median (u-r) color in different ranges of HI mass and stellar mass, finding very similar behaviors, with galaxies becoming redder with increasing both gas and stellar contents (the dependence is supralinear). In contrast, the ratio between the masses of neutral gas and stars is strongly anti-correlated with color. Given that u-band light indicates recent star formation and that r-band light is a very good tracer of dynamical mass, this result indicates that gas-rich galaxies are less efficient at forming stars and, likely, more dark matter dominated.
- The morphological analysis of the ALFALFA galaxies have revealed an unsurprisingly very large fraction ($< 90\%$) of spiral types. This analysis has been extended to the bar fraction, which we find is strongly anti-correlated with the HI content, thus corroborating results of earlier studies. Again this is a finding that may have important implications on how star formation is regulated in intermediate-mass disc galaxies.

The present statistical analysis shows only 'the tip of the iceberg' of the kind of interesting research on galaxies that can be done with our multi-wavelength dataset. In addition, some extra parameters such as local density and HI-mass deficiency estimators, will be added soon, further increasing the versatility of the catalog. It is hard to find a better tool to delve into how galaxies form and evolve.

Acknowledgments

I would specially like to thank my advisor, Prof. Josep Maria Solanes Majúa, for giving me the opportunity to do this project as well as for his continued professional guidance throughout it.

I would also like to thank my family and friends, for their support and encouragement during all these years.

[1] http://egg.astro.cornell.edu/index.php/
[2] M. C. Toribio et al., "H_I content and optical properties of field galaxies from the ALFALFA survey. I. selection of a control sample". The Astrophysical Journal, 732:92 (15pp), 2011 May 10.
[3] http://skyserver.sdss.org/dr7/en/
[4] H. Domínguez Sánchez et al., "Improving galaxy morphologies for SDSS with Deep Learning". Monthly Notices of the Royal Astronomical Society. MNRAS 476, 3661–3676 (2018). Advance Access publication 2018 February 9.
[5] http://www.nsatlas.org/data
[6] http://www.star.bris.ac.uk/ mbt/topcat/
[7] Iskra Strateva et al., "Color separation of galaxy types in the Sloan Digital Sky Survey Imaging Data". The Astronomical Journal, 122:1861–1874, 2001 October.
[8] Ivan K. Baldry et al., "Quantifying the bimodal color-magnitude distribution of galaxies". Astrophys. J. 600:681–694, 2004.
[9] Preethi B. Nair, Roberto G. Abraham, "A catalog of detailed visual morphological classifications for 14034 galaxies in the Sloan Digital Sky Survey", 14 Jan 2010.
[10] Karen L. Masters "GalaxyZoo and ALFALFA: atomic gas and the regulation of star formation in barred disc galaxies". Monthly Notices of the Royal Astronomical Society, Volume 424, Issue 3, 11 August 2012, 2180–2192.