Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **505,** 304–324 (2021) doi:10.1093/mnras/stab1282
Advance Access publication 2021 May 7

The H I and stellar mass bivariate distribution of centrals and satellites for all, late-, and early-type local galaxies

A. R. Calette [1]★ Vladimir Avila-Reese [1] Aldo Rodríguez-Puebla [1] Claudia del P. Lagos [2,3] and Barbara Catinella [2,3]

[1]*Instituto de Astronomía, Universidad Nacional Autónoma de México, A. P. 70-264, 04510 Ciudad de México, México*
[2]*International Centre for Radio Astronomy Research (ICRAR), M468, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia*
[3]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)*

Accepted 2021 April 26. Received 2021 March 30; in original form 2021 January 1

ABSTRACT

We characterize the conditional distributions of the H I gas-to-stellar mass ratio, $R_{\rm HI} \equiv M_{\rm HI}/M_*$, given the stellar mass, M_*, of local galaxies from $M_* \sim 10^7$–10^{12} M$_\odot$ separated into centrals and satellites as well as into late- and early-type galaxies (LTGs and ETGs, respectively). To do so, we use (1) the homogeneous 'eXtended GALEX Arecibo SDSS Survey', xGASS (Catinella et al. 2018), by re-estimating their upper limits and taking into account them in our statistical analysis; and (2) the results from a large compilation of H I data reported in Calette et al. (2018). We use the $R_{\rm HI}$ conditional distributions combined with the Galaxy Stellar Mass Function to infer the bivariate $M_{\rm HI}$ and M_* distribution of all galaxies as well of the late/early-type and central/satellite subsamples and their combinations. Satellites are on average less H I gas-rich than centrals at low and intermediate masses, with differences being larger for ETGs than LTGs; at $M_* > 3$–5×10^{10} M$_\odot$ the differences are negligible. The differences in the H I gas content are much larger between LTGs and ETGs than between centrals and satellites. Our empirical H I Mass Function is strongly dominated by central galaxies at all masses. The empirically constrained bivariate $M_{\rm HI}$ and M_* distributions presented here can be used to compare and constrain theoretical predictions as well as to generate galaxy mock catalogues.

Key words: methods: statistical – galaxies: fundamental parameters – galaxies: general – galaxies: ISM – galaxies: luminosity function, mass function.

1 INTRODUCTION

The evolution of galaxies depends on the interplay of many complex processes. Among them: gas cooling within dark matter haloes, transformation of the cool atomic hydrogen (H I) gas into cold dense molecular hydrogen (H_2) clouds, the formation of stars in the densest regions of these clouds, and the ulterior feedback that stars and their explosions exert on the interstellar medium (for a review, e.g. Mo, van den Bosch & White 2010). Therefore, the amounts of H I and H_2 gas relative to the stellar mass, morphological type, optical colours, and other galaxy properties, are crucial for understanding the evolutionary stage of local galaxies (see e.g. Lagos et al. 2011, 2014). It is also well known that the environment, in particular whether a galaxy is central or satellite (e.g. Kauffmann et al. 2004; Boselli & Gavazzi 2006; Davies et al. 2019), plays a role in the evolution of galaxies, so that information on the gas fractions of galaxies as a function of environment is also relevant (e.g. Brown et al. 2017; Stevens et al. 2019, and more references therein).

Although H I gas is the dominant component in the interstellar medium of local galaxies (Fukugita, Hogan & Peebles 1998), its detection is not easy because of its weak 21-cm emission line. Great efforts have been made to build large radio H I surveys as the H I Parkes All-Sky Survey (HIPASS; Meyer et al. 2004) and Arecibo Fast Legacy ALFA Survey (ALFALFA; Giovanelli et al. 2005; Haynes et al. 2011). However, these blind radio surveys are not yet as deep and do not cover such large areas as the optical/infrared extragalactic surveys, and are affected by strong selection effects. Thus, the inferred H I gas scaling relations, H I velocity function, as well as other correlations and H I spatial distributions, result biased if based on detections only (cf. Meyer et al. 2007; Haynes et al. 2011; Huang et al. 2012; Papastergis et al. 2013; Maddox et al. 2015; Guo et al. 2017; Calette et al. 2018). So, volume corrections or strategies like H I spectral stacking (e.g. Brown et al. 2015) are required to infer approximations to the intrinsic relations and distribution functions. Another way to attempt to overcome the strong selection effects of blind H I radio surveys is to construct 'well controlled' H I samples by means of radio follow-up observations of optically selected galaxy samples (e.g. Wei et al. 2010; Papastergis et al. 2012; Catinella et al. 2013, 2018; Kannappan et al. 2013; Boselli, Cortese & Boquien 2014a; Eckert et al. 2015; Stark et al. 2016; van Driel et al. 2016; Masters et al. 2019). These samples were designed for a variety of scientific goals, and as a result they are diverse and heterogeneous, covering different stellar mass ranges, distances, and

★ E-mail: ar.calette.morin@gmail.com

H I flux detection limits, and commonly they are far from complete in stellar mass.

1.1 The HI conditional distributions of late and early-type galaxies

In Calette et al. (2018, and with updates in Rodríguez-Puebla et al. 2020, hereafter Papers I and II, respectively), we undertook the task of compiling and homogenizing from the literature many H I galaxy samples such the ones listed above (including most of them), with the additional requirement of information on galaxy morphology being available. The latter was done as the H I gas content of galaxies strongly depends on morphology, hence it is more appropriate to analyse it separately for galaxies of at least two broad morphological groups. We took into account the reported upper limits for the radio non-detections, and after homogenizing them to a distance of ∼25 Mpc and similar signal-to-noise ratio detection limit we applied a survival analysis to determine gas correlations. As a result, we were able to constrain not only the mean $M_{\rm HI}$–M_* relation for late- and early-type galaxies (LTG and ETG, respectively) down to $M_* \sim 10^7$ $M_\odot$, but the respective conditional probability density distribution functions (PDFs) of $M_{\rm HI}$ given M_*, $P(M_{\rm HI}|M_*)$. From these PDFs, one can calculate any moment of the distributions, in particular the standard deviation around the mean relation, as well as the percentiles.

In Paper II, we used the well-constrained Galaxy Stellar Mass Function (GSMF) for all, LTG and ELTs down to $\sim 10^7$ $M_\odot$ computed there, and combined them with the $P(M_{\rm HI}|M_*)$ distributions to generate the bivariate (joint) M_* and $M_{\rm HI}$ distribution function. By projecting this bivariate distribution into the H I axis, we obtained the H I MFs, for LTGs and ETGs, as well as for all galaxies. We have shown that our empirical H I MF (corresponding to a volume-limited sample complete above $M_{\rm HI} \sim 10^8$ $M_\odot$) agrees well with those measured from blind radio surveys.

In Paper I, we showed that the conditional PDFs of the H I-to-stellar mass ratio, $R_{\rm HI} \equiv M_{\rm HI}/M_*$, given M_* can be well described by a Schechter-type function for LTGs (see also Lemonias et al. 2013) and a (broken) Schechter-type function plus a top-hat function for ETGs, having the latter significantly lower values of H I gas content than the former at fixed stellar mass. In Fig. 1, we reproduce these $P(R_{\rm HI}|M_*)$ distributions as a function of M_*, left-hand and medium panels, along with the respective logarithmic mean $R_{\rm HI}$–M_* relations and standard deviations (first and second moments of the log $R_{\rm HI}$ distributions), solid lines and shaded regions, respectively. The thick dashed lines are the corresponding medians. While for LTGs, both the mean and median $R_{\rm HI}$–M_* relations are similar, for ETGs, they differ, specially at the high-mass side. The right-hand panel shows the resulting $R_{\rm HI}$ conditional PDFs for all galaxies as well as the respective first and second moments. We infer the $R_{\rm HI}$ conditional distribution for all galaxies by using the fractions of ETGs and LTGs as a function of M_* from the Sloan Digital Sky Survey (SDSS) based on the Huertas-Company et al. (2011) morphological classification, corrected for volume completeness (see Paper II for details).

1.2 The HI gas content of central and satellite galaxies

The $P(R_{\rm HI}|M_*)$ distributions and the main relations shown in Fig. 1 do not distinguish between central and satellite galaxies. Though it is not clear whether the H I gas fraction of galaxies correlates directly or not with the large-scale environment (see Paper I for a discussion, and the references therein), at the level of central and satellite galaxies, there are differences with latter having lower H I gas contents at a given stellar mass than the former (e.g. Brown et al. 2016; Stark et al. 2016, but see Lu et al. 2020).

The goal of this paper is to introduce adequate functions to our empirical H I conditional PDFs for LTGs and ETGs in such a way that they can be separated into central and satellite galaxies. For this, we will use the recent H I observational survey eXtended GALEX Arecibo SDSS Survey (xGASS; Catinella et al. 2018). xGASS is an homogeneously constructed H I, ultraviolet, and optical galaxy sample with well defined limits in $R_{\rm HI}$, M_*, and volume. Since this survey was constructed from SDSS, most of the galaxies can be separated into centrals and satellites making use of the Yang et al. (2007, 2012) halo-based group definition applied to SDSS. Thus, from xGASS, we calculate the ratios of central and satellite to total $R_{\rm HI}$ conditional PDFs as a function of M_* for both LTGs and ELTs. These ratios are applied to our empirical LTG and ETG $R_{\rm HI}$ PDFs to separate them into centrals and satellites.

This paper is organized as follows. In Section 2, we describe the xGASS survey and our processing, in particular for the upper limits. Section 3 presents the results of our statistical analysis of xGASS: the H I-to-stellar mass relations for LTGs and ETGs separated into centrals and satellites, as well as the respective H I conditional PDFs and join fits of analytic functions to these. In Section 4, we use the xGASS H I conditional PDFs to separate the distributions constrained in Papers I and II into centrals and satellites. By combining these distribution with the GSMF, we construct the full bivariate (joint) M_* and $M_{\rm HI}$ distributions of all galaxies as well as of subsamples of centrals/satellites, LTGs/ETGs, and their combinations. Section 5 is devoted to discussing the caveats and implications of our results. Finally, in Section 6, we present a summary of the paper and the conclusions.

2 ANALYSIS OF THE xGASS SURVEY

The survey xGASS (Catinella et al. 2018) is an $R_{\rm HI}$-limited census of 1179 galaxies selected by redshift and M_* in the ranges $0.01 \leq z \leq 0.05$ and 10^9 $M_\odot \leq M_* \leq 10^{11.5}$ $M_\odot$, respectively. The sample galaxies were drawn from the intersection of SDSS *DR7* (Abazajian et al. 2009), *GALEX* (Martin et al. 2005), and projected *ALFALFA* footprints (Haynes et al. 2011). The xGASS consists of two samples: (1) GASS (Catinella et al. 2010, 2012, 2013), a sample of galaxies with $M_* > 10^{10}$ $M_\odot$ and redshift $0.025 \leq z \leq 0.05$, and (2) the low-stellar mass extension of GASS (hereafter low-GASS Catinella et al. 2018), a sample of galaxies with stellar masses in the range 10^9 $M_\odot$ $\leq M_* \leq 10^{10.2}$ $M_\odot$ and redshift $0.01 \leq z \leq 0.02$. Both samples were constructed in such a way that the stellar mass distribution of the targets is roughly flat. The xGASS survey is the most complete H I observational study of a local optically based representative galaxy sample to date.

In xGASS, the H I mass is obtained from the H I observations of ALFALFA $\alpha.40$ or the Cornell H I digital archive (Springob et al. 2005). For galaxies with no H I information, observations were performed using the *Arecibo Radio Telescope* with the strategy of observing the targets until detected or until a limit of a few percent in $R_{\rm HI}$ ratio is reached. The detection limits for each sample are:

(i) GASS: $R_{\rm HI} > 0.015$ for galaxies with $M_* > 10^{10.5}$ $M_\odot$ and a constant H I mass limit of $M_{\rm HI} = 10^{8.7}$ $M_\odot$ for galaxies with lower stellar masses.

(ii) low-GASS: $R_{\rm HI} > 0.02$ for galaxies with $M_* > 10^{9.7}$ $M_\odot$ and a constant H I mass limit of $M_{\rm HI} = 10^8$ $M_\odot$ for lower mass galaxies.

The detection limits in $R_{\rm HI}$ considered mainly the telescope sensitivity, integration time, and the redshift range of the surveys.

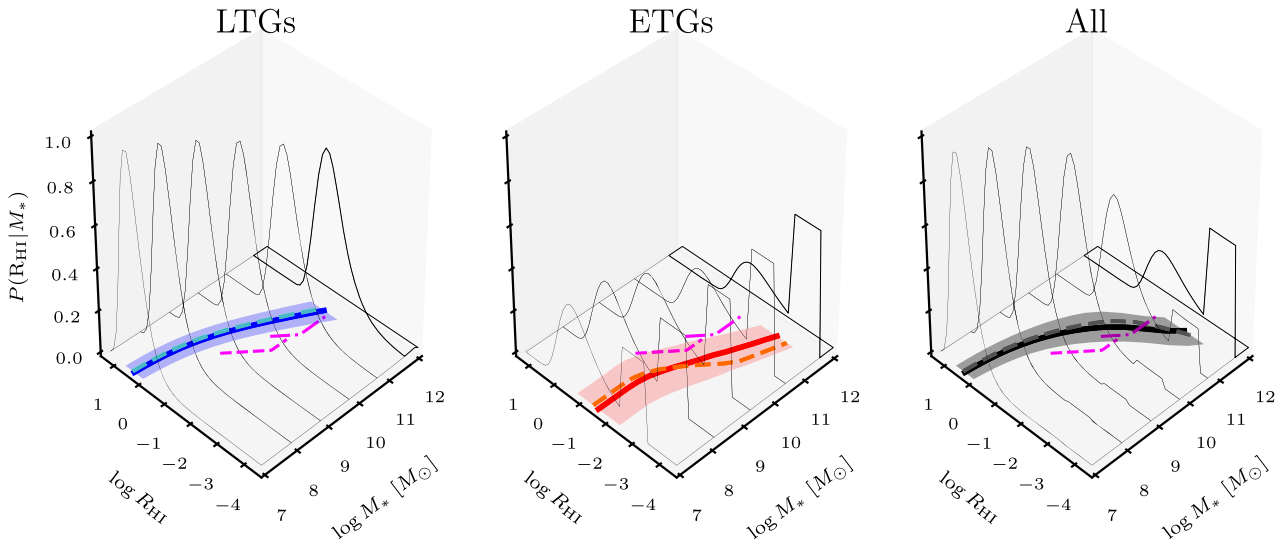

Figure 1. The $R_{\rm HI}$ gas conditional PDFs of LTGs, ETGs, and all galaxies as a function of M_* from Calette et al. (2018) and updated in Rodríguez-Puebla et al. (2020). In the projected log $R_{\rm HI}$-log M_* planes, the logarithmic mean relations and their standard deviation are shown with thick solid lines and shaded areas, respectively. The dashed lines correspond to the medians instead of the logarithmic means. The magenta dashed and dot–dashed lines show the xGASS detection limits, see Section 2.

2.1 Morphology and central/satellite designations for xGASS galaxies

At fixed M_*, the gas content in galaxies varies significantly with morphology (e.g. Kannappan et al. 2013; Boselli et al. 2014b; Calette et al. 2018). Thus, we introduce a morphological characterization for xGASS galaxies that complements the dependence on stellar mass. Here, we use the Huertas-Company et al. (2011) automated morphological classification for ∼700 000 galaxies from the SDSS DR7 spectroscopic sample, where each galaxy has a probability of being elliptical, S0, Sab, and Scd by means of support vector machines method and the Fukugita et al. (2007) sample as a training set.[1] On the other hand, Meert, Vikram & Bernardi (2015) calibrated Huertas-Company et al. (2011) probabilities to T-types using a simple linear model given by,

$$T = -4.6 \cdot P(\mathrm{Ell}) - 2.4 \cdot P(\mathrm{S0}) + 2.5 \cdot P(\mathrm{Sab}) + 6.1 \cdot P(\mathrm{Scd}). \quad (1)$$

The latter was constrained using the visual classification of Nair & Abraham (2010) by a linear regression. Using equation (1) and the probability classification from Huertas-Company et al. (2011), we assign T-types to xGASS galaxies. Of the 1179 galaxies in the xGASS sample, we find that 1150 are in the Huertas-Company et al. (2011) morphology catalogue and only consider these for our analysis.

We separate xGASS galaxies into two broad morphological groups: LTGs and ETGs. We consider ETGs as those galaxies with $T < 0.5$ and LTGs as those with $T \geq 0.5$ following Meert et al. (2015). The above corresponds respectively to S0 or earlier and Sab or later morphologies, see their equation (8) for details.

To segregate galaxies into centrals and satellites, we use the xGASS flag env_code_B defined as[2]:

$$\texttt{env\_code\_B} = \begin{cases} 0: & \text{satellite} \\ 1: & \text{isolated central} \\ 2: & \text{group central} \\ -1: & \text{not in group catalogue} \end{cases}$$

We consider centrals those galaxies with env_code_B=1 or 2. The term isolated central does not imply what typically is known in the literature as an isolated environment but it refers to the presence of only one galaxy within the halo. Satellites are those with env_code_B=0. As described in Janowiecki et al. (2017), for determining whether a galaxy is central or satellite in xGASS, the authors used the Yang et al. (2007) halo-based group catalogue updated to the SDSS DR7. For xGASS, the 'modelB' group catalogue was adopted, and cases of 'galaxy shredding' and false pairs have been resolved by visual inspection (see details in Janowiecki et al. 2017). Fortunately, only a small fraction of xGASS galaxies, 2 per cent, are not in the Yang et al. (2007) 'modelB' catalogue or suffer from galaxy shredding and false pairs. Approximately 30 per cent of xGASS galaxies are classified as satellites in groups, ∼ 50 per cent as isolated centrals, and ∼ 20 per cent as centrals (the most massive member) in groups. The central/satellite designation adopted for the xGASS survey has been used in several works for studying the effects of environment on the gas content of galaxies (e.g. Janowiecki et al. 2017, 2020; Stevens et al. 2019; Cortese et al. 2020; Watts et al. 2020). Nevertheless, it should be stressed that galaxy group finders like the Yang et al. (2005, 2007) halo-based finder may suffer from membership allocation and central/satellite designation errors. In Section 5.2.2, we discuss this caveat and how it can affect the results obtained in this paper.

[1] In Section 5.2.1, we discuss how our our results do change when applying an alternative morphological classification scheme.

[2] xGASS data description: https://xgass.icrar.org/assets/data/xGASS_representative_sample.readme

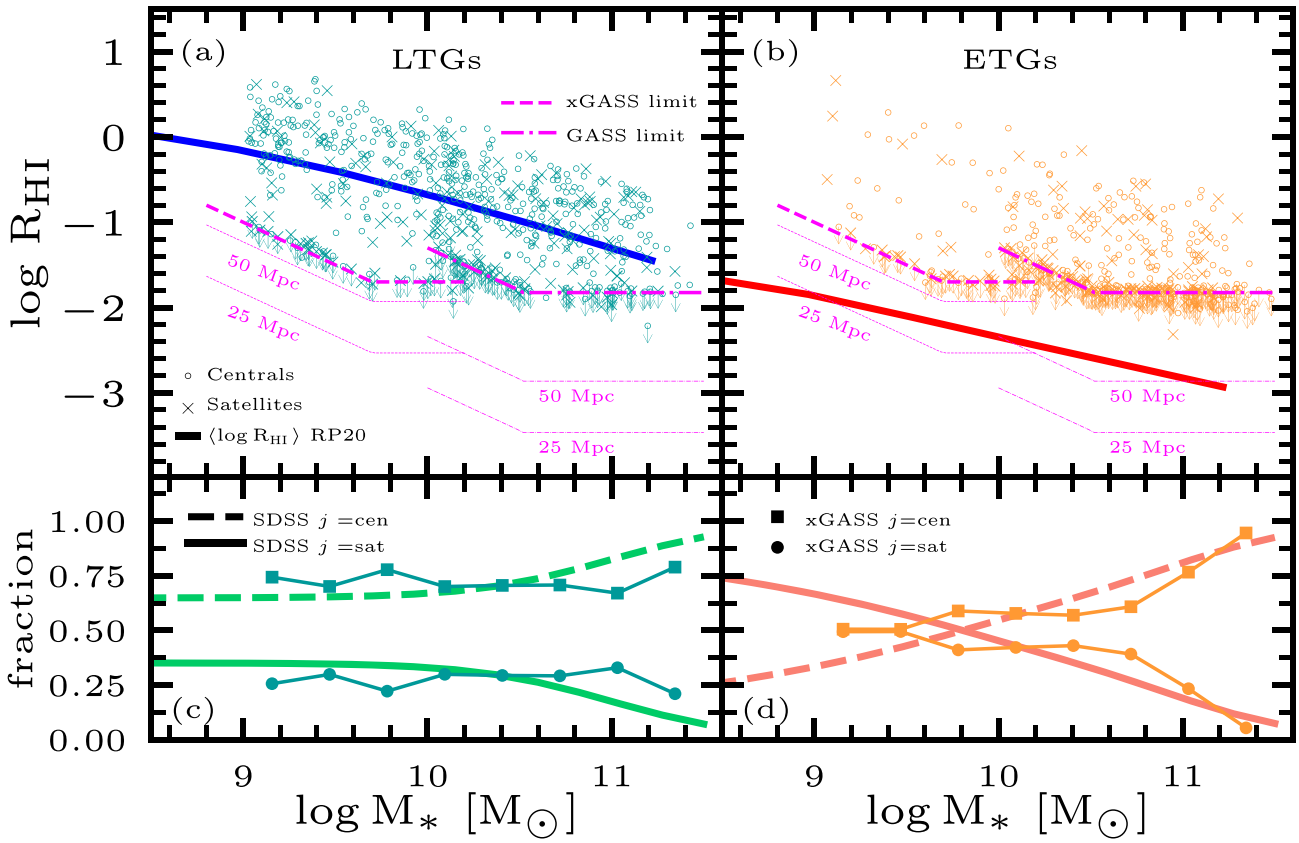

Figure 2. Presentation of the xGASS sample. Panel (a)*:* LTGs in the log $R_{\rm H\,I}$-log M_* diagram, with centrals and satellites plotted as empty circles and crosses, respectively. The downward arrows indicate the reported upper limits for non detected galaxies in radio. Dot–dashed and dashed lines show the imposed limit detection in the GASS and the low-GASS samples, respectively. We reproduce the logarithmic mean of LTGs obtained in Papers I and II with blue solid line (see also Fig. 1). Panel (b)*:* Same as panel (a) but for ETGs. Panel (c)*:* Fraction of LTGs that are satellites (circles) or centrals (squares) as a function of M_*. The respective fractions as inferred from SDSS DR7 based on the Huertas-Company et al. (2011) morphological classification and the Yang et al. (2012) central/satellite division are plotted with the dashed and solid lines, respectively. Panel (d)*:* Same as panel (c) but for ETGs.

2.2 Final xGASS sample and selection effects

The final sample of xGASS galaxies with morphology and central/satellite classifications includes 1134 objects. In panels (a) and (d) of Fig. 2, we present these galaxies in the $R_{\rm H\,I}$–M_* plane separated into LTGs and ETGs, respectively. In each panel, central and satellite galaxies are plotted with open circles and crosses, respectively, and upper limits are shown with downward arrows. The number of galaxies with upper limits is significant, 55 per cent for ETGs and 17 per cent for LTGs. The dot–dashed and dashed lines show the imposed detection limit in the GASS and GASS-low samples, respectively. Most of the upper limits pile up close to these lines. However, since galaxies are at different distances the distribution of the upper limits is somewhat scattered. In the same panels, we reproduce the logarithmic means of LTGs and ETGs obtained in Papers I and II. Their corresponding $R_{\rm H\,I}$ conditional distributions at different stellar masses are shown respectively in panels (b) and (e). In these panels, we also reproduce the xGASS detection limits. Clearly the empirical distribution of $R_{\rm H\,I}$ is truncated by the xGASS detection limits. This truncation is abrupt for ETGs, which are even above the first moments of the empirical $R_{\rm H\,I}$ PDFs (the red solid line in panel d).

An upper limit in H I mass is reported when a galaxy in a given survey has not been detected in the 21-cm line for the defined integration time and above a given signal-to-noise ratio. The H I mass upper limit is calculated using the respective H I flux detection limit and the distance to the galaxy, $M_{\rm H\,I}^{\rm u.l.} \propto D(z)^2$. When inferring any correlation or probability distribution from $M_{\rm H\,I}$, it is mandatory to account for upper limits. In Section 2.4, we describe the survival analysis we follow to do so. In addition, it is important to note that the xGASS upper limits are high and notably truncate the low-side $R_{\rm H\,I}$ distribution, specially for ETGs. This is due to the large distances in this survey, in particular, for GASS. In fact, in galaxy samples at closer distances than xGASS, a fraction of their galaxies were detected in H I with $R_{\rm H\,I}$ values below the xGASS detection limits, for instance, in ATLAS^{3D} (Cappellari et al. 2011; Serra et al. 2012) and Herschel Reference Survey (HRS; Boselli et al. 2010, 2014a). On the other hand, the H I detection limits of these closer galaxy samples, after taking into account the differences in the observational and instrumental settings, result in much lower upper limits than those from xGASS, in particular, for the GASS sample. Thus, the upper limits from xGASS are biased high due to distance selection effect. Following Paper I and based on some assumptions, in Section 2.3, we attempt to correct for this bias in the upper limits.

Panels (c) and (f) of Fig. 2 present respectively the xGASS fraction of LTGs and ETGs that are satellites, circles, or centrals, squares, as a function of M_*. In the same panels, the solid lines correspond

to fit to the satellite fractions for LTGs and ETGs from the Yang et al. (2012) SDSS DR7 galaxy group catalogue (the dashed lines are the respective central fractions and they are by definition the complements of the solid lines; see Appendix A). At this point, it is important to ask ourselves if xGASS suffers of selection effects that could bias the sample by morphology (for the morphological classification adopted here, i.e. Huertas-Company et al. 2011) and by environment. A bias in the morphology is not relevant when the inferred $R_{\rm H\,I}$–M_* relations and $R_{\rm H\,I}$ distributions are determined separately for LTGs or ETGs. However, this possible bias is expected to affect the relations and distributions for all, central, and satellite galaxies when averaging among LTGs and ETGs.

In Fig. A1 in Appendix A, we compare the ETG and satellite fractions as a function of M_* from xGASS with those measured from SDSS DR7 (panels (a) and (d), respectively). As seen, the xGASS fraction of satellites as a function of M_* roughly agree with that from the whole SDSS DR7 (the fraction of centrals is the complement). However, this is not the case for the fraction of ETGs (the fraction of LTGs is the complement): xGASS selects systematically a higher fraction of ETGs than SDSS up to $M_* \sim 10^{11}$ M$_\odot$. Obviously, the differences remain when considering only central or satellite galaxies, but they are larger for satellites, compare panels (b) and (c). For $M_* \gtrsim 10^{11}$ M$_\odot$, the difference inverts. Note that the flat distribution in mass of xGASS is not an issue in Fig. 2 given that the comparisons between fractions are at a given M_*.

For the inferences in Section 3 of the $R_{\rm H\,I}$–M_* relations and $R_{\rm H\,I}$ distributions given M_* corresponding to all galaxies (LTGs + ETGs), to all centrals (LTGs + ETGs), and to all satellites (LTGs + ETGs), we introduce weights for the xGASS galaxies in order to be consistent with the fractions of ETGs as a function of M_* for both the samples of centrals and satellites from the SDSS DR7 (panels (b) and (c) of Fig. A1). The weighting procedure is described in Appendix A.

2.3 Reestimating the H I upper limits

As mentioned above, when comparing the distribution of xGASS galaxies in Fig. 2 with the respective empirical H I conditional PDFs, shown in Fig. 1, we note that the xGASS detection limits truncate significantly the H I conditional PDFs of ETGs (the corresponding $R_{\rm H\,I}$–M_* relation lies even below the detection limits). In contrast, for LTGs the truncation is not significant given the high H I gas contents for most of these galaxies. We ask ourselves: where would non-detected ETGs in xGASS appear in the $R_{\rm H\,I}$–M_* plane if they were observed with the same instrument, observational setup, and allowed signal-to-noise ratio but at lower distances? The thin dash-dotted and dotted lines in Fig. 2, labelled respectively as 50 and 25 Mpc, show the shift that the GASS and GASS-low detection limits would have at these distances.[3] We see that at a distance of ∼25 Mpc, the detection limits lie now below the $R_{\rm H\,I}$–M_* relation of ETGs. Fortunately, there are close samples of ETGs with radio observations. As mentioned above, this is the case of the ATLAS$^{\rm 3D}$ survey (median distance of 25 Mpc), which has been used in Paper I for reestimating the $M_{\rm H\,I}$ upper limits of GASS ETGs, and eventually, for assigning detection values to a fraction of these upper limits.

Here, we follow a procedure similar as in Paper I for reestimating the $M_{\rm H\,I}$ upper limits of xGASS. We emphasize that the procedure in Paper I is based on the assumption that the H I gas content at a fixed M_* of galaxies at distances ∼25 Mpc (the median distance of ATLAS$^{\rm 3D}$) is statistically the same as that of galaxies up to 100–200 Mpc (the distances of GASS galaxies). Under this assumption, the H I observations for ATLAS$^{\rm 3D}$ (and also HRS) galaxies allowed us to re-estimate *in a statistical sense* the $R_{\rm H\,I}$ upper limits of GASS galaxies and to assign (detected) $M_{\rm H\,I}$ values to a fraction of them. Of course, only future deeper radio observations for each galaxy could provide a measure of its true H I mass or a new improved upper limit. Performing a similar analysis to GASS-low will require information of a survey such as ATLAS$^{\rm 3D}$. Unfortunately, this survey extends only down to stellar masses slightly smaller than $\sim 10^{10}$ M$_\odot$, making the extension to GASS-low impractical at this point.

In the case of LTGs, most of them are detected in GASS despite their relatively shallow H I detection limit. On the other hand, for LTGs there is not a closer and homogeneous sample similar to ATLAS$^{\rm 3D}$. Thus, in Paper I, we did not attempt to correct the upper limits of LTGs from GASS by the distance effect. For GASS-low, the fraction of radio-detected LTGs from closer samples below the GASS-low detection limit is slightly larger than in GASS. The overall fraction of upper limits for LTGs in xGASS is 17 per cent. Therefore, following the above argument for ETGs, it would be desirable to attempt to also re-estimate the upper limits of LTGs.

As mentioned above, there are not close samples, as ATLAS$^{\rm 3D}$, with more or less well-defined detection limits in $R_{\rm H\,I}$ for $M_* < 10^{10}$ M$_\odot$, both for ETGs and LTGs. However, we can use the empirically constrained $R_{\rm H\,I}$ distributions in Papers I and II to reestimate the reported GASS-low upper limits due to their bias by distance. Even more, to homogenize our procedure, we decided here to use these empirical distributions for both xGASS ETGs and LTGs. For GASS ETGs, the re-estimation of upper limits obtained here are very similar to those in Paper I. Following the discussion above, in Appendix B, we describe in detail our procedure to re-estimate the upper limits of ETGs and LTGs for xGASS.

2.4 Statistical analysis including H I upper limits

In order to estimate from xGASS, the $R_{\rm H\,I}$–M_* relations separated into central and satellite galaxies or, even more, the full $R_{\rm H\,I}$ conditional PDFs given M_*, as in Paper I, the upper limits should be taken into account adequately. In observational Astrophysics, we are often interested on particular astronomical objects (e.g. stars, galaxies, etc) and in order to design samples to study them, we set a selection criteria based on a given property, P_1, to construct such observational samples (for example stellar mass or luminosity). But there are situations when we are also interested in another property, P_2 (for example H I content). Nevertheless, due to instrumental limitations, we cannot always measure the property P_2 in all objects, instead we assign upper limits or 'censored' data values. In such situation, it is necessary to build a parent sample based on a well studied property P_1 and then examine for the property of interest P_2 from property P_1. The above description is exactly the case for the xGASS sample, in which $P_1 = M_*$ and $P_2 = M_{\rm H\,I}$.

To use both detections and upper limits from xGASS, in this work, we rely on Kaplan-Meier (KM) non-parametric estimator (Kaplan & Meier 1958) specifically developed for the analysis of censored data in clinical research, but properly adapted to astronomical data by Feigelson & Nelson (1985). For a given sample, the KM estimator allows us to obtain the cumulative distribution function (CDF) when including censored data and from different statistical estimators can be calculated. However, to obtain reliable results, it is recommended that the fraction of censored data (upper limits) be less than ∼ 50 per cent. We construct the $R_{\rm H\,I}$ CDFs at different

[3]Notice that the GASS and the GASS-low samples are at a median distance of 165 and 65 Mpc, respectively.

stellar mass bins. After applying the corrections to the ETG upper limits (see Section 2.3), the minimum $R_{\rm HI}$ values (censored data) used in the KM estimator are around $-3.0 < \log R_{\rm HI} < -3.5$, and the CDFs at these values start with fractions typically of 0.3–0.4. This means that around 30-40 per cent of ETGs have $R_{\rm HI}$ upper limits. As mentioned in the footnote of Appendix B0.1, in Paper I, we assigned real values (detections) to these galaxies by assuming they follow a top-hat function of width ∼1 dex below the minimum upper limit value of the given mass bin. Our main argument was that even quiescent ETGs should have H I gas fractions larger than a few 10^{-5}, taking into account stellar mass loss and some eventual gas capture from minor mergers and cosmic accretion.

3 RESULTS FROM xGASS

3.1 Correlations for all, central, and satellite galaxies

In the upper panels of Fig. 3, we plot again the xGASS data as in Fig. 2 but after applying to the upper limits the procedure described in Section 2.3; we added a third panel showing the whole sample (LTGs+ ETGs). For each M_* bin of width $\Delta \log M_* = 0.31$ dex, we use the procedure based on the KM estimator described in Section 2.4 to calculate the mean logarithmic value of $R_{\rm HI}$ and the standard deviation at each M_* bin. The results are plotted with circles and error bars. For comparison, the thick solid line in each panel is the respective logarithmic mean relation as obtained in Paper II and also reproduced in Figs 1 and 2 above. For LTGs, xGASS is in very good agreement with our empirical relation from Paper II. In the case of ETGs, the averages of xGASS galaxies (after re-scaling the upper limits by the distance bias) are slightly above than the corresponding relation from Paper II but within the standard deviations. Note that these upper limits lie now around the GASS and low-GASS detection limits shifted to a distance of 25 Mpc.

In the right-hand panel of Fig. 3, corresponding to all galaxies, we reproduce the logarithmic mean $R_{\rm HI}$ values reported by Catinella et al. (2018), violet squares. These authors calculated the means (i) setting the H I mass of non-detections to their upper limit values (this leads to overestimate the mean), and (ii) applying weights to correct for the stellar mass bias of the sample, that is, to make the sample mass complete in volume. Regarding (ii), it is not expected to be relevant for the means calculated in small mass bins since the weights are roughly the same for similar masses. At low and intermediate stellar masses, our means are in good agreement with those from Catinella et al. (2018) but at the highest masses, where ETGs dominate, our means are lower than those reported by these authors. This is due to the special treatment we applied to adequately include the upper limits of ETGs. Recall that we also weighted xGASS galaxies by morphology and environment to agree with the SDSS DR7 fractions of ETGs and satellites as a function of M_*, see Section 2.1. The weights correct mainly the excess of ETGs in xGASS with respect to SDSS up to $M_* \sim 10^{11}$ M$_\odot$ and the lack at larger masses (the latter specially applies for satellites), see Fig. A1. Therefore, the average values plotted in Fig. 3 for all galaxies are weighted towards LTGs up to $M_* \sim 10^{11}$ M$_\odot$ and against them at higher masses.

In Appendix C, we present results for xGASS without taking into account our procedure for the upper limits, nor the correction by morphology/environment. For LTGs, the results are almost indistinguishable from those presented here but for ETGs, for which the fraction of non-detections is high, for $M_* > 5 \times 10^9$ M$_\odot$, the mean $R_{\rm HI}$ values and their standard deviations obtained with the KM estimator are very uncertain and can be taken just as an upper bound. For the whole sample, combining LTGs and ETGs, we show that the weights by morphology slightly increase the mean $R_{\rm HI}$ values for masses below $M_* \sim 5 \times 10^{10}$ M$_\odot$, while for the highest masses, the weights decrease the mean $R_{\rm HI}$ by ∼0.3 dex.

The middle panels of Fig. 3 show $\langle \log R_{\rm HI} \rangle$ and the errors of the mean, this time for central and satellite galaxies, separately. The solid lines connect the respective means showed in the upper panels. Centrals have on average slightly higher H I gas fractions than the average. For satellites, the differences are more pronounced especially towards lower stellar masses. Overall, centrals have higher H I gas contents than satellites, in particular at lower masses.

In the lower panels of Fig. 3, we plot the logarithmic standard deviations for centrals and satellites at each mass bin for LTG, ETGs, and all galaxies. The population of ETGs presents larger scatter around the $R_{\rm HI}$–M_* relations for centrals and satellites than LTGs. In each of the lower panels of Fig. 3, we plot also the relative differences between the corresponding central and satellite means, $\Delta\langle \log R_{\rm HI} \rangle_{\rm cen-sat} \equiv \langle \log R_{\rm HI,cen} \rangle - \langle \log R_{\rm HI,sat} \rangle$ (thick solid lines), plotted in the medium panels. As seen, these differences tend to be smaller than the corresponding standard deviations, both for LTGs and ETGs, specially at larger masses. On average, satellite galaxies have lower H I gas contents than centrals, specially at low masses. Finally, in panel (i), corresponding to all galaxies, we reproduce the relative differences between the central and satellite medians reported in Stevens et al. (2019) for xGASS (long-dashed line). Despite them measuring medians and us logarithmic means and them setting non-detections to their upper limit values, the agreement is reasonable.

3.2 Conditional H I distributions for all, central, and satellite galaxies

In Fig. 4, we compare the $R_{\rm HI}$ conditional CDFs of LTGs and ETGs from the processed xGASS sample (thick solid lines) with those inferred empirically in Papers I and II (thin solid lines). The fits were averaged within the width of the M_* bin. The cumulative distributions for xGASS ETGs start at fractions around 0.3−0.4. These are the fractions of the remaining upper limits after our corrections of Sections 2.3 and 2.4. If we proceed as in Paper I, we should assign $R_{\rm HI}$ values following a top-hat function of width ∼1 dex below the lowest upper limit value in each mass bin for undetected ETGs. This is shown in Fig. 4 with dotted lines. The xGASS H I conditional CDFs for LTGs agree well with the analytical fits constrained in Paper I. For ETGs, the CDFs from xGASS tend to be somewhat shifted to higher $R_{\rm HI}$ values than those determined in Paper I.[4] Differences are seen also in the respective logarithmic mean values plotted in Fig. 3.

In Appendix C, we compare the $R_{\rm HI}$ conditional CDFs shown in Fig. 4 with those obtained without correcting the xGASS upper limits, Fig. C2. From this comparison, it is evident that without this correction, the CDFs for ETGs result poorly constrained.

Fig. 5 presents the $R_{\rm HI}$ conditional CDFs in different M_* bins calculated as described in Section 2.4 for the whole xGASS sample (black lines), and for centrals (dark grey lines) and satellites (purple lines) only, that is, $P^i(> R_{\rm HI}|M_*)$, where i refers to all, central or satellite, respectively. We find that the lower M_* the larger the difference in the distributions between central and satellite galaxies,

[4]In Paper I, to infer the H I conditional distributions, (i) we used not only the GASS survey but other samples, and (ii) for converting to detections a fraction of the ETG GASS upper limits, a uniform $R_{\rm HI}$ distribution was used while here the empirical $R_{\rm HI}$ conditional PDFs for ETGs constrained in Paper I are used, see Appendix B0.1. Therefore, we expect differences between the H I conditional CDFs of ETGs in Paper I and those estimated here for xGASS.

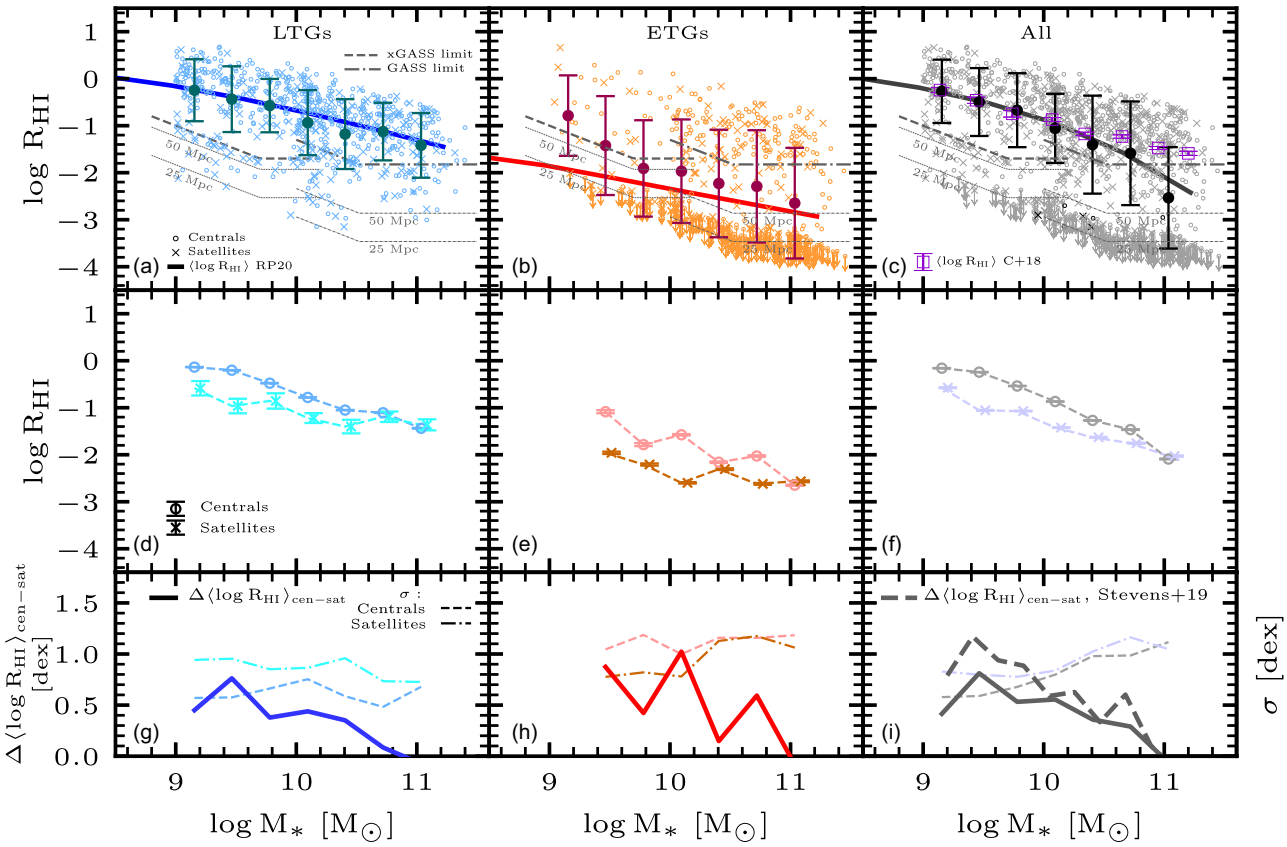

Figure 3. Upper panels: xGASS galaxies in the log $R_{\rm HI}$ − log M_* diagram, as in Fig. 2 but after applying corrections to the upper limits (see text). The symbols with error bars are the logarithmic means and standard deviations in M_* bins obtained with the KM estimator for taking into account upper limits (the data are presented in tabulated form in the Supplementary Material). The solid lines show the mean $R_{\rm HI}$–M_* relations from Paper II. In panel (c), the violet empty squares are the logarithmic means as reported in Catinella et al. (2018). Middle panels: Logarithmic means and their error on the mean in M_* bins obtained with the KM estimator for the subpopulations of central (open circles with error bars) and satellite (crosses with error bars) galaxies, for LTG, ETG, and all galaxies from left to right. Lower panels: Second moments of the log $R_{\rm HI}$ distributions from the KM estimator for the subpopulations of central and satellite galaxies (dashed lines) showed in the upper panels. The solid lines are the relative differences between the means of central and satellite subpopulations showed in the medium panels. The long-dashed line in panel (i) corresponds to the relative differences between the medians of centrals and satellites as reported for xGASS in Stevens et al. (2019).

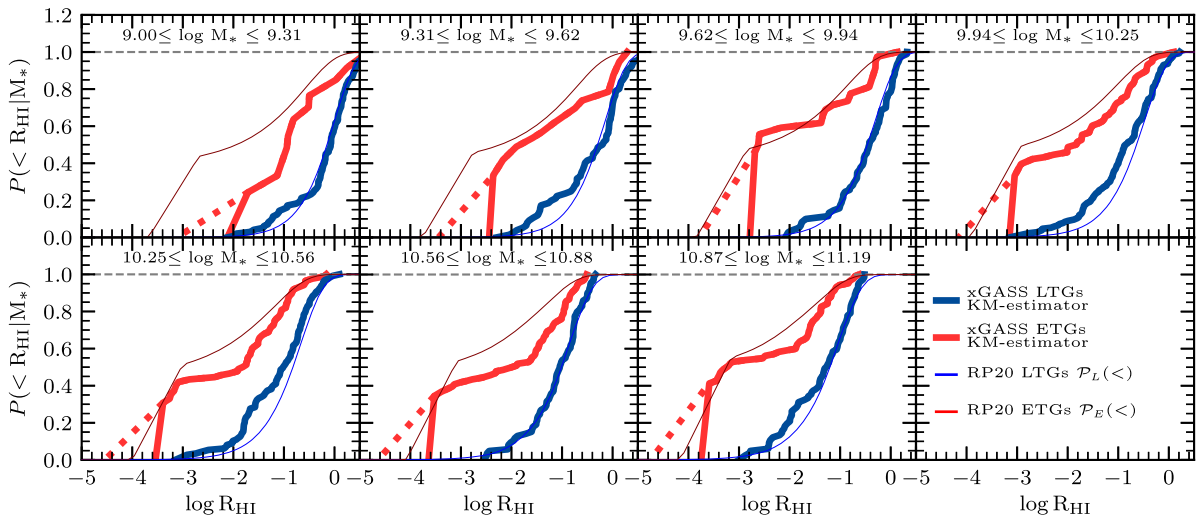

Figure 4. Cumulative histograms of LTG (blue lines) and ETG (red lines) H I conditional distributions (CDFs) at different M_* bins from the processed xGASS sample. For comparison, fits to the respective CDFs from Paper II are shown with thin lines. For a correct comparison, these fits were averaged within the mass ranges of the bins.

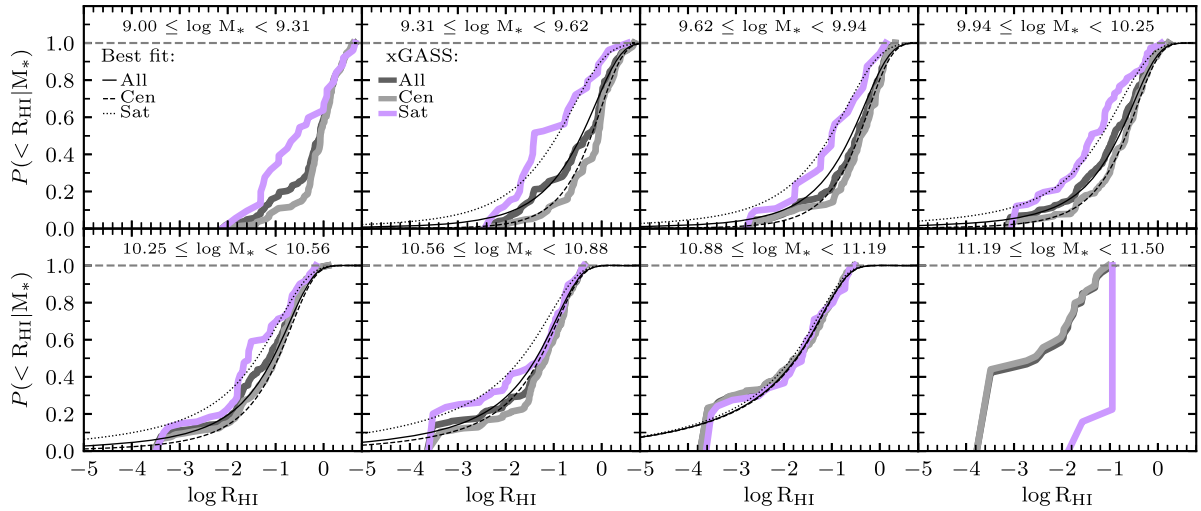

Figure 5. Cumulative H I conditional distributions (CDFs) at different M_* bins from the processed xGASS sample of all galaxies and for only centrals and satellites, see colour notation in the first panel. The solid, dashed, and dotted lines are our best joint fits to the different subpopulations shown in this figure and in Figs 6 and 7, see text.

with the latter having lower $R_{\rm HI}$. Recall that for calculating these distributions, the xGASS sample has been weighted by morphology and environment to agree with the SDSS DR7 fractions as a function of M_*, see Section 2.1 and Appendix A. The main bias of xGASS galaxies is actually by morphology; the bias by environment is small and mainly due to the former. Figs 6 and 7 are as Fig. 5 but now for LTGs and ETGs from xGASS, respectively.

3.3 Corrections from xGASS to calculate H I distributions for centrals and satellites

We would like to obtain from the xGASS analysis presented above a way to estimate the H I conditional CDFs of central and satellite galaxies when only the average CDFs (among centrals and satellites) are known. If $\mathcal{P}_j(>R_{\rm HI}|M_*)$, $j=$ LTG or ETG, are the H I conditional CDFs from Paper II, then the corresponding CDFs for central and satellites can be calculated as:

$$\mathcal{P}^i_j(>R_{\rm HI}|M_*)=\left[\frac{P^i_j(>R_{\rm HI}|M_*)}{P_j(>R_{\rm HI}|M_*)}\right]_{\rm xGASS}\times\mathcal{P}_j(>R_{\rm HI}|M_*), \quad (2)$$

where i refers to either central or satellite galaxy, and the subindex xGASS refers to analytic fits to the H I CDFs constrained above. Thus, our goal now is to (i) perform a continuous analytic fit to the different xGASS H I CDFs given M_* entering in equation (2), and (ii) to be able to extrapolate the fits to lower stellar masses than those of the xGASS sample.

The H I conditional CDFs from the processed xGASS data presented in Figs 5–7 are for the whole sample as well as for different subsamples. In many cases, the numbers of objects in a given M_* bin, specially for subsamples containing ETGs and satellites, are low. Then, the CDFs are poorly defined and may suffer of strong sample variance. In view of this, performing fits independently to each CDF is not viable. Besides, it is important that the fitted functions describing the CDFs obey by construction the law of total probability. According to this law applied to our context, the relation of the total conditional probability distribution of $R_{\rm HI}$ given M_*, $P_T(<R_{\rm HI}|M_*)$, with, for example, two subsamples A and B, with their respective conditional probability distributions $P_A(<R_{\rm HI}|M_*)$ and $P_B(<R_{\rm HI}|M_*)$, is given by:

$$P_T(<R_{\rm HI}|M_*)=P_A(M_*)P_A(<R_{\rm HI}|M_*)+P_B(M_*)P_B(<R_{\rm HI}|M_*), \quad (3)$$

where $P_A(M_*)$ and $P_B(M_*)$ are the marginalized probability distributions of these subsamples. In our case, the marginalized probabilities are the fractions of galaxies in the samples A and B as a function of stellar mass, $\phi_A(M_*)/\phi_T(M_*)$ and $\phi_B(M_*)/\phi_T(M_*)$, respectively. In Appendix D, we present the different equations that should be obeyed according to the law of total probability for the whole sample of galaxies and different subsamples of LTGs/ETGs, centrals/satellites, and their combinations. In these 'probability conservation' equations enter different fractions of subsamples (the marginalized probability distributions) as a function of M_*. In Appendix A, we obtain analytic fits to these fractions using the volume-complete SDSS survey. As discussed in Section 2.1, the fractions of ETGs (centrals or satellites) as a function of M_* in xGASS are different to those from SDSS. This is why we decided to weight the xGASS sample to agree with the SDSS DR7 morphological fractions. Having done this, we can use then the SDSS fractions in the 'probability conservation' equations mentioned above.

Based on the considerations discussed above, we implement the following strategy for obtaining the fits to the $R_{\rm HI}$ conditional CDFs of the whole xGASS sample as well as of different subsamples:

(i) Propose parametric functions that describe the $R_{\rm HI}$ conditional CDFs given M_* of the following four galaxy subsamples: all LTGs, all ETGs, central LTGs, and central ETGs.

(ii) Calculate the $R_{\rm HI}$ conditional CDFs given M_* for: the whole sample of galaxies, and the four subsamples of centrals, satellites, satellite LTGs, and satellite ETGs, from the CDFs of the previous item by means of the equations of total probability (see Appendix D).

(iii) Implement a continuous joint fitting procedure to the $R_{\rm HI}$ conditional CDFs given M_* of the whole sample and the different subsamples mentioned above as obtained from xGASS after our processing (Figs 5–7) in order to constrain the parameters of the functions mentioned in the first item.

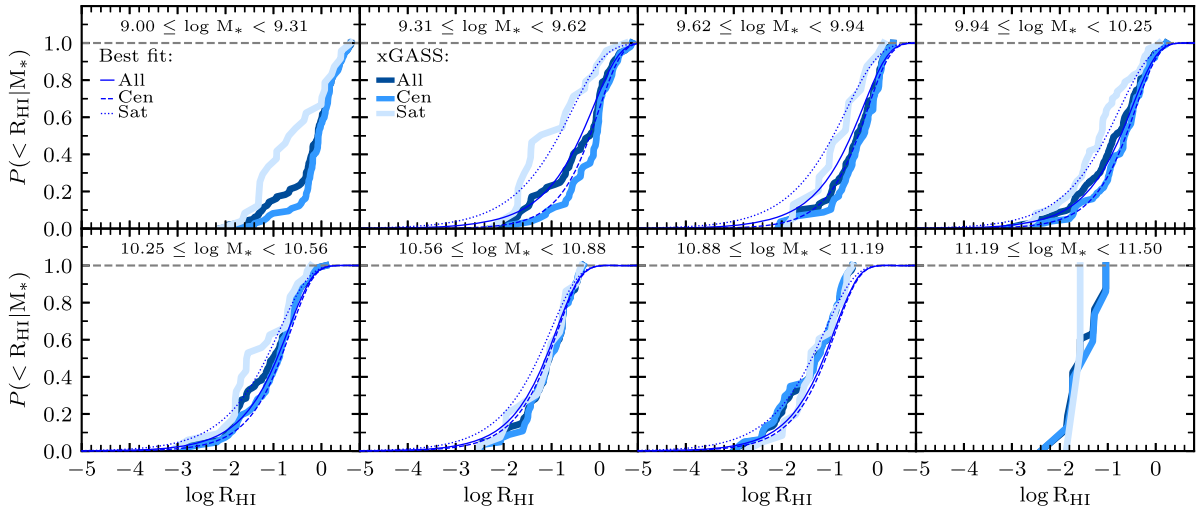

Figure 6. As Fig. 5 but for the subsample of LTGs.

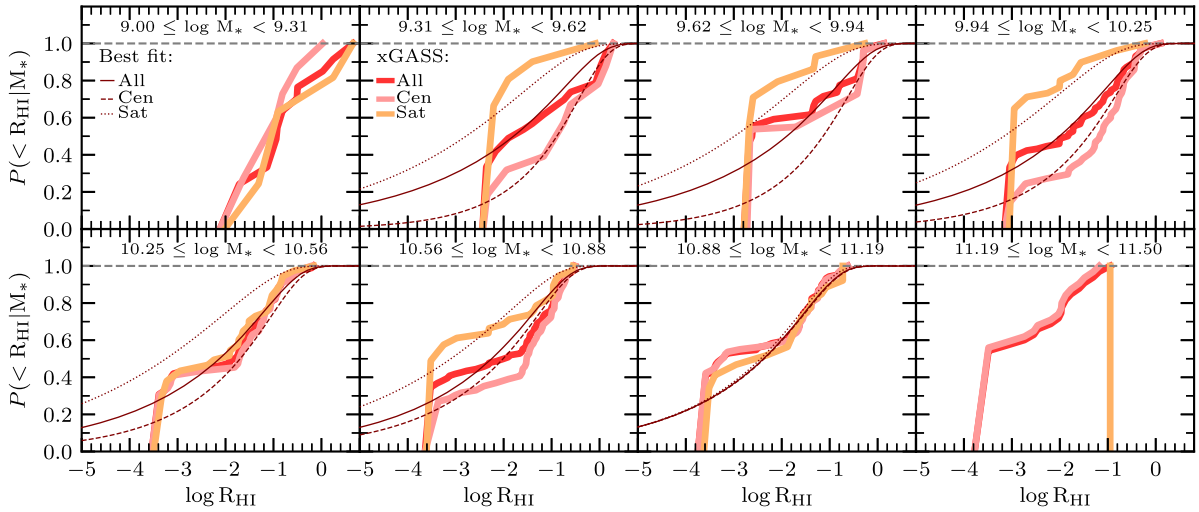

Figure 7. As Fig. 5 but for the subsample of ETGs.

For item (i), we propose a generic function for the four subsets of H I CDFs, the incomplete gamma function[5]:

$$\mathcal{P}(< x|M_*) = \frac{1}{\Gamma(\alpha)} \int_0 e^{-x} x^{\alpha-1} \mathrm{d}x, \tag{4}$$

where Γ is the gamma function, $x \equiv R_{\rm HI}/{\rm R}_0$, and the parameters α and $\rm R_0$ depend on M_*. We parametrize these dependencies as:

$$\alpha(M_*) = a(\log M_* - 10) + b, \tag{5}$$

where a and b are the slope and normalization of the power law, respectively, and

$${\rm R}_0(M_*) = \frac{c}{\left(\frac{M_*}{M_{\rm tr}}\right)^d + \left(\frac{M_*}{M_{\rm tr}}\right)^e}. \tag{6}$$

Here, c is a normalization coefficient, $M_{\rm tr}$ is the transition mass where the double power law changes its slope, d and e are the slopes for the low- and high-mass ends, respectively. In fact, for the mass range of xGASS galaxies, a single power law is enough to describe $R_0(M_*)$. However, since we will extrapolate the fits of xGASS $R_{\rm HI}$ CDFs to lower stellar masses, the second power law is necessary. We have found that the values of d and $M_{\rm tr}$ can be fixed, and not left as free parameters. These values were constrained in Paper II from the H I CDFs of LTGs and ETGs for the compilation and processing presented in Paper I in a large M_* range; we fix these parameters to the values constrained therein: $d = -0.018$ and $\log(M_{\rm tr}/{\rm M}_\odot) = 8.646$ for LTGs; $d = -0.820$ and $\log(M_{\rm tr}/{\rm M}_\odot) = 8.354$ for ETGs. Thus, in equation (4–6) there are four free parameters, a, b, c, and e

[5]We have shown in Paper I that the H I conditional PDFs given M_* can be described by Schechter-like functions. Thus, it is reasonable to propose the incomplete gamma function for describing the respective cumulative PDFs. On the other hand, given the low numbers and non-regular variations in the $R_{\rm HI}$ CDFs with mass of some subsamples from xGASS, it is impractical to search for functions with more parameters.

Table 1. Best-fitting parameters to four sets of H I CDFs.

CDFs	*a*	*b*	*c*	*e*
LTGs	0.005 ± 0.09	0.53 ± 0.09	0.79 ± 0.18	0.67 ± 0.12
LTGs Centrals	-0.21 ± 0.15	0.71 ± 0.15	0.67 ± 0.12	0.60 ± 0.11
ETGs	0.07 ± 0.05	0.22 ± 0.07	0.86 ± 0.09	0.65 ± 0.09
ETGs Centrals	-0.004 ± 0.11	0.31 ± 0.15	1.09 ± 0.13	0.75 ± 0.12

that remain. The above function equation (4) is proposed to describe each one of the four subsamples of CDFs mentioned in (i). Therefore, we have 16 free parameters in all.

We constrain the 16 free parameters by jointly fitting the nine sets of $R_{\rm HI}$ conditional CDFs from xGASS mentioned in (i) and (ii) above, and plot them in Figs 5–7. To do so we use a Monte Carlo Markov Chain method described in detail in Rodríguez-Puebla, Avila-Reese & Drory (2013). We did not use the information from the largest and lowest stellar mass bins in all the cases because the data in these bins are scarce and the corresponding CDFs are poorly determined. In Table 1, we present the best constrained values for the 16 free parameters. With these values, the four xGASS $R_{\rm HI}$ conditional CDFs mentioned in (i) above are fully described. By using the equations from Appendix D, the other five $R_{\rm HI}$ CDFs mentioned in (ii) are also described. Thus, any xGASS H I conditional CDF given M_* is described analytically by the fits, in particular those CDFs in the brackets in equation (2). However, we remark that our aim here is not to determine the $R_{\rm HI}$ conditional distributions for the xGASS survey but to capture the trends with stellar mass of the central- and satellite-to-total ratios as a function of $R_{\rm HI}$ for LTGs and ETGs, that is, the term in the brackets of equation (2). This term combined with our previous accurate inferences of the $R_{\rm HI}$ conditional distributions of LTGs and ETGs (the second term in equation 2) will allow us to estimate the respective $R_{\rm HI}$ distributions of central and satellite galaxies.

The obtained best fits from the continuous joint fitting procedure are shown in Figs 5–7 with thin solid, dashed, and dotted lines. The fits capture the main systematic trends of the different conditional CDFs with $R_{\rm HI}$ and M_*. For some mass bins of ETGs (Fig. 7), the fits depart from the data. However, note that the differences between central and satellite galaxy CDFs in these cases move away from the observed overall systematic trend with mass. Recall that the fits are designed to capture the continuous trends for all, late-, and early-type samples *jointly*. While we might propose functions with more parameters, the uncertainties and scarcity of the data for describing the CDFs as a function of M_* of the whole sample as well as of the different subsamples do not warranty statistically significant improvements in the fits.

Finally, note that the stellar mass range over which our best-fitting models are constrained for central and satellite galaxies by the xGASS data is at $10^9 \lesssim M_*/{\rm M}_\odot \lesssim 10^{11.5}$. Conservatively, in the next sections, we will assume that our best-fitting models are still valid no more than 0.5 dex above and below the above M_* range of the xGASS data, as indicated in the figures. Nonetheless, our previous empirical determinations for the $R_{\rm HI}$ conditional CDFs (not including the separation between centrals and satellites) extend down to $M_* \sim 10^7$ M$_\odot$. Thus, to estimate these CDFs separated into centrals and satellites at low masses using equation (2), we extrapolate the best-fitting models constrained by the xGASS data. For this, we extrapolate to low masses the constrained mass-dependent functions given in equations (5) and (6) as well as the fractions and subfractions as a function of M_* entering in the equations of 'probability conservation' presented in Appendix D. We use the fits to these fractions and subfractions to the SDSS data presented in Appendix A to extrapolate them down to $M_* \sim 10^7$ M$_\odot$. Unfortunately, information on the H I gas content of dwarf galaxies that have been separated into centrals and satellites is very limited. Such information can be found in the UNGC catalogue of very local galaxies (Karachentsev, Makarov & Kaisina 2013), used in Paper I. Fig. 8 shows the differences of the logarithmic mean $R_{\rm HI}$ values between centrals and satellites from UNGC (calculated taking into account upper limits) along with these differences as calculated from our $R_{\rm HI}$ conditional distributions and the extrapolations of our best-fitting models to the xGASS data. The comparison shows that our extrapolation provides results that are consistent within the uncertainties with the UNGC observational data.

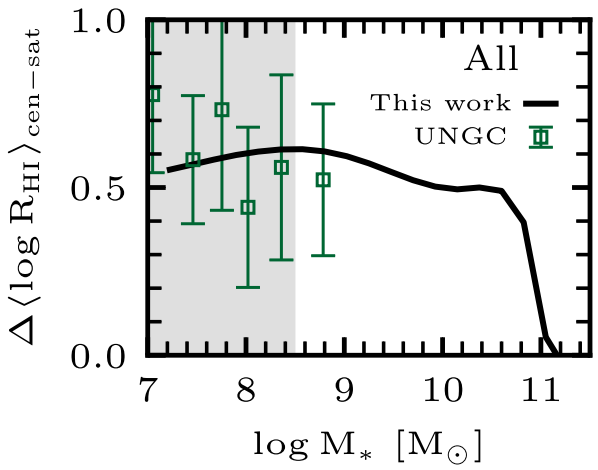

Figure 8. Difference of the logarithmic mean $R_{\rm HI}$ between central and satellite galaxies (in dex). The black solid line corresponds to this difference as a function of M_* from our results. Green squares with error bars are differences from the UNGC catalogue for $M_* < 10^9$ M$_\odot$. Error bars result from propagating the errors of the mean of central and satellites in the given mass bins. The shaded grey area indicates the extrapolation to lower masses of our empirically constrained model.

4 THE BIVARIATE $M_{\rm HI}$ AND M_* DISTRIBUTIONS OF CENTRAL AND SATELLITE GALAXIES

We are now in position to apply the xGASS-based functions found in the previous section (and their extrapolations to lower masses) to the $R_{\rm HI}$ conditional distributions of LTGs and ETGs from Paper II to obtain the corresponding distributions for central and satellite galaxies, see equation (2). The above is the main goal of this paper. From these $R_{\rm HI}$ distributions as a function of M_*, we can calculate any statistical estimator, for example the first and second moments, that is, the $R_{\rm HI}$–M_* relations and their scatters for both central and satellite galaxies. In Figs S1-S3 from the supplementary material, we show our empirically determined $R_{\rm HI}$ conditional PDFs for different masses, and for all, LTG, and ETGs separated into centrals and satellites, including our extrapolations to low stellar masses. In these figures, the xGASS PDFs as obtained in the previous section are also shown. Note than in the latter case, they correspond actually to averages within the given mass bins.

Following, we extend the results showed in Paper II regarding the joint or bivariate M_* and $R_{\rm HI}$ distribution for all galaxies but now separating them into centrals and satellites. As discussed in that

paper, by combining the $M_{\rm H\,I}$ (or $R_{\rm H\,I}$) conditional PDFs given M_* and the GSMF, $\phi_*(M_*)$, the bivariate distribution function, $\Phi(R_{\rm H\,I}, M_*)$, can be calculated. This function is defined as the bivariate number of galaxies within the mass ranges $\log M_* \pm {\rm d}\log M_*/2$ and $\log R_{\rm H\,I} \pm {\rm d}\log R_{\rm H\,I}/2$ in a given volume V, and it has units of dex^{-2} Mpc^{-3}.

In the left-hand panels of Fig. 9, from top to bottom, we show the bivariate M_* and $R_{\rm H\,I}$ distribution for all, LTGs, and ETGs, respectively. The coloured isocountours correspond to different intervals of bivariate number densities, $\Phi(R_{\rm H\,I}, M_*)$, as indicated in the palette (notice that they display four orders of magnitude). To construct these bivariate distributions, we used the $R_{\rm H\,I}$ conditional PDFs given M_* for LTGs and ETGs, the GSMF, and the fractions of LTGs and ETGs as a function of M_* reported in Paper II. The solid lines show the logarithmic mean relations, $\langle \log R_{\rm H\,I}\rangle$-$\log M_*$. As extensively discussed in Papers I and II, since LTGs dominate in number density at low masses, the $\langle \log R_{\rm H\,I}\rangle$–$\log M_*$ relation of all galaxies is similar to the one of LTGs up to $M_* \sim 10^{10}$ M$_\odot$. At higher masses, the fraction of ETGs, which have much lower H I gas contents (compare the medium and bottom left panels of Fig. 9), increase and then the relation of all galaxies strongly falls to be finally similar to the one of ETGs at $M_* \gtrsim 10^{11.7}$ M$_\odot$. Note that the $R_{\rm H\,I}$ distribution for ETGs is non-regular, with a second concentration of galaxies at very low values of $R_{\rm H\,I}$. The above is due to the top-hat component of the $R_{\rm H\,I}$ conditional PDFs (see Fig. 1).

The new results from this paper are the bivariate distributions for the galaxies separated into *centrals and satellites*, both for the LTG and ETG subsamples as well as for the total galaxy population. The left-hand panels of Fig. 9 also show the logarithmic mean relations for the central and satellite subsamples, respectively. The middle and right-hand panels present the bivariate M_* and $R_{\rm H\,I}$ distribution of the central and satellite subsamples with their respective logarithmic mean relations. The dashed and dotted lines in these panels show the arithmetic mean relations, $\langle R_{\rm H\,I}\rangle$-$M_*$, and the relations using the median of $R_{\rm H\,I}$, respectively.

For LTGs, satellites have on average a lower H I gas content than centrals. In particular, H I gas-rich galaxies with $R_{\rm H\,I} > 5$ are all centrals (there are no gas-rich satellites). On the other hand, the gas-poor low-mass LTGs are mostly satellites. At $M_* \gtrsim 5 \times 10^{10}$ M$_\odot$, central and satellite LTGs have approximately similar $R_{\rm H\,I}$ gas distributions.

For ETGs, the difference in the $R_{\rm H\,I}$ distribution between centrals and satellites is more significant than for LTGs. At $M_* < 10^9$ M$_\odot$, among the ETGs, satellites are much more common than centrals. The H I gas contents of these satellite ETGs is strongly bimodal, with a subpopulation of galaxies with $R_{\rm H\,I}$ values close to those of the central ETGs and another subpopulation with very low $R_{\rm H\,I}$ values. For central ETGs of masses $\lesssim 10^{10}$ M$_\odot$, there is a small fraction with relatively high values of $R_{\rm H\,I}$. They probably correspond to the so-called blue ETGs, some of which are even star forming (Lacerna et al. 2016, 2020). The blue/star-forming ETGs are typically very isolated galaxies and they indeed are expected to have relatively high gas fractions. At $M_* > 5 \times 10^9$ M$_\odot$, centrals are more common than satellites. The difference in the $R_{\rm H\,I}$ distribution of the centrals and satellites ETGs is small.

4.1 The H I mass functions

As shown in Paper II, the integration (marginalization) of the bivariate $M_{\rm H\,I}$ and M_* distribution over M_* results in the H I MF. The panel (a) of Fig. 10 presents the above distribution, $\Phi(M_{\rm H\,I}|M_*)$, for all galaxies and the projected H I MF (right rotated subpanel). We also plot the logarithmic mean of $M_{\rm H\,I}$ as a function of M_* for all, central, and satellite galaxies, as well as the decomposition of the H I MF into centrals and satellites. For completeness, the GSMFs of all, central, and satellite galaxies are plotted in the upper sub-panel; these functions are actually input in our approach along with the H I conditional PDFs given M_*.

In Paper II, it was shown that our empirical H I MF agrees well with those measured from the blind radio surveys ALFALFA and HIPASS, down to the completeness of our inference, $M_{\rm H\,I} \sim 10^8$ M$_\odot$, which results from the completeness limit of the input GSMF, $M_* \sim 10^7$ M$_\odot$. As seen in Fig. 10, the H I MF is dominated by central galaxies at all masses. The fraction of centrals (satellites) is ~ 90 per cent (~ 10 per cent) or more (less) for $M_{\rm H\,I} \gtrsim 10^9M_\odot$. For masses down to $\sim 10^8$ M$_\odot$, the fraction of centrals (satellites) decreases down to ~ 70 per cent (increases up to ~ 30 per cent). The differences in number density between central and satellites are larger for $M_{\rm H\,I}$ than for M_*. In panels (b) and (c) of Fig. 10 we present the bivariate distributions and their projections, the H I MF and GSMF, as in the panel (a), but for the subsamples of LTGs and ETGs. Since LTGs dominante in abundance, their mass functions are similar to those of the whole galaxy population.

5 DISCUSSION

5.1 On the H I gas fraction of central and satellite galaxies

There are several pieces of evidence that the H I gas fraction of galaxies tends to be lower in higher density environments (e.g. Haynes & Giovanelli 1984; Gavazzi et al. 2005; Cortese et al. 2011; Rasmussen et al. 2012; Catinella et al. 2013; Boselli et al. 2014b). Studies of the H I gas content of member galaxies within clusters have shown that galaxies in most massive clusters are H I deficient, especially towards the centre (e.g. Haynes & Giovanelli 1984; Bravo-Alfaro et al. 2000; Solanes et al. 2001; Rasmussen et al. 2012; Serra et al. 2012; Taylor et al. 2012; Gavazzi et al. 2013). However, the above can be in part due to the morphology–density relation; that is, ETGs, which have exhausted their gas efficiently and early and are *intrinsically* gas-poorer, are more abundant in the higher density regions of groups and clusters than LTGs. On the other hand, the H I gas content in very isolated LTGs is on average higher than in cluster LTGs, however, the differences tend to be within the 1σ scatter, see Paper I and references therein. The differences between these two opposite environments are larger for ETGs.

Other authors, rather than exploring environmental effects in specific clusters or for very isolated galaxies, used statistical samples to study the effects of the cluster/group mass and richness on the H I gas content of galaxies, mainly the satellite ones (e.g, Hess & Wilcots 2013; Yoon & Rosenberg 2015; Stark et al. 2016; Brown et al. 2017; Lu et al. 2020). Once a galaxy becomes a satellite inside a halo, the local environmental effects (ram pressure and viscous stripping, starvation, harassment, tidal interactions) work in the direction of lowering the gas content of the galaxy, more efficiently in more massive and rich haloes (see e.g. Stark et al. 2016; Stevens et al. 2019, and references therein). It is worth mentioning that in simulations (Wright et al. 2019), it was found that what matters most for the quenching time-scale of satellites is not the halo mass, but the ratio between the satellite galaxy mass to the halo mass, with smaller ratios being associated to faster quenching.

By means of the H I statistical stacking technique applied to an overlap between the ALFALFA survey and the SDSS Yang et al. (2007) halo-based group catalogue, Brown et al. (2017) found that satellites in more massive haloes have on average lower H I content

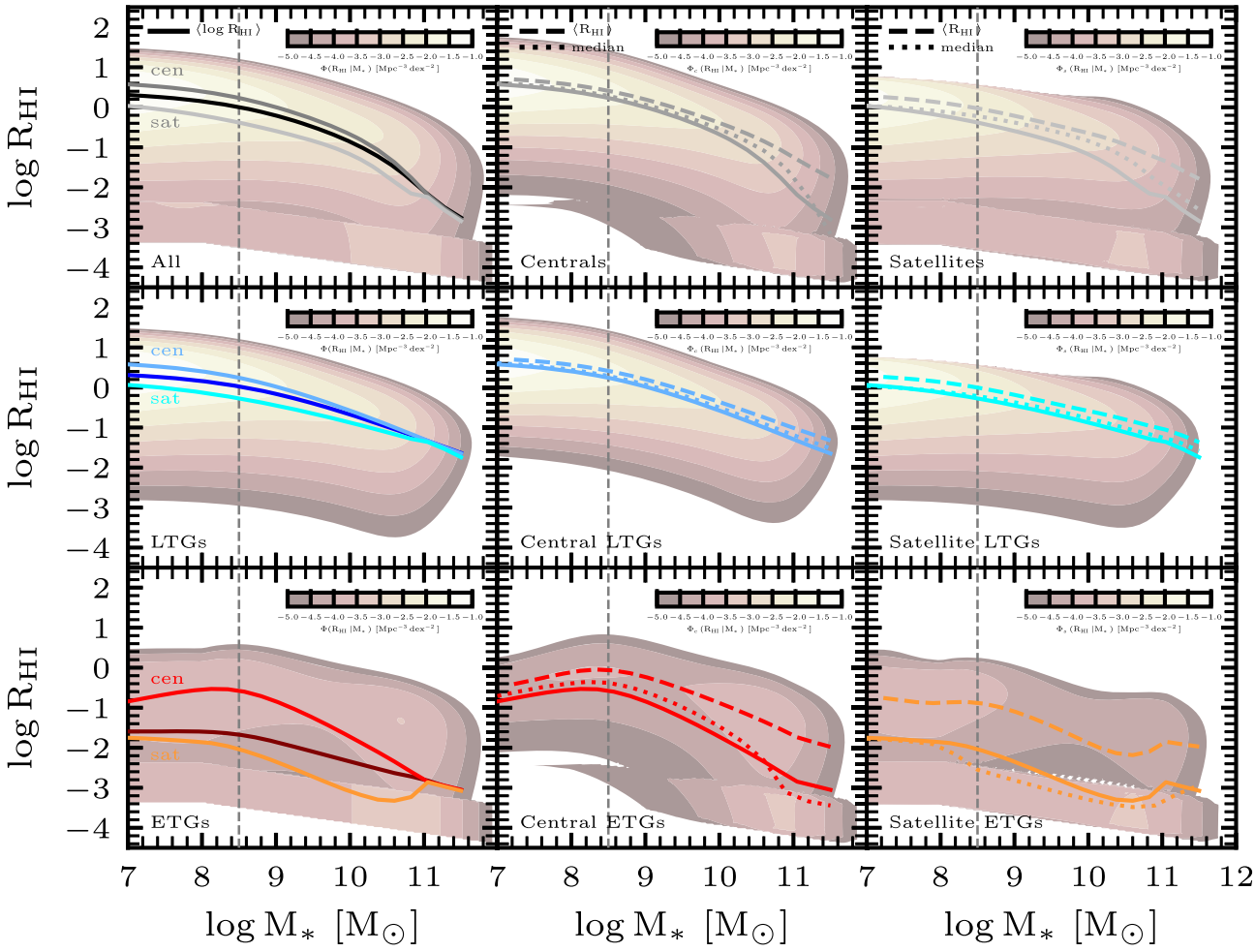

Figure 9. Empirical bivariate $R_{\rm HI}$ and M_*distributions, $\Phi(R_{\rm HI}|M_*)$. Upper panels: From left to right, the distributions for all, central, and satellite galaxies. The solid black, dark grey, and light grey lines in the left-hand panel compare the logarithmic means, $\langle \log R_{\rm HI}\rangle(M_*)$, of all, central, and satellite galaxies, respectively. Satellite galaxies have lower H I gas contents than centrals. The dark and bright lines are reproduced in the medium and right-hand panels, respectively. In these two last panels are also shown the arithmetic means ($\langle R_{\rm HI}\rangle(M_*)$, dashed line) and the medians (dotted line). Middle panels: as the upper panels but now for only LTGs. Lower panels: as the upper panels but now for only ETGs. The distribution for ETGs is highly bimodal. Hence, the different statistical estimators differ significantly among them. The dashed grey lines indicate extrapolations to lower stellar masses of our empirically constrained model for centrals and satellites.

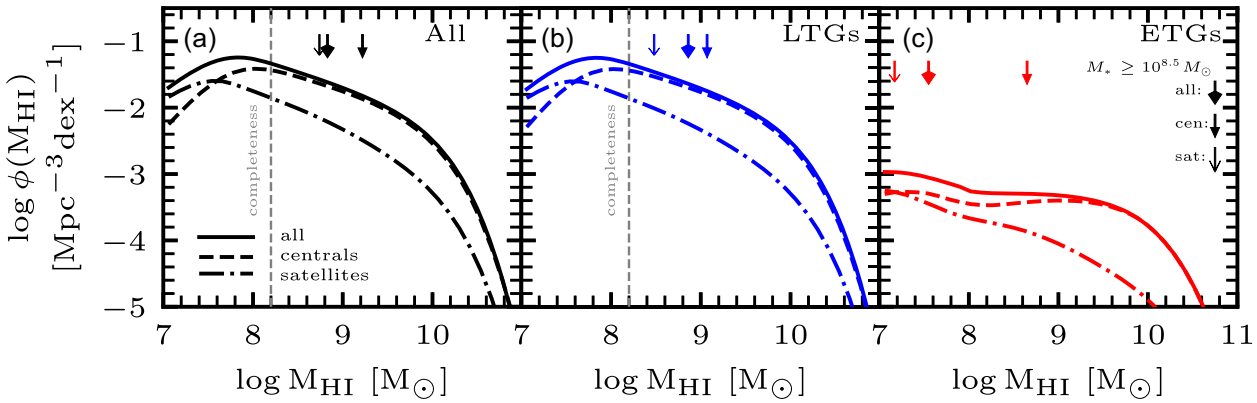

Figure 10. Panel (a): H I MF for all, central, and satellite galaxies when integrating the bivariate M_* and $M_{\rm HI}$ distributions over M_*. The shaded green area represents extrapolations for the H I MF of all galaxies. Panel (b): As panel (a) but only for LTGs. Panel (c): As panel (a) but only for ETGs. Downward arrows indicate the H I masses corresponding to $M_* \sim 10^{8.5}$, the mass below which our model for centrals and satellites are extrapolations. The vertical dashed lines indicate the completeness limits of our H I MFs. Due to the low H I-to-stellar mass ratio of ETGs, note that the H I completeness limit is below 10^7 M$_\odot$.

at fixed M_* and specific SFR than those hosted by haloes of lower mass. According to their analysis, the systematic environmental suppression of H I at both fixed M_* and fixed specific SFR in satellites begins in halo masses typical of the massive group regime ($>10^{13}$ $M_\odot$), and fast-acting mechanisms such as ram-pressure stripping are suggested to explain their results. Stark et al. (2016) use RESOLVE, a volume-limited multiwavelength census of ∼1500 local galaxies, to study the H I-to-stellar mass ratio, $R_{\rm H\,I}$, of satellite galaxies as a function of the halo (group) mass. They found that at fixed M_*, satellites have decreasing $R_{\rm H\,I}$ values with increasing halo mass at $M_{\rm h} \gtrsim 10^{12}$ $M_\odot$. The analogous relationship for centrals is uncertain and due to the poor overlap in stellar masses between centrals and satellites in the selected halo mass bins, it is not clear how different the $R_{\rm H\,I}$ values of centrals and satellites are at a fixed M_*. Their results for satellites suggest the presence of starvation and/or stripping mechanisms associated with halo gas heating in intermediate-mass groups.

The question that we address in this section is how different the H I gas content between centrals and satellites is at a fixed stellar mass separated explicitly into LTGs and ETGs. In Section 3, we presented the respective results for the xGASS survey. Upper limits were corrected for the distance bias (Section 2.3) and included into our survival statistical analysis (Section 2.4). The $R_{\rm H\,I}$ conditional distributions plotted in Figs 5–7 show that they are different among central and satellite galaxies at masses lower than $\sim 3 \times 10^{10}$ $M_\odot$. Fig. 3 shows the corresponding $\langle \log R_{\rm H\,I} \rangle - \log M_*$ relations of centrals and satellites. At fixed M_*, satellites have on average lower H I gas content than centrals with the differences increasing as M_* decreases. For LTGs, these differences at $M_* \sim 10^9$ $M_\odot$ are of ∼0.6 dex, decreasing to 0 at masses $M_* \sim 10^{11}$ $M_\odot$. For ETGs, the differences are of ∼1 dex at masses $M_* \lesssim 10^{10}$ $M_\odot$. However, it should be noted that the scatter (standard deviation) around the mean relations of centrals and satellites is large and the differences between the corresponding relations of both populations is smaller than their standard deviation.

By using the xGASS measurements to the H I conditional distributions, in Section 3.3, we constrained a set of proposed functions that allow us to project the $R_{\rm H\,I}$ conditional PDFs for LTGs and ETGs presented in Paper II into their corresponding distributions of centrals and satellites. The obtained bivariate (joint) $R_{\rm H\,I}$ and M_* distributions are shown in Fig. 9 along with their respective relations using different statistical estimators. As discussed in Section 4, the bivariate distributions of centrals and satellites are different for both LTGs and ETGs, and consequently for all galaxies. The differences depend on mass and for ETGs they are not easy to quantify by statistical estimators due the non-regular distribution of $R_{\rm H\,I}$.

To dig deeper into the differences in our empirically constructed H I distributions of central and satellite galaxies, we apply a two-sample Kolgomorov–Smirnov test to the obtained $R_{\rm H\,I}$ conditional PDFs given M_* of centrals and satellites for LTGs, ETGs, and all galaxies (Figs S1–S3 in the supplementary material). Quantitatively, the central and satellite H I distributions are different at the 95 per cent or higher level ($p < 0.05$) for $M_* \lesssim 3 \times 10^{10}$ $M_\odot$ in all the cases. For larger masses, the differences are smaller and both centrals and satellites are consistent with being drawn from the same distribution of H I gas content.

In the upper panels of Fig. 11, the relative differences in $\langle \log R_{\rm H\,I} \rangle$ (solid lines) and median $R_{\rm H\,I}$ (dotted lines) between centrals and satellites as a function of mass are shown for LTGs, ETGs, and all galaxies. We show also the arithmetic mean, $\langle R_{\rm H\,I} \rangle$. For LTGs, the relative difference between centrals and satellites is negligible at masses around 10^{11} $M_\odot$ and it increases up to ∼0.55 dex at $M_* \sim 5 \times 10^8$ $M_\odot$, remaining similar at lower masses. The relative differences for the arithmetic mean are slightly smaller than for the logarithmic mean or the medians. For ETGs, the relative differences between centrals and satellites are larger than for LTGs. Since for ETGs, and for both centrals and satellites, the $R_{\rm H\,I}$ conditional distributions given M_* are non-regular, the statistical estimators (geometric or arithmetic mean and median) significantly differ among each other, and consequently, also different is the relative difference among these estimators for centrals and satellites. Our results suggest that the relative difference in $\langle \log R_{\rm H\,I} \rangle$ is negligible for $M_* > 10^{11}$ $M_\odot$, but at lower masses, satellites are much more H I gas-poor than centrals, by ∼1.2 dex at $M_* \sim 3 \times 10^8 - 5 \times 10^9$ $M_\odot$. The relative difference in the medians, is larger than in the logarithmic means, specially at the range $M_* \sim 3 \times 10^8 - 10^{10}$ $M_\odot$. For the arithmetic means, the relative difference is significantly lower at all masses. The arithmetic means minimize the contribution of galaxies with very low $R_{\rm H\,I}$ values, which in the case of ETGs, as already discussed, distribute in a dominant second mode both for central and satellite galaxies (see their $R_{\rm H\,I}$ conditional PDFs in Fig. S3 from the supplementary material). For the combined population of LTGs and ETGs, the relative differences in $\langle \log R_{\rm H\,I} \rangle$ between centrals and satellites are 0.4−0.6 dex for $M_* < 5 \times 10^{10}$ $M_\odot$. The differences are slightly larger for the medians and smaller for the arithmetic means.

Finally, from Fig. 9, we note that the H I distributions of late- and early-types (left-hand panels) differ much more than the distribution of centrals and satellites (top panels). The lower panels of Fig. 11 show the relative differences in $\langle \log R_{\rm H\,I} \rangle$ and median $R_{\rm H\,I}$ between LTGs and ETGs as a function of mass for central, satellite and all galaxies, from left to right, respectively. We also show the respective differences but for the arithmetic mean, $\langle R_{\rm H\,I} \rangle$, as dashed line. The relative differences in the lower panels are much higher than in upper panels. Overall, the above can be interpreted as the present-day H I gas content of galaxies depending more on their internal nature, that is, whether they are of late or early-type morphology, than on external conditions associated to whether the galaxy is central or satellite. Nevertheless, this claim should be taken with caution. As mentioned above, there is evidence of the H I gas content of satellite galaxies being lower in massive haloes than in less massive ones at fixed stellar mass. It is interesting to mention that internal galaxy properties such as colour or specific star-formation rate could correlate even better with the H I gas content than morphology. For instance, Cook et al. (2019) showed that selecting only the subset of star-forming galaxies in the xGASS sample, the observed dependence at a fixed M_* of H I gas content on bulge-to-total ratio (a proxy for morphology) tends to disappear. The dependence of H I gas content on either internal properties, such as morphology, or on external conditions, such as the galaxy being central or satellite, could be related to both if the environment is responsible for reducing the gas content – and consequently quenching the star formation – and morphologically transforming galaxies. However, while common environmental effects such as ram pressure and starvation drain the gas and quench the satellites, a morphological transformation is not expected (e.g. van den Bosch et al. 2008; Weinmann et al. 2009). There is evidence that perhaps low-mass discs can be transformed into S0 gas-poor galaxies when they fall into clusters of galaxies, while the formation of massive S0 galaxies seems to be more related to high-redshift dissipational processes (Fraser-McKelvie et al. 2018, and more references therein).

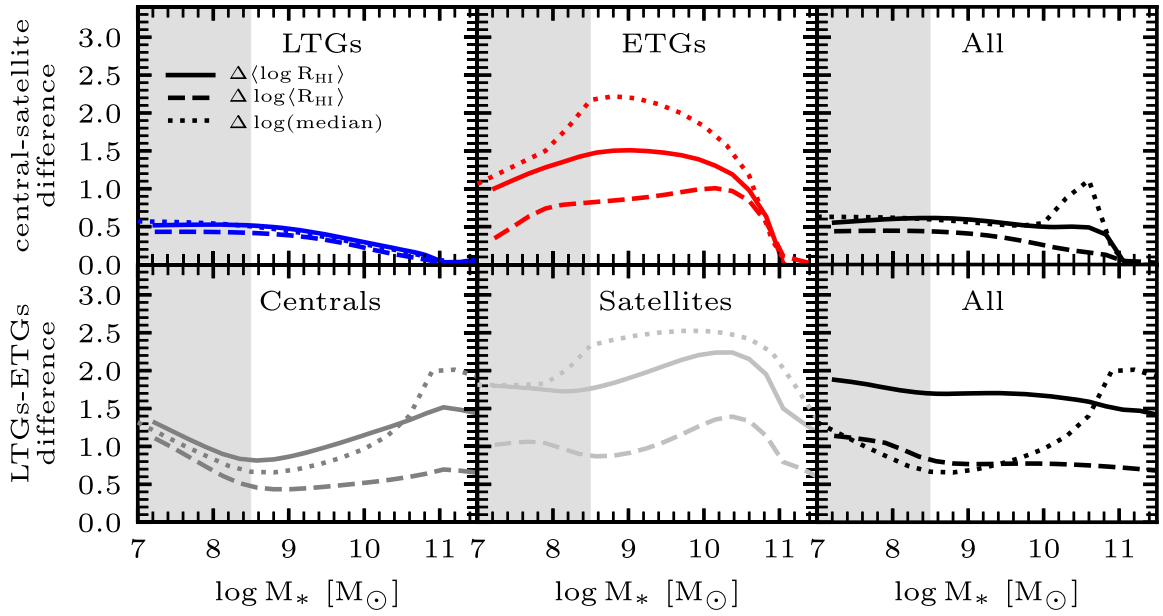

Figure 11. Upper panels: relative differences (in dex) between the logarithmic means (solid line), arithmetic means (dashed line), and medians (dotted line) of centrals and satellites as a function of M_*. From left to right, these relative differences are shown for the LTG, ETG, and whole galaxy populations. In all the cases, centrals have higher H I gas contents than satellites, but at the largest masses, these differences become very small. Lower panels: as the upper panels but in this case the relative differences are between LTGs and ETGs for the central, satellite, and whole galaxy population (from left to right, respectively). LTGs have much higher H I gas contents than ETGs at all masses. In both, upper and lower panels shaded grey areas indicate extrapolations to lower stellar masses of our empirically constrained model for centrals and satellites.

5.2 Caveats

5.2.1 Effects of different morphological classifications

The results presented here partially depend on the adopted criteria to morphologically classify galaxies as LTGs or ETGs. According to the above, we have separated the xGASS sample into LTGs and ETGs, and estimated the different fractions and subfractions as a function of M_* (Appendix A) required for our fitting procedure, by using the automatic morphological classification of Huertas-Company et al. (2011) implemented for SDSS galaxies. Next, we explore how much our results are affected by using an alternative morphological classification. Domínguez Sánchez et al. (2018) applied an automatic classification method to determine the morphology of the SDSS galaxies. We use their results to separate the xGASS sample into LTGs and ETGs, by employing the same morphological division criterion as we did in the case of the Huertas-Company et al. (2011) classification. Recall that elliptical and S0 galaxies were defined as ETGS, and from Sa to later types as LTGs.

The Domínguez Sánchez et al. (2018) morphological classification finds more ETGs than the one from Huertas-Company et al. (2011) at all masses, see Appendix A. As a consequence, the fractions of the different subpopulations change in xGASS, and also change the H I conditional CDFs corresponding to these subpopulations. We have repeated the whole analysis presented in Section 3 but for the Domínguez Sánchez et al. (2018) morphological classification, and obtained different functions for the $[P_i^j(> R_{\rm HI}|M_*)/P_i(> R_{\rm HI}|M_*)]_{\rm xGASS}$ ratios appearing in equation (2). By using these new functions, we calculated the corresponding H I CDFs of central and satellite galaxies for the LTG and ETG populations. Notice that the weights applied to xGASS were changed accordingly.

Fig. 12 compares the resulting mean $\langle \log R_{\rm HI}\rangle$–$\log M_*$ relations of centrals and satellites for the LTG and ETG populations from the Huertas-Company et al. (2011) and the Domínguez Sánchez et al. (2018) morphological classifications. The differences in the $\langle \log R_{\rm HI}\rangle$–$\log M_*$ relations of centrals and satellites introduced by the uncertainty in morphological classification are negligible for LTGs. These differences for ETGs range from ∼0.35 to 0.05 dex at masses $\sim 2 \times 10^8$ and $M_* \sim 2 \times 10^{10}$ M$_\odot$, respectively with Domínguez Sánchez et al. (2018) classification giving less separation into centrals and satellites than the Huertas-Company et al. (2011) one. At higher masses, differences between one or another classification scheme are negligible. The total relations shown in the right-hand panel are the weighted averages of LTGs and ETGs. Recall that the weights applied to xGASS depend on the morphological classification scheme, see Appendix A. This is why the total $\langle \log R_{\rm HI}\rangle$–$\log M_*$ relation is different when using one or the other morphological classification. Since for the Domínguez Sánchez et al. (2018) classification the fraction of ETGs is larger than for the Huertas-Company et al. (2011) classification, and because ETGs are H I gas poorer than LTGs, the

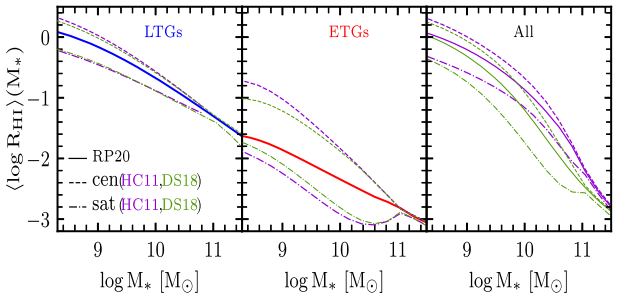

Figure 12. Logarithmic mean $R_{\rm HI}$–M_* relations for LTGs, ETGs, and all galaxies (solid lines) and their respective decomposition into central (short-dashed lines) and satellite (dot–dashed lines) using the Huertas-Company et al. (2011, magenta; see also Fig. 9) and Domínguez Sánchez et al. (2018, green) morphological classifications.

mean $\langle \log R_{\rm H\,I} \rangle$–$\log M_*$ relation in the former case is below than in the latter case.

In conclusion, variations in the morphological classification affect weakly our inferences of the difference between the H I gas content of centrals and satellites and only for ETGs. Adopting the Domínguez Sánchez et al. (2018) morphological classification instead of Huertas-Company et al. (2011) leads to a smaller separation in the mean relations of central and satellite ETGs than adopting the latter.

5.2.2 *Effects of membership and central/satellite designation errors*

In this paper, we have used the xGASS survey for modeling the H I gas content of central and satellites galaxies. As mentioned in Section 2.1, in xGASS, the central/satellite assignation comes from the SDSS group catalogue of Yang et al. (2005, 2007). This group catalogue, as others, may suffer of membership allocation and central/satellite designation errors. For example, Campbell et al. (2015) used a group catalogue constructed based on the Yang et al. (2005) group finder in a galaxy mock sample and estimated that the fraction of satellites that are truly satellites in the mock (purity) is around 70 per cent, while for centrals, the purity decreases from ~ 95 per cent at low group masses, $\sim 10^{12}\,h^{-1}{\rm M}_\odot$, to below 60 per cent at masses $\gtrsim 10^{14}\,h^{-1}{\rm M}_\odot$. On the other hand, the fraction of satellites in haloes that are correctly assigned to groups (completeness) is ~ 80 per cent independent of the halo mass, while for centrals the completeness decreases from ~ 90 per cent at low halo masses to ~ 60 per cent at the largest masses. The main source of confusion for centrals at large group masses is the central inversion problem, when the most luminous or massive galaxy is a satellite rather than the true central (van den Bosch et al. 2005; Skibba et al. 2011).

Thus, the differences in the H I gas fractions between centrals and satellites inferred with xGASS (see Section 3) could be larger. The above also implies that the differences in the overall H I distributions of central and satellites reported in Section 4 could be larger. Note, however, that for the xGASS sample that we use here, Janowiecki et al. (2017) improved the group membership given by Yang et al. (2007) by visually inspecting false pairs and galaxy shredding.

6 SUMMARY AND CONCLUSIONS

We have analysed the multiwavelength xGASS survey (Catinella et al. 2018), applying the same procedure as in Paper I to (i) re-scale their upper limits on the basis of samples observed in radio at lower distances, and (ii) treat the corrected upper limits with a survival analysis in order to infer full statistical distributions of the H I gas content of galaxies.

We have found that for LTGs, the $R_{\rm H\,I}$–M_* relation and the full $R_{\rm H\,I}$ conditional distributions as a function of M_* from xGASS agree very well with those empirically determined in Paper I for a larger stellar mass range sample (Fig. 3). For ETGs, the $R_{\rm H\,I}$ distributions from xGASS galaxies imply slightly higher values of $R_{\rm H\,I}$ than our previous determinations. For xGASS LTGs, centrals are on average more H I gas-rich than satellites of the same stellar mass. These differences are negligible for $\log(M_*/{\rm M}_\odot) > 10.8$, while at the lowest masses, $9.0 < \log(M_*/{\rm M}_\odot) \lesssim 9.7$, these differences are 0.5–0.7 dex, on average. For ETGs, the differences between centrals and satellites are larger than for LTGs. However, in both cases, the 1σ scatter around the $R_{\rm H\,I}$–M_* relations of centrals and satellites is larger than the difference between their means.

By means of a continuous fitting procedure to the processed xGASS data, we determined a set of functions that allowed us to project our empirical H I conditional cumulative distributions given M_* of both LTGs and ETGs into central and satellite galaxies. In other words, xGASS provides the information required to estimate the H I conditional distributions of centrals and satellites from the overall H I conditional distributions for both LTGs and ETGs. We use the above mentioned functions to extrapolate to stellar masses lower than those of the xGASS survey. By combining the $R_{\rm H\,I}$ conditional distributions given M_* with the corresponding GSMFs, the bivariate M_* and $R_{\rm H\,I}$ distribution functions, $\Phi(M_*, R_{\rm H\,I})$, for LTGs, ETGs, and all galaxies, separated into centrals and satellites, were calculated (Fig. 9). The main results obtained from this exercise are summarized below:

(i) For LTGs, satellites have on average less H I than centrals. Up to $M_* \sim 10^9$ M$_\odot$, the relative difference is $\sim$0.5 dex and all the gas-rich dwarf LTGs are centrals. For higher masses, this relative difference decreases up to $M_* \sim 3 \times 10^{10}$ M$_\odot$, above which there is no difference between centrals and satellites. Since the bivariate distribution is regular for LTGs, even for centrals and satellites separately, the $R_{\rm H\,I}$–M_* relations calculated with different statistical estimators are roughly similar.

(ii) For ETGs, the bivariate distributions for centrals and satellites differ more than for LTGs, satellites being on average more devoid of H I than centrals up to $M_* \sim 5 \times 10^{10}$ M$_\odot$. However, the $R_{\rm H\,I}$ distribution of satellite ETGs is strongly bimodal, with a fraction of them having $R_{\rm H\,I}$ values close to those of central ETGs and another fraction with very low $R_{\rm H\,I}$ values. At $M_* \gtrsim 5 \times 10^{10}$ M$_\odot$, central ETGs are already more abundant than satellite ETGs but both have statistically similar H I gas content.

(iii) Since the bivariate distributions for ETGs, both centrals and satellites, are non-regular, the $R_{\rm H\,I}$–M_* relations calculated with different statistical estimators are different. In particular, the relation based on arithmetic means, $\langle R_{\rm H\,I} \rangle$, is significantly above the relations based on logarithmic means or medians.

(iv) The projection of the bivariate distribution when integrating it over M_* is the H I MF and agrees well with those measured in blind radio surveys. We show here that the H I MF is completely dominated by central galaxies at all masses, both for LTGs and ETGs (Fig. 11).

Overall, our results show that the difference in the bivariate $R_{\rm H\,I}$ and M_* distribution between LTGs and ETGs is significantly larger than between central and satellite galaxies. This suggests that the H I gas content of galaxies depends more on their internal nature, that is, whether they are of late- or early-type morphology, than on external conditions associated to whether the galaxy is central or satellite.

In this paper, we presented a full statistical description of the H I gas content of local galaxies as a function of their stellar mass and separated into late- and early-type and into central and satellites. These results can be used for comparisons with theoretical predictions of galaxy evolution, and for adding the H I gas component in empirical approaches aimed to model the local galaxy population. In particular, our results can be used to establish the M_*-$M_{\rm H\,I}$-M_h connection from the outcome of large N-body cosmological simulations, where complete mock galaxy catalogues can be generated. In a forthcoming paper, we will present results of this connection including predictions on the spatial clustering of galaxies using both their stellar and H I masses.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewer for her/his comments and suggestions that helped to improve the presentation of this paper. ARC acknowledges CONACyT for a PhD fellowship. ARP and VAR acknowledge financial support from CONACyT through 'Ciencia

Basica' grant 285721, and from DGAPA-UNAM through PAPIIT grant IA104118. CL and BC acknowledge partial funding from the ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

DATA AVAILABILITY

The xGASS, the automated morphological classification for SDSS DR7, and the Yang et al. (2012, see also Yang et al. 2007) galaxy group catalogues are publicly available.[6] The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Boselli A., Gavazzi G., 2006, PASP, 118, 517
Boselli A. et al., 2010, PASP, 122, 261
Boselli A., Cortese L., Boquien M., 2014a, A&A, 564, A65
Boselli A., Cortese L., Boquien M., Boissier S., Catinella B., Lagos C., Saintonge A., 2014b, A&A, 564, A66
Bravo-Alfaro H., Cayatte V., van Gorkom J. H., Balkowski C., 2000, AJ, 119, 580
Brown T. et al., 2017, MNRAS, 466, 1275
Brown T., Catinella B., Cortese L., Kilborn V., Haynes M. P., Giovanelli R., 2015, MNRAS, 452, 2479
Brown G. M., Johnston K. G., Hoare M. G., Lumsden S. L., 2016, MNRAS, 463, 2839
Calette A. R., Avila-Reese V., Rodríguez-Puebla A., Hernández-Toledo H., Papastergis E., 2018, Rev. Mex. Astron. Astrofis, 54, 443
Campbell D., van den Bosch F. C., Hearin A., Padmanabhan N., Berlind A., Mo H. J., Tinker J., Yang X., 2015, MNRAS, 452, 444
Cappellari M. et al., 2011, MNRAS, 413, 813
Catinella B. et al., 2010, MNRAS, 403, 683
Catinella B. et al., 2012, A&A, 544, A65
Catinella B. et al., 2013, MNRAS, 436, 34
Catinella B. et al., 2018, MNRAS, 476, 875
Cook R. H. W., Cortese L., Catinella B., Robotham A., 2019, MNRAS, 490, 4060
Cortese L., Catinella B., Boissier S., Boselli A., Heinis S., 2011, MNRAS, 415, 1797
Cortese L., Catinella B., Cook R. H. W., Janowiecki S., 2020, MNRAS, 494, L42
Davies L. J. M. et al., 2019, MNRAS, 483, 1881
Domínguez Sánchez H., Huertas-Company M., Bernardi M., Tuccillo D., Fischer J. L., 2018, MNRAS, 476, 3661
Eckert K. D., Kannappan S. J., Stark D. V., Moffett A. J., Norris M. A., Snyder E. M., Hoversten E. A., 2015, ApJ, 810, 166
Feigelson E. D., Nelson P. I., 1985, ApJ, 293, 192
Fraser-McKelvie A., Aragón-Salamanca A., Merrifield M., Tabor M., Bernardi M., Drory N., Parikh T., Argudo-Fernández M., 2018, MNRAS, 481, 5580
Fukugita M., Hogan C. J., Peebles P. J. E., 1998, ApJ, 503, 518
Fukugita M. et al., 2007, AJ, 134, 579
Gavazzi G., Boselli A., van Driel W., O'Neil K., 2005, A&A, 429, 439
Gavazzi G., Fumagalli M., Fossati M., Galardo V., Grossetti F., Boselli A., Giovanelli R., Haynes M. P., 2013, A&A, 553, A89
Giovanelli R. et al., 2005, AJ, 130, 2598
Guo H., Li C., Zheng Z., Mo H. J., Jing Y. P., Zu Y., Lim S. H., Xu H., 2017, ApJ, 846, 61

[6]xGASS: https://xgass.icrar.org/data.html, Automated morphological classification for SDSS DR7: http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/525/A157, and Yang et al. (2012) galaxy group catalogues: https://gax.sjtu.edu.cn/data/Group.html

Haynes M. P., Giovanelli R., 1984, AJ, 89, 758
Haynes M. P. et al., 2011, AJ, 142, 170
Hess K. M., Wilcots E. M., 2013, AJ, 146, 124
Huang S., Haynes M. P., Giovanelli R., Brinchmann J., 2012, ApJ, 756, 113
Huertas-Company M., Aguerri J. A. L., Bernardi M., Mei S., Sánchez Almeida J., 2011, A&A, 525, A157
Janowiecki S., Catinella B., Cortese L., Saintonge A., Brown T., Wang J., 2017, MNRAS, 466, 4795
Janowiecki S., Catinella B., Cortese L., Saintonge A., Wang J., 2020, MNRAS, 493, 1982
Kannappan S. J. et al., 2013, ApJ, 777, 42
Kaplan E. L., Meier P., 1958, J. Am. Stat. Assoc., 53, 457
Karachentsev I. D., Makarov D. I., Kaisina E. I., 2013, AJ, 145, 101
Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, MNRAS, 353, 713
Lacerna I., Hernández-Toledo H. M., Avila-Reese V., Abonza-Sane J., del Olmo A., 2016, A&A, 588, A79
Lacerna I., Ibarra-Medel H., Avila-Reese V., Hernández-Toledo H. M., Vázquez-Mata J. A., Sánchez S. F., 2020, A&A, 644, A117
Lagos C. D. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011, MNRAS, 418, 1649
Lagos C. D. P., Baugh C. M., Zwaan M. A., Lacey C. G., Gonzalez-Perez V., Power C., Swinbank A. M., van Kampen E., 2014, MNRAS, 440, 920
Lemonias J. J., Schiminovich D., Catinella B., Heckman T. M., Moran S. M., 2013, ApJ, 776, 74
Lu Y., Yang X., Liu C., Guo H., Xu H., Katsianis A., Wang Z., 2020, preprint (arXiv:2008.09804)
Maddox N., Hess K. M., Obreschkow D., Jarvis M. J., Blyth S.-L., 2015, MNRAS, 447, 1610
Martin D. C. et al., 2005, ApJ, 619, L1
Masters K. L. et al., 2019, MNRAS, 488, 3396
Meert A., Vikram V., Bernardi M., 2015, MNRAS, 446, 3943
Meyer M. J. et al., 2004, MNRAS, 350, 1195
Meyer M. J., Zwaan M. A., Webster R. L., Brown M. J. I., Staveley-Smith L., 2007, ApJ, 654, 702
Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution, Cambridge University Press, Cambridge
Nair P. B., Abraham R. G., 2010, ApJS, 186, 427
Papastergis E., Cattaneo A., Huang S., Giovanelli R., Haynes M. P., 2012, ApJ, 759, 138
Papastergis E., Giovanelli R., Haynes M. P., Rodríguez-Puebla A., Jones M. G., 2013, ApJ, 776, 43
Rasmussen J., Mulchaey J. S., Bai L., Ponman T. J., Raychaudhury S., Dariush A., 2012, ApJ, 757, 122
Rodríguez-Puebla A., Avila-Reese V., Drory N., 2013, ApJ, 767, 92
Rodríguez-Puebla A., Calette A. R., Avila-Reese V., Rodriguez-Gomez V., Huertas-Company M., 2020, PASA, 37, e024
Serra P. et al., 2012, MNRAS, 422, 1835
Skibba R. A., van den Bosch F. C., Yang X., More S., Mo H., Fontanot F., 2011, MNRAS, 410, 417
Solanes J. M., Manrique A., García-Gómez C., González-Casado G., Giovanelli R., Haynes M. P., 2001, ApJ, 548, 97
Springob C. M., Haynes M. P., Giovanelli R., Kent B. R., 2005, ApJS, 160, 149
Stark D. V. et al., 2016, ApJ, 832, 126
Stevens A. R. H. et al., 2019, MNRAS, 483, 5334
Taylor R., Davies J. I., Auld R., Minchin R. F., 2012, MNRAS, 423, 787
van den Bosch F. C., Weinmann S. M., Yang X., Mo H. J., Li C., Jing Y. P., 2005, MNRAS, 361, 1203
van den Bosch F. C., Aquino D., Yang X., Mo H. J., Pasquali A., McIntosh D. H., Weinmann S. M., Kang X., 2008, MNRAS, 387, 79
van Driel W. et al., 2016, A&A, 595, A118
Watts A. B., Catinella B., Cortese L., Power C., 2020, MNRAS, 492, 3672
Wei L. H., Kannappan S. J., Vogel S. N., Baker A. J., 2010, ApJ, 708, 841
Weinmann S. M., Kauffmann G., van den Bosch F. C., Pasquali A., McIntosh D. H., Mo H., Yang X., Guo Y., 2009, MNRAS, 394, 1213
Wright R. J., Lagos C. d. P., Davies L. J. M., Power C., Trayford J. W., Wong O. I., 2019, MNRAS, 487, 3740

Yang X., Mo H. J., van den Bosch F. C., Jing Y. P., 2005, MNRAS, 356, 1293
Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, ApJ, 671, 153
Yang X., Mo H. J., van den Bosch F. C., Zhang Y., Han J., 2012, ApJ, 752, 41
Yoon I., Rosenberg J. L., 2015, ApJ, 812, 4

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Figure S1. $R_{\rm HI}$ PDFs of all galaxies as a function of stellar mass; the $\log(M_*/{\rm M}_\odot)$ is shown in parenthesis at each panel.
Figure S2. Same as Fig. S1 but for LTGs.
Figure S3. Same as Fig. S1 but for ETGs.
Table S1. xGASS fractions as presented in panels (c) and (d) of Fig. 2.
Table S2. xGASS $\langle \log R_{\rm HI} \rangle - \log M_*$ relations for all galaxies.
Table S3. xGASS $\langle \log R_{\rm HI} \rangle - \log M_*$ relations for LTGs.
Table S4. xGASS $\langle \log R_{\rm HI} \rangle - \log M_*$ relations for ETGs.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: CORRECTING xGASS TO THE MORPHOLOGY AND ENVIRONMENT DISTRIBUTIONS OF SDSS

In this Appendix, we first define the fractions corresponding those galaxy subsamples required to perform the joint analytic fitting to the xGASS H I conditional CDFs in Section 3.3. Then, we compare these fractions from xGASS to those from the volume-corrected SDSS DR7. Finally, we explain our procedure for weighting xGASS galaxies in order they reproduced the fractions of ETGs and of satellites from SDSS. Following, in the definition of the different fractions, for simplicity, we omit the dependence on M_*.

(i) *Fraction of ETGs/LTGs:*
Defined as the ratio of ETG to total mass functions, $f^E \equiv \phi^E/\phi$. The fraction of LTGs, f_L, is the complement, $f^L = 1 - f^E$.
(ii) *Fraction of centrals/satellites*:
Defined as the ratio of central to total mass functions, $f^c \equiv \phi^c/\phi$. The fraction of satellites, f^s, is the complement, $f^s = 1 - f^c$.
(iii) *Fraction of ETGs/LTGs for centrals*:
For the subsample of centrals described by the central mass function ϕ^c, $f_E^c \equiv \phi_E^c/\phi^c$ is the fraction of ETGs for centrals. The respective fraction of LTGs is the complement, $f_L^c = 1 - f_E^c$.
(iv) *Fraction of ETGs/LTGs for satellites*:
For the subsample of satellites described by the satellite mass function ϕ^c, $f_E^s \equiv \phi_E^s/\phi^c$ is the fraction of ETGs for satellites. The respective fraction of LTGs is the complement, $f_L^s = 1 - f_E^s$.
(v) *Fraction of satellites/centrals for LTGs*:
For the subsample of LTGs described by the LTG mass function ϕ^L, $f_s^L \equiv \phi_s^L/\phi^L$ is the fraction of satellites for LTGs. The respective fraction of centrals is the complement, $f_c^L = 1 - f_s^L$.
(vi) *Fraction of satellites/centrals for ETGs*:
For the subsample of ETGs described by the ETG mass function ϕ^E, $f_s^E \equiv \phi_s^E/\phi^E$ is the fraction of satellites for ETGs. The respective fraction of centrals is the complement, $f_c^E = 1 - f_s^E$.

Fig. A1 shows the fractions defined above for xGASS, solid black squares connected with solid lines. From left to right, the upper panels show the fraction of satellites for all galaxies (ii), and the fractions of ETGs for the subsamples of central and satellite galaxies (iii and iv). The lower panels show the fraction of ETGs for all galaxies (i), and the fractions of satellites for the subsamples of LTGs and ETGs (v and vi). In these panels, the respective fractions measured from the volume-complete SDSS DR7 are also plotted (black circles) along with analytical fits to these fractions (black lines; see below). We use Meert et al. (2015) photometry and an average stellar mass from five different mass-to-luminosity prescriptions, updated galaxy group catalogues from Yang et al. (2007), Yang et al. (2012), and the Huertas-Company et al. (2011) morphological classification (see Paper II for details and for the corrections applied to obtain a volume complete sample). As seen in panels (b) and (c), the fractions of ETGs in the central and satellite subsamples are systematically larger up to $M_* \sim 10^{11}$ M$_\odot$ for xGASS than for SDSS; at larger masses, the difference inverts for the subsample of satellite galaxies. In the insets of these panels, we plot the ratios of the respective fractions of SDSS to xGASS.

As mentioned in Section 2.1, to infer from xGASS the $R_{\rm HI}$–M_* relations and $R_{\rm HI}$ distributions given M_* corresponding to all galaxies, as well as to the subsamples of central and satellite galaxies, the biases of xGASS with respect to SDSS in morphology and environment should be corrected. To do so, we adopt a methodology similar as in Catinella et al. (2018) for recovering a volume complete sample. When we compute the above mentioned $R_{\rm HI}$–M_* relations or the whole $R_{\rm HI}$ conditional distributions given M_*, we apply weights to xGASS galaxies to recover the volume-compete SDSS fractions of ETGs in the central and satellite subsamples. The weights are the ratios shown in the insets of panels (b) and (c) of Fig. A1. This automatically also recovers the overall SDSS fraction of ETGs and the overall fraction of satellites. In any case, note that the relevant bias of the xGASS sample with respect to SDSS is by morphology; the bias in selecting central/satellite galaxies is small and mainly due to the former.

For the above procedure and for extrapolating the fits to the $R_{\rm HI}$ conditional CDFs from xGASS to masses lower than $M_* = 10^9$ M$_\odot$, we use actually analytical fits to the SDSS fractions. The fits are performed to the overall fraction of satellites (panel (a)) and the fractions of ETGs for the central and satellite subsamples, panels (b) and (c), respectively. For the SDSS fractions $f_E^c(M_*)$ and $f_E^s(M_*)$, we perform MCMC multiparametric fits to a composition of two analytic Sigmoid functions, following the procedure described in Rodríguez-Puebla et al. (2013). The final analytic function is:

$$f_E^j(M_*) = \frac{1-A}{1+e^{-\gamma_1(x_{C,1}+x_{0,1})}} + \frac{A}{1+e^{-\gamma_2(x_{C,2}-x_{0,2})}}, \tag{A1}$$

where $j = c$ or s, $x_{C,i} = M_*/\mathcal{M}_{C,i}$, with $i = 1, 2$. For the overall fraction $f^s(M_*)$, we use an analytic function composed of a Sigmoid and constant function given by

$$f^s(M_*) = 1 - \left[A \cdot \frac{1}{1+e^{-\gamma(x_C - x_0)}} + H \right], \tag{A2}$$

where H is the constant function. Here, the Sigmoid normalization factor is defined as $A \equiv 1 - H$.

The obtained fits are shown in Fig. A1 with the solid grey lines. The fractions in the lower panels were calculated from the fractions of the upper panels. The dashed grey lines in all the panels are just the respective complementary fractions.

Finally, in Section 5.2.1 we explore the effects on our results of using different morphological classification than the one used here.

The alternative classification was that of (Domínguez Sánchez et al. 2018). In Fig. A1, we show with brown colours the same fractions defined in this Appendix but using the (Domínguez Sánchez et al. 2018) morphologies for the xGASS and SDSS galaxies. Interestingly, now the excess in the xGASS fraction of ETGs with respect to SDSS at masses lower than $\sim 3 \times 10^{10}$ $M_\odot$ is less than when using the Huertas-Company et al. (2011) classification, while for masses larger than this, there is now a lack of ETGs in xGASS.

APPENDIX B: PROCEDURE FOR RE-ESTIMATING THE UPPER LIMITS OF xGASS

B0.1 Upper limits of ETGs

In Paper I, based on ATLAS^{3D} results, we re-scaled by distance the GASS upper limits of ETGs to use these valuable data along with those of ATLAS^{3D} and other samples. To do so, we decreased the upper limits by $(D_i(z)/\overline{D}_{\rm ATLAS^{3D}})^2$, being D_i the luminosity distance of each GASS ETG and $\overline{D}_{\rm ATLAS^{3D}} = 25$ Mpc the median luminosity distance of ATLAS^{3D}. The key assumption behind this exercise is that ETGs of similar masses from GASS and ATLAS^{3D} follow the same $M_{\rm HI}$ statistical distribution despite their slightly different ages. For ~ 25 per cent of the ETGs upper limits in GASS, we actually assigned them a detection taking into account that in between the GASS detection limit and this limit shifted to 25 Mpc, ~ 25 per cent of galaxies in the ATLAS^{3D} sample were detected. For the remaining 75 per cent of GASS upper limits, we re-calculated them using the distance of 25 Mpc. That is, even for such a small distance, yet a significant fraction of GASS ETGs would remain as non-detected but their re-scaled upper limits to those of ATLAS^{3D} result much lower than the reported ones. These upper limits along with those from other ETG samples compiled in Paper I, pile up around values in $R_{\rm HI}$ of $10^{-3}-10^{-4}$. The larger the mass, the smaller these values. From the performed continuous fit to the observed $R_{\rm HI}$ distributions in M_* bins, the $R_{\rm HI}$ values where the upper limits pile up were constrained by the function $\mathcal{R}_1(M_*)$, see equation (11) in Paper II. The values of $\mathcal{R}_1(M_*)$ correspond roughly to those where the top-hat functions start in the conditional PDFs for ETGs shown in Fig. 1. The fraction of galaxies in the top-hat functions correspond to the fractions of non detections.[7] As expected, for $M_* \gtrsim 10^{10}$ $M_\odot$, the values of $\mathcal{R}_1(M_*)$ are close to the upper limits of ATLAS^{3D}. However, have in mind that in Paper I we included other galaxy samples besides GASS and ATLAS^{3D}.

Based on the analysis of Paper I, we proceed here as follows in order to re-estimate the xGASS upper limits of ETGs:

(i) From the empirical ETG $R_{\rm HI}$ conditional PDFs reported in Paper II, calculate the fraction of galaxies that lie in each stellar mass bin in between the GASS and ATLAS^{3D} $R_{\rm HI}$ detection limits (as done in Paper I), and in between the $R_{\rm HI}$ detection limit of the of low-xGASS and $\mathcal{R}_1(M_*)$ (recall that in ATLAS^{3D} there are not low-mass galaxies).

(ii) Assign $R_{\rm HI}$ values to a fraction of the xGASS upper limits at each M_* bin equal to the respective fraction as calculated in (i). To do so, pick randomly $R_{\rm HI}$ values from the empirical ETG $R_{\rm HI}$ conditional PDFs in the $R_{\rm HI}$ ranges determined in (i; in Paper I, for GASS galaxies, a uniform distribution was assumed).

(iii) For the (large) fraction of galaxies with upper limits that were not assigned an $R_{\rm HI}$ value, lower their upper limits by a factor $(D_i(z)/25\ {\rm Mpc})^2$, where $D_i(z)$ is in Mpc. This is equivalent to say that these galaxies, with similar observational setups and signal-to-noise ratios as used in xGASS and GASS, will remain undetected in HI at the distance of 25 Mpc, but their upper limits are re-calculated accordingly to this distance.

It is worth of mentioning that for $M_* > 10^{10}$ $M_\odot$, the values of the fractions calculated in (i) are around $30-40$ per cent, larger than

[7]To estimate the $R_{\rm HI}$ distributions of ETGs, in Paper I we assumed that the true $R_{\rm HI}$ values should be up to $\sim$1 dex below the upper limit values after corrections and survival analysis, following a uniform distribution. This is why the $R_{\rm HI}$ conditional PDFs shown in Fig. 1 have a top-hat distribution of $\sim$1 dex width at their low-$R_{\rm HI}$ ends; see Paper I for arguments in favour of this assumption and for a discussion.

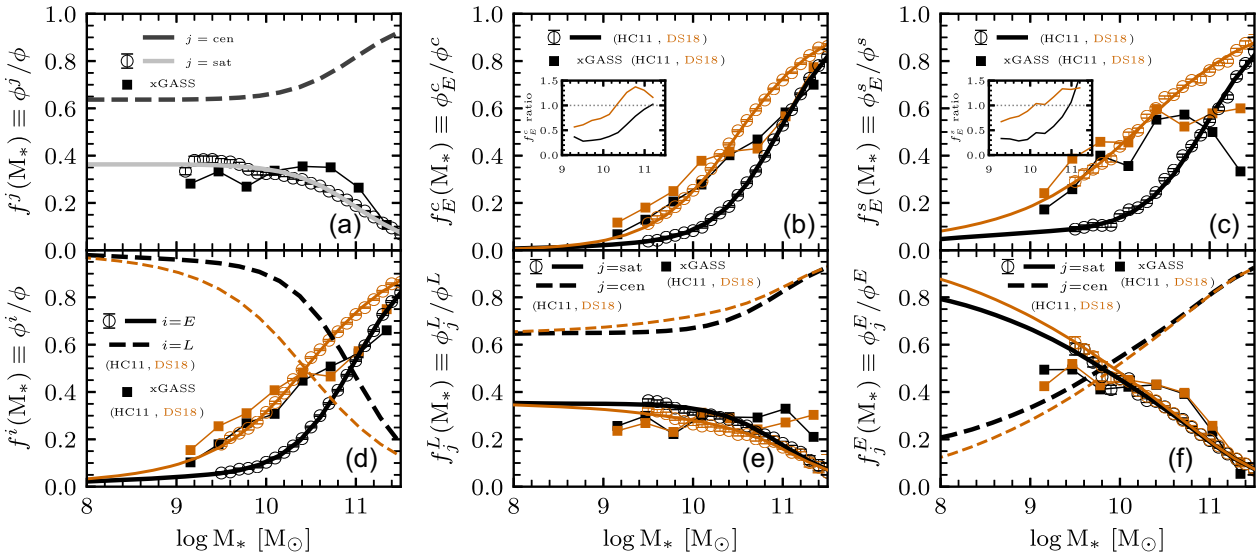

Figure A1. Different fractions of subsamples calculated from the volume-complete SDSS using the Yang et al. (2012) group catalogue for defining centrals and satellites. Black and brown colours are inferences using the Huertas-Company et al. (HC11, 2011) and Domínguez Sánchez et al. (DS18, 2018) morphological classifications, respectively. The fractions corresponding to xGASS are shown with filled squares connected by solid lines. The solid lines are fits to the SDSS data; the dashed lines show the respective complementary fractions. The insets in panels (b) and (c) are the ratios of SDSS to xGASS fractions.

the $\sim 25-30$ per cent fraction of galaxies detected by ATLAS^{3D} in between the detection limit of this survey and the one of GASS (see Paper I).

B0.2 Upper limits of LTGs

From Fig. 2, we see that the xGASS detection limits lie in the very low end of our empirical $R_{\rm H\,I}$ conditional PDFs of LTGs shown in Fig. 1. The fraction of LTGs with upper limits that pile up around these limits is relatively small. Note that if these galaxies were closer, then they likely would have been detected in H I, as is the case for galaxies from the closer HRS sample, see Paper I. Thus, we convert the upper limit of a given LTG to a *detection* with the $R_{\rm H\,I}$ value randomly picked out from the tail of the empirical $R_{\rm H\,I}$ conditional PDF given M_* from Paper II.

APPENDIX C: RESULTS WITHOUT TAKING INTO ACCOUNT CORRECTIONS TO xGASS

The upper panels of Fig. C1 are as the upper panels of Fig. 3 but without taking into account our procedure for the upper limits of xGASS, nor the correction by morphology/environment (the respective data are presented in tabulated form in the Supplementary Material). Here, instead of the standard deviation, we plot the error of the mean. For ETGs, in the stellar mass bins above $M_* \sim 5 \times 10^9$ $\rm M_\odot$ the obtained means with the KM estimator are shown with an arrow. This is because the fraction of non-detections are higher than 50 per cent in these mass bins, in which cases the KM estimator provides uncertain results; the means should be taken as upper bounds while the error on the means (or standard deviations) are meaningless. In the middle panels, we compare the $R_{\rm H\,I}$ means and errors on the mean obtained with the KM estimator with and without including our corrections to upper limits and morphology/environment bias of xGASS. For LTGs, the results are almost indistinguishable from each other. For ETGs less massive than $M_* \sim 5 \times 10^9$ $\rm M_\odot$, the means are slightly higher when our procedure for the upper limits is not taken into account. For $M_* \gtrsim 5 \times 10^9$ $\rm M_\odot$, the means are only an estimate of the upper bound. For the whole sample, combining LTGs and ETGs, the KM results without taking into account the procedure for upper limits are only slightly below to those reported in Catinella et al. (2018), who assigned $R_{\rm H\,I}$ values to non-detections equal to their upper limit values. Finally, in the lower panels of Fig. C1, we compare the $R_{\rm H\,I}$ means obtained with the KM estimator taking into account our procedure for the upper limits but applying and not applying the weights by morphology/environment (the respective data are presented in tabulated form in the Supplementary Material). The weights (mainly by morphology) slightly increase the mean $R_{\rm H\,I}$ values for masses below $M_* \sim 5 \times 10^{10}$ $\rm M_\odot$, while for the highest masses, the weights decrease $R_{\rm H\,I}$ by $\sim$0.3 dex.

In Fig. C2, we reproduce the $R_{\rm H\,I}$ conditional CDFs plotted in Fig. 5 and compare them with those without taking into into account our procedure for the upper limits of xGASS. For LTGs, the CDFs in both cases are very similar, excepting at the low-$R_{\rm H\,I}$ end in the most massive bins. For ETGs, when the procedure for the upper limits is not taken into account, the CDFs undergo a sharp cut at relatively high values of $R_{\rm H\,I}$. In this case, we can not constrain any reliable $R_{\rm H\,I}$ conditional CDF.

APPENDIX D: CONSERVATION EQUATIONS

As discussed in Section 3.3, performing fits to xGASS CDFs must obey the law of total probability. Here, we present the 'probability conservation equations' in order to satisfy such requirement for the whole set of galaxies, different subsets of LTGs/ETGs, centrals/satellites, and their combinations.

First, to describe the H I conditional CDFs of all LTGs and ETGs, and central LTGs and ETGs (four sets of CDFs), we propose the analytic incomplete gamma function given by equation (4) for each one of these populations.

The remaining five sets of H I CDFs to be used also for the fitting procedure are described by the below listed five equations that obey the law of total probability, and that allow us to use the above mentioned four sets of CDFs for calculating these five sets of CDFs. Such equations require information on different fractions of populations and subpopulations of galaxies as a function of M_*. In Appendix A, we discuss how we estimate these fractions. For simplicity, we do not show the dependence of these fractions on M_* in the following equations:

(i) *H I CDFs of the whole sample:*

$$P^T(<R_{\rm H\,I}|M_*) = f^L \cdot P^L(<R_{\rm H\,I}|M_*) + f^E \cdot P^E(<R_{\rm H\,I}|M_*) \tag{D1}$$

where f^E and f^L are the fractions of ETGs and LTGs, respectively; $f^E + f^L = 1$.

(ii) *H I CDFs of the subsample of centrals:*

$$P^c(<R_{\rm H\,I}|M_*) = f^c_L \cdot P^c_L(<R_{\rm H\,I}|M_*) + f^c_E \cdot P^c_E(<R_{\rm H\,I}|M_*) \tag{D2}$$

where f^c_E and f^c_L are the fractions of centrals that are ETGs and LTGs, respectively; $f^c_E + f^c_L = f^c$

(iii) *H I CDFs of the subsample of satellites*

$$P^s(<R_{\rm H\,I}|M_*) = \frac{1}{f^s}\left[P^T(<R_{\rm H\,I}|M_*) - f^c \cdot P^c(<R_{\rm H\,I}|M_*)\right] \tag{D3}$$

where $P^T(<R_{\rm H\,I}|M_*)$ and $P^c(<R_{\rm H\,I}|M_*)$ are the total and centrals CDFs given by equations (D1) and (D2) respectively. f^c and f^s are the fraction of centrals and satellites, $f^c + f^s = 1$.

(iv) *H I CDFs of the subsample of satellites that are LTGs*

$$P^s_L(<R_{\rm H\,I}|M_*) = \frac{1}{f^L_s}\left[P^T_L(<R_{\rm H\,I}|M_*) - f^L_c \cdot P^c_L(<R_{\rm H\,I}|M_*)\right] \tag{D4}$$

where $P^T_L(<R_{\rm H\,I}|M_*)$ and $P^c_L(<R_{\rm H\,I}|M_*)$ are the total LTGs and LTGs centrals CDFs analytic fits given by equation (4) respectively. f^L_c and f^L_s are the fractions of LTGs that are centrals and satellites, $f^L_c + f^L_s = f^L$

(v) *H I CDFs of the subsample of satellites that are ETGs*

$$P^s_E(<R_{\rm H\,I}|M_*) = \frac{1}{f^E_s}\left[P^T_E(<R_{\rm H\,I}|M_*) - f^E_c \cdot P^c_E(<R_{\rm H\,I}|M_*)\right] \tag{D5}$$

where $P^T_E(<R_{\rm H\,I}|M_*)$ and $P^c_E(<R_{\rm H\,I}|M_*)$ are the ETGs and ETG centrals CDFs analytic fits given by equation (4), respectively. f^E_c and f^E_s are the fractions of ETGs that are centrals and satellites, $f^E_c + f^E_s = f^E$.

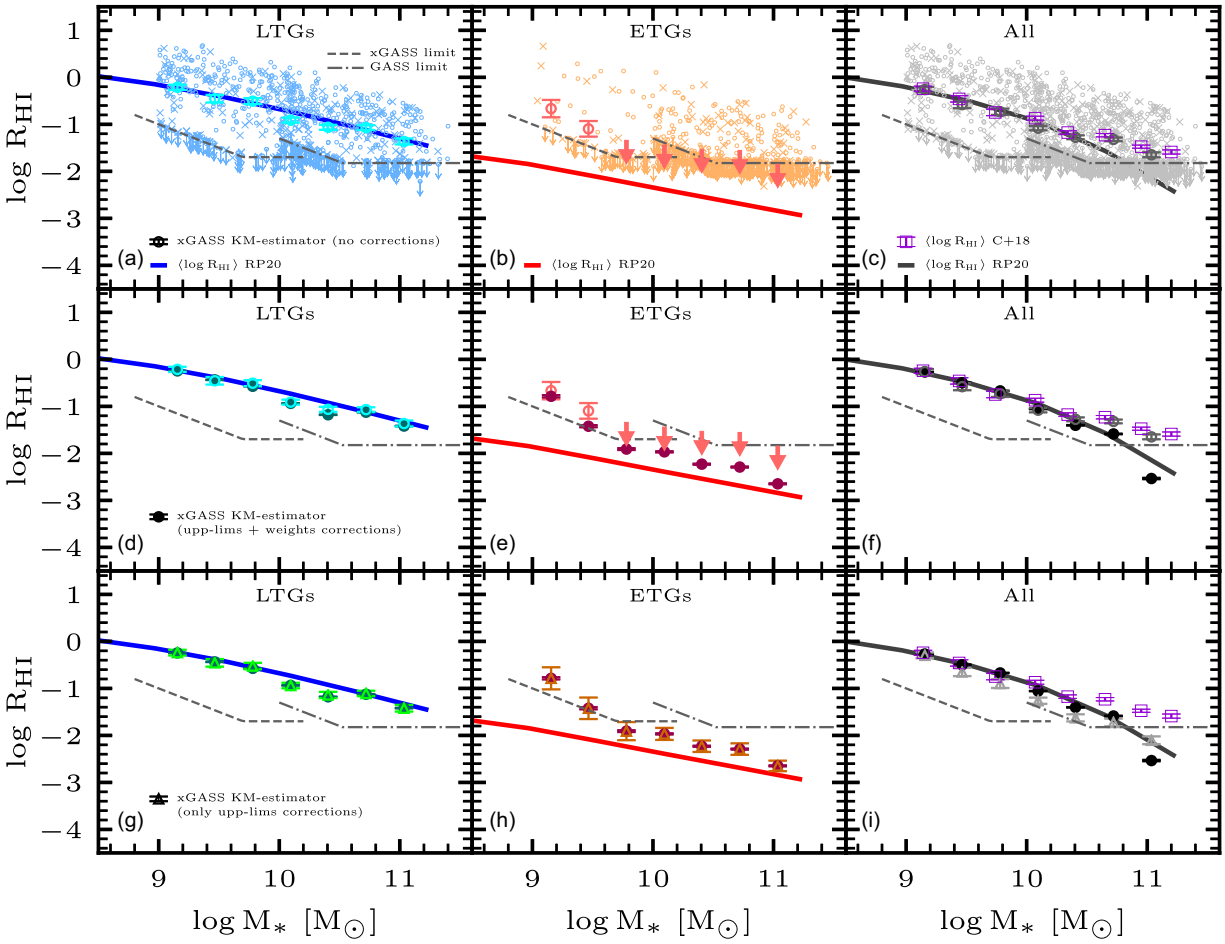

Figure C1. Upper panels: xGASS galaxies in the log $R_{\rm H\,I}$ − log M_* diagram, as in Fig. 2. Large empty circles with error bars are the logarithmic means and the error of the mean in M_* bins obtained with the KM estimator without taking into account our procedure for the upper limits of xGASS, nor the correction by morphology/environment. The solid lines show the mean $R_{\rm H\,I}$–M_* relations from Paper II. In panel (b), means above $M_* \sim 5 \times 10^9$ M$_\odot$ are shown as arrows given that the fraction of non-detections are $>$ 50 per cent in these mass bins (see text). In panel (c), the violet empty squares are the logarithmic means and error of the mean as reported in Catinella et al. (2018). Middle panels: logarithmic means and their error on the mean using the KM estimator with (filled circles as in Fig. 3) and without (empty circles or arrows, as in the upper panels) including our corrections to upper limits and morphology/environment bias of xGASS. Lower panels: logarithmic means and their errors on the mean obtained with the KM estimator taking into account our procedure for the upper limits and weighting by morphology/environment (filled circles, as in the middle panels) and not weighting by morphology/environment (empty triangles).

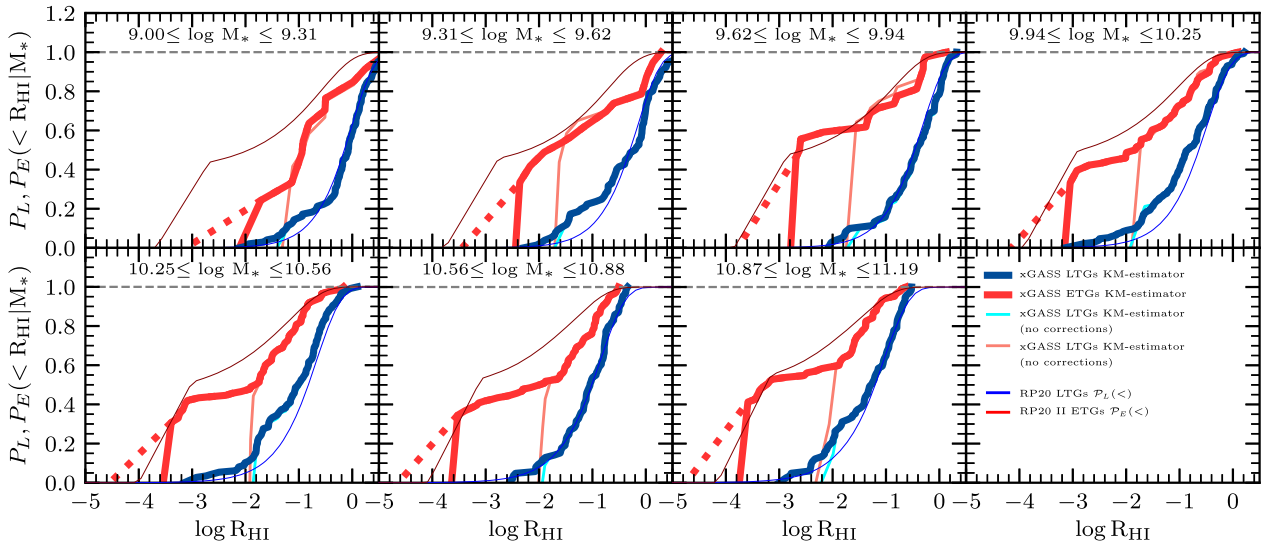

Figure C2. Same R_{HI} conditional CDFs from Fig. 5 and results without taking into into account our procedure for the upper limits of xGASS nor the correction by morphology/environment (lighter colours).

This paper has been typeset from a TEX/LATEX file prepared by the author.