

The Arecibo Galaxy Environment Survey (AGES)

XI. The expanded Abell 1367 field: Data catalogue and H I census over the surveyed volume^{★,★★}

Boris Deshev¹, Rhys Taylor¹, Robert Minchin², Tom C. Scott³, and Elias Brinks⁴

¹ Astronomical Institute, Czech Academy of Sciences, Boční II 1401, 14131 Prague, Czech Republic
e-mail: deshev@asu.cas.cz

² Stratospheric Observatory for Infrared Astronomy/USRA, NASA Ames Research Center, 232-12, Moffett Field, CA 94035, USA

³ Institute of Astrophysics and Space Sciences (IA), Rua das Estrelas, 4150-762 Porto, Portugal

⁴ Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK

Received 13 January 2022 / Accepted 16 June 2022

ABSTRACT

Context. Many galaxy properties are known to correlate with the environment in which the galaxies are embedded. Their cold, neutral gas content, which is usually assessed through 21 cm HI observations, is related to many other galaxy properties as it is the underlying fuel for star formation. With its high sensitivity and broad sky coverage the blind Arecibo Galaxy Environment Survey (AGES) brings significant improvement to the census of HI properties of galaxies in a wide range of environments, from voids to the core of a massive cluster. Here we present an HI census over a volume of $\sim 44\,000\text{ Mpc}^3$ towards the merging cluster Abell 1367 and extending well beyond.

Aims. We measure the effects that different environments have on the HI content of their constituent galaxies.

Methods. We use AGES- a deep, blind, HI survey carried out with the Arecibo radio telescope, which covers 20 square degrees on the sky centred on the merging cluster Abell 1367, mapping the large-scale structure (LSS) surrounding the cluster out to $cz = 20\,000\text{ km s}^{-1}$. The survey is sensitive down to a column density of $N_{\text{HI}} = 1.5 \times 10^{17}\text{ cm}^{-2}$ for emission filling the beam and a line width of 10 km s^{-1} . As an approximate mass sensitivity limit, a member of A1367 (at a distance of 92 Mpc), containing $M_{\text{HI}} = 2.7 \times 10^8 M_{\odot}$ distributed over a top-hat profile of 50 km s^{-1} width would be detected at 4σ . The results are analysed in combination with optical spectroscopy data from SDSS which we use to estimate the local galaxy density based on the Voronoi-Delaunay method.

Results. We present the results of the complete AGES survey of the A1367 field. In total, we detect 457 HI sources, 213 of which are detected for the first time by the AGES survey, and 134 of which are presented in this article for the first time. Of the 457 detections, 225 are in the cluster and 232 are in the remaining volume surveyed. Here we present the full catalogue of HI detections and their basic properties, including optical ones. We concentrate on the difference between the cluster and the foreground and background LSS, revealing a continuous correlation of HI-detected fraction (and HI deficiency) with local galaxy density, independent of global environment.

Key words. surveys – ISM: atoms – galaxies: evolution – galaxies: clusters: individual: A1367

1. Introduction

In recent decades, the availability of data from blind, all-sky surveys like the Sloan Digital Sky Survey (SDSS, [York et al. 2000](#)) and The Arecibo Legacy Fast ALFA Survey (ALFALFA, [Giovannelli et al. 2005](#)) have enabled an ever growing appreciation of the effects of environment on the evolution of galaxies ([Boselli & Gavazzi 2006, 2014; Boselli et al. 2021](#)). The presence of ram pressure (RPS, [Gunn & Gott 1972](#)), which actively removes gas from galaxies falling onto large clusters, has been unequivocally proven ([Kenney et al. 2004](#)) and the fine details of its workings have now been observed ([Arrigoni Battaia et al. 2012; Boselli et al. 2018; Jáchym et al. 2019](#)). This is now known to be an important factor in explaining the gas content

and star formation properties of cluster galaxies ([Cortese et al. 2021](#)).

On the other hand, the vast majority of galaxies do not reside in clusters, which have halo masses of $\log M_{\text{h}} \geq 14 M_{\odot}$, where RPS could realistically have detectable effects on a large galaxy's HI disc. Taking all galaxies with $z \leq 0.1$ from the Max Planck Institute for Astrophysics – Johns Hopkins University (MPA-JHU) value-added galaxy catalogue¹ based on SDSS shows that $\sim 40\%$ of them show little-to-no ongoing star formation. If we combine the same data set with the cluster and group catalogue of [Tempel et al. \(2017\)](#), we see that clusters with $M_{\text{h}} \geq 14$ host only around 8% of all galaxies. Understanding what determines the properties of non-cluster galaxies has proven much more difficult ([Cortese et al. 2021](#)).

For a long time it has been known that the closed box model for galaxy evolution does not match the observed metallicity of Milky way stars ([Schmidt 1963; Tinsley 1980](#)), which suggests a continuous accretion of low-metallicity gas from the

* Tables with properties of HI detections and their optical counterparts are only available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](ftp://cdsarc.u-strasbg.fr) (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/665/A155>.

** Throughout this article, clicking on the GitHub logo will open the Python script used for executing the calculation or producing the figure being discussed.

¹ https://www.sdss.org/dr12/spectro/galaxy_mpa_jhu/

circumgalactic medium. This is supported by the gas depletion timescales, which for the $z = 0$ galaxy population are significantly shorter than the Hubble time (Saintonge et al. 2017). Gas accretion must therefore be the dominant environmental effect that impacts all galaxies. The absence of accretion is termed starvation or strangulation and has been suggested as an environmental effect by Larson et al. (1980) and confirmed to be affecting most passive galaxies in the local Universe (Peng et al. 2015). This prevention of gas accretion leads to a slow decline in star formation rate until its complete cessation ~ 4 Gyr later.

Studies of the effects of environment use several different definitions of environment that are sensitive to the working of different physical processes. On one side is the global environment, which is usually the membership of the galaxies to a cluster or a group with a subdivision between central and satellite galaxies within those units. There are a number of catalogues of groups, clusters, and superclusters based on SDSS (e.g. Yang et al. 2007; Tago et al. 2010; Liivamägi et al. 2012; Tempel et al. 2017). The local environment is usually assessed by some measure of the local galaxy number density, be it projected on the sky or in 3D.

The cluster A1367 is an ongoing merger of multiple structures (Bechtold et al. 1983; Donnelly et al. 1998; Girardi et al. 1998; Sakai et al. 2002; Gavazzi et al. 2003a; Cortese et al. 2004). Its properties have been extensively studied and have revealed the presence of a plethora of physical processes related to cluster infall, such as merger shocks (Ge et al. 2019), X-ray tails (Ge et al. 2021), multiple jelly-fish galaxies (Scott et al. 2010, 2012; Roberts et al. 2021), merger induced star formation (Gavazzi et al. 2003a), ram pressure stripping and tidal interactions (Gavazzi et al. 2001; Scott et al. 2018), and preprocessing (Cortese et al. 2006). The cluster is also part of the ‘Great wall’ filament that connects it to the Coma cluster (Zabludoff et al. 1993).

For consistency with previous AGES publications, throughout this article we use standard Λ CDM with $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Thus, $1'' = 0.44 \text{ kpc}$ at the distance of A1367 (92.8 Mpc), and the full width at half maximum (FWHM) of the resampled Arecibo beam ($3.4'$) at the same distance is $\sim 90 \text{ kpc}$.

2. Observations and data reduction

The AGES used the seven-beam Arecibo *L*-Band Feed Array (ALFA) instrument which has a hexagonal footprint on the sky. The FWHM of the central beam is $\sim 3.3 \times 3.8'$. The observing strategy is described in detail in Auld et al. (2006), Cortese et al. (2008), and Davies et al. (2011). Briefly, AGES is a drift-scan survey, with the telescope driven to the start position and the sky allowed to drift overhead for the duration of the scan. Each scan therefore covers a swathe of right ascension (RA), with subsequent scans staggered by approximately one-third of a beam width ($\sim 1'$) in declination (Dec) to form a Nyquist-sampled map. A source above Arecibo takes 12 s to cross the beam. With seven beams in the ALFA instrument, each point is scanned 25 times, giving an on-source integration time of 300 s. Each beam integrates data from two polarisations, once per second, over 4096 channels, with a spectral resolution equivalent to 5.5 km s^{-1} per channel which is three-point Hanning smoothed to 10 km s^{-1} for greater sensitivity. The bandwidth spans the velocity range -2000 to $+20\,000 \text{ km s}^{-1}$ using the WAPP spectrometers. The output cube is the average of the two polarisations.

Observations for the A1367 field were taken in April–May 2006, April 2007, November 2011, January–February 2012,

May–June 2012, April–May 2013, December 2013, February–March 2015, May–June 2015, December 2016 – February 2017, June 2017, January–June 2018, and December 2018 – February 2019. Observations were prioritised to begin with a central 5×1 degree strip centred on the main body of the Abell 1367 cluster. These initial results were fully described in Cortese et al. (2008).

Here we present the full 5×4 degree field, spanning RA (11:33:32–11:54:39) and Dec (17:45:00–21:54:00). Each drift scan covered the full 5 degrees of RA (20 min of time). The data were reduced with LIVEDATA and GRIDZILLA (described in Barnes et al. 2001). LIVEDATA is used to measure and subtract the bandpass and GRIDZILLA combines the scans to produce the final data cube with a circular Gaussian beam with FWHM of $3.4'$. Following Taylor et al. (2014), we found it to be beneficial to further process the data in Python, typically applying a second-order polynomial to the full spectral baseline. This greatly improves the baseline fit (especially with regards to the effects of strong continuum sources) and, as the baseline is large and the polynomial is of low order, it does not introduce any artefacts that could be potential false-positive detections.

We also apply a MEDMED-style estimator to the spatial bandpass. As described in Putman et al. (2002), this measures the spatial bandpass (the scan in RA) in five equal-sized boxes, and subtracts the median of the medians. Although the median of the entire scan is usually an adequate measurement of the bandpass, particularly bright, extended sources can exaggerate the median value above the true value of the bandpass; therefore, using the median of the medians gives a more robust value (see Minchin et al. 2010 for an example; further discussion of the combined effects of this processing is given in Taylor et al. 2020).

From measurements of individual sources (see following section), the typical rms of individual Hanning smoothed spectra is ~ 0.6 – 0.7 mJy , equivalent to a 1σ column density sensitivity of $N_{\text{HI}} = 1.5 \times 10^{17} \text{ cm}^{-2}$ for emission filling the beam and with 10 km s^{-1} line width (Keenan et al. 2016). As an approximate mass sensitivity limit, at the 92.8 Mpc distance of the cluster (following Cortese et al. 2008 for consistency with the previous AGES analysis), a 4σ , 50 km s^{-1} velocity width source with a top-hat profile would have $M_{\text{HI}} = 2.7 \times 10^8 M_{\odot}$.

2.1. Source extraction

In this work, we have adopted a source-extraction strategy similar to all previous AGES papers. We catalogue detections using a combination of visual (by eye) and automatic techniques. The visual methods use the visualisation tool FRELLED (Taylor 2015). This is a two-stage process. First, FRELLED renders the data as a three-dimensional volumetric cube, allowing the user to interactively create masks to identify and measure potential sources. The advantage of a 3D rendering is that the user can instantly see the full extent of a source in both the spatial and spectral axes, making the masks easy to define. The disadvantage is that the volumetric rendering requires substantial clipping in order to minimise the visibility of the noise, making it unsuitable for finding the faintest sources. We therefore also use a second stage, whereby, after masking the brightest sources in 3D, we scan the cube as a series of 2D slices. This visualisation method does not require such heavy clipping but makes it more difficult (i.e. slower) to define the source masks. The combination of 3D and 2D inspection is therefore designed to optimise the speed and completeness for both bright and faint sources (further details are given in Taylor 2015). The whole identification process was performed independently by two of the authors

of the present article (BD and RT). Once complete, FRELLED uses the masks to generate input parameter files for the Miriad task MBSPECT (Sault et al. 1995), which we used to measure the spectral parameters of each source. These input parameters are included in the HI spectrum files available through the AGES online database².

All source-extraction methods suffer from limited completeness and reliability. We supplement our visual inspection with the automatic program GLADoS, which is described in Taylor et al. (2012). In brief, GLADoS searches each spectrum for pixels above a specified S/N threshold (4σ) which are contained within spectral features above a velocity width threshold (25 km s^{-1}). When a candidate is found in the standard averaged cube, GLADoS repeats the search in the same spectrum within each polarisation. As the noise level is slightly higher in the separate polarisations, the S/N threshold is slightly softened to 3.5σ and additional Hanning smoothing is applied to each individual cube.

We re-wrote the original GLADoS code in Python and added substantial new functionality (to be described in detail in a future paper). One challenge for automatic algorithms, as described quantitatively by Taylor et al. (2013), is their low reliability; it is no advantage that the process is automatic if the initial candidate source list numbers are in the tens of thousands. GLADoS mitigates this through the use of multiple polarisations, and in the new version, by a GUI which allows the user to set the candidate status (accept, reject, or uncertain) interactively. This allows a much more rapid inspection and re-inspection of the automated catalogue. In addition, as with FRELLED, it also outputs an optical image of the SDSS field at the HI coordinates alongside the HI spectrum. Although we would prefer the HI identification to be entirely independent of the optical information, in practise the presence of an optical galaxy at the position of an HI candidate source (even a weak one) has been found to significantly improve the probability that the candidate is real. Finally, we further reduce the burden of the automatic extractor by first applying the masks generated by the visual search, so that GLADoS avoids repeating the detections already found visually.

All sources extracted with both methods are measured using MBSPECT, assuming they are unresolved. A complete list of the 457 HI detections and their measured properties is given in the tables at the CDS. We assign a descriptive flag to all sources in the catalogue ranging from 0 to 4. These are based on visual inspection performed individually by each member of the collaboration. Inspected were the noise pattern in the data cube around each detection, as well as its spectrum. The presence or absence of an optical counterpart was not part of the flagging process as we did not want to introduce bias into the HI catalogue stemming from the limitations of the available optical catalogues. The flags are as follows:

- 0 – certain detection without any additional flags;
- 1 – uncertain detections. Those are usually low-S/N sources or sources detected close to a region with known Radio-Frequency Interference (RFI). All other flags indicate certain detections;
- 2 – indicates HI detections with more than one optical galaxy likely contributing to the measured signal;
- 3 – the source is at the edge of the cube and part of the flux is likely missed;
- 4 – indicates sources that are likely resolved by the $3.4'$ beam of Arecibo, and therefore the tabulated total flux and M_{HI} should be treated as lower limits.

Table 1. Breakdown of the HI catalogue by method of detection and flag.

Extraction method	Flag					Total
	0	1	2	3	4	
FRELLED	358	9	19	5	21 ^(a)	412
GLADoS	9	35	0	1	0	45
Total	367	44	19	6	21	457

Notes. ^(a)9 objects have flag 4 in addition to their primary flag. Only primary flags are in the table.

The flags are listed in the tables at the CDS. A breakdown of the catalogue by method of detection (visual/automatic) and flag is given in Table 1. The visual inspection was based on an atlas presenting the sources and their parameters. Example of this atlas, showing sources with a range of S/N and flag, is shown in Appendix A. The full atlas is available for download³ (0.5 GB).

Following the above search for detections, an HI spectrum was extracted at the position of every SDSS spectroscopic target within the AGES volume. Those were visually inspected but did not yield any additional detections.

In previous works, such as Cortese et al. (2008), we used Arecibo to perform follow-up observations to verify uncertain detections. During the course of this analysis, this became impossible; however, we obtained 5 h of observing time on the Five-hundred-metre Aperture Spherical Telescope (FAST) to re-observe uncertain detections. As we are concerned with the large-scale distribution of sources rather than individual detections, we leave the analysis of these observations to a subsequent paper and omit uncertain detections from further analysis.

2.2. Optical counterparts

We searched for optical counterparts of the HI detections among the galaxies with SDSS spectroscopic redshifts. The search region around the centre of each HI detection was $1.7'$ (FWHM/2) projected on the sky, and in velocity range $|V_{\text{HI}} - V_{\text{opt}}| \leq W_{20}/2 + \Delta W_{20}$, where W_{20} is the velocity width of the line profile at 20% of the peak intensity and ΔW_{20} its uncertainty. Similar criteria were used in all previous AGES papers. This method confirmed the optical counterparts of 322 HI detections, with a median distance between the centroid of the HI emission and the centre of the optical counterpart of $12.5''$ and velocity difference of 7 km s^{-1} . When more than one galaxy satisfies those criteria, the closest one on the sky is listed as the primary optical counterpart. Those galaxies have flag = 2. Many of the resolved galaxies with flag = 4 also have more than one optical counterpart, and therefore the total number of HI-detected galaxies in this field is likely significantly higher than 457. Additional searches for optical counterparts were carried out in the NASA/IPAC Extragalactic Database (NED), yielding a further 13 objects with spectroscopic redshift, bringing the total to 335. In addition, the table at the CDS lists the most likely optical counterpart for 106 of the 122 remaining HI sources. Those galaxies lack optical redshifts but lie within the HI beam. We could not find optical counterparts for 16 of our HI detections. Those are almost all low S/N detections, 13 of which are flagged as uncertain (flag = 1).

³ https://www.dropbox.com/s/xffshfom5h0m3f0/AGES_A1367_atlas.tar.gz?dl=0

² <http://www.naic.edu/~ages/public/>

Some additional properties of the optical counterparts of the HI detections needed for the analysis presented in the following sections were taken from publicly available data sets. The stellar mass is from MPA-JHU value-added catalogue based on SDSS DR8 and assume a (Chabrier 2003) initial mass function. The details of the methods used to derive the stellar mass are given in Kauffmann et al. (2003), Tremonti et al. (2004), and Brinchmann et al. (2004). The catalogue provides stellar mass estimates for 296 of the 335 optical galaxies which we detect in HI. Where relevant, this is the sample used in the following analysis. The r -band size of the stellar discs is from SDSS (Ahumada et al. 2020).

2.3. Optical sample and upper limits on M_{HI} for non-detections

SDSS provides, through the MPA-JHU catalogue, the precise positions, redshifts, and stellar masses for 830 galaxies within the AGES survey volume. We use this as our optical sample in the following analysis. Of those, 296 objects are directly detected in HI and catalogued. For the remaining 534, we need to estimate the maximum amount of HI that could be hosted by them without being detected. We base our estimate on the rms noise measured per pixel in the data cube. Assuming that those galaxies are not spatially resolved, we need to estimate the velocity width of their signal. We do that using the i -band Tully–Fisher relation (TF, Tully & Fisher 1977) published by Masters et al. (2006) and SDSS i -band Petrosian magnitudes. A two sample Kolmogorov–Smirnov test between the profile widths of non-detections, derived as described above, and the measured W_{50} of HI-detected galaxies returned a p -value of 4×10^{-6} , indicating that the two samples can be drawn from the same parent distribution.

3. Results

The complete list of all HI detections and their measured properties- which are used in the following analysis- is given in the tables at the CDS. The list with the optical counterparts of the HI detections and their properties is also provided.

The distribution of measured total HI mass (M_{HI}) and overall distribution of all AGES detections as a function of redshift z are shown in Fig. 1 in comparison to the distribution of SDSS spectroscopic galaxies and ALFALFA detections (Haynes et al. 2018) within the AGES survey volume. The AGES catalogue contains 457 HI detections, 413 of which are considered certain, almost doubling the source count compared to ALFALFA (212). The density of HI detections follows the distribution of galaxy concentrations in velocity as traced by SDSS from optical spectroscopy. Also visible is the relative absence of HI detections in ALFALFA at high- z .

We have a number of HI detections for which there is no spectroscopically confirmed optical counterpart (even if most have an obvious candidate, which are also listed in the tables at the CDS), and there are many optical galaxies without a corresponding HI detection. At the high- z end, the AGES survey contains $\sim 50\%$ fewer galaxies than SDSS. This fraction rises steadily towards the low- z end of the survey, with a noticeable decrease in HI detections at the velocity range of A1367. The growth of the fraction of HI-detected galaxies towards lower redshift is explained by the higher HI mass sensitivity shown in the bottom panel of Fig. 1. The black dashed curve shows the expected M_{HI} of a galaxy with $W_{20} = 200 \text{ km s}^{-1}$ detected at 6.5σ . The fraction of optical galaxies detected in HI is analysed

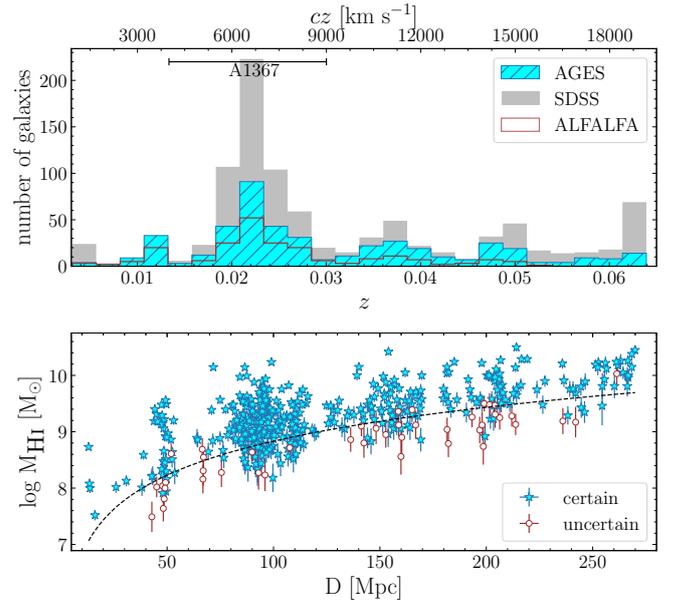


Fig. 1. *Top:* redshift distribution of the AGES HI detections (cyan hatched histogram) compared to SDSS spectroscopic galaxies (solid grey histogram), and ALFALFA detections (empty brown histogram) within the AGES volume. The velocity range of A1367 is indicated with the horizontal bar. *Bottom:* distribution of M_{HI} as a function of distance to the source, assuming Hubble flow distances. The black, dashed line shows the mass of a galaxy with velocity width of 200 km s^{-1} detected at 6.5σ . Certain and uncertain detections are shown with blue stars and brown circles, respectively. ☺

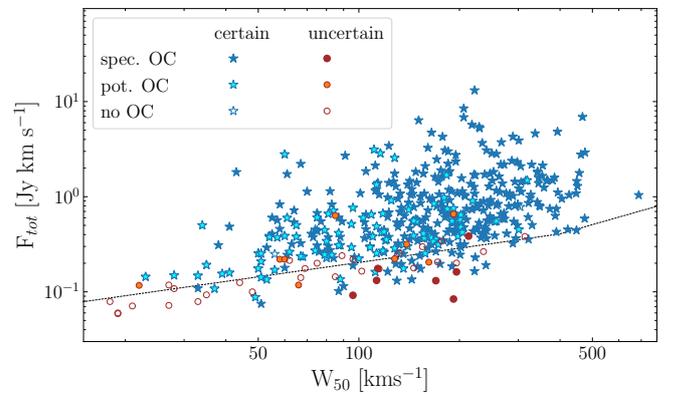


Fig. 2. Distribution of total flux in the 21 cm line as a function of line width for the certain detections (stars) and uncertain ones (circles). Filled and empty symbols show detections with and without an optical counterpart (OC), respectively, with the different shade fill segregating the spectroscopically confirmed and the potential optical counterparts. The dashed black line shows the ALFALFA reliability limit (Saintonge 2007) at 6.5σ . ☺

in detail in Sect. 4. The percentage of sources below the 6.5σ limit from the five separate flags listed in Table 1 is 10, 77, 5, 0, 0, respectively (qualitatively this information is also conveyed in Fig. 2).

Figure 2 shows the distribution of the 457 HI detections in the plane defined by their total flux in the line and the width of the line measured at 50% of the maximum of the flux (W_{50}). The sources are split into certain and uncertain ones, according to their detection flag. Also indicated in the table is the presence or absence of a spectroscopically confirmed or photometric

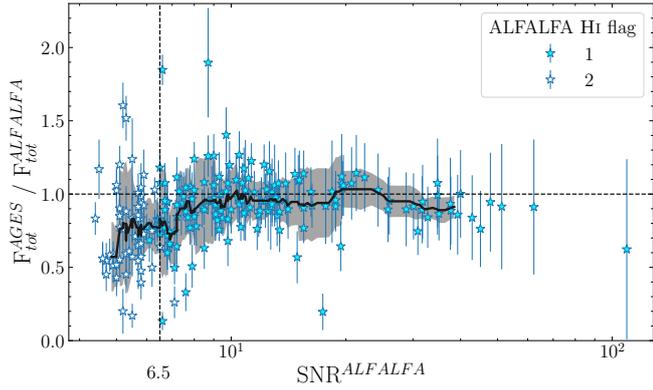


Fig. 3. Comparison between the measured total line flux of HI sources common to both ALFALFA and AGES surveys. The flux ratio is plotted as a function of the ALFALFA S/N. The black curve and shaded region represent a running median and $\pm 1\sigma$. The detection reliability limit is indicated with the vertical dashed line. ☹

optical counterpart. The dashed black line shows the distribution of galaxies at constant $S/N = 6.5$ following [Saintonge \(2007\)](#) and [Cortese et al. \(2008\)](#). As concluded by the latter, this is a reliable limit for separating the certain and uncertain detections. Of all AGES detections, 13% (72) are below this line, 47% (34) of which are flagged as uncertain. The median uncertainty in the total flux in the line for all AGES sources is $\Delta F \sim 0.1 \text{ Jy km s}^{-1}$. Within one ΔF above and below the 6.5σ line, the fraction of uncertain detections changes from 0 to $>50\%$. As pointed out by [Cortese et al. \(2008\)](#), the lack of sources at the high-velocity-width, low-flux regime is a result of the fact that weak sources have a better chance of being detected if their flux is distributed over fewer resolution elements, which gives preference to face-on oriented galaxies. This is reflected in the slope of the dashed black line.

Figure 3 shows the ratio of the total flux in the line measured from the AGES and ALFALFA surveys for the 212 HI sources common to both surveys. This is presented as a function of the S/N as measured by the ALFALFA survey and published by [Haynes et al. \(2018\)](#). The error bars are a quadrature sum between the tabulated uncertainties and those published by [Haynes et al. \(2018\)](#). The thick black curve and shaded area represent a running median over a window of 15 sources ($\sqrt{N} = 15$, where $N = 212$ is the total number of sources common to the two catalogues). The vertical dashed line shows the reliability limit for detecting an HI source in the ALFALFA survey. ALFALFA offers two quality flags on the published HI flux, based on S/N and the availability of optical counterpart, which are indicated with different symbols. As a result of the greater integration time employed by AGES, none of these sources are flagged as uncertain. The increased scatter at low S/N is expected ([Haynes et al. 2011](#)). The systematic drop in the flux ratio at low S/N can be explained as a manifestation of Eddington bias ([Eddington 1913](#)) where faint sources close to and below the detection limit are only detected if their flux is artificially boosted by fortuitous noise distribution. AGES detects those same sources at higher S/N and shows lower, more accurate flux.

3.1. Local galaxy density ☹

The top two panels of Fig. 4 show the distribution of M_{HI} for all AGES detections and the stellar mass for all galaxies in the MPA-JHU catalogue within the AGES volume. The white areas

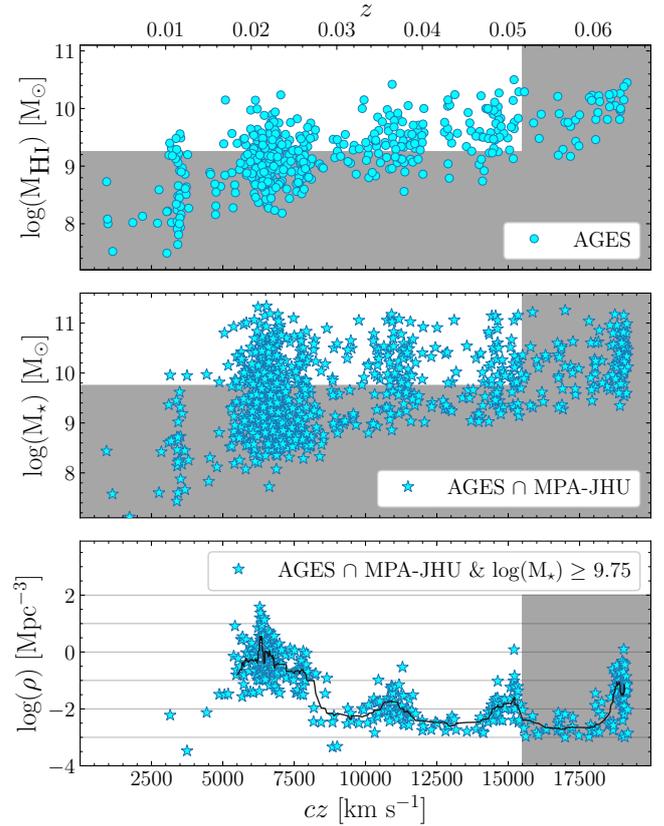


Fig. 4. Distribution of HI mass for all AGES detections (*top*) and stellar mass for all optical galaxies in the intersection between AGES and MPA-JHU (*middle*) as a function of recession velocity. The shaded regions indicate the areas not complete in M_{HI} and M_* , respectively. *Bottom panel*: local galaxy density for the mass-complete part of the middle sample. The black line shows the running median. ☹

on those two panels define the HI mass and stellar mass complete samples, used below to calculate the local density of galaxies. The limits of $M_* \leq 9.75 M_\odot$ and $M_{\text{HI}} \leq 9.25 M_\odot$ are defined as the peaks of the corresponding mass distributions in a narrow redshift bin centred at $15\,500 \text{ km s}^{-1}$. In the following analysis, we did not use the higher redshift part of the survey due to the presence of RFI.

The local galaxy density is calculated from the volume of a Voronoi cell centred on every galaxy and with a size equal to half the distance to the nearest galaxy in all directions, similar to the Voronoi-Delaunay method ([Voronoi 1908](#); [Delaunay et al. 1934](#); [van de Weygaert 1994](#); [Marinoni et al. 2002](#); [Pereira et al. 2017](#)). This is calculated in a 3D space of comoving physical distances for the mass-complete sample, and is therefore only available for galaxies with $\log M_* \geq 9.75 M_\odot$. In this case, we used the redshift to calculate the distance to the objects, including cluster galaxies. This allows us to calculate and compare the local galaxy density in the same way throughout the surveyed volume. Estimating distance from redshift for virialised structures is affected by the Fingers-of-God (FoG effect, [Jackson 1972](#); [Tully & Fisher 1978](#)) and Kaiser ([Kaiser 1987](#)) effects. The two work in opposite directions, and in general the Kaiser effect is expected to be weaker than FoG ([Kuutma et al. 2017](#)). By comparing the distribution on the sky of SDSS members of A1367 with their distribution along the line of sight, both approximated with a Gaussian, we determined that the cluster is stretched along the latter by a factor of 3.7. Therefore, we reduce

all cluster-centric distances for A1367 members along the line of sight by this amount. We note that this does not completely remove the FoG effect, which is impossible without precise redshift independent distances; instead it minimises the artificial lowering of 3D densities in the cluster.

The calculation of the Voronoi tessellations and their volumes was done in Python using SCI`PY` packages `VORONOI` and `DELAUNAY`. The distribution of local density is shown in the bottom panel of Fig. 4, and further analysed in Sect. 4.

3.2. Observed phase space view of the surveyed volume

Figure 5 presents the HI detections in the context of the large-scale structure in which they are embedded, where the surveyed volume is shown with its true axis ratio. The recession velocity range has been split in two for clarity. The volume is represented in two planes, collapsing the Dec or RA ranges, respectively. The figure shows the local density of optical galaxies from the stellar-mass-complete sample, colour coded according to the colour bar. The density clearly shows the position of the cluster A1367, centred at around 6500 km s^{-1} , but also a number of other structures with a range of densities. The potential optical counterparts are shown in Fig. 5 as small brown points plotted at their measured RA or Dec, and the recession velocity of the HI signal. The optical galaxies with stellar mass below the completeness limit, for which local density is not available, are shown as small black points. The figure also shows the areas of the survey affected by RFI or with increased noise levels at the edges of the data cube. Those areas are marked with light-grey diamonds and show RA or Dec columns with rms noise deviating from the mean rms of the cube by more than 4.5σ . We note that in some cases (particularly with the presence of RFI), this does not need to affect the entire column, hence the presence of some HI detections in an area marked in grey. However, the plot clearly shows entire redshift bands where HI detections are almost completely absent due to RFI, namely around $16\,000$ and $17\,000 \text{ km s}^{-1}$, but also a narrow stripe at around 7300 km s^{-1} running through the cluster.

3.3. $H\text{I}$ deficiency

It is instructive to compare the amount of HI-detected in individual galaxies with a baseline average of the entire population in order to assess the presence of any physical mechanisms that can alter the gas content of galaxies, be it through gas removal or through prevention of accretion. We employed a formula published by Haynes & Giovanelli (1984) which links the total gas mass to the size of the optical disc of the host galaxy.

$$\text{DEF}_{\text{HI}} = \log_{10} \left(\frac{M_{\text{HI}}^{\text{expected}}}{M_{\odot}} \right) - \log_{10} \left(\frac{M_{\text{HI}}^{\text{observed}}}{M_{\odot}} \right). \quad (1)$$

This relies on the relatively tight correlation between the size of the HI disc and its total mass on one hand and the size of the optical disc on the other. Haynes & Giovanelli (1984) provided two separate formulas to be used with galaxies with early- and late-type morphologies, respectively. We found 149 galaxies among the HI-detected sample with optical counterparts for which the Galaxy Zoo project contains morphologies determined with certainty (Lintott et al. 2011). Of those, only 7 were classified as elliptical. We therefore decided to only use the formula for spiral galaxies when calculating HI deficiencies. In particular, we took the Petrosian radius, which contains 90% of their r -band luminosity. In order to bring the Petrosian radius offered by the SDSS pipeline into agreement with the

size estimates used by Haynes & Giovanelli (1984), we applied a correction calculated by comparing petroR90 with the radii of galaxies measured at surface brightness of $25 \text{ mag arcsec}^{-1}$ (D_{25}) by the GOLDmine project (Gavazzi et al. 2003b). We found 47 galaxies in common between the Goldmine data base and the catalogue of optical counterparts of the HI detections presented here. The ratios of the two size estimates were normally distributed with a mean of 1.4 and standard deviation of 0.3. Thus, the expected HI mass of all optical galaxies is: $M_{\text{HI}}^{\text{expected}} = 7.195 + 0.851 \times \log(2 \times \text{PR90} \times 1.4)^2$, where PR90 is the Petrosian radius in kiloparsecs. The deficiency is defined as the difference in the logarithms of the expected and observed M_{HI} .

The results are presented in Fig. 6 and analysed further in Sect. 4.2. Figure 6 shows all HI detections with optical counterparts demonstrating the known, relatively tight relation between the HI mass fraction, M_{HI}/M_{\star} , and M_{\star} . All points are colour coded according to their HI deficiency calculated as described above. The small grey points and down-pointing arrows show the upper limit on gas fraction for galaxies not directly detected in HI, calculated as explained in Sect. 2.3. Also shown is the relation from the xGASS survey, published by Catinella et al. (2018), which is to be compared with the median of the AGES data shown with the black line and error bars.

4. Discussion

The part of the AGES survey presented here has an on-the-sky extent of $5 \times 4^{\circ}$, which corresponds to 18.6×23.2 comoving megaparsecs, at the distant end of the survey. In contrast, the distance between the nearest and most distant HI detections presented in this work is 257 Mpc. This large distance defines the significant variation in sensitivity over the surveyed volume. For example, with the very stable noise level across the volume, a single resolution element with 3σ signal would be equivalent to $\log M_{\text{HI}} = 6.0 M_{\odot}$ at the near end of the survey and 8.6 at the far end. This large difference in sensitivity is expressed with the black dashed curve on Fig. 1. For this reason, we analysed the surveyed volume in three bins which also represent different environments. The low- z foreground part of the cluster A1367 spans up to $cz = 4000 \text{ km s}^{-1}$ and has a total volume of 361 Mpc^3 and average 6.5σ sensitivity for a galaxy with 200 km s^{-1} velocity profile of $\log M_{\text{HI}} = 7.8 M_{\odot}$. The second part contains the cluster A1367 and is defined as the volume between $4000 \leq cz \leq 9000 \text{ km s}^{-1}$, with a total volume of 3719 Mpc^3 and an average sensitivity of $\log M_{\text{HI}} = 8.7 M_{\odot}$. The last part lies at $cz > 9000 \text{ km s}^{-1}$ and has a total volume of $39\,940 \text{ Mpc}^3$ and an average sensitivity of $\log M_{\text{HI}} = 9.4 M_{\odot}$.

4.1. $H\text{I}$ below the detection threshold

Taking the reliability threshold shown in the bottom panel of Fig. 1, namely 6.5σ for an HI profile over 200 km s^{-1} , means that we are sensitive to $M_{\text{HI}} \leq 9$ and ≤ 8 over 10% and 4% of the total surveyed volume of $44\,021 \text{ Mpc}^3$, respectively. The ‘by eye’ strategy for source finding that we adopted here, plus the occasionally more favourable velocity distribution of the HI signal, allows us to find a significant number of detections lying below this threshold – $\sim 16\%$ of all detections, with approximately half of those considered certain. Despite this, we expect that some amount of HI emission is present in the data cube at very low S/N and is therefore not catalogued. Traditionally, the HI community uses spectral stacking to recover such emission lying at

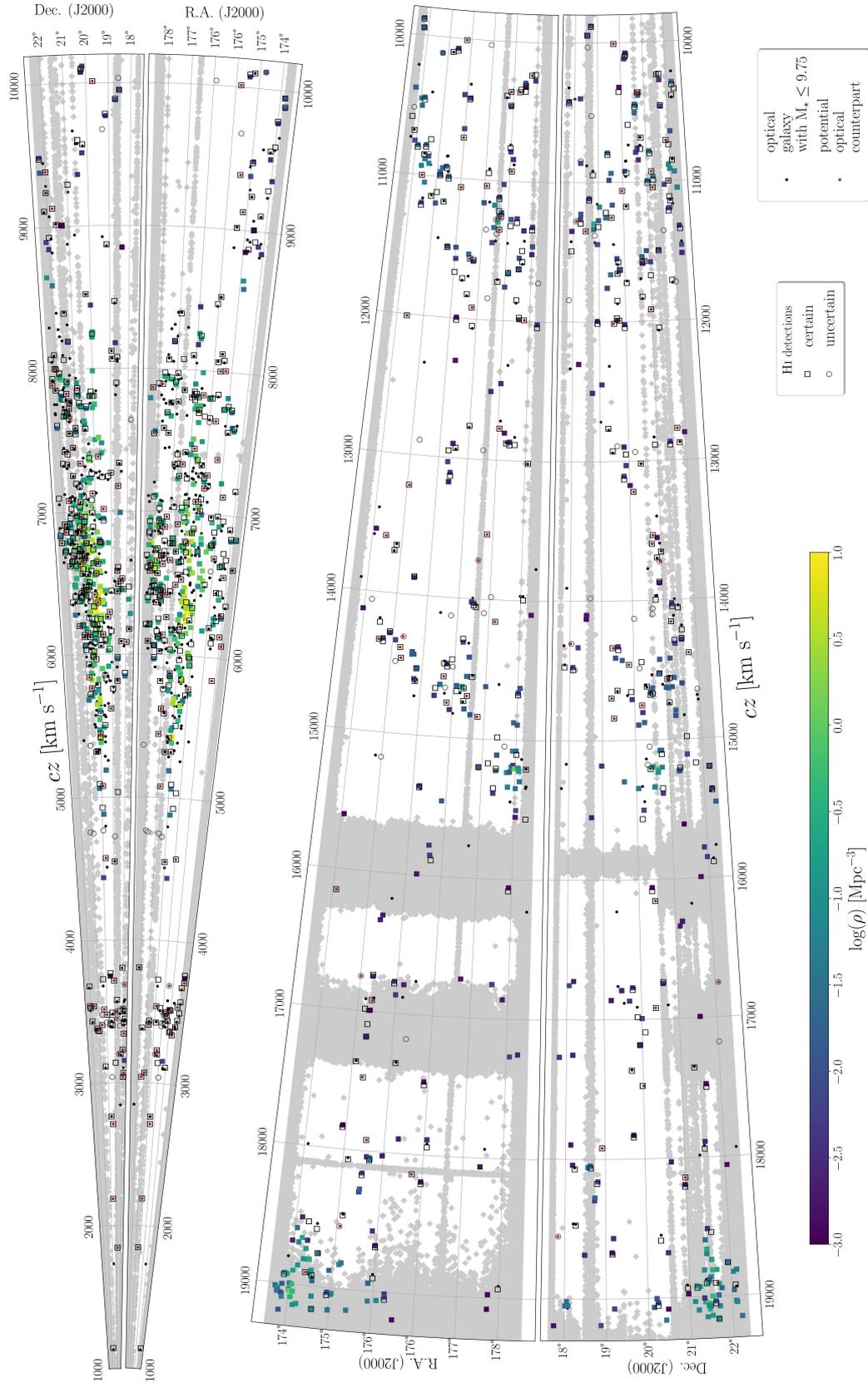


Fig. 5. Polar plot showing the distribution of the HI-detected and optical galaxies in the RA - cz and Dec - cz planes. The z range of the survey is split in two for clarity. Large squares of different colour show galaxies with spectroscopic redshift from SDSS and with stellar mass $M_* \geq 9.75 M_\odot$ with the colour indicating their local density according to the colour bar. Small black points show optical galaxies with stellar mass below this limit. Small brown points mark potential optical counterparts without optical redshift. Certain HI detections are shown with empty, black squares, uncertain ones with empty, dark-grey circles. The light-grey shaded areas of the map show areas where the ability to detect HI signal is reduced (see text for details). The figure is meant to be viewed on a (large) screen.

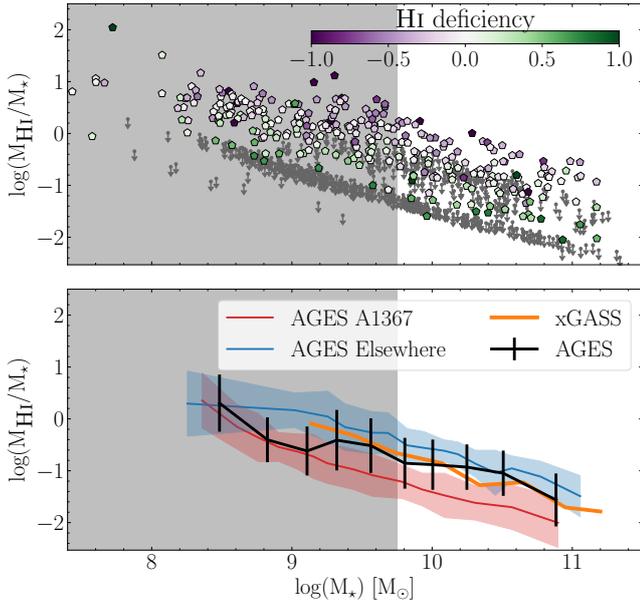


Fig. 6. Ratio of total mass in atomic hydrogen to total stellar mass for all AGES detections with known optical counterparts (pentagons) and upper limits for non-detections (grey, down-pointing arrows). The pentagons are colour coded according to their HI deficiency. The thick black line shows the median of all AGES galaxies. The orange solid line shows the median values for xGASS galaxies (Catinella et al. 2018). The grey shaded region encompasses the data excluded from the stellar-mass-limited sample. The red and blue curves and shaded regions show the median relations for AGES galaxies that are members of the A1367 cluster and non-members, respectively. ☉

low S/N (Fabello et al. 2012; Geréb et al. 2015; Hu et al. 2021). This method uses the availability of known on-the-sky positions and the redshift of galaxies (usually from optical spectroscopy) to extract their HI spectra. Averaging these spectra together will average out the Gaussian noise while adding the signal to produce one HI detection representing all the otherwise undetected galaxies.

Of the 830 galaxies in the optical sample, 296 are individually detected and catalogued, leaving 534 available for stacking. Figure 7 shows the results from such a stacking experiment. For every such galaxy, we extracted a spectrum from the HI data cube in the same way as the spectra of individual detections were extracted. Those spectra were then converted to units of $M_{\odot}/\text{km s}^{-1}$ using the formula listed in Sect. 3. Following that, weights were calculated based on the rms noise in the outer parts of the spectra, which are presumed not to contain any signal.

The galaxies were split into three bins of recession velocity: A1367 members, foreground, and background. Additionally, all galaxies lying in the RFI affected regions at $cz \geq 15\,500 \text{ km s}^{-1}$ were removed from the background bin. This left a total of 10, 325, and 113 galaxies in the three bins, respectively. Subsequently, 10 000 iterations were done with jackknife resampling where 90% of the available spectra were randomly selected, without replacement, and stacked. This was done to detect the presence of highly deviant spectra. The total M_{HI} in each stacked spectrum was integrated in the velocity range between the two vertical grey lines in Fig. 7. The thick lines show the spectra that gave the median of the distribution of M_{HI} and the thin ones show the 95% confidence interval.

The average HI content of the galaxies in each stack as measured from these spectra is given in Table 2. The observed

increase in the average HI mass in the directly detected galaxies and in the stack of non-detections is likely driven by the sensitivity curve shown in Fig. 1 which, as the redshift increases, removes low-mass galaxies from the stack with detections and moves them into the stack with non-detections. To confirm the presence of the signal in the stacked spectra, we also show a reference stack in grey in each panel; these are extracted at the positions of the optical galaxies but with randomised redshifts and stacked in the same way. The bottom row of panels on Fig. 7 shows the same stacked spectra but in flux units, thus avoiding the distance² factor used in the conversion of flux to HI mass.

A comparison between the stacked and reference spectra in the foreground bin shows that the measured HI mass from integrating the spectra is consistent with a non-detection. The small number of galaxies in this stack prevents us from reaching sufficiently low noise levels to draw any conclusions on the HI content of galaxies with M_{HI} below $10^7 M_{\odot}$.

Previous studies aimed at detecting galaxies with very low gas content (e.g. Popping 2010) show that objects with HI masses below $10^7 M_{\odot}$ are very rare. On the contrary, Jones et al. (2018) find that the HI mass function does not show a decline until its measured limit of $M_{\text{HI}} = 10^{6.5} M_{\odot}$. The reason for this may lie in a combination of the semi-constant column density distribution of HI within the galaxies, expressed in the mass – size relation for HI discs (Prochaska & Wolfe 2009; Wang et al. 2016), and the higher susceptibility of HI to photoionisation at low column densities (Schaye 2001). Interferometric observations confirm the constancy of the relation down to $\log M_{\text{HI}} \sim 7$ (Wang et al. 2016). Beam dilution, as a consequence, also plays a role, especially for single-dish surveys, as this relation drives the sizes of the HI discs well below the size of the beam.

4.2. What goes into HI deficiency

As defined by Haynes & Giovanelli (1984) and used by many researchers (e.g. Cortese et al. 2011; Stark et al. 2016; Catinella et al. 2018; Loni et al. 2021), the calculation of HI deficiency relies on additional data and a calibration of the relation between optical diameter and M_{HI} . This calibration is based on measurements of the optical diameter in a particular filter and limiting magnitude, and is known to depend on galaxy morphology. It also relies on a sample of galaxies with ‘normal’ gas content that is unaltered by external factors and observed through a telescope capable of measuring the emission present at all the spatial frequencies. Obviously these requirements and additional data are not going to be available for all HI data sets and might prove problematic in the future when we measure the HI content of galaxies at higher redshifts. It is possible that the local calibration of the mass – size relation does not hold for high-redshift galaxies. All of the above-mentioned issues and the specifics of our deficiency estimate drive a relatively high uncertainty. For galaxies with $\text{DEF} < 0.5$, that is, those that within the uncertainty have the expected HI content, we find a mean $\text{DEF} = -0.005 \pm 0.33$. A comparison with Cortese et al. (2011) shows good agreement with their estimate of 0.06 ± 0.28 .

The top panel of Fig. 8 presents the distribution of HI deficiency for all galaxies with optical redshift measurements within the HI survey. Those are split in two groups: members of A1367 and all others (which we term ‘Elsewhere’), which have means of 0.41 ± 0.49 and -0.02 ± 0.41 , respectively. A comparison between the HI deficiencies of the cluster and the Elsewhere sample show that $\text{DEF} = 0.3$ is a good dividing point between normal and deficient galaxies, in agreement with Dressler (1986), Solanes et al. (2001), Loni et al. (2021). This

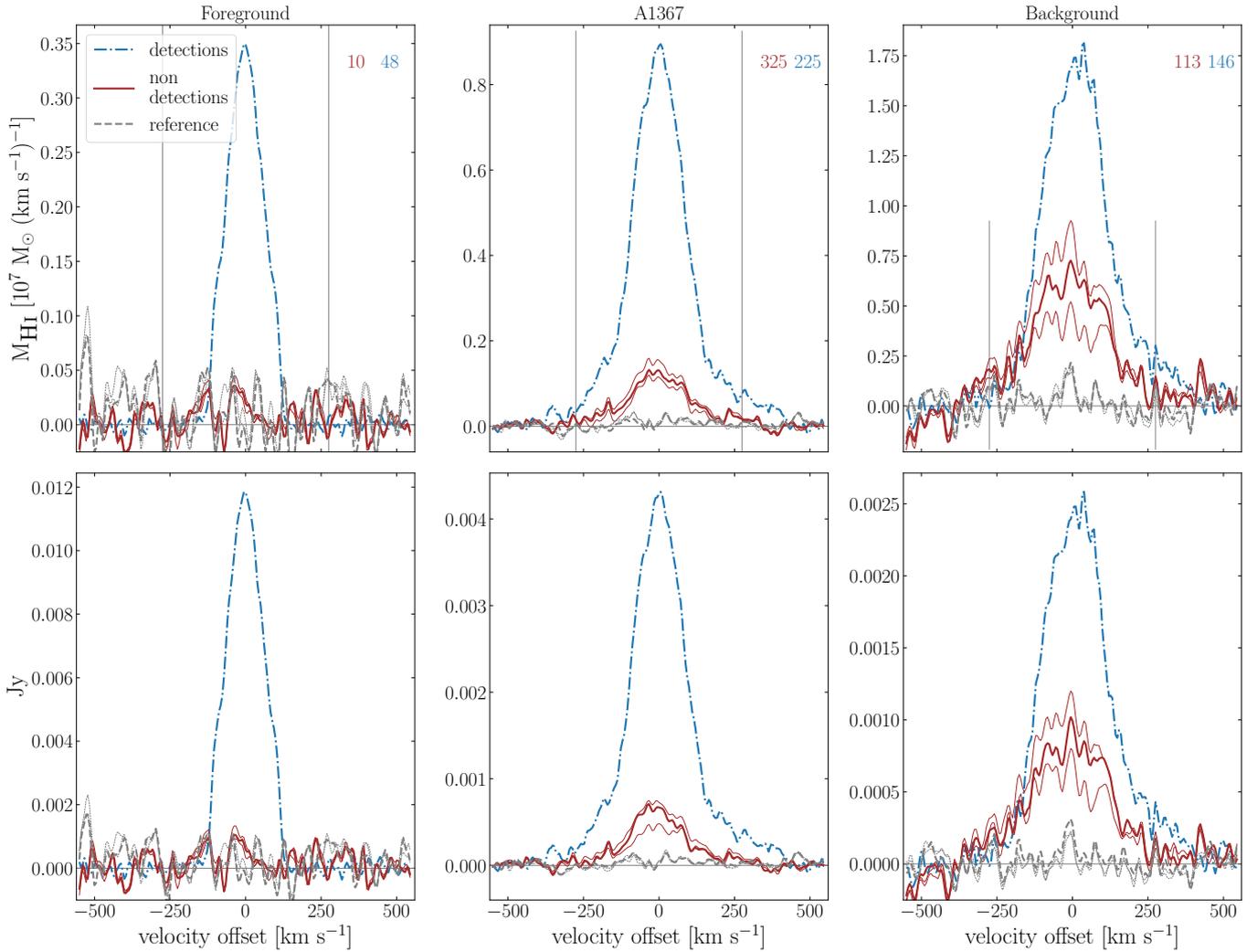


Fig. 7. HI signal present in the data below the detection threshold (brown curves) recovered by stacking spectra from galaxies in A1367, those in front of and behind it. The thick lines show the median and the thin ones the 95% confidence interval of the distribution of total M_{HI} in the stacked spectra, from jackknife resampling. For comparison, the dash-dotted blue lines shows the average spectrum of all individually detected galaxies. The dashed grey curves show the signal distribution in a reference stack (see text for details). The vertical grey lines show the velocity range over which the curves are integrated. The total number of galaxies in each stack is indicated at the top, coloured in blue for detections and red for non-detections. The *bottom row* shows the same spectra but in flux units. \odot

Table 2. Average HI mass in Fig. 7 in units of $10^8 M_{\odot}$.

	Foreground	A1367	Background
Detected	7.5	19.8	47.1
Non-detected	$0.41^{+0.16}_{-0.15}$	$3.19^{+0.64}_{-0.63}$	$19.9^{+4.89}_{-5.15}$

means that 87% of the Elsewhere sample can be considered HI normal, as opposed to 38% for the cluster sample.

It is obvious that the most HI-deficient galaxies would not be detected at all at 21 cm. The bottom panel of Fig. 8 shows the distribution of HI deficiency only for HI-detected galaxies. Removing the non-detections from the sample drastically reduces the difference between the means of the two environmental bins, from 0.43 to 0.12. The standard method for accounting for non-detected galaxies is to compute an upper limit on their detectable gas content to find the lower limit on their deficiency, as we did in Sect. 3.3 and present in Fig. 6 and in the upper panel of Fig. 8. This offers the advantage of giving information on indi-

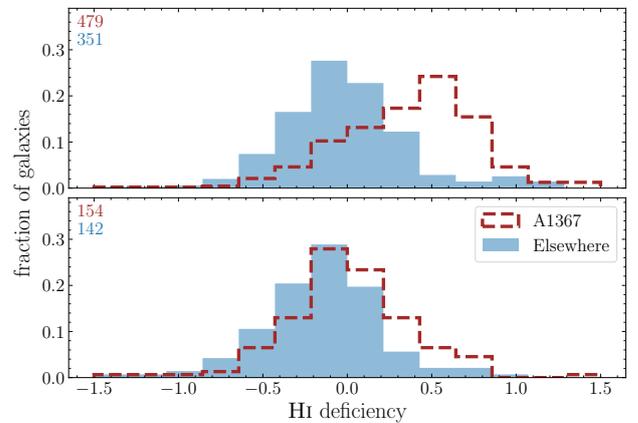


Fig. 8. Distribution of HI deficiency among all optical galaxies within the surveyed volume (*top*), and the HI detections (*bottom*). Those are split into members of A1367 (brown, dashed histogram) and all the rest (blue, filled histogram). The total number of galaxies in each histogram is printed in the corresponding colour in the *top left corner*. \odot

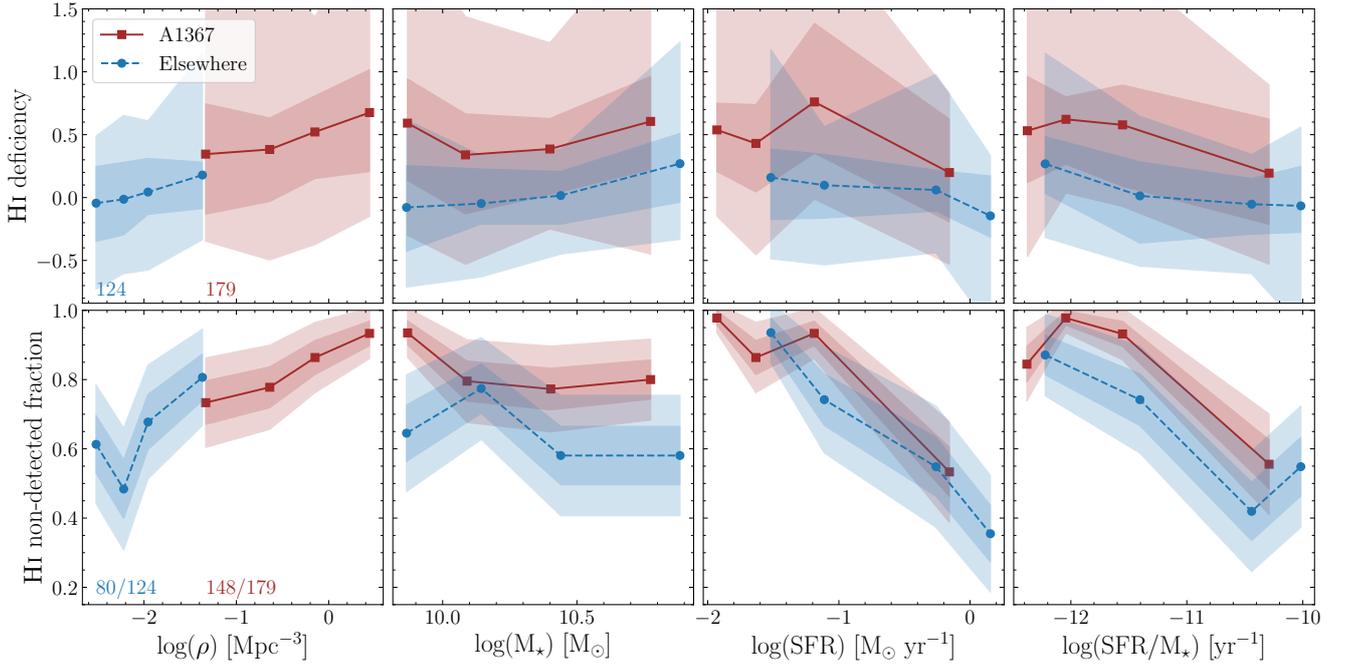


Fig. 9. HI deficiency (*top*) and HI ‘non-detected fraction’ (*bottom*), as a function of (*left to right*) local galaxy density, stellar mass, star formation rate, and specific star formation rate. Cluster galaxies are shown with brown squares and solid lines, galaxies outside the cluster are shown with blue dots and dashed lines. The number of galaxies in the two environments is printed in the *top left panel* in the corresponding colour. The overall number of ‘non-detected’/all galaxies is printed in the *bottom left panel*. The shaded regions show 1σ and 2σ assuming normal distribution. ☺

vidual objects, but relies on the calibration of the expected HI mass from external data sets. We note that the $\sim 70\%$ reduction in the difference between the means of the two distributions can be accounted for by simply counting the non-detected fraction. The use of the non-detected fraction may offer a simpler alternative to the deficiency limits, and we explore the advantages and disadvantages of this parameter below.

4.3. HI ‘non-detected fraction’ and HI deficiency

Because any observations will have a sensitivity limit varying with redshift, the ‘non-detected fraction’ needs to be calculated for a complete sample. In Fig. 4 we outline the stellar and HI mass completeness limits as $\log M_{\text{HI}}^{\text{lim}} = 9.25 M_{\odot}$ and $\log M_{\star}^{\text{lim}} = 9.75 M_{\odot}$. In this, section we analyse the ‘non-detected fraction’ as calculated for this stellar and HI mass complete sample. In this sense, ‘non-detected fraction’ should be understood as an alias, meaning the fraction of galaxies with stellar mass above 9.75 and HI mass below 9.25. We use the alias for brevity; it will be precisely equal to the true non-detected fraction for observations with uniform sensitivity. As Fig. 4 shows, many of the galaxies considered as ‘non-detections’ are in fact detected but have M_{HI} below the completeness limit. As described earlier, these completeness limits apply to the AGES and SDSS data for sources with recession velocity $cz \leq 15\,500 \text{ km s}^{-1}$. Beyond this redshift, RFI pollution affects the completeness of the HI survey.

The top panel of Fig. 9 shows the distribution of HI deficiency for the stellar-mass-complete sample. The bottom row shows the HI ‘non-detected fraction’ for the M_{HI} and M_{\star} complete sample. The number of galaxies in the two samples is printed in the first panels. These are presented as a function of local galaxy density, stellar mass, star formation rate and specific star formation rate. The shaded areas on the top row show

the 2.5, 17, 83, and 97.5 percentile of the distribution of HI deficiency (1σ and 2σ for a Gaussian distribution). The uncertainties shown on the bottom row are binomial 66% and 95% confidence intervals calculated using normal approximation. The points are binned along the x -axis in bins of constant population number to guarantee uniform significance along the relations. Along the x -axis the points are plotted at the median value for all galaxies in the bin. We choose this over the mean because of its reduced susceptibility to outliers, even though it makes the distribution appear narrower.

As already apparent from Fig. 5 and the bottom panel of Fig. 4, the cluster galaxies are distributed at higher local densities than the rest. There is, in fact, a small overlap between the two distributions of local densities which is hidden by the binning in the left-most panels of Fig. 9. The Elsewhere sample contains a number of groups as well as void-like regions. The highest density region outside A1367 present in our data is at the high-redshift end of the survey which is heavily polluted by RFI from the radar at the nearby Luis Muñoz Marín International Airport (see Fig. 5). As discussed above, this part is excluded from this analysis. We would like to point out that while this particular problem is caused by ground-based facilities, these types of problems for radio astronomy, in general, are set to become much more severe in the near future with the introduction of constellation satellites.

The left-most panels show that, qualitatively, HI deficiency and the ‘non-detected fraction’ show the same trends. Both show a continuous increase with local density, irrespective of the global environment, that is, whether or not the galaxy belongs to a cluster. The HI ‘non-detected fraction’ outside the cluster matches very well the average non-detected fraction published by Huang et al. (2012) based on the 40% ALFALFA catalogue. As clusters of galaxies are very rare objects in the Universe and represent only a tiny fraction of the volume surveyed by

ALFALFA, this is not surprising. Dénes et al. (2014) published maps of HI deficiency as a function of local galaxy density, showing essentially the same result as that shown in the left-most column of Fig. 9, despite the different method of measuring local galaxy density.

The second column of Fig. 9 again shows a qualitative agreement between HI deficiency and HI ‘non-detected fraction’. At a given stellar mass, cluster galaxies are more gas deficient than galaxies outside the cluster. This result is also in good qualitative agreement with the results by Cortese et al. (2011) and the results for satellites only by Stark et al. (2016). Our sample does not allow segregation of satellites and centrals but we can assume that at least the first three mass bins are dominated by satellites and this defines the observed relations.

There is little difference between the third and fourth columns of Fig. 9, which is expected given the lack of correlation with stellar mass. However, these panels appear to show a qualitative difference between HI deficiency and the HI ‘non-detected fraction’. While the former shows only dependence on global environment but not on SFR, the latter shows no significant dependence on global environment but a strong correlation with SFR. However, it should be noted that because of the large scatter in the distribution of HI deficiency it could also be considered consistent with the trend shown by the ‘non-detected fraction’.

As we have already stressed, it is imperative that this analysis be based on a stellar and HI mass complete sample. We verified that this is the case and also verified that the exact value at which completeness is achieved by the survey does not change the slope of the relation but only its intercept.

The ‘non-detected fraction’ is essentially a ratio between the integrals of the HI mass function and stellar mass function above a certain limit. The exact limits used do not change the shape of the relation, only its normalisation, at least within this survey. Studies of the variations in the HI mass function with environment point at a small increase in the knee of the Schechter function fit as the local density increases (Jones et al. 2016) or of the slope of the low-mass end of the fit (Zwaan et al. 2005). Studies of the variations in the stellar mass function with environment are similarly inconclusive but generally agree that the high-mass slope increases in high-density environments while the low-mass slope likely decreases (Etherington et al. 2017; Papovich et al. 2018). Establishing the physics driving this relation requires larger data sets and will be the subject of a future paper. However, these results show the importance of mechanisms related to groups and clusters, like tidal interactions, mergers and/or ram pressure stripping. The local density, in this case, is measured as the inverse of the volume available for each galaxy. This is a parameter which could be related to the gas accretion rate from the circumgalactic medium. Availability of data for groups and clusters with a range of properties will be needed to establish whether it is active gas stripping or lack of gas accretion that shapes this relation.

The ‘non-detected fraction’ is in effect a low-resolution HI deficiency which results in significantly reduced uncertainties compared to HI deficiency. In the case of the correlation between SFR and the ‘non-detected fraction’ (Fig. 9), it is obvious that the large scatter in HI deficiency prevents us from seeing its correlation with SFR. Trading resolution for sensitivity is common practice when low S/Ns are involved; see for example Sect. 4.1. The main problem with using the ‘non-detected fraction’ as a parameter for studying environmental effects is that it can only be defined for an ensemble of galaxies, while HI deficiency can be calculated for individual galaxies. As such it is applicable to

large blind surveys like AGES, while deficiency is superior for targeted, high-S/N, observations.

5. Summary

We present the complete Arecibo Galaxy Environment Survey (AGES) survey of Abell 1367 and its surroundings. This is one of the most sensitive blind HI surveys performed to date. We detect 457 HI sources over the imaged 44 000 Mpc³. Within the SDSS and NED, we found 335 optical counterparts with spectroscopic redshift. Searching through the NED database, we found another 106 potential optical counterparts with unknown optical redshift. The presented HI catalogue contains 16 sources without optical counterparts.

These HI sources are distributed over a wide range of local densities and are therefore well suited for studying the effects of environment on galaxy evolution. Analysis of the data with respect to local and global environment reveals that the dependence of HI properties on global environment is part of a continuous dependence on local density.

Acknowledgements. This work was supported by the Czech Science Foundation grant 19-18647S and the institutional project RVO 67985815. This research made use of Astropy (<http://www.astropy.org>), a community-developed core Python package for Astronomy (Astropy Collaboration 2013, 2018). This research made use of NASA’s Astrophysics Data System Bibliographic Services. This research has made use of the NASA/IPAC Extragalactic Database (NED), which is funded by the National Aeronautics and Space Administration and operated by the California Institute of Technology. This research made use of APLpy, an open-source plotting package for Python (Robitaille & Bressert 2012; Robitaille 2019). This publication uses data generated via the Zooniverse.org platform, development of which is funded by generous support, including a Global Impact Award from Google, and by a grant from the Alfred P. Sloan Foundation.

References

- Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, *ApJS*, 249, 3
 Arrigoni Battaia, F., Gavazzi, G., Fumagalli, M., et al. 2012, *A&A*, 543, A112
 Astropy Collaboration (Robitaille, T. P., et al.) 2013, *A&A*, 558, A33
 Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, 156, 123
 Auld, R., Minchin, R. F., Davies, J. I., et al. 2006, *MNRAS*, 371, 1617
 Barnes, D. G., Staveley-Smith, L., de Blok, W. J. G., et al. 2001, *MNRAS*, 322, 486
 Bechtold, J., Forman, W., Giacconi, R., et al. 1983, *ApJ*, 265, 26
 Boselli, A., & Gavazzi, G. 2006, *PASP*, 118, 517
 Boselli, A., & Gavazzi, G. 2014, *A&ARv*, 22, 74
 Boselli, A., Fossati, M., Cuillandre, J. C., et al. 2018, *A&A*, 615, A114
 Boselli, A., Fossati, M., & Sun, M. 2021, *A&ARv*, 30, A3
 Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *MNRAS*, 351, 1151
 Catinella, B., Saintonge, A., Janowiecki, S., et al. 2018, *MNRAS*, 476, 875
 Chabrier, G. 2003, *PASP*, 115, 763
 Cortese, L., Gavazzi, G., Boselli, A., Iglesias-Paramo, J., & Carrasco, L. 2004, *A&A*, 425, 429
 Cortese, L., Gavazzi, G., Boselli, A., et al. 2006, *A&A*, 453, 847
 Cortese, L., Minchin, R. F., Auld, R. R., et al. 2008, *MNRAS*, 383, 1519
 Cortese, L., Catinella, B., Boissier, S., Boselli, A., & Heinis, S. 2011, *MNRAS*, 415, 1797
 Cortese, L., Catinella, B., & Smith, R. 2021, *PASA*, 38, e035
 Davies, J. I., Auld, R., Burns, L., et al. 2011, *MNRAS*, 415, 1883
 Delaunay, B., et al. 1934, *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk*, 7, 1
 Dénes, H., Kilborn, V. A., & Koribalski, B. S. 2014, *MNRAS*, 444, 667
 Donnelly, R. H., Markevitch, M., Forman, W., et al. 1998, *ApJ*, 500, 138
 Dressler, A. 1986, *ApJ*, 301, 35
 Eddington, A. S. 1913, *MNRAS*, 73, 359
 Etherington, J., Thomas, D., Maraston, C., et al. 2017, *MNRAS*, 466, 228
 Fabello, S., Kauffmann, G., Catinella, B., et al. 2012, *MNRAS*, 427, 2841
 Gavazzi, G., Boselli, A., Mayer, L., et al. 2001, *ApJ*, 563, L23
 Gavazzi, G., Cortese, L., Boselli, A., et al. 2003a, *ApJ*, 597, 210
 Gavazzi, G., Boselli, A., Donati, A., Franzetti, P., & Scodreggio, M. 2003b, *A&A*, 400, 451

- Ge, C., Sun, M., Liu, R.-Y., et al. 2019, *MNRAS*, **486**, L36
- Ge, C., Sun, M., Yagi, M., et al. 2021, *MNRAS*, **508**, L69
- Geréb, K., Morganti, R., Oosterloo, T. A., Hoppmann, L., & Staveley-Smith, L. 2015, *A&A*, **580**, A43
- Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2005, *AJ*, **130**, 2598
- Girardi, M., Giuricin, G., Mardirossian, F., Mezzetti, M., & Boschin, W. 1998, *ApJ*, **505**, 74
- Gunn, J. E., & Gott, J. R., III 1972, *ApJ*, **176**, 1
- Haynes, M. P., & Giovanelli, R. 1984, *AJ*, **89**, 758
- Haynes, M. P., Giovanelli, R., Martin, A. M., et al. 2011, *AJ*, **142**, 170
- Haynes, M. P., Giovanelli, R., Kent, B. R., et al. 2018, *ApJ*, **861**, 49
- Hu, W., Cortese, L., Staveley-Smith, L., et al. 2021, *MNRAS*, **507**, 5580
- Huang, S., Haynes, M. P., Giovanelli, R., & Brinchmann, J. 2012, *ApJ*, **756**, 113
- Jáchym, P., Kenney, J. D. P., Sun, M., et al. 2019, *ApJ*, **883**, 145
- Jackson, J. C. 1972, *MNRAS*, **156**, 1P
- Jones, M. G., Papastergis, E., Haynes, M. P., & Giovanelli, R. 2016, *MNRAS*, **457**, 4393
- Jones, M. G., Haynes, M. P., Giovanelli, R., & Moorman, C. 2018, *MNRAS*, **477**, 2
- Kaiser, N. 1987, *MNRAS*, **227**, 1
- Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, **341**, 33
- Keenan, O. C., Davies, J. I., Taylor, R., & Minchin, R. F. 2016, *MNRAS*, **456**, 951
- Kenney, J. D. P., van Gorkom, J. H., & Vollmer, B. 2004, *AJ*, **127**, 3361
- Kuutma, T., Tamm, A., & Tempel, E. 2017, *A&A*, **600**, L6
- Larson, R. B., Tinsley, B. M., & Caldwell, C. N. 1980, *ApJ*, **237**, 692
- Liivamägi, L. J., Tempel, E., & Saar, E. 2012, *A&A*, **539**, A80
- Lintott, C., Schawinski, K., Bamford, S., et al. 2011, *MNRAS*, **410**, 166
- Loni, A., Serra, P., Kleiner, D., et al. 2021, *A&A*, **648**, A31
- Marinoni, C., Davis, M., Newman, J. A., & Coil, A. L. 2002, *ApJ*, **580**, 122
- Masters, K. L., Springob, C. M., Haynes, M. P., & Giovanelli, R. 2006, *ApJ*, **653**, 861
- Minchin, R. F., Momjian, E., Auld, R., et al. 2010, *AJ*, **140**, 1093
- Papovich, C., Kawinwanichakij, L., Quadri, R. F., et al. 2018, *ApJ*, **854**, 30
- Peng, Y., Maiolino, R., & Cochrane, R. 2015, *Nature*, **521**, 192
- Pereira, S., Campusano, L. E., Hitschfeld-Kahler, N., et al. 2017, *ApJ*, **838**, 109
- Popping, A. 2010, PhD Thesis, University of Groningen, The Netherlands
- Prochaska, J. X., & Wolfe, A. M. 2009, *ApJ*, **696**, 1543
- Putman, M. E., de Heij, V., Staveley-Smith, L., et al. 2002, *AJ*, **123**, 873
- Roberts, I. D., van Weeren, R. J., McGee, S. L., et al. 2021, *A&A*, **650**, A111
- Robitaille, T. 2019, <https://doi.org/10.5281/zenodo.2567476>
- Robitaille, T., & Bressert, E. 2012, Astrophysics Source Code Library [record ascl:1208.017]
- Saintonge, A. 2007, *AJ*, **133**, 2087
- Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, *ApJS*, **233**, 22
- Sakai, S., Kennicutt, R. C., Jr., van der Hulst, J. M., & Moss, C. 2002, *ApJ*, **578**, 842
- Sault, R. J., Teuben, P. J., & Wright, M. C. H. 1995, in *Astronomical Data Analysis Software and Systems IV*, eds. R. A. Shaw, H. E. Payne, & J. J. E. Hayes, *ASP Conf. Ser.*, **77**, 433
- Schaye, J. 2001, *ApJ*, **562**, L95
- Schmidt, M. 1963, *ApJ*, **137**, 758
- Scott, T. C., Bravo-Alfaro, H., Brinks, E., et al. 2010, *MNRAS*, **403**, 1175
- Scott, T. C., Cortese, L., Brinks, E., et al. 2012, *MNRAS*, **419**, L19
- Scott, T. C., Brinks, E., Cortese, L., Boselli, A., & Bravo-Alfaro, H. 2018, *MNRAS*, **475**, 4648
- Solanes, J. M., Manrique, A., García-Gómez, C., et al. 2001, *ApJ*, **548**, 97
- Stark, D. V., Kannappan, S. J., Eckert, K. D., et al. 2016, *ApJ*, **832**, 126
- Tago, E., Saar, E., Tempel, E., et al. 2010, *A&A*, **514**, A102
- Taylor, R. 2015, *Astron. Comput.*, **13**, 67
- Taylor, R., Davies, J. I., Auld, R., & Minchin, R. F. 2012, *MNRAS*, **423**, 787
- Taylor, R., Davies, J. I., Auld, R., Minchin, R. F., & Smith, R. 2013, *MNRAS*, **428**, 459
- Taylor, R., Minchin, R. F., Herbst, H., et al. 2014, *MNRAS*, **443**, 2634
- Taylor, R., Köppen, J., Jáchym, P., et al. 2020, *AJ*, **159**, 218
- Tempel, E., Tuvikene, T., Kipper, R., & Libeskind, N. I. 2017, *A&A*, **602**, A100
- Tinsley, B. M. 1980, *Fund. Cosm. Phys.*, **5**, 287
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, **613**, 898
- Tully, R. B., & Fisher, J. R. 1977, *A&A*, **54**, 661
- Tully, R. B., & Fisher, J. R. 1978, in *Large Scale Structures in the Universe*, eds. M. S. Longair, & J. Einasto, 79, 214
- van de Weygaert, R. 1994, *A&A*, **283**, 361
- Voronoi, G. 1908, *J. Reine Angew. Math.*, **133**, 97
- Wang, J., Koribalski, B. S., Serra, P., et al. 2016, *MNRAS*, **460**, 2143
- Yang, X., Mo, H. J., van den Bosch, F. C., et al. 2007, *ApJ*, **671**, 153
- York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, *AJ*, **120**, 1579
- Zabludoff, A. I., Geller, M. J., Huchra, J. P., & Ramella, M. 1993, *AJ*, **106**, 1301
- Zwaan, M. A., Meyer, M. J., Staveley-Smith, L., & Webster, R. L. 2005, *MNRAS*, **359**, L30

Appendix A: Atlas of HI detections

In this Appendix, we show five example HI detections spanning a range of galaxy properties and S/Ns. The complete atlas showing all HI detections presented in this article can be downloaded from https://www.dropbox.com/s/xffshfom5h0m3f0/AGES_A1367_atlas.tar.gz?dl=0.

Description of the individual panels. Top row, from left to right:

1 - Moment 0 map. The velocity integral is highlighted in red in the bottom panel. The cyan circle has a diameter equal to the FWHM of Arecibo's beam and is centred at the catalogue coordinates of the source.

2 - Renzogram. For every channel, the contour is drawn at a significance level indicated at the top. Contours are colour coded according to their velocity as indicated in the colour bar. The contours are overlaid on a RGB image made from r, g, and i band images from the SDSS. The cyan circle is the same as in the previous panel. The image has the same sky coverage as panel 1.

3 - The same optical image used in panel 2. The blue circles show SDSS spectroscopic extragalactic targets with the numbers indicating their recession velocity. Orange pentagons indicate other HI detections in the field. Their catalogue name and recession velocity are indicated. The cyan circle is as in panel 1. The panel has the same extent on the sky as the previous panels.

4 - The optical image shown in panels 2 and 3 but showing only the beam area indicated by the circle on the previous panels.

5 - A white light optical image showing the same part of the sky as panel 4. This was produced by integrating the three channels and with 'sqrt' stretch to boost the visibility of low-surface-brightness features. Different symbols and colours indicate optical counterparts. The cyan square indicates the primary optical counterpart (as listed in the Table A.1). Magenta triangles show the secondary optical counterpart. Yellow hexagons show potential optical counterparts. The recession velocity is indicated, when available.

Bottom row:

HI spectrum of the source. The thin blue line shows the spectrum used for measurements. This is extracted from Hanning 3 smoothed cube. The thick blue line shows a spectrum extracted from a Hanning 5 smoothed cube used when searching for detections. The thin grey lines show reference spectra extracted at eight locations two beams away from the centre of the source. The solid vertical line shows the velocity centre of the source. The two dashed-dotted lines show the W_{50} width of the source. The two dotted lines show the W_{20} of the source. On the left, a number of measured parameters of the source are written, many of which are also given in the tables at the CDS. From top to bottom: total HI mass, total flux, S/N of the peak, central velocity, and redshift of the source, flag of the HI source, asymmetry of the line profile, S/N of the source.

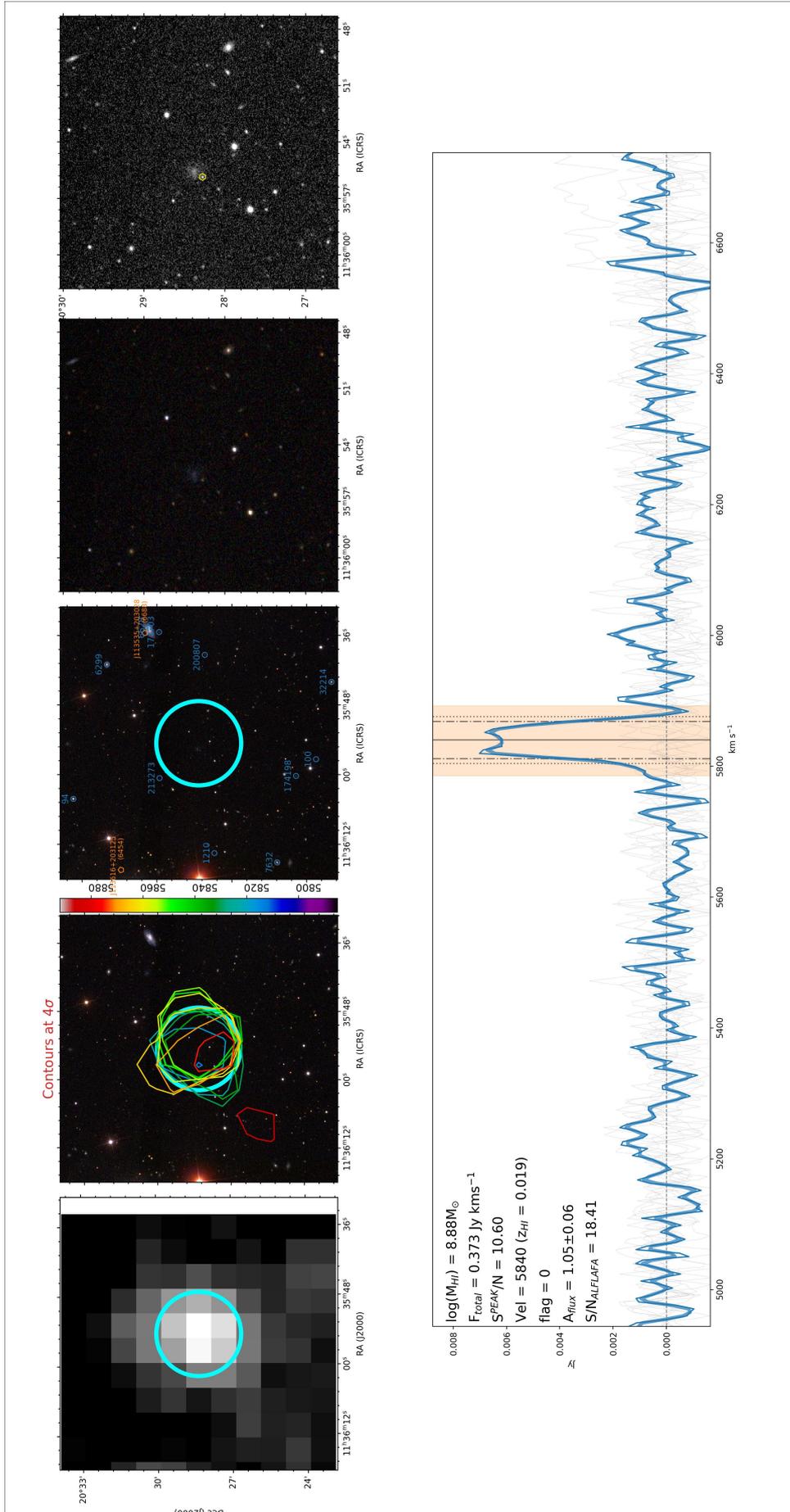


Fig. A.3. J113554+202819.

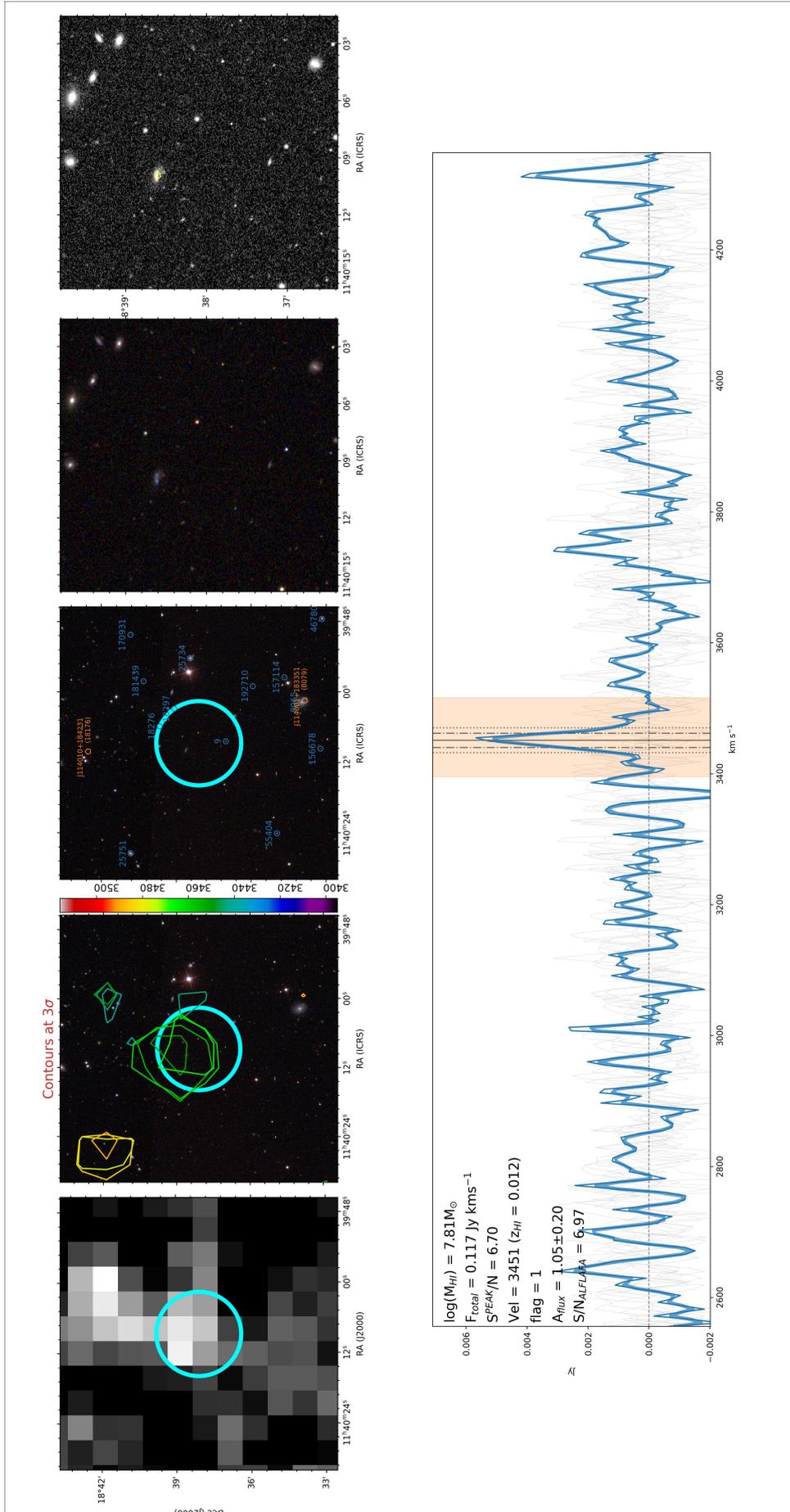


Fig. A.5. J114008+183805.