Revista Mexicana de Astronomía y Astrofísica, **54**, 443–483 (2018)

THE HI- AND H_2-TO-STELLAR MASS CORRELATIONS OF LATE- AND EARLY-TYPE GALAXIES AND THEIR CONSISTENCY WITH THE OBSERVATIONAL MASS FUNCTIONS

A. R. Calette[1], Vladimir Avila-Reese[1], Aldo Rodríguez-Puebla[1,2,3], Héctor Hernández-Toledo[1], and Emmanouil Papastergis[4,5]

Received March 16 2017; accepted July 19 2018

ABSTRACT

We compile and homogenize local galaxy samples with available information on morphology, and stellar, HI and/or H_2 masses. After taking into account non gas detections, we determine the HI- and H_2-to-stellar mass relations and their 1σ scatter for late- and early-type galaxies. These relations are fitted to single or double power laws. Late-type galaxies are significantly gas richer than early-type ones, especially at high masses. The H_2-to-HI mass ratios as a function of M_* are discussed. We constrain the distribution functions of the HI- and H_2-to-stellar mass ratios. We find that they can be described by a Schechter function for late types and a (broken) Schechter + uniform function for early types. Using the observed galaxy stellar mass function and the volume-complete late-to-early-type galaxy ratio as a function of M_*, these distributions are mapped into HI and H_2 mass functions. The mass functions are consistent with those inferred from large surveys. The results presented here can be used to constrain models and simulations of galaxy evolution.

RESUMEN

Compilamos y homogeneizamos muestras de galaxias locales con información sobre la morfología, la masa estelar, y la de HI y/o H_2.Tomando en cuenta las no-detecciones en gas determinamos las relaciones masa estelar a masa de HI y H_2 y sus dispersiones, para galaxias tardías y tempranas. Las relaciones se ajustan con leyes de potencia simple o doble. Las galaxias tardías son más ricas en gas que las tempranas. Se discuten los cocientes de masa H_2 a HI en función de M_*. Constreñimos las distribuciones de los cocientes de masa de HI y H_2 a masa estelar, y encontramos que se describen bien por una función de Schechter (galaxias tardías) o una función Schechter (cortada) + uniforme (galaxias tempranas). Usamos la función de masa estelar y el cociente de galaxias tempranas a tardías en función de M_* para mapear estas distribuciones a funciones de masa de HI y H_2 las cuales concuerdan con las inferidas de los grandes catastros. Los resultados que presentamos pueden usarse para constreñir modelos y simulaciones de evolución de galaxias.

Key Words: galaxies: general — galaxies: ISM — galaxies: mass functions — galaxies: statistics

1. INTRODUCTION

Galaxies are complex systems, formed mainly from the cold gas captured by the gravitational potential of dark matter halos and transformed into stars, but also reheated and eventually ejected from the galaxy by feedback processes (see for a recent review Somerville & Davé 2015). Therefore, the content of gas, stars, and dark matter of galaxies provides key information to understand their evolution and present-day status, as well as to constrain

[1]Instituto de Astronomía, Universidad Nacional Autónoma de México.
[2]Department of Astronomy & Astrophysics, University of California at Santa Cruz, USA.
[3]Center for Astronomy and Astrophysics, Shanghai Jiao Tong University, China.
[4]Kapteyn Astronomical Institute, University of Groningen, The Netherlands.
[5]Credit Risk Modeling Department, Coöperative Rabobank U.A., The Netherlands.

models and simulations of galaxy formation (see e.g., Zhang et al. 2009; Fu et al. 2010; Lagos et al. 2011; Duffy et al. 2012; Lagos et al. 2015).

Local galaxies fall into two main populations, according to the dominance of the disk or bulge component (late- and early-types, respectively; a strong segregation is also observed by color or star formation rate). The main properties and evolutionary paths of these components are different. Therefore, the present-day stellar, gaseous, and dark matter fractions are expected to be different among late-type/blue/star-forming and early-type/red/passive galaxies of similar masses. The above demands that the gas-to-stellar mass relations be determined separately for each population. Morphology, color and star formation rate correlate among them, though there is a fraction of galaxies that skips the correlations. In any case, when only two broad groups are used to classify galaxies, the segregation in the resulting correlations for each group is expected to be similar for any of these criteria. Here we adopt the morphology as the criterion for classifying galaxies into two broad populations.

With the advent of large homogeneous optical/infrared surveys, the statistical distributions of galaxies, for example the galaxy stellar mass function (GSMF), are now very well determined. In the last years, using these surveys and direct or statistical methods, the relationship between the stellar, M_*, and halo masses has been constrained (e.g., Mandelbaum et al. 2006; Conroy & Wechsler 2009; More et al. 2011; Behroozi et al. 2010; Moster et al. 2010; Rodríguez-Puebla et al. 2013; Behroozi et al. 2013; Moster et al. 2013; Zu & Mandelbaum 2015). Recently, the stellar-to-halo mass relation has been even inferred for (central) galaxies separated into blue and red ones by Rodríguez-Puebla et al. (2015). These authors have found that there is a segregation by color in this relation (see also Mandelbaum et al. 2016). The semi-empirical stellar-to-halo mass relation and its scatter provide key constraints to models and simulations of galaxy evolution. These constraints would be stronger if the relations between the stellar and atomic/molecular gas contents of galaxies were included. With this information, the galaxy baryonic mass function can be also constructed and the baryonic-to-halo mass relation can be inferred, see e.g, Baldry et al. (2008).

While the stellar component is routinely obtained from large galaxy surveys in optical/infrared bands, the information about the cold gas content is much more scarce due to the limits in sensitivity and sky coverage of current radio telescopes. In fact, the few blind HI surveys, obtained with a fixed integration time per pointing, suffer from strong biases, and for H_2 (CO) there are no surveys. For instance, the HI Parkes All-Sky Survey (HIPASS; Barnes et al. 2001; Meyer et al. 2004) or the Arecibo Legacy Fast ALFA survey (ALFALFA; Giovanelli et al. 2005; Haynes et al. 2011; Huang et al. 2012a), miss galaxies with low gas-to-stellar mass ratios, specially at low stellar masses. Therefore, the HI-to-stellar mass ratios inferred from the crossmatch of these surveys with optical ones should be regarded as an upper limit envelope (see e.g., Baldry et al. 2008; Papastergis et al. 2012; Maddox et al. 2015). In the future, facilities such as the Square Kilometre Array (SKA; Carilli & Rawlings 2004; Blyth et al. 2015), or precursor instruments such as the Australian SKA Pathfinder (ASKAP; Johnston et al. 2008) and the outfitted Westerbork Synthesis Radio Telescope (WSRT), will bring extragalactic gas studies more in line with optical surveys. Until then, the gas-to-stellar mass relations of galaxies can be constrained: (i) from limited studies of radio follow-up observations of large optically-selected galaxy samples or by cross-correlating some radio surveys with optical/infrared surveys (e.g., Catinella et al. 2012; Saintonge et al. 2011; Boselli et al. 2010; Papastergis et al. 2012); and (ii) from model-dependent inferences based, for instance, on the observed metallicities of galaxies or from calibrated correlations with photometrical properties (e.g., Baldry et al. 2008; Zhang et al. 2009).

While this paper does not present new observations, it can be considered as an extension of previous efforts to attempt to determine the HI-, H_2- and cold gas-to-stellar mass correlations of local galaxies over a wide range of stellar masses. Moreover, here we separate galaxies into at least two broad populations, late- and early-type galaxies (hereafter LTGs and ETGs, respectively). These empirical correlations are fundamental benchmarks for models and simulations of galaxy evolution. Our main goal here is to constrain these correlations by using and uniforming large galaxy samples of good quality radio observations with confirmed optical counterparts. Moreover, the well determined local GSMF combined with these correlations can be used to construct the galaxy HI and H_2 mass functions, GHIMF and GH_2MF, respectively. As a test of consistency, we compare these mass functions with those reported in the literature for HI and CO (H_2).

Many of the samples compiled here suffer from incompleteness and selection effects or, in many cases, the radio observations provide only upper limits to

the flux (non-detections). To provide reliable determinations of the HI- and H_2-to-stellar mass correlations, for both LTGs and ETGs, here we homogenize as much as possible the data, check them against selection effects that could affect the calibration of the correlations, and take into account adequately the upper limits. We are aware of the limitations of this approach. Note, however, that in absence of large homogeneous galaxy surveys reporting gas scaling relations over a wide dynamical range and separated into late- and early-type galaxies, our approach is well supported as well as its fair use.

The plan of the paper is as follows. In § 2 and Appendices A and B, we present our compilation and homogenization of local galaxy samples with available information on stellar mass, morphological type, and HI and/or H_2 masses from the literature. In § 3, we test the different compiled samples against possible biases in the gas content due to selection effects. In § 4, we describe the strategy to infer the gas-to-stellar mass correlations taking into account upper limits, and present the determination of these correlations for the LTG and ETG populations (mean and standard deviations). Further, in § 5 we constrain the full distributions of the gas-to-stellar mass ratios as a function of M_*. In § 6 we explore the consistency of the determined correlations with the observed HI and H_2 mass functions, by using the GSMF as an interface. In § 7.1 we discuss the H_2-to-HI mass ratios of LTGs and ETGs inferred from our correlations; § 7.2 is devoted to a discussion on the role of the environment, and § 7.3 presents comparisons with some previous attempts to determine the gas scaling relations. A summary of our results and the conclusions are presented in § 8. Finally, Table 1 lists all the acronyms used in this paper, including the ones of the surveys/catalogs used here.

TABLE 1

LIST OF ACRONYMS USED IN THIS PAPER

BCD	Blue compact dwarf
ETG	Early-type galaxy
GHIMF	Galaxy HI Mass Function
GH_2MF	Galaxy H_2 Mass Function
GSMF	Galaxy Stellar Mass Function
IMF	Initial Mass Function
LTG	Late-type galaxy
MW	Milky Way
$R_{\rm HI}$ and $R_{\rm H_2}$	HI- and H_2-to stellar mass ratio
SB	Surface brightness
SFR	Star formation rate
ALFALFA	Arecibo Legacy Fast ALFA survey
ALLSMOG	APEX Low-redshift Legacy Survey for MOlecular Gas
AMIGA	Analysis of the interstellar Medium of Isolated GAlaxies
ASKAP	Australian SKA Pathfinder
$ATLAS^{3D}$	(A volume-limited survey of local ETGs)
COLD GASS	CO Legacy Database for GASS
FCRAO	Five College Radio Astronomy Observatory
GALEX	Galaxy Evolution EXplorer
GAMA	Galaxy And Mass Assembly
GASS	GALEX Arecibo SDSS Survey
HERACLES	HERA CO-Line Extragalactic Survey
HIPASS	HI Parkes All-Sky Survey
HRS	Herschel Reference Survey
NFGS	Nearby Field Galaxy Catalog
NRTA	Nancay Radio Telescope
SDSS	Sloan Digital Sky Survey
SINGS	Spitzer Infrared Nearby Galaxies Survey
SKA	Square Kilometre Array
THINGS	The HI Nearby Galaxy Survey
UNAM-KIAS	UNAM-KIAS survey of SDSS isolated galaxies
UNGC	Updated Nearby Galaxy Catalog
WRST	Westerbork Synthesis Radio Telescope

2. COMPILATION OF OBSERVATIONAL DATA

The main goal of this section is to present our extensive compilation of observational studies (catalogs, surveys or small samples) that meet the following criteria:

- Include HI and/or H_2 masses from radio observations, and luminosities/stellar masses from optical/infrared observations.
- Provide the galaxy morphological type or a proxy of it.
- Describe the selection criteria of the sample and provide details about the radio observations, flux limits, etc.
- Include individual distances to the sources and corrections for peculiar motions/large-scale structures for the nearby galaxies.
- In the case of non-detections, provide estimates of upper limits for HI or H_2 masses.

The observational samples that meet the above criteria are listed in Table 2. In Appendices A and B, we present a summary of each one. We have found information on colors ($g-r$ or $B-K$) for most of the samples. For $M_* > 10^9\ M_\odot$, the galaxies in the color–mass diagram segregate into the so-called red sequence and blue cloud. Excluding those more inclined than 70 degrees, we find that $\approx 83\%$ of LTGs ($\approx 80\%$ of ETGs) have colors that can be classified as blue (red) by using a mass-dependent ($g-r$) criterion to define blue/red galaxies. At masses lower than $M_* \approx 10^9\ M_\odot$, the overwhelming majority of galaxies are of late types and are classified as blue.

2.1. *Systematic Effects on the* HI- *and* H_2*-to-Stellar Mass Correlations*

To reduce potential systematic effects that can bias how we derive the HI- and H_2-to-stellar mass correlations we homogenize all the compiled observations to the same basis. Following, we discuss some potential sources of bias/segregation and the calibration that we apply to the observations. It is important to stress that to infer scaling correlations,

TABLE 2

OBSERVATIONAL SAMPLES

Sample	Selection	Environment	HI	Detections / Total	H_2	Detections / Total	IMF	Category
UNGC	ETG+LTG	local 11 Mpc	Yes	407 / 418	No	–	diet-Salpeter	Gold
GASS/COLD GASS	ETG+LTG	no selection	Yes	511 / 749	Yes	229 / 360	Chabrier (2003)	Gold
HRS–field	ETG+LTG	no selection	Yes	199 / 224	Yes	101 / 156	Chabrier (2003)	Gold
ATLAS^{3D}–field	ETG	field	Yes	51 / 151	Yes	55 / 242	Kroupa (2001)	Gold
NFGS	ETG+LTG	no selection	Yes	163 / 189	Yes	27 / 31	Chabrier (2003)	Silver
Stark et al. (2013) compilation*	LTG	no selection	Yes	62/62	Yes	14 / 19	diet-Salpeter	Silver
Leroy+08 THINGS/HERACLES	LTG	nearby	Yes	23 / 23	Yes	18 / 20	Kroupa (2001)	Silver
Dwarfs-Geha+06	LTG	nearby	Yes	88 / 88	No	–	Kroupa et al. (1993)	Silver
ALFALFA dwarf	ETG+LTG	no selection	Yes	57 / 57	No	–	Chabrier (2003)	Silver
ALLSMOG	LTG	field	No	–	Yes	25 / 42	Kroupa (2001)	Silver
Bauermeister et al. (2013) compilation	LTG	field	No	–	Yes	7 / 8	Kroupa (2001)	Silver
ATLAS^{3D}–Virgo	ETG	Virgo core	Yes	2 / 15	Yes	4 / 21	Kroupa (2001)	Bronze
AMIGA	ETG+LTG	isolated	Yes	229 / 233	Yes	158 / 241	diet-Salpeter	Bronze
HRS–Virgo	ETG+LTG	Virgo core	Yes	55 / 82	Yes	36 / 62	Chabrier (2003)	Bronze
UNAM-KIAS	ETG+LTG	isolated	Yes	352 / 352	No	–	Kroupa (2001)	Bronze
Dwarfs-NSA	LTGs	isolated	Yes	124 / 124	No	–	Chabrier (2003)	Bronze

*From this compilation, we considered only galaxies that were not in GASS, COLD GASS and ATLAS^{3D} samples.

as those of the gas fraction as a function of stellar mass, it is important to have a statistically representative and unbiased population of galaxies *in each mass bin*. Thus, there is no need to have mass limited volume-complete samples (see also § 4.1). However, a volume-complete sample assures that possible biases of the measure in question due to selection functions in galaxy type, color, environment, surface brightness, etc., are not introduced. The main expected bias in the gas content at a given stellar mass is due to the galaxy type/color; this is why we need to separate the samples at least into two broad populations, LTGs and ETGs.

2.1.1. *Galaxy Type*

The gas content of galaxies, at a given M_*, segregates significantly with galaxy morphological type (e.g., Kannappan et al. 2013; Boselli et al. 2014c). Thus, information on morphology is necessary in order to separate galaxies at least into two broad populations, LTGs and ETGs. Apart of its physical basis, this separation is important to avoid introducing biases in the obtained correlations due to selection effects related to the morphology of the different samples used here. For example, some samples are only of late-type or star-forming galaxies, others only of early-type galaxies, etc., so that combining them without a separation by morphology would yield correlations that are not statistically representative. We consider as ETGs those classified as ellipticals (E), lenticulars (S0), dwarf E, and dwarf spheroidals or with $T < 1$, and as LTGs those classified as spirals (S), irregulars (Irr), dwarf Irr, and blue compact dwarfs or with $T \geq 1$. The morphological classification criteria used in the different samples are diverse, ranging from individual visual evaluation to automatic classification methods, as the one by Huertas-Company et al. (2011). We are aware of the high level of uncertainty introduced by using different morphological classification methods. However, in our case the morphological classification is used to separate galaxies just into two broad groups. Therefore, such an uncertainty is not expected to affect significantly any of our results. It is important to highlight that the terms LTG and ETG are useful only as qualitative descriptors. These descriptors *should not be applied to individual galaxies, but instead to two distinct populations of galaxies in a statistical sense.*

2.1.2. *Environment*

The gas content of galaxies is expected to depend on the environment (e.g., Zwaan et al. 2005; Geha et al. 2012; Jones et al. 2016; Brown et al. 2017). In this study we are not able to study in detail such a dependence, though our separation into LTG and ETG populations partially takes into account this dependence because these populations segregate by environment (e.g., Dressler 1980; Kauffmann et al. 2004; Blanton et al. 2005a; Blanton & Moustakas 2009, and references therein). In any case, in our compilation we include three samples specially selected to contain very isolated galaxies and one subsample of galaxies from the Virgo Cluster central regions. We will check whether or not their HI and H_2 mass fractions significantly deviate from the mean relations.

2.1.3. *Systematical Uncertainties on the Stellar Masses*

There are many sources of systematic uncertainty in the inference of stellar masses related to the choices of: initial mass function (IMF), stellar population synthesis and dust attenuation models, star

TABLE 3

NUMBER OF GALAXIES WITH DETECTIONS AND UPPER LIMITS BY MORPHOLOGY

Morphology(%)	Detections(%)	Upper limits(%)	Total
HI data			
LTG (78%)	1975 (94%)	121 (6%)	2096
ETG (22%)	292 (50%)	288 (50%)	580
H_2 data			
LTG (63%)	533 (75%)	180 (25%)	713
ETG (37%)	124 (29%)	298 (71%)	422

TABLE 4

NUMBER OF GALAXIES WITH DETECTIONS AND UPPER LIMITS BY CATEGORY

Category (%)	Detections (%)	Upper limits (%)	Total
HI data			
Golden (58%)	1168 (76%)	374 (24%)	1542
Silver (16%)	391 (94%)	26 (6%)	417
Bronze (26%)	708 (99%)	9 (1%)	717
H_2 data			
Golden (67%)	385 (51%)	373 (49%)	758
Silver (10%)	91 (76%)	29 (24%)	120
Bronze (23%)	181 (70%)	76 (30%)	257

formation history parametrization, metallicity, filter setup, etc. For inferences from broad-band spectral energy distribution fitting and using a large diversity of methods and assumptions, Pforr et al. (2012) estimate a maximal variation in stellar mass calculations of ≈ 0.6 dex. The major contribution to these uncertainties comes from the IMF. The IMF can introduce a systematic variation of up to ≈ 0.25 dex (see e.g., Conroy 2013). For local normal galaxies and from UV/optical/IR data (as it is the case of our compiled galaxies), Moustakas et al. (2013) find a mean systematic difference between different mass-to-luminosity estimators (fixed IMF) of less than 0.2 dex. We have seen that in most of the samples compiled here, the stellar masses are calculated using roughly similar mass-to-luminosity estimators, but the IMF are not always the same.Therefore, we homogenize the reported stellar masses in the different compiled samples to the mass corresponding to a Chabrier (2003) initial mass function (IMF), and neglect other sources of systematic differences.

2.1.4. *Other Effects*

We also homogenize the distances to the value of $H_0 = 70$ kms^{-1} Mpc^{-1}. In most of the samples compiled here (at least the most relevant ones for our study), *distances were corrected for peculiar motions and large-scale structure effects*. When the authors included helium and metals to their reported HI and H_2 masses, we take care of subtracting these contributions. When we calculate the total cold gas mass, then helium and metals are explicitly taken into account.

2.1.5. *Categories*

The different HI and H_2 samples used in this paper are widely diverse, in particular they were obtained with different selection functions, radio telescopes, exposure times, etc. We have divided the different samples into three categories according to the feasibility of determining from each one robust and statistically representative HI- or H_2-to-stellar mass correlations for the LTG and ETG populations. We will explore whether or not the less feasible categories should be included for determining these correlations. The three categories are:

1. **Golden:** It includes datasets based on volume-complete (above a given luminosity/mass) samples or on representative galaxies selected from volume-complete samples. The Golden datasets, by construction, are unbiased samples of the distribution of galaxy properties.

2. **Silver:** It includes datasets from galaxy samples that are not volume complete, but that are intended to be statistically representative at least for their morphological groups, i.e., these samples do not present obvious or strong selection effects.

3. **Bronze:** This category includes samples selected deliberately by environment, and it will be used to explore the effects of environment on the LTG and ETG HI- or H_2-to-stellar mass correlations.

2.2. *The Compiled* HI *Sample*

Appendix A presents a summary of the HI samples compiled in this paper (see also Table 2). Table 3 lists the total numbers and fractions of compiled galaxies with detection and non-detection for each galaxy population. Table 4 lists the number of detected and non-detected galaxies for the golden, silver, and bronze categories listed above (§ 2.1.5).

Figure 1 shows the mass ratio $R_{\rm HI} \equiv M_{HI}/M_*$ vs. M_* for the compiled samples. Note that we have applied some corrections to the reported samples (see

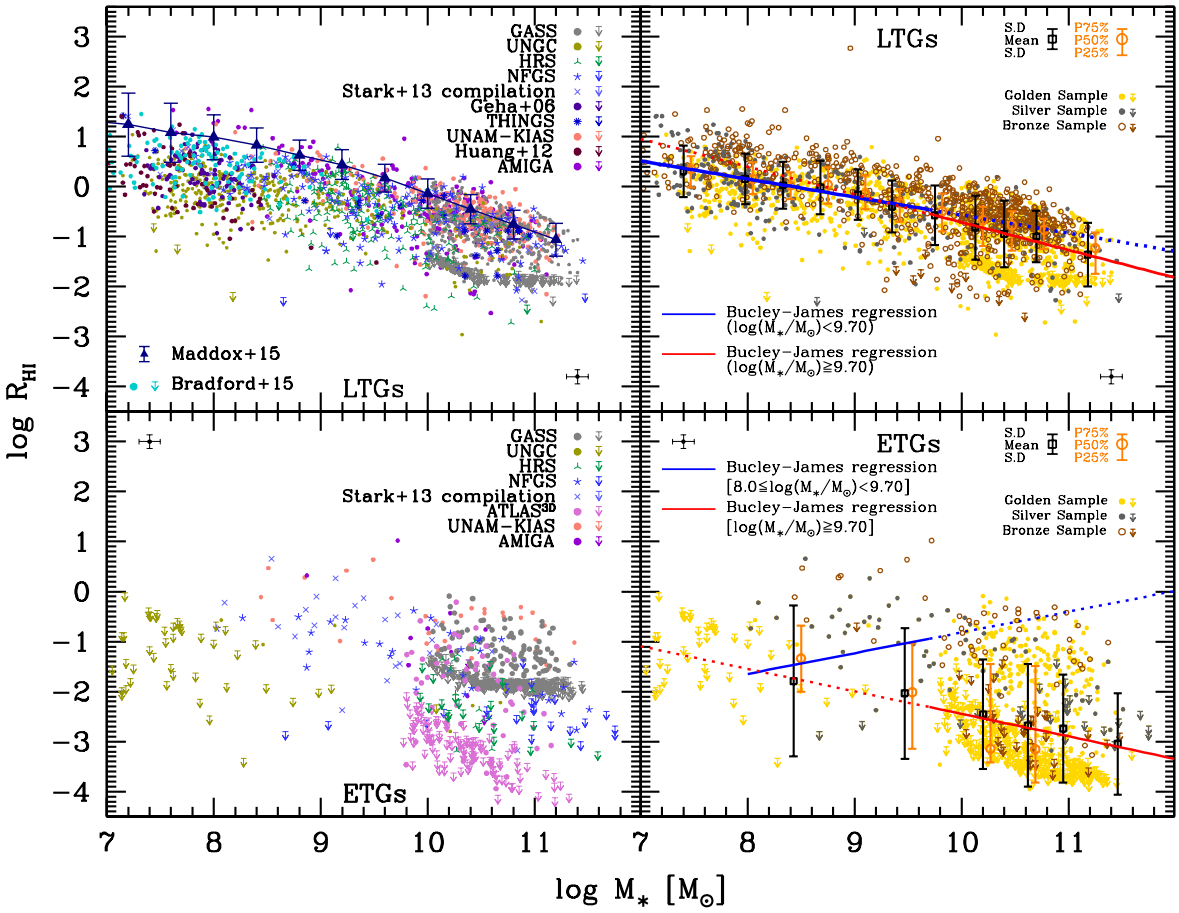

Fig. 1. Atomic gas-to-stellar mass ratio as a function of M_*. *Upper panels:* Compiled and homogenized data with information on $R_{\rm HI}$ and M_* for LTGs (the different sources are indicated inside the left panel; see Appendix A for the acronyms and authors); downward arrows show the reported upper limits for non-detections. The blue triangles with thin error bars are mean values and standard deviations from the v.40 ALFALFA and SDSS crossmatch according to Maddox et al. (2015); the ALFALFA galaxies are biased toward high values of $R_{\rm HI}$ (see text). Right panel is the same as left one, but with the data separated into three categories: Golden, Silver, and Bronze (yellow, gray, and brown symbols, respectively). The red and blue lines are Buckley-James linear regressions (taking into account non-detections) for the high- and low-mass regions, respectively; the dotted lines show extrapolations from these fits. Squares with error bars represent the mean and standard deviation of the data in different mass bins, taking into account non-detections by means of the Kaplan-Meier estimator. Open circles with error bars show the corresponding median and 25-75 percentiles. Estimates of the observational uncertainties are shown in the panel corners (see text). *Lower panels:* Same as upper panels but for ETGs. In the right panel, we have corrected for distance the galaxies with upper limits from GASS to make them consistent with the distances of the ATLAS^{3D} sample (see text); the upper limits from the latter were increased by a factor of two to homogenize them to the ALFALFA instrument and signal-to-noise criteria. For the bins where more than 50% of the data are upper limits, the median and percentiles are not calculated. The color figure can be viewed online.

above) to homogenize all the data. The upper and bottom left panels of Figure 1 show, respectively, the compilations for LTGs and ETGs. The different symbols indicate the source reference of the data and the downward arrows are the corresponding upper limits of the HI-flux for non-detections. We also reproduce the mean and standard deviation of different mass bins as reported in Maddox et al. (2015) for a cross-match of the ALFALFA and SDSS surveys. As mentioned in the Introduction, the ALFALFA survey is biased to high $R_{\rm HI}$ values, specially towards the low mass side. Note that the small ALFALFA subsample of dwarf galaxies by Huang et al. (2012b, dark purple dots) was selected mainly as an attempt to take into account low-HI mass galaxies at the low-mass end.

2.3. *The Compiled* H_2 *Sample*

Since the emission of cold H_2 in the ISM is extremely weak, a tracer of the H_2 abundance should be used. The best tracer from the observational point of view is the CO molecule due to its relatively high abundance and its low excitation energy. The H_2 mass is related to the CO luminosity through a CO-to-H_2 conversion factor: $M_{H_2} = \alpha_{\rm CO} L_{\rm CO}$. This factor has been determined in molecular clouds in the Milky Way (MW), $\alpha_{\rm CO,MW} = 3.2$ (K km s^{-1} pc^{-1})$^{-1}$, with a systematic uncertainty of 30%. It was common to assume that this conversion factor was the same for all galaxies. However, several pieces of evidence show that $\alpha_{\rm CO}$ is not constant, and that it depends mainly on the gas-phase metallicity, increasing as the galaxy metallicity decreases (e.g., Boselli et al. 2002; Schruba et al. 2012; Narayanan et al. 2012; Bolatto et al. 2013, and references therein). At first-order, $\alpha_{\rm CO}$ changes slowly for metallicities larger than $12 + \log_{10}({\rm O/H}) \approx 8.4$ (approximately half the solar one) and increases considerably as the metallicity decreases. Here, we combine the dependence of $\alpha_{\rm CO}$ on metallicity given by Wolfire et al. (2010) and the observed mass–metallicity relation to obtain an approximate estimation of the dependence of $\alpha_{\rm CO}$ on M_* for LTGs; see Appendix C for details. We are aware that the uncertainties involved in any metallicity-dependent correction remain substantial (Bolatto et al. 2013). Note, however, that our aim is to introduce and explore at a statistical level a reasonable mass-dependent correction to the CO-to-H_2 factor, which must be better than ignoring it. In any case, we present results both for $\alpha_{\rm CO} = \alpha_{\rm CO,MW}$ and our inferred mass-dependent $\alpha_{\rm CO}$ factor. In fact, the mass-dependent factor is important only for LTGs with $M_* <\lesssim 3 \times 10^{10}\ M_\odot$; for larger masses and for all ETGs, $\alpha_{\rm CO} \approx \alpha_{\rm CO,MW}$[6].

Appendix B presents a description of the CO (H_2) samples that we utilize in this paper. Table 3 lists the number of galaxies with detections and upper limits in the compilation sample in terms of morphology. Table 4 lists the number of detections and upper limits for the golden, silver, and bronze categories mentioned above (§ 2.1.5).

Figure 2 shows the mass ratio $R_{\rm H_2} \equiv M_{H_2}/M_*$ vs. M_* for the compiled samples. Similarly to the $R_{\rm HI}$ vs. M_* relation, we applied some corrections to observations in order to homogenize our compiled sample and to allow a more consistent comparison between the different samples. The upper and bottom left panels of Figure 2 show, respectively, the compiled datasets for LTGs and ETGs.

[6]This is well justified since massive LTGs are metallic with typical values larger than $12 + \log_{10}({\rm O/H}) \approx 8.7$, while ETG have large metallicities at all masses.

3. TESTS FOR SELECTION EFFECTS AND PRELIMINARY RESULTS

In this section we check the gas-to-stellar mass correlations from the different compiled samples for possible selection effects. We also introduce, when possible, a homogenization at the upper limits of ETGs. The reader interested only in the main results can skip to § 4.

As seen in Figures 1 and 2 there is a significant fraction of galaxies with no detections in radio, for which the authors report an upper limit for the flux (converted into an HI or H_2 mass). *The non-detection of observed galaxies gives information that we cannot ignore,* otherwise a bias towards high gas fractions would be introduced in the gas-to-stellar mass relations to be inferred. To take into account the upper limits in the compiled data, we resort to survival analysis methods for combining censored and uncensored data (i.e., detections and upper limits for non-detections; see e.g., Feigelson & Babu 2012). We will use two methods: the Buckley-James linear regression (Buckley & James 1979) and the Kaplan-Meier product limit estimator (Kaplan & Meier 1958). Both are survival analysis methods commonly applied in astronomy.[7] The former is useful for obtaining a linear regression from the censored and uncensored data. Alternatively, for data that cannot be described by a linear relation, we can bin them by mass, use the Kaplan-Meier estimator to calculate the mean, standard deviation,[8] median, and 25-75 percentiles in each stellar mass bin, and fit

[7]We use the Asurv (Astronomy SURVival analysis) package developed by T. Isobe, M. LaValley and E. Feigelson in 1992 (see also Feigelson & Nelson 1985), and implemented in the Stsdas package (Space Telescope Science Science Data Analysis) in Iraf. In particular, we make use of the Buckleyjames (Buckley-James linear regression) and Kmestimate (Kaplan-Meier estimator) routines.

[8]The Iraf package provides actually the standard error of the mean, $SEM = s/\sqrt{n}$, where $s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$ is the sample standard deviation, n is the number of observations, and $\bar{x}$ is the sample mean. In fact, s is a biased estimator of the (true) population standard deviation σ. For small samples, the former underestimates the true population standard deviation. A commonly used rule of thumb to correct the bias when the distribution is assumed to be normal, is to introduce the term $n - 1.5$ in the computation of s instead of n. In this case, $s \to \sigma$. Therefore, an approximation to the population standard deviation is $\sigma = (n/\sqrt{n - 1.5}) \times SEM$. This is the expression we use to calculate the reported standard deviations.

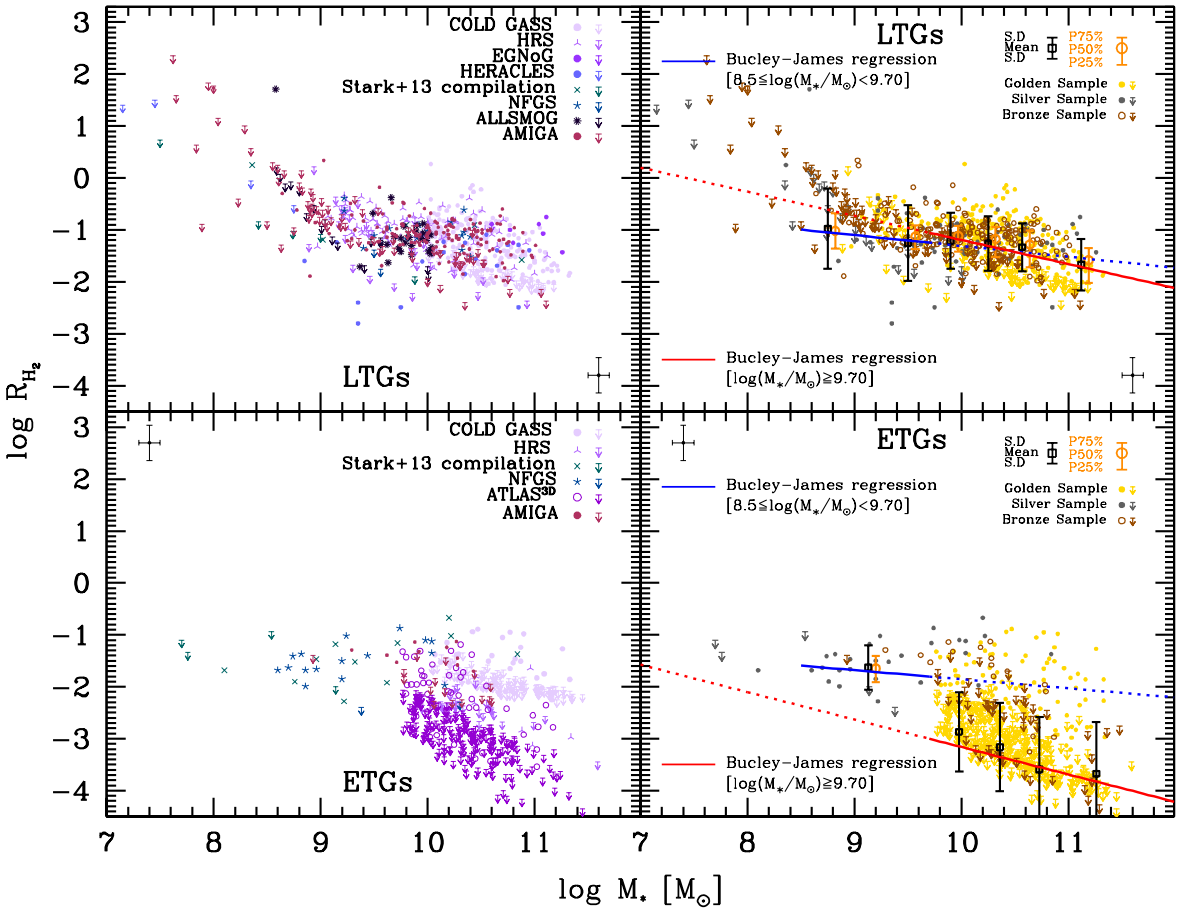

Fig. 2. Molecular gas-to-stellar mass ratio as a function of M_*. *Upper panels:* Compiled and homogenized data with information on $R_{\rm H_2}$ and M_* for LTGs (see inside the panels for the different sources; see Appendix B for the acronyms and authors); downward arrows show the reported upper limits for non-detections. Right panel is the same, but with the data separated into three categories: Golden, Silver, and Bronze (yellow, gray, and brown symbols, respectively). The red and blue lines are Buckley-James linear regressions (taking into account non-detections). The dotted lines show extrapolations from these fits. The green dashed line shows an estimate for the $R_{\rm H_2}$–M_* relation inferred from combining the empirical SFR–M_{H_2} and SFR–M_* correlations for blue/star-forming galaxies (see text for details). Squares with error bars are the mean and standard deviation of the data in different mass bins, taking into account non-detections by means of the Kaplan-Meier estimator. Open circles with error bars are the corresponding median and 25-75 percentiles. Estimates of the observational/calculation uncertainties are shown in the panel corners (see text). *Lower panels:* The same as upper panels but for ETGs. In the right panel, we have corrected for distance the galaxies with upper limits from COLD GASS to make them consistent with the distances of the ATLAS^{3D} sample (see text). For the bins where more than 50% of the data are upper limits, the median and percentiles are not calculated. The color figure can be viewed online.

these results to a function using conventional methods, e.g., the Levenberg-Marquardt algorithm. For the latter case, the binning in $\log M_*$ is started with a width of ≈ 0.25 dex but if the data are too scarce in the bin, then its width is increased so as to have no less than 25% of galaxies in the most populated bins. Note that, for detection fractions smaller than 50%, the median and percentiles are very uncertain or impossible to be calculated with the Kaplan-Meier estimator (Lee & Wang 2003), while the mean can still be estimated for fractions as small as $\approx 20\%$, though with a large uncertainty. In the case of the Bukley-James linear regression, reliable results are guaranteed for detection fractions larger than $70-80\%$.

When the fraction of non-detections is significant, the inferred correlations could be affected by *selection effects in the upper limits* reported in the different samples. This is the case for ETGs, where a clear systematical segregation between the upper limits of the GALEX Arecibo SDSS Survey (GASS)

and ATLAS^{3D} or Herschel Reference Survey (HRS) surveys is observed in the $\log R_{\rm HI} - \log M_*$ plane (see the gap in the lower left panel of Figure 1), as well as between the CO Legacy Database for GASS (COLD GASS) and ATLAS^{3D} or HRS surveys in the $\log R_{\rm H_2} - \log M_*$ plane (see the gap in the lower left panel of Figure 2). The determination of the upper limits depends on distance and instrumental/observational constraints (telescope sensitivity, integration time, spatial coverage, signal-to-noise threshold, etc.). The HI observations of GASS and ATLAS^{3D} were carried out with different radio telescopes: the single-dish Arecibo Telescope and the Westerbork Synthesis Radio Telescope (WRST) interferometer array, respectively. Serra et al. (2012) discussed the differences between detections by single- and multiple-beam observations. From some galaxies from ATLAS^{3D} that they observed also with the Arecibo telescope, they concluded that the upper limits should be increased by a factor of ≈ 2 in order to agree with the ALFALFA survey sensitivity and the signal-to-noise threshold they used for declaring non-detections in their multiple-beam observations. Thus, to homogenize the upper limits, we corrected the ATLAS^{3D} upper limits by this factor. In the case of $R_{\rm H_2}$, the CO observations in the ATLAS^{3D} and COLD GASS samples were obtained with the same radio telescope (IRAM).

The GASS (COLD GASS) samples are selected to include galaxies at distances between ≈ 109 and 222 Mpc, while the ATLAS^{3D} and HRS surveys include only nearby galaxies, with average distances of 25 and 19 Mpc, respectively. Since the definition of the upper limits *depends on distance*, for the same radio telescope and integration time, more distant galaxies have systematically higher upper limits than nearer galaxies. This introduces a clear selection effect. When we have information for a sample of galaxies nearer than another sample, and under the assumption that both samples are roughly representative of the same local galaxy population, a distance-dependent correction to the upper limits of the non-detected galaxies in the more distant sample should be introduced. In Appendix D, we describe our approach to apply such a correction to GASS (COLD GASS) ETG upper limits with respect to the ATLAS^{3D} ETGs. We test our corrections by using a mock catalog. This correction for distance is an approximation based on the assumption that the (COLD)GASS and ATLAS^{3D} ETGs are statistically similar populations. In any case, we will present the correlations for ETGs for both cases, with and without this correction.

Note that after our corrections for distance and instrumental effects, the upper limits of the massive ETGs in the GASS/COLD GASS sample are now consistent with those in the ATLAS^{3D} (as well as HRS) samples, as seen in the right panels of Figures 1 and 2 to be described below, and in Figure 17 in Appendix D. In the case of LTGs, there is no evidence of much lower values of $R_{\rm HI}$ and $R_{\rm H_2}$ than the upper limits given in GASS and COLD GASS for galaxies nearer than those in these samples.

In the right panels of Figures 1 and 2, all the compiled data shown in the left panels are again plotted with dots and arrows for detections and non-detections, respectively. The yellow, dark gray, and brown colors correspond to galaxies from the Golden, Silver, and Bronze categories, respectively (see § 2.1.5). The above mentioned corrections to the upper limits of GASS/COLD GASS and ATLAS^{3D} ETG samples were applied. Note that the large gaps in the upper limits between the GASS/COLD GASS and ATLAS^{3D} (or HRS) samples tend to disappear after the corrections.

We further group the data in logarithmic mass bins and calculate in each mass bin the mean and standard deviation of $\log R_{\rm HI}$ and $\log R_{\rm H_2}$ (black circles with error bars), taking into account the upper limits with the Kaplan-Meier estimator as described above. The orange squares with error bars show the corresponding medians and 25-75 percentiles, respectively. In some mass bins the fraction of detections is smaller than 50% for ETGs. Therefore, the medians and percentiles cannot be estimated (see above). However, the means and standard deviations can still be calculated, though they are quite uncertain.

As seen in the right panels of Figures 1 and 2, the logarithmic mean and median values tend to coincide and the 25-75 percentiles are roughly symmetric in most of the cases. Both facts suggest that the scatter around the mean relations (at least for the LTG population) tends to follow a nearly symmetrical distribution, for instance, a normal distribution in the logarithmic values (for a more detailed analysis of the scatter distributions see § 5).

In the following, we check whether each one of the compiled and homogenized samples deviate significantly from the mean trends. This could be due to selection effects in the sample. For example, we expect systematical deviations in the gas contents for the Bronze samples, because they contain galaxies in extreme environments. As a first approximation, we apply the Buckle-James linear regression to each one

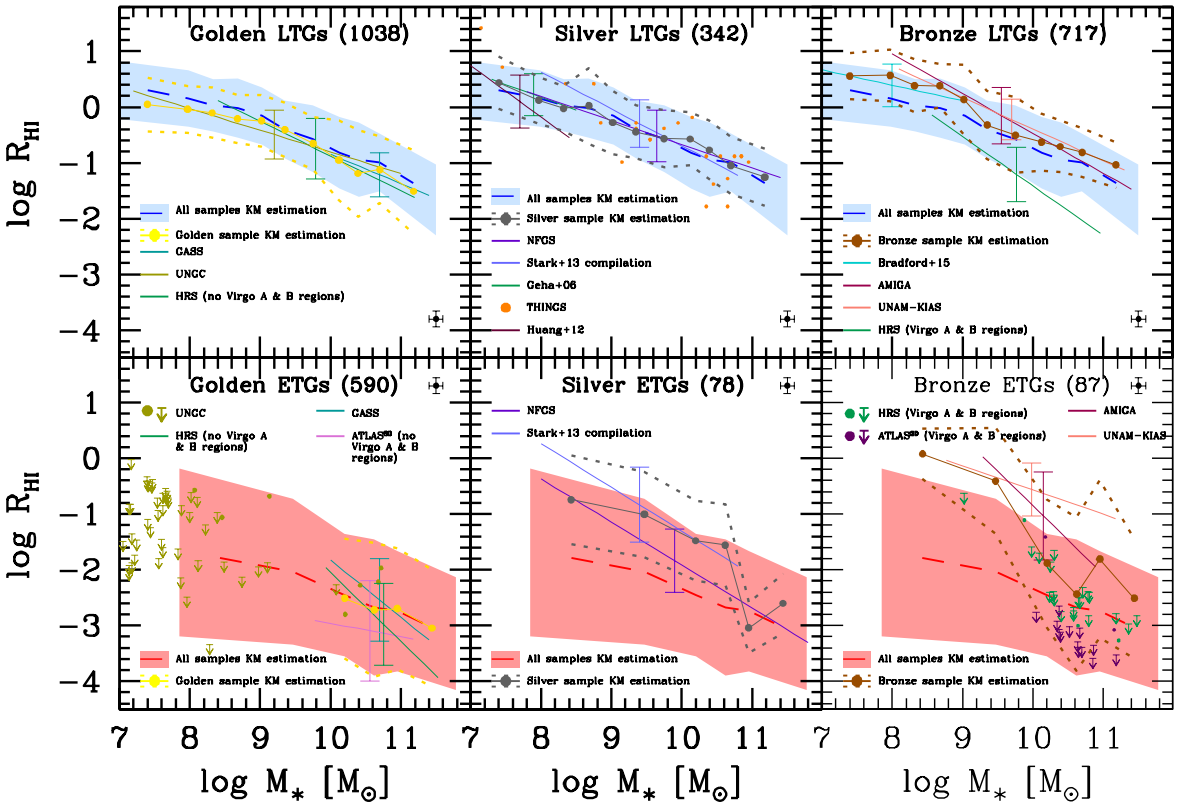

Fig. 3. Atomic gas-to-stellar mass ratio as a function of M_* for the Golden, Bronze, and Silver LTGs (upper panels) and ETGs (lower panels). The mean and standard deviation in different mass bins, taking into account upper limits by means of the Kaplan-Meier estimator, are plotted for each case (filled circles connected by a dotted line and dotted lines around, respectively). For comparison, the mean and standard deviation (dashed lines and shaded area) from all the LTG (ETG) samples are reproduced in the corresponding upper (lower) panels. For each sample compiled and homogenized from the literature, the Buckley-James linear regression is applied, taking into account upper limits. The lines show the result, covering the range of the given sample; the error bars show the corresponding standard deviations obtained from the regression. When the data are too scarce and dominated by upper limits, the linear regression is not applied but the data are plotted. The numbers of LTG and ETG objects in each category are indicated in the respective panel. The color figure can be viewed online.

of the compiled individual samples, taking into account in this way the upper limits. When the data in the sample are too scarce and/or are dominated by non-detections, the linear regression is not performed, but the data are plotted.

3.1. $R_{\rm HI}$ *vs.* M_*

In Figure 3, results for $\log R_{\rm HI}$ vs. $\log M_*$ are shown for LTGs (upper panels) and ETGs (lower panels). From left to right, the regressions for samples in the Golden, Silver, and Bronze categories are plotted. The error bars correspond to the 1σ scatter of the regression. Each line covers the mass range of the corresponding sample. The blue/red dashed lines and shaded regions in each panel correspond to the mean and standard deviation values calculated with the Kaplan-Meier estimator in mass bins for *all* the compiled LTG and ETG samples, previously plotted in Figures 1 and 2, respectively. However, the yellow, gray, and brown dots connected with thin solid lines in each panel are the mean values in each mass bin calculated *only* for the Golden, Silver, and Bronze samples, respectively. The standard deviations are plotted as dotted lines. In the following, we discuss the results shown in Figure 3.

Golden category: For LTGs, the three samples grouped in this category agree well among themselves in the mass ranges where they overlap; even the 1σ scatter of each sample does not differ sig-

nificantly[9]. Therefore, as expected, these samples provide unbiased information for determining the $R_{\rm HI}$–M_* relation of LTGs from $\log(M_*/M_\odot)\approx$ 7.3 to 11.4. For ETGs, the deviations of the Golden linear regressions among themselves and compared to all galaxies are within the 1σ scatter, which is actually large. If no corrections to the upper limits of the GASS and ATLAS^{3D} are applied, then the regression for the former would be significantly above the regression for the latter. Within the large scatter, the three Golden samples of ETGs seem not to be particularly biased, and they cover a mass range from $\log(M_*/M_\odot)\approx$ 8.5 to 11.5. At smaller masses, the Updated Nearby Galaxy Catalog (UNGC) sample provides mostly only upper limits to $R_{\rm HI}$.

Silver category: The LTG and ETG samples in this category, as expected, show more dispersed distributions in their respective $R_{\rm HI}$–M_* planes than those from the Golden category. However, the deviations of the Silver linear regressions among themsleves and compared to all the galaxies are within the corresponding 1σ scatter. If any, there is a trend of the Silver samples to have mean $R_{\rm HI}$ values above the mean values of all galaxies especially for ETGs. Since the samples in this category are volume-complete, (they were specially constructed to study HI gas content), a selection effect towards objects with non-negligible or higher than the mean HI content can be expected. In any case, the biases are small. Thus, we decided to include the Silver samples to infer the $R_{\rm HI}$–M_* correlations in order to slightly increase the statistics (the number of galaxies in this category is actually much smaller than in the Golden category), specially for ETGs of masses smaller than $\log(M_*/M_\odot)\approx$ 9.7 (see Table 4).

Bronze category and effects of the environment: The very isolated LTGs (from the UNAM-KIAS and Analysis of the interstellar Medium of Isolated GAlaxies -AMIGA- samples) have HI contents higher than the mean of all the galaxies, especially at lower masses: $\log R_{\rm HI}$ is $0.1-0.2$ dex larger than the average at $\log(M_*/M_\odot)\gtrsim 10$ and these differences increase up to $0.6-0.3$ dex for $8<\log(M_*/M_\odot)<9$, though the number of galaxies at these masses is very small. The HI content of the Bradford et al. (2015) isolated dwarf galaxies is also larger than the mean of all the galaxies but not by a factor larger than 0.4 dex. For isolated ETGs, the differences can attain an order of magnitude and are at the limit of the upper standard deviations around the means of all the ETGs. Thus, while isolated LTGs have somewhat larger $R_{\rm HI}$ ratios on average than galaxies in other environments, in the case of isolated ETGs, this difference is very large; isolated ETGs can be almost as gas rich as LTGs. In the Bronze group we have included also galaxies from the central regions of the Virgo Cluster, as reported in HRS and ATLAS^{3D} (only ETGs for the latter). According to Figure 3, the LTGs in this high-density environment are clearly HI deficient with respect to LTGs in less dense environments. For ETGs, the HI content is very low but only slightly lower on average than the HI content of all ETGs. It should be noted that ETGs, in particular the massive ones, tend to be located in high-density environments.

We conclude that the HI content of galaxies is affected by the effects of extreme environments. The most remarkable effect occurs for ETGs, which in a very isolated environment can be as rich in HI as LTGs. Therefore, we decided not to include galaxies from the Bronze category to determine the $R_{\rm HI}$–M_* correlation of ETGs. In fact, our compilation in the Golden and Silver categories includes galaxies from a range of environments (for instance, in the largest compiled catalog, UNGC, 58% of the galaxies are members of groups and 42% are field galaxies, see Karachentsev et al. 2014) in such a way that the $R_{\rm HI}$–M_* correlation determined below should represent an average of different environments. Excluding the Bronze category for the ETG population, we avoid biases due to effects of the most extreme environments. For LTGs, the inclusion of the Bronze category does not introduce significant biases in the $R_{\rm HI}$–M_* correlation of all galaxies but it helps to improve the statistics. The mean values of $R_{\rm HI}$ in mass bins above $\approx 10^9\ M_\odot$ are actually close to the mean values of the entire sample (compare the brown solid and blue dashed lines); at smaller masses the deviation increases, but the differences are well within the 1σ dispersion.

3.2. $R_{\rm H_2}$ *vs.* M_*

In Figure 4, we present plots similar to Figure 3 but for $\log R_{\rm H_2}$ vs. $\log M_*$. The symbol and line codes are the same in both figures. In the following, we discuss the results shown in Figure 4.

Golden category: For LTGs, the two samples grouped in this category agree well among themsleves and with the overall sample, though for masses $<10^{10}\ M_\odot$, where the Golden galaxies are only those from the HRS sample, the average $R_{\rm H_2}$ values are

[9]Note also that the 1σ scatter provided by the Buckle-James linear regression is consistent with the standard deviations in the mass bins obtained with the Kaplan-Meier estimator.

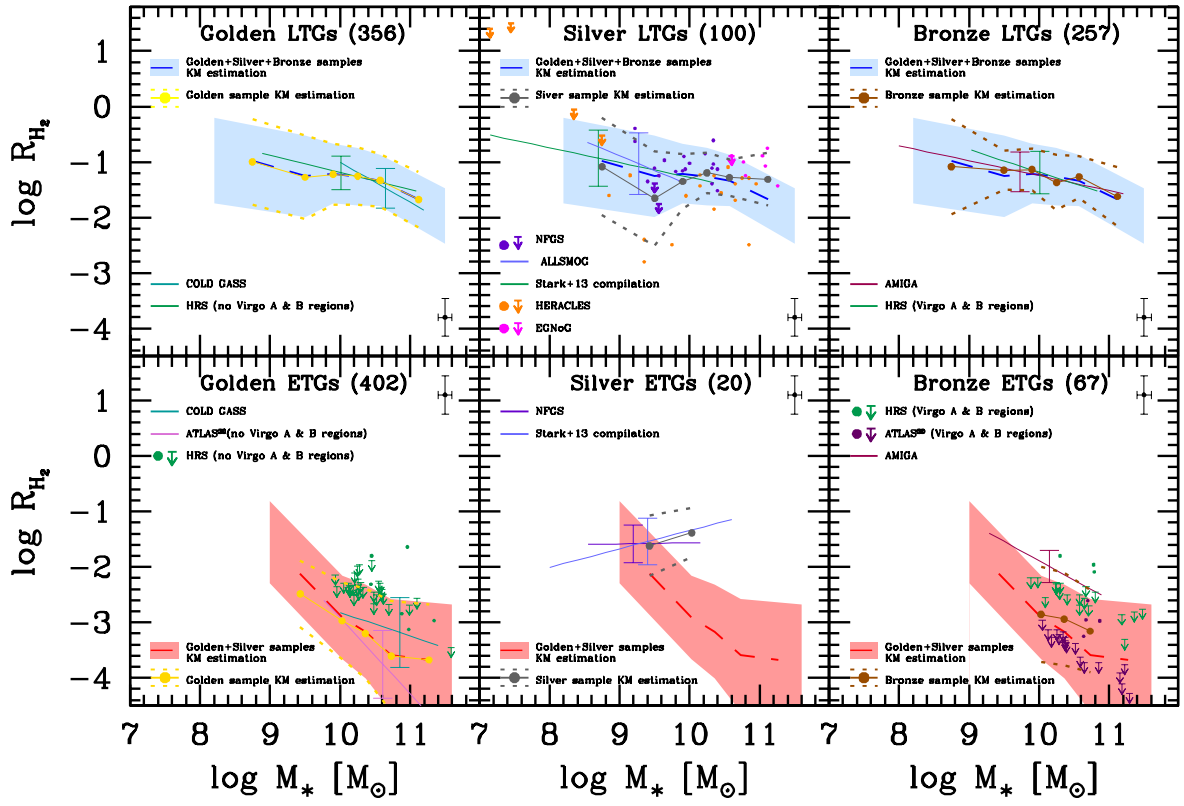

Fig. 4. Same as Figure 3 but for the molecular gas-to-stellar mass ratio. The color figure can be viewed online.

slightly larger than those from the overall LTG sample (compare the solid yellow and dashed blue lines), but still well within the 1σ scatter (shaded area). For ETGs, the deviations of the linear regressions of the Golden samples among themselves, and compared to all ETGs, are within the respective 1σ scatters, which are actually large. If no corrections to the upper limits of the GASS and ATLAS^{3D} were applied, then the regression for the former would be significantly above the regression for the latter. Summarizing, the Golden samples of LTGs and ETGs do not show particular shifts in their respective $R_{\rm H_2}$– M_* correlations. Therefore, the combination of them is expected to provide reliable information for determining the respective $R_{\rm H_2}$– M_* correlations: for LTGs in the $\approx 10^{8.5} - 10^{11.5}$ $M_\odot$ mass range, and for ETGs, only for $M_* \gtrsim 10^{10}$ $M_\odot$.

Silver category: The LTG samples present a dispersed distribution in the log$R_{\rm H_2}$–logM_* plane but well within the 1σ scatter of the overall sample (shaded area). The mean values in mass bins from samples of the Silver category are in reasonable agreement with the mean values from all the samples (compare the gray solid and blue dashed lines). Therefore, the Silver samples, though scattered and not complete in any sense, seem not to exhibit a clear systematical shift in their H_2 content. We include these samples to infer the $R_{\rm H_2}$-M_* correlation of LTGs. For ETGs, the two Silver samples provide information for masses below $M_* \approx 10^{10}$ $M_\odot$, and both are consistent with each other. Therefore, we include these samples to infer the ETG $R_{\rm H_2}$-M_* correlation down to $M_* \approx 10^{8.5}$ $M_\odot$.

Bronze category and the effects of environment: The isolated (from the AMIGA sample) and Virgo central (from the HRS catalog) LTGs have H_2 contents similar to the mean in different mass bins of all the galaxies. If any, the Virgo LTGs have on average slightly higher values of $R_{\rm H_2}$ than the isolated LTGs, especially at masses smaller than $M_* \approx 10^{10}$ $M_\odot$. Given that LTGs in extreme environments do not segregate from the average $R_{\rm H_2}$ values at different masses of all galaxies, we include them for calculating the $R_{\rm H_2}$–M_* correlation of LTGs. For ETGs, the AMIGA isolated galaxies have on average values of $R_{\rm H_2}$ significantly higher than the mean of other galaxies, while those ETGs from the Virgo central regions (from HRS and ATLAS^{3D}; mostly upper limits), seem to be on average consistent with the mean of all the galaxies, though the

scatter is large. Given the strong deviation of isolated ETGs from the mean trend, we prefer to exclude galaxies from the Bronze category to determine the ETG R_{H_2}–M_* correlation. We conclude that the H_2 content of LTGs is weakly dependent on the environment of galaxies, but in the case of ETGs, very isolated galaxies have systematically higher R_{H_2} values than galaxies in more dense environments.

4. THE GAS-TO-STELLAR MASS CORRELATIONS OF THE TWO MAIN GALAXY POPULATIONS

4.1. *Strategy for Constraining the Correlations*

In spite of the diversity in the compiled samples and their different selection functions, the exploration presented in the previous section shows that the HI and H_2 contents as a function of M_* for most of the samples compiled here do not segregate significantly among them. The exception are the Bronze samples for ETGs. Therefore, *the Bronze ETGs are excluded from our analysis.* The strong segregation is actually by morphology (or color, or star formation rate), and this is why we have separated from the beginning the compiled data into two broad galaxy groups, LTGs and ETGs.

To determine gas-to-stellar mass ratios as a function of M_* we need (1) to take into account the upper limits of undetected galaxies in radio, and (2) to evaluate the correlation independently of the number of data points in each mass bin. If we have many data points in some mass bins and only a few ones in other mass bins (as would happen if we use, for instance, a mass-limited volume-complete sample, with many more data points at smaller masses than at larger masses), then the overall correlation of R_{HI} or R_{H_2} with M_* would be dominated by the former, probably giving incorrect values of R_{HI} or R_{H_2} at other masses. In view of these two requirements, our strategy to determine the $\log R_{HI}$–$\log M_*$ and $\log R_{H_2}$–$\log M_*$ correlations is as follows:

1. Calculate the logarithmic means and standard deviations (scatter) in stellar mass bins obtained from the compiled data taking into account the non-detections (upper limits) by means of the Kaplan-Meier estimator.

2. Obtain an estimate of the intrinsic standard deviations (scatter), taking into account estimates of the observational errors.

3. Propose a function to describe the relation given by the mean and intrinsic scatter as a function of mass (e.g., a single or double power law).

4. Constrain the parameters of this function by performing a formal fit to the mean and scatter calculated for each mass bin; note that in this case *the fitting gives the same weight to each mass bin, irrespective of the number of galaxies in each bin.*

4.2. *The* HI*-to-Stellar Mass Correlations*

In the upper left panel of Figure 5, along with the data from the Golden, Silver, and Bronze LTG samples, the mean and standard deviation (squares and black error bars) calculated for each mass bin with the Kaplan-Meier method are plotted. In the lower left panel, the same is plotted but for the Golden and Silver ETG samples (recall that the Bronze samples are excluded in this case). We see that the total standard deviations in $\log R_{HI}$, σ_{dat}, do not evidence a systematical dependence on mass both for LTGs and ETGs. Then, we can use a constant value for each case. For LTGs, the standard deviations have values around 0.45–0.65 dex with an average of $\sigma_{dat} \approx 0.53$ dex. For ETGs, the standard deviations are much larger and more dispersed than for LTGs (see § 4.4 below for a discussion on why this could be). We assume an average value of $\sigma_{dat} = 1$ dex for ETGs.

The *intrinsic* standard deviation (scatter) can be estimated as $\sigma^2_{intr} \approx \sigma^2_{dat} - \sigma^2_{err}$ (this is valid for normal distributions), where σ_{err} is the mean statistical error in the $\log R_{HI}$ determination due to the observational uncertainties. In Appendix E we present an estimate of this error, $\sigma_{err} \approx 0.14$ dex. Therefore, $\sigma_{intr} \approx 0.52$ and 0.99 dex for LTGs and ETGs, respectively. These estimates should be taken only as indicative values given the assumptions and rough approximations involved in their calculations. For example, we will see in § 5 that the distributions of $\log R_{HI}$ (detections and non-detections) in different mass bins tend to deviate from a normal distribution, in particular for ETGs

Next, we propose that the HI-to-stellar mass relations can be described by the general function:

$$y(M_*) = \frac{C}{\left(\frac{M_*}{M_*^{tr}}\right)^a + \left(\frac{M_*}{M_*^{tr}}\right)^b} \quad (1)$$

where $y = R_{HI}$, C is the normalization factor, a and b are the low- and high-mass slopes of the function and M_*^{tr} is the transition mass. This function is continuous and differentiable. If $a = b$, then equation (1) describes a single power law, or a linear relation in logarithmic scales. In this case, the equation remains as $y(M_*) = C'(M_*/M_\odot)^{-a}$. For $a \neq b$, the function corresponds to a double power law.

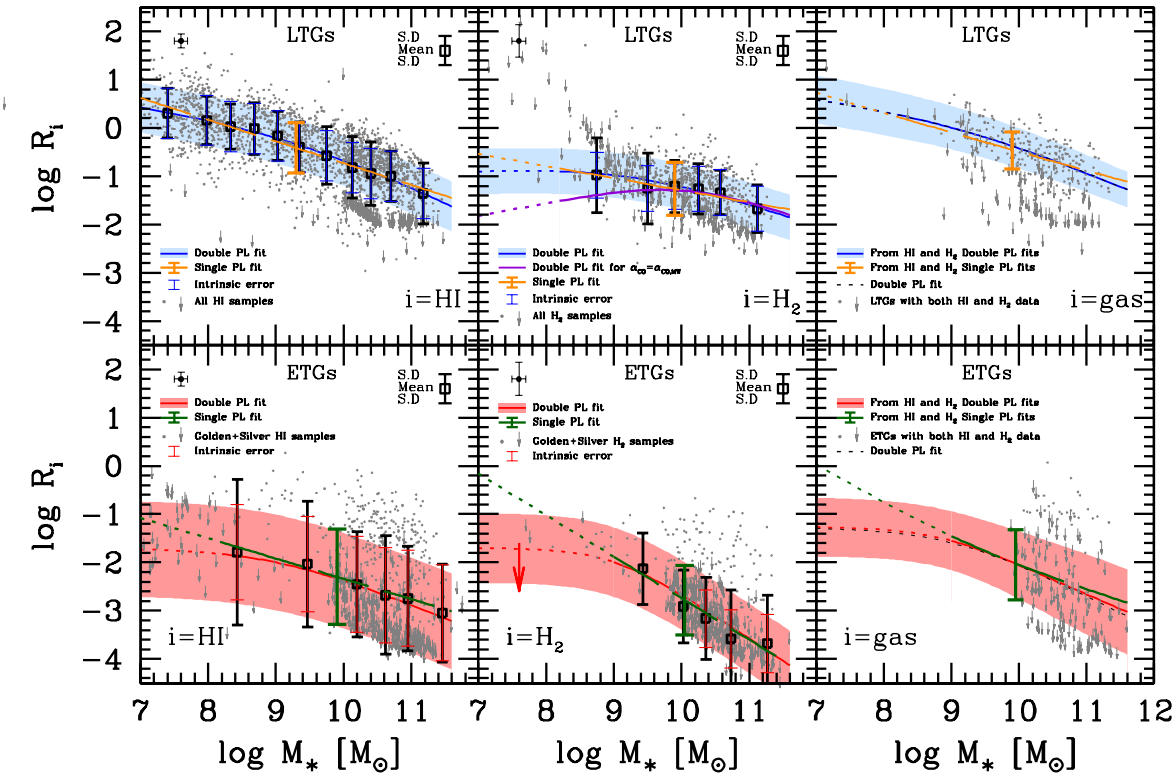

Fig. 5. *Left panels:* The $R_{\rm HI}$–M_* correlation for LTGs (upper panel) and ETGs (lower panel). Dots are detections and arrows are upper limits for non-detections (for ETGs the Bronze sample was excluded). The squares and error bars show the mean and standard deviation in different mass bins calculated by means of the Kaplan-Meier estimator for censored and uncensored data. The thin error bars correspond to our estimate of the *intrinsic* scatter after taking into account the observational errors (shown in the panel corners). The solid and long-dashed lines in each panel are respectively the best double- and single-power law fits. The shaded areas show the intrinsic scatter; to avoid overcrowding, for the single power-law fit, the intrinsic scatter is plotted only at one point. The dotted lines are extrapolations of the correlations to low masses, where the data are scarce and dominated by upper limits. *Middle panels:* Same as in the left panels but for $R_{\rm H_2}$. For the ETG population, the double power-law fit was performed with the conservative constraint that below $M_* = 10^9$ $M_\odot$, the low-mass slope is zero. *Right panels:* The $R_{\rm gas}$–M_* correlations for LTGs and ETGs as calculated from combining the respective double- and single-power law $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* correlations and taking into account helium and metals (see text). The shaded area and error bar are the (1σ) intrinsic scatter obtained by error propagation of the intrinsic scatter around the corresponding $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* relations. For completeness, the data from our compilation that have determinations of both HI and H_2 masses are also plotted (the obtained correlations are not fits to these data). Dotted lines are extrapolations of the inferred relations to smaller masses. The short dashed lines show the best fits using the double power-law function. The color figure can be viewed online.

We fit the logarithm of function, equation (1), to the mean values of $\log R_{\rm HI}$ as a function of mass (squares in the left panels of Figure 5) with the corresponding (constant) intrinsic standard deviation as estimated above (thin blue/red error bars). For LTGs, the fit is carried out in the range $7.3 \lesssim \log(M_*/M_\odot) \lesssim 11.2$, and for ETGs in the range $8.5 \lesssim \log(M_*/M_\odot) \lesssim 11.5$. The Levenberg-Marquardt method is used for the fit (Press et al. 1996). First, we perform the fits to the binned LTG and ETG data using a single power law, i.e., we fix $a = b$. The dashed orange and green lines with an error bar in the left panels of Figure 5 show the results. The fit parameters are given in Table 5. We note that *these fits and those of the Buckley-James linear regression for all the data (not binned) in logarithm are very similar.*

Then, we fit to the binned data the logarithm of the double power-law function given in equation (1). The corresponding best-fit parameters are presented in Table 6. We note that the fits are almost the same if the total mean standard deviation, $\sigma_{\rm dat}$, is used instead of the intrinsic one. The reduced $\chi^2_{\rm red}$ are 0.01 and 0.03, respectively. The fits are actually

TABLE 5

BEST FIT PARAMETERS TO THE SINGLE POWER LAW (EQUATION 1, $a = b$)

	$\log C'$	a	$\sigma_{\rm dat}$	$\sigma_{\rm intr}$
		$R_{\rm HI}$-M_*		
LTG	3.77 ± 0.22	-0.45 ± 0.02	0.53	0.52
ETG	1.88 ± 0.33	-0.42 ± 0.03	1.00	0.99
ETG$^{\rm ndc}$	1.34 ± 0.46	-0.37 ± 0.05	1.35	1.34
		$R_{\rm H_2}$-M_*		
LTG	1.21 ± 0.53	-0.25 ± 0.05	0.58	0.47
ETG	5.86 ± 1.45	-0.86 ± 0.14	0.80	0.72
ETG$^{\rm ndc}$	5.27 ± 1.78	-0.80 ± 0.17	0.95	0.88
		$R_{\rm gas}$-M_*		
LTG	4.76 ± 0.05	-0.52 ± 0.03	–	0.44
ETG	3.70 ± 0.07	-0.58 ± 0.01	–	0.68

• The suffix "ndc" indicates that for the ETG correlations, no distance correction was applied to the upper limits in the (COLD) GASS samples.
• $\sigma_{\rm dat}$ and $\sigma_{\rm intr}$ are given in dex.

TABLE 6

BEST FIT PARAMETERS TO THE DOUBLE POWER LAW (EQUATION 1, $a \neq b$)

	C	a	b	$\log(M_*^{\rm tr}/M_\odot)$	$\sigma_{\rm dat}$	$\sigma_{\rm intr}$
			$R_{\rm HI}$-M_*			
LTG	0.98 ± 0.06	0.21 ± 0.04	0.67 ± 0.03	9.24 ± 0.04	0.53	0.52
ETG	0.02 ± 0.01	0.00 ± 0.15	0.58 ± 0.03	9.00 ± 0.30	1.00	0.99
ETG$^{\rm ndc}$	0.02 ± 0.01	0.00 ± 0.55	0.51 ± 0.05	9.00 ± 0.60	1.35	1.34
			$R_{\rm H_2}$-M_*			
LTG	0.19 ± 0.02	-0.07 ± 0.18	0.47 ± 0.04	9.24 ± 0.12	0.58	0.47
ETG	0.02 ± 0.01	0.00 ± 0.00	0.94 ± 0.15	9.01 ± 0.12	0.80	0.72
ETG$^{\rm ndc}$	0.02 ± 0.03	0.00 ± 0.00	0.88 ± 0.18	9.01 ± 0.15	0.95	0.88
			$R_{\rm gas}$-M_*			
LTG	1.69 ± 0.02	0.18 ± 0.01	0.61 ± 0.02	9.20 ± 0.04	–	0.44
ETG	0.05 ± 0.02	0.01 ± 0.03	0.70 ± 0.01	9.02 ± 0.05	–	0.68

• The suffix "ndc" indicates that for the ETG correlations, no distance correction was applied to the upper limits in the (COLD) GASS samples.
• $\sigma_{\rm dat}$ and $\sigma_{\rm intr}$ are given in dex.

performed to a low number of points (the number of mass bins) with large error bars; this is why the $\chi^2_{\rm red}$ are smaller than 1. Note, however, that the error bars are not related to measurement uncertainties but correspond to the population scatter of the data. Therefore, in this case $\chi^2_{\rm red} < 1$ implies that while the best fit is good, other fits could be also good within the scatter of the correlations. In the case of the single power-law fits, the $\chi^2_{\rm red}$ are 0.03 and 0.01, respectively for LTG and ETG.

The double power-law $R_{\rm HI}$–M_* relations and the estimated intrinsic (1σ) scatter for the LTG (ETG) population are plotted in the left upper (lower) panel of Figure 5 with solid lines and shaded areas, respectively. From the fits, we find for LTGs a transition mass $M_*^{\rm tr} = 1.74 \times 10^9\ M_\odot$, with $R_{\rm HI} \propto {M_*}^{-0.21}$ and ${M_*}^{-0.67}$ at masses much smaller and larger than this, respectively. For ETGs, $M_*^{\rm tr} = 1 \times 10^9\ M_\odot$, and $R_{\rm HI} \propto {M_*}^{0.0}$ and ${M_*}^{-0.58}$, at masses much smaller and larger than this, respectively.

Both the double and single power laws describe well the HI-to-stellar mass correlations. However, the former could be more adequate than the latter. In Figure 1 we plot the Buckley-James linear regressions to the $R_{\rm HI}$ vs. M_* data for the low- and high-mass regions (below and above $\log(M_*/M_\odot) \approx 9.7$; for ETGs the regression is applied only for masses above $10^8\ M_\odot$); the dotted lines show the extrapolation of the fits. The slope at low masses for LTGs, −0.36, is shallower than the one at high masses, −0.55. For ETGs, there is even evidence of a change in the slope sign at low masses. A flattening of the overall (late + early type galaxies) correlation at low masses has been also suggested by Baldry et al. (2008), who have used the empirical mass–metallicity relation coupled with a metallicity-to-gas mass fraction relation (which can be derived from a simple chemical evolution model) to obtain a gas-to-stellar mass correlation in a large mass range. Another evidence that at low masses the $R_{\rm HI}$–M_* relation flattens is shown in the work by Maddox et al. (2015) already mentioned (see also Huang et al. 2012a). While the sample used by these authors does not allow to infer the $R_{\rm HI}$–M_* correlation of galaxies due to its bias towards high $R_{\rm HI}$ values (see above), the upper envelope of this correlation can be actually constrained; the high-$R_{\rm HI}$ envelope does not suffer from selection limit effects. As seen for the data from Maddox et al. (2015) reproduced in the left upper panel of our Figure 1, this envelope tends to flatten at $M_* \lesssim 2 \times 10^9\ M_\odot$,[10] which suggests (but does not demonstrate) that the mean relation can also exhibit such a flattening. Another piece of evidence in favor of the flattening can be found in Huang et al. (2012b), and more recently in Bradford et al. (2015) for their sample of low-mass galaxies combined with larger mass galaxies from the ALFALFA survey.

4.3. *The* H_2*-to-Stellar Mass Correlations*

In the upper middle panel of Figure 5, along with the data from the Golden, Silver, and Bronze LTG samples, the mean and standard deviation (error

[10] In Huang et al. (2014), the $SDSS - GALEX - \alpha.40$ common sample was weighted by $V/V_{\rm max}$ to correct for incompleteness and mimic then the scaling relations derived from a volume-limited sample. However, only galaxies with $M_{HI} \gtrsim 10^{8.2}\ M_\odot$ are included in their plot of $R_{\rm HI}$ vs. M_* (Figure 1); at lower masses, the correlation likely continues being biased to high values of $R_{\rm HI}$. Even that a weak flattening below $M_* \approx 10^9\ M_\odot$ is observed in their average curve.

bars) calculated in each mass bin with the Kaplan-Meier method are plotted. In the lower panel, the same is plotted but for the Golden and Silver ETG samples (recall that the Bronze samples are excluded in this case). The poor observational information at stellar masses smaller than $\approx 5 \times 10^8$ $M_\odot$ does not allow us to constrain the correlations at these masses, both for LTG and ETGs. Regarding the total standard deviations, for both LTGs and ETGs, they vary from mass bin to mass bin but without a clear trend. Then we can use a constant value for both cases. For LTGs, the total standard deviations have values around 0.5–0.8 dex with an average of $\sigma_{\rm dat} \approx 0.58$ dex. For ETGs, the average value is roughly 0.8 dex. As in the case of HI (previous subsection), we further estimate indicative values for the *intrinsic* population standard deviations (scatter). For this, we present in Appendix E an estimate of the the mean observational error of the $\log R_{\rm H_2}$ determination, $\sigma_{\rm err} \approx 0.34$ dex. Therefore, the estimated mean intrinsic scatters in $\log R_{\rm H_2}$ are $\sigma_{\rm intr} \approx 0.47$ and 0.72 dex for LTGs and ETGs, respectively. Given the assumptions and approximations involved in these estimates, they should be taken with caution. For example, we will see in § 5 that the distributions of $\log R_{\rm H_2}$ (detections and non-detections) in different mass bins tend to deviate from a normal distribution, in particular for the ETGs.

We fit the logarithm of function equation (1), $y = R_{\rm H_2}$, to the mean values of $\log R_{\rm H_2}$ as a function of mass (squares in the left panels of Figure 5) with their corresponding scatter as estimated above (thin blue/red error bars), assumed to be the individual standard deviations for the fit. Again, the Levenberg-Marquardt method is used to perform the fit. The fits extend only down to $M_* \approx 5 \times 10^8$ $M_\odot$. First, the fits are performed for a singe power law, i.e., we fix $a = b$. The dashed orange and green lines in the middle panels of Figure 5 show the results. The parameters of the fit and their standard deviations are given in Table 5. *The fits are very similar to those obtained using the Buckley-James linear regression to all (not binned) logarithmic data.*

Then, we fit the binned LTG and ETG data to the double power-law function equation (1). In the case of the ETG population, we impose an extra condition to the fit: that the slope of the relation at masses below $\approx 10^9$ $M_\odot$ be flat. The few data at these masses clearly show that $R_{\rm H_2}$ does not increase for smaller M_*; it is likely that it even decreases, so that our assumption of a flat slope is conservative. The corresponding best-fit parameters are presented in Table 6. As in the case of the $R_{\rm HI} - M_*$ correlations, the reduced $\chi^2_{\rm red}$ are smaller than 1 (0.04 and 0.10, respectively), which implies that while the best fits are good, other fits could describe reasonably well the scattered data. In the case of the single power-law fits, $\chi^2_{\rm red}$ were 0.04 and 0.07, respectively for LTG and ETG. The double power-law $R_{\rm H_2}$–M_* relations and their (1σ) *intrinsic* scatter for the LTG (ETG) population are plotted in the middle upper (lower) panel of Figure 5 with solid lines and shaded areas, respectively. We note that the fits are almost the same if the total mean standard deviation, $\sigma_{\rm dat}$, is used instead of the intrinsic one.

From these fits, we find for LTGs, $M_*^{\rm tr} = 1.74 \times 10^9$ $M_\odot$, with $R_{\rm H_2} \propto {M_*}^{-0.07}$ and ${M_*}^{-0.47}$ at much smaller and larger masses than this, respectively. For ETGs, $M_*^{\rm tr} = 1.02 \times 10^9$ $M_\odot$, with $R_{\rm H_2} \propto {M_*}^{0.00}$ and ${M_*}^{-0.94}$ at much smaller and larger masses than this, respectively. In the middle upper panel of Figure 5, we plot also the best double power-law fit to the $R_{\rm H_2}$–M_* correlation of LTGs when the $\alpha_{\rm CO}$ factor is assumed constant and equal to the MW value (purple dashed line).

Both the single and double power-law functions describe equally well the $R_{\rm H_2}$ –M_* correlations for the LTG and ETG population, but there is some evidence of a change of slope at small masses. In Figure 2, the Buckley-James linear regressions to the $R_{\rm H_2}$ vs. M_* data below and above $\log(M_*/M_\odot) \approx 9.7$ are plotted (in the former case the regressions are applied for masses only above 10^8 $M_\odot$); the dotted lines show the extrapolation of the fits. The slopes in the small mass range at low masses for LTGs/ETGs are shallower than those at high masses. Besides, in the case of ETGs, if the single power-law fit shown in Figure 5 is extrapolated to small masses, ETGs of $M_* \approx 10^7$ $M_\odot$ would be dominated in mass by H_2 gas. Red/passive dwarf spheroidals are not expected to contain significant fractions of molecular gas. Recently, Accurso et al. (2017) have also reported a flattening in the H_2-to-stellar mass correlation at stellar masses below $\approx 10^{10}$ $M_\odot$.

4.4. *The Cold Gas-to-Stellar Mass Correlations*

Combining the $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* relations presented above, we can obtain now the $R_{\rm gas}$–M_* relation, for both the LTG and ETG populations. Here, $R_{\rm gas} = M_{\rm gas}/M_* = 1.4(R_{\rm HI} + R_{\rm H_2})$, where $M_{\rm gas}$ is the galaxy cold gas mass, including helium and metals (the factor 1.4 accounts for these components). The intrinsic scatter around the gas-to-stellar mass relation can be estimated by propagating the intrinsic scatter around the HI- and H_2-to-stellar mass relations. Under the assumption of

null covariance, the logarithmic standard deviation around the $\log R_{\rm gas}$–$\log M_*$ relation is given by

$$\sigma_{\rm intr,R_{gas}} = \frac{1}{R_{\rm H_I}+R_{\rm H_2}}\left(R^2_{\rm H_I}\sigma^2_{\rm intr,R_{H_I}} + R^2_{\rm H_2}\sigma^2_{\rm intr,R_{H_2}}\right)^{\frac{1}{2}}. \quad (2)$$

The obtained cold gas-to-stellar mass correlations for the LTG and ETG populations are plotted in the right panels of Figure 5. The solid lines and shaded bands (intrinsic scatter given by the error propagation) were obtained from the double power-law correlations, while the solid green lines and the error bars were obtained from the single power-law correlations. For completeness, we also plot in Figure 5 those galaxies from our compilation that have determinations for *both* the HI and H_2 masses. Note that a large fraction of our compilation has no determinations for both quantities at the same time. We fit the results obtained for the single (double) power-law fits, taking into account the intrinsic scatter, to the logarithm of the single (double) power-law function given in equation (1) with $y = R_{\rm gas}$ and report in Table 5 (Table 6) the obtained parameters for both the LTGs and ETGs. The fits for the double power-law are shown as dotted lines in Figure 5. The standard deviations $\sigma_{\log R_{\rm gas}}$ change slightly with mass; we report an average value for them in Tables 5 and 6. Both for LTGs and ETGs, the mass at which the $R_{\rm gas}$–M_* correlations change slope is $M_*^{\rm tr} \approx 1.7\times10^9\ M_\odot$, *the mass that roughly separates dwarf from normal galaxies.*

According to Figure 5, the LTG and ETG $R_{\rm gas}$–M_* correlations are significantly different. The gas content in the former is at all masses larger than in the latter, the difference being maximal at the largest masses. For the LTG population, $M_{\rm gas} \approx M_*$ on average at $\log(M_*/M_\odot) \approx 9$, and at smaller masses, these galaxies are dominated by cold gas; at stellar masses around $2\times10^7\ M_\odot$, $M_{\rm gas}$ is on average three times larger than M_*. For ETGs, there is a hint that at $\approx 10^9\ M_\odot, R_{\rm gas}$ changes from increasing as M_* is smaller to decreasing at larger masses.

5. THE DISTRIBUTIONS OF THE SCATTER AROUND THE GAS-TO-STELLAR MASS RELATIONS

To determine the correlations presented above, we have made use only of the mean and standard deviation of the data in different mass bins. It is also of interest to learn about the scatter distributions around the main relations. Furthermore, in the next section we will require the full distributions of $R_{\rm HI}(M_*)$ and $R_{\rm H_2}(M_*)$ in order to generate a mock galaxy catalog through which the HI and H_2 mass functions will be calculated. The Kaplan-Meier estimator provides information to construct the probability density function (PDF) at a given stellar mass including the uncensored data. By using these PDFs we explore the distribution of the $R_{\rm HI}$ and $R_{\rm H_2}$ data (detections + upper limits). Given the heterogeneous nature of our compiled data, these "scatter" distributions should be taken just as a rough approximation. However, when the uncensored data dominate (this happens in most of the mass bins for the ETG samples), the Kaplan-Meier estimator cannot predict well the distribution of the uncensored data.

Late-type galaxies.- Figures 6 and 7 present the $R_{\rm HI}$ and $R_{\rm H_2}$ PDFs in different M_* bins for LTGs. Based on the bivariate HI and stellar mass function analysis of Lemonias et al. (2013), who used the GASS sample for (all-type) massive galaxies, we propose that the PDFs of $R_{\rm HI}$ and $R_{\rm H_2}$ for LTGs can be described by a Schechter (Sch) function (equation 3 below; x denotes either $R_{\rm HI}$ or $R_{\rm H_2}$). By fitting this function to the $R_{\rm HI}$ data in each stellar mass bin we find that the power-law index α weakly depends on M_*, with most of the values being around -0.15 (see also Lemonias et al. 2013), while the break parameter x^* varies with M_*. A similar behavior was found for $R_{\rm H_2}$, with most of the values of α around -0.10. We then perform for each case ($R_{\rm HI}$ and $R_{\rm H_2}$) a continuous fit across the range of stellar-mass bins rather than fits within independent bins. The general function proposed to describe the $R_{\rm HI}$ and $R_{\rm H_2}$ PDFs of LTGs, at a fixed M_* and within the range $\log x \pm d\log x/2$, is:

$$P_{\rm Sch}(x|M_*) = \frac{\phi^*}{\log e}\left(\frac{x}{x^*}\right)^{\alpha+1}\exp\left(-\frac{x}{x^*}\right), \quad (3)$$

with the normalization condition, $\phi^* = 1/\Gamma(1+\alpha)$, where Γ is the complete gamma function, which guarantees that the integration over the full space in x is 1. The parameters α and x^* depend on M_*. We propose the following functions for these dependences:

$$\alpha(M_*) = c + d\log M_*, \quad (4)$$

and

$$x^*(M_*) = \frac{x_0}{\left(\frac{M_*}{m_{\rm tr}}\right)^e + \left(\frac{M_*}{m_{\rm tr}}\right)^f}. \quad (5)$$

The parameters $c, d, x_0, m_{\rm tr}, e$, and f are constrained from a *continuous fit across all the mass bins* using a Markov Chain Monte Carlo method, following Rodríguez-Puebla et al. (2013). Since the stellar mass bins from the data have a finite width, for a more precise determination we convolve the

TABLE 7

BEST FIT PARAMETERS TO THE FULL DISTRIBUTIONS

	c	d	x_0	$\log(m_*^{\rm tr}/M_\odot)$	e	f	g	h	i	j
$P(R_{\rm HI}\|M_*)$ distributions										
LTG	1.11±0.35	-0.11±0.04	2.45±0.76	8.77±0.45	0.002±0.10	0.61±0.07	–	–	–	–
ETG	-0.42±0.80	-0.02±0.08	2.15±0.55	8.30±0.38	-0.43±1.10	0.52±0.09	-0.22±0.37	0.07±0.04	-1.62±1.08	-0.13±0.11
$P(R_{\rm H_2}\|M_*)$ distributions										
LTG	0.70±1.28	-0.07±0.13	0.15±0.03	10.37±0.31	0.19±0.17	0.19±0.16	–	–	–	–
ETG	-0.52±1.19	-0.01±0.11	0.71±0.27	7.90±1.09	0.42±0.50	0.21±0.28	0.24±0.97	0.04±0.09	5.74±3.17	-0.86±0.29

For LTGs the distributions are given by equation (3), for ETGs, by equation (7).

PDF with the GSMF over a given bin. Therefore, the PDF of x averaged within the bin $\Delta M_* = [{\rm M}_{*1}, {\rm M}_{*2}]$ is:

$$\langle P_{\rm Sch}(x|\Delta M_*)\rangle = \frac{\int_{{\rm M}_{*1}}^{{\rm M}_{*2}} P_{\rm Sch}(x|M_*)\Phi_{late}(M_*)dM_*}{\int_{{\rm M}_{*1}}^{{\rm M}_{*2}} \Phi_{late}(M_*)dM_*}, \quad (6)$$

where $\Phi_{late}(M_*)$ is the GSMF for LTGs (see § 6). The constrained parameters are reported in Table 7. The obtained mass-dependent PDFs are plotted in the panels of Figures 6 and 7. The solid blue line corresponds to the number density-weighted distribution within the given mass bin (equation 6), while the red dotted line shows the function equation (3) evaluated at the mass corresponding to the logarithmic center of each bin. As seen, the Kaplan-Meier PDFs obtained from the data (gray histograms) are well described by the proposed Schechter function averaged within the different mass bins (blue lines), both for $R_{\rm HI}$ and $R_{\rm H_2}$.

Early-type galaxies.- We present the $R_{\rm HI}$ and $R_{\rm H_2}$ PDFs for ETGs in Figures 8 and 9, respectively. The distributions are very extended, implying a large scatter in the $R_{\rm H_2}$–M_* correlations, as discussed in § 4.2 and 4.3.[11] The distributions seem to be bimodal, with a significant fraction of ETGs having gas fractions around a low limit ($\approx 10^{-4}$) and the remaining galaxies with larger gas fractions, following an asymmetrical distribution. The low limit is given by the Kaplan-Meier estimator and it is associated with the reported upper limits of non-detections. We should keep in mind that when non-detections dominate, the Kaplan-Meier estimator cannot provide a reliable PDF at the low end of the distribution. From a physical point of view, we know that ETGs are in general quiescent galaxies that likely exhausted their cold gas reservoirs and did not accrete more gas. However, small amounts of gas can still be available from the winds of old/intermediate-age stars. For instance, Sun-like stars can lose $\approx 10^{-4} - 10^{-5}$ of their masses in 1 Gyr; more massive stars lose higher fractions. A fraction of the ejected material is expected to cool efficiently and ends up as HI and/or H_2 gas. However, those ETGs that have larger fractions of cold gas could acquire it by radiative cooling from their hot halos, or by accretion from the cosmic web, and/or by accretion from recent mergers (see for a discussion Lagos et al. 2014, and references therein). The amount of gas acquired depends on the halo mass, the environment, the gas mass of the colliding galaxy, etc. The range of possibilities is large; hence the scatter around the ETG $R_{\rm HI} - M_*$ and $R_{\rm H_2}$–M_* relations is expected to be large, as semi-analytic models show (Lagos et al. 2014).

To describe the PDFs seen in Figures 8 and 9, we propose a (broken) Schechter function plus a uniform distribution. The value of $R_{\rm HI}$ or $R_{\rm H_2}$ where the Schechter function breaks and the uniform distribution starts, x_2, seems to depend on M_* (see Figures 8 and 9). The lowest values where the distributions end, x_1, are not well determined by the Kaplan-Meier estimator, as mentioned above. To avoid unnecessary sophistication, we just fix x_1 as one tenth of x_2. This implies physical lowest values for $R_{\rm HI}$ and $R_{\rm H_2}$ of $10^{-4\div-5}$, which are plaussible according to our discussion above. The value of the Schechter parameter α shows a weak dependence on M_* for both HI and H_2. However, the fraction of galaxies between x_1 and x_2, F, seems to depend on M_*. For the uniform distribution, this fraction is given by $F = P(< x_2|M_*) - P(< x_1|M_*) = \int_{x_1}^{x_2} C d\log x$, where $C = F/(\log x_2 - \log x_1)$; given our assumption of $\log x_2 - \log x_1 = 1$ dex, then $C = F(M_*)$. We parametrize all these dependences on M_* and perform a continuous fit across the range of stellar-mass

[11] Given this large scatter, previous works, for small samples of massive galaxies, have suggested that red or early-type galaxies do not follow a defined correlation between M_{HI} and M_* (or luminosity; e.g., Welch et al. 2010; Serra et al. 2012) and between M_{H_2} and M_* (e.g., Saintonge et al. 2011; Lisenfeld et al. 2011; Young et al. 2011).

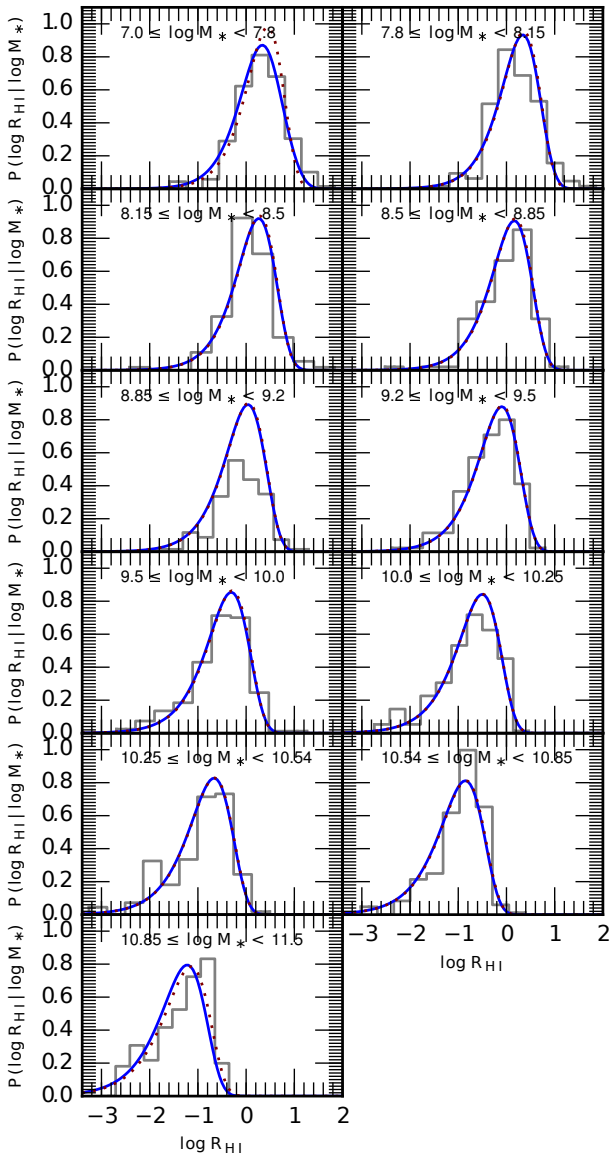

Fig. 6. Distributions (PDFs) of the LTG HI-to-stellar mass ratios in different stellar mass bins (indicated inside the panels). The gray histograms show results from the Kaplan-Meier estimator applied to the data (detections + upper limits), and the solid blue line corresponds to the best fitted number density-weighted distribution within the given mass bin (equation 6); the constrained parameters of the mass-dependent PDF (equation 3) are given in Table 7. The red dotted line shows the constrained function equation (3) evaluated at the mass corresponding to the logarithmic center of each mass bin. The color figure can be viewed online.

bins, both for the $R_{\rm HI}$ and $R_{\rm H_2}$ data. The general function proposed to describe the PDFs of ETGs as a function of M_* within the range $\log x \pm d\log x/2$ is the sum of a Schechter function, $P_{\rm Sch}(x|M_*)$,

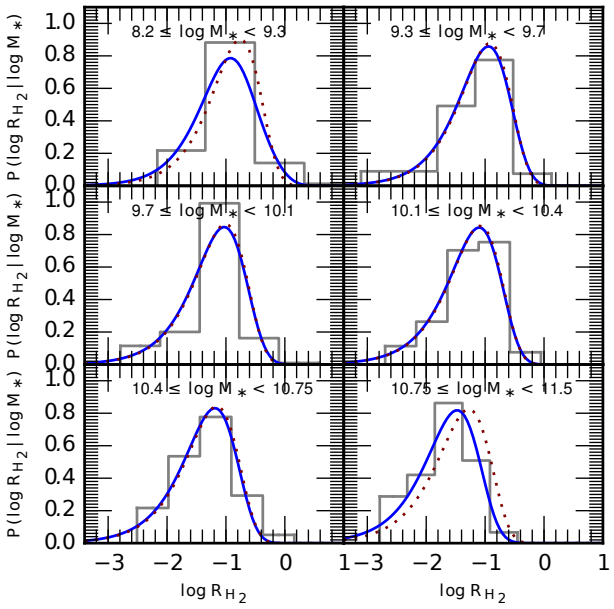

Fig. 7. Same as Figure 6 but for the H_2-to-stellar mass ratios. The color figure can be viewed online.

and a uniform function in x but dependent on M_*, $C = F(M_*)$:

$$F(M_*) = g + h\log M_*, \quad x_1 \le x < x_2(M_*), \qquad (7)$$
$$x_2(M_*) = i + j\log M_*,$$
$$P_{\rm Sch}(x|M_*), \quad x \ge x_2(M_*),$$

where the parameters x^* and α in $P_{\rm Sch}(x|M_*)$ are described by equation (3) with the normalization condition $\phi^* = (1-F)/\Gamma(1+\alpha)$, and $\log x_1 = \log x_2 - 1$. The parameters x_0, $m_{\rm tr}$, e, and f of the broken Schechter function and the parameters g, h, i, and j of the uniform distribution are constrained as described for LTGs above, from a continuous fit accross all the mass bins using the number density-weighted PDFs at each stellar mass bin:

$$\langle P_{\rm Sch}(x|\Delta M_*) + C\rangle = \frac{\int_{\rm M_{*1}}^{\rm M_{*2}} (P_{\rm Sch}(x|M_*) + C)\cdot \Phi_{early}(M_*)dM_*}{\int_{\rm M_{*1}}^{\rm M_{*2}} \Phi_{early}(M_*)dM_*}, \qquad (8)$$

where $\Phi_{early}(M_*)$ is the GSMF for ETGs (see § 6). The constrained parameters are reported in Table 7, both for $R_{\rm HI}$ and $R_{\rm H_2}$. The mass-dependent distribution function obtained is plotted in the panels of Figures 8 and 9 The solid red line corresponds to the number density-weighted distribution within

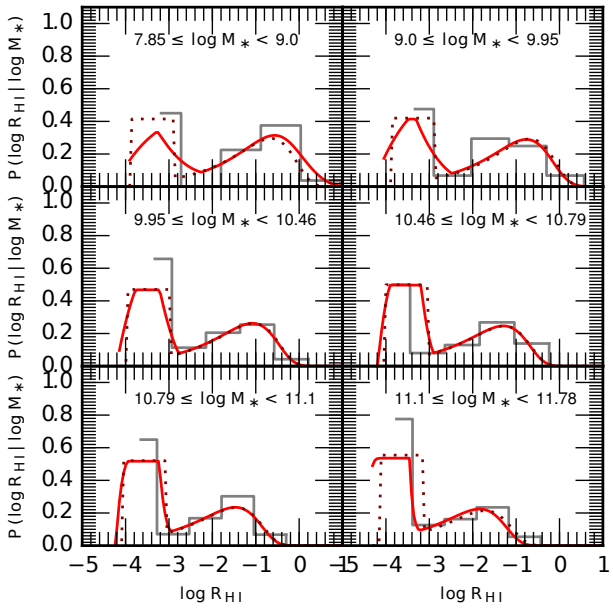

Fig. 8. Distributions (PDFs) of the ETG HI-to-stellar mass ratios in different stellar mass bins (indicated inside the panels). The gray histograms show results from the Kaplan-Meier estimator applied to the data (detections + upper limits), and the solid red line corresponds to the best fitted number density-weighted distribution within the given mass bin (equation 8); the constrained parameters of the mass-dependent PDF (equation 7) are given in Table 7. The red dotted line shows the constrained function equation (7) evaluated at the mass corresponding to the logarithmic center of each mass bin. The color figure can be viewed online.

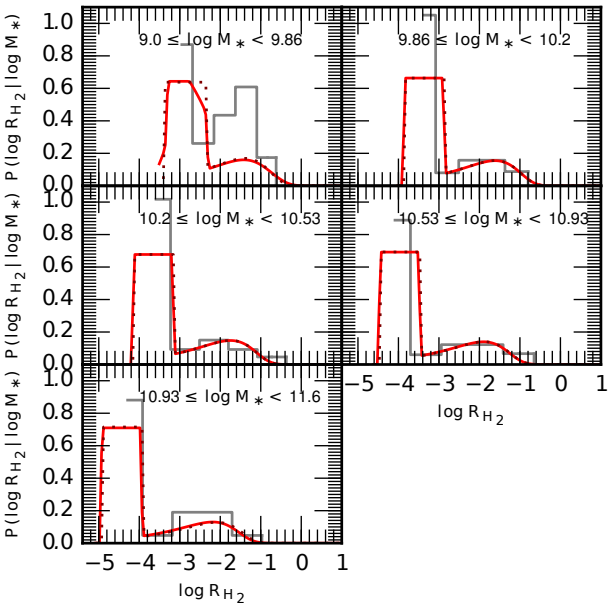

Fig. 9. Same as Figure 8 but for the H_2-to-stellar mass ratios. The color figure can be viewed online.

the given mass bin (equation 8), while the red dotted line shoes the proposed broken Schechter + uniform function evaluated at the mass corresponding to the logarithmic center of each bin. As seen, the Kaplan-Meier PDFs obtained from the data (gray histograms) are reasonably well described by the proposed function (equation 7) averaged within the different mass bins (red lines), both for $R_{\rm HI}$ and $R_{\rm H_2}$.

Finally, in Figures 10 and 11 we reproduce from Figure 5 the means and standard deviations obtained with the Kaplan-Meier estimator in different M_* bins (gray dots and error bars) for LTG and ETGs, respectively, and compare them with the means and standard deviations of the general mass-dependent distributions functions given in equations (3) and (7) and constrained with the data (black solid line and the two dotted lines surrounding it). The agreement is rather good in the log-log $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* diagrams both for LTGs and ETGs. Black dashed lines are extrapolations of the mean and standard deviation inferences from the distributions mentioned above, assuming they are the same as in the last mass bin with available gas observations. We also plot in these figures the respective mean double power-law relations determined in § 4.2 and 4.3 (dashed blue or red lines, for LTGs and ETGs respectively; dotted blue or red lines are extrapolations.).

In conclusion, the $R_{\rm HI}$ and $R_{\rm H_2}$ distributions as a function of M_* described by equations (3) and (7) (with the parameters given in Table 7) for LTGs and ETGs, respectively, are fully consistent with the corresponding $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* correlations determined in § 4.2 and 4.3. Therefore, *equations (3) and (7) provide a consistent description of the* HI- *and* H_2*-to-stellar mass relations and their scatter distributions for LTGs and ETGs, respectively.*

6. CONSISTENCY OF THE GAS-TO-STELLAR MASS CORRELATIONS WITH THE OBSERVED GALAXY GAS MASS FUNCTIONS

The HI- and H_2-to-stellar mass relations can be used to map the observed GSMF into the HI and H_2 mass functions (GHIMF and GH_2MF, respectively). In this way, we can check whether the correlations we have inferred from observations in § 4.2 and 4.3 are consistent with the GHIMF and GH_2MF obtained from HI and CO (H_2) surveys, respectively. In order to carry out this check of consistency, we need, on the one hand, a GSMF defined in a volume large enough to include massive galaxies and to minimize cosmic variance, and on the other hand, complete down to very small masses. As a first approximation

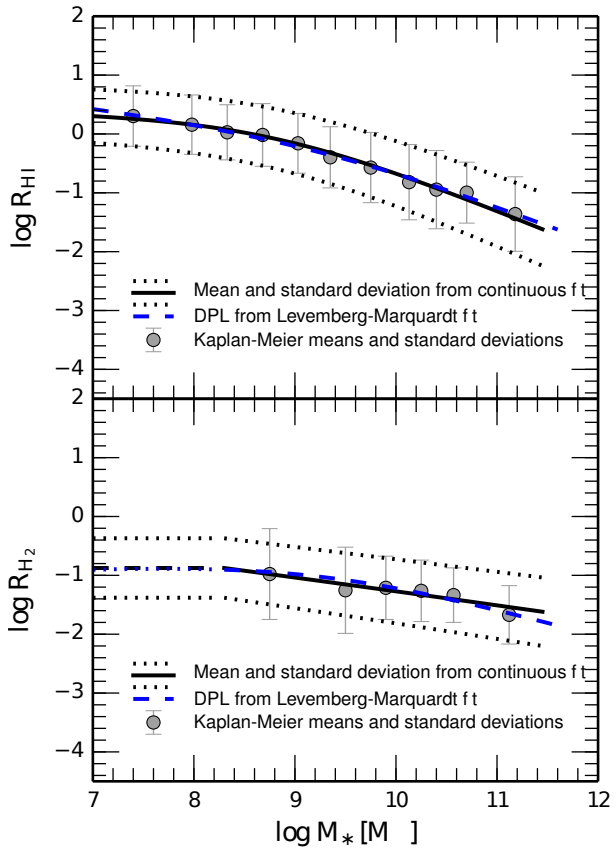

Fig. 10. Mean and standard deviation as a function of stellar mass (solid and dotted black lines) from the distributions of $R_{\rm HI}$ (upper panel) and $R_{\rm H_2}$ (lower panel) for LTGs as given by equation (3) (see Table 7 for the constrained parameters). When the data are insufficient at small masses, the distributions are assumed to be the same as in the last mass bin (dashed black lines). The gray dots with error bars are the mean and standard deviation obtained with the Kaplan-Meier estimator applied to the data (detections + upper limits) in different mass bins, as shown in Figure 5. The double-power law fits to these data as reported in § 4 are reproduced with the blue dashed lines (the blue dotted lines are extrapolations of these fits). The color figure can be viewed online.

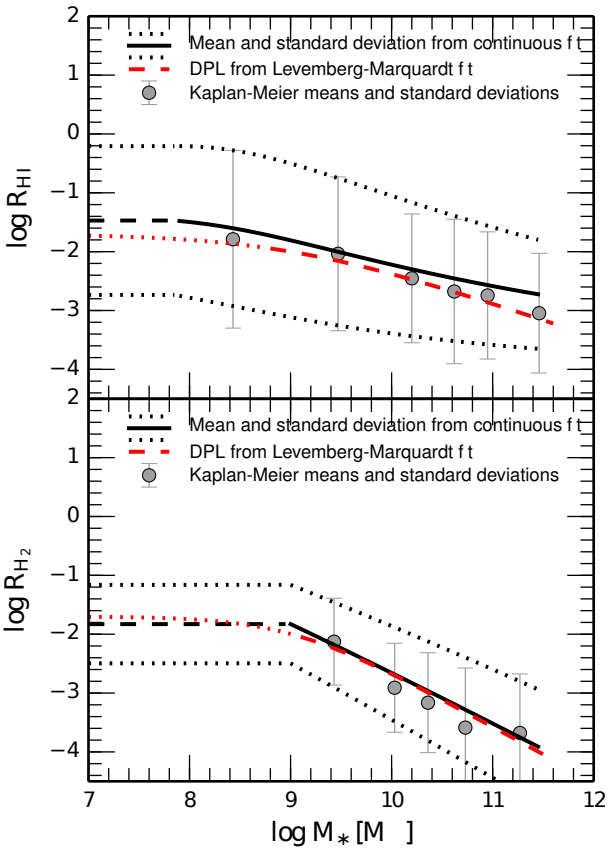

Fig. 11. Same as in Figure 10 but for ETGs. The color figure can be viewed online.

to obtain this GSMF, we follow here a procedure similar to Kravtsov et al. (2014, see their Appendix A). We use the combination of two GSMFs: Bernardi et al. (2013) for the large SDSS volume (complete from $M_* \approx 10^9\ M_\odot$), and Baldry et al. (2012) for a local small volume but nearly complete down to $M_* \approx 10^7\ M_\odot$ (GAMA). In Appendix F we describe how we apply some corrections and homogenize both samples to obtain an uniform GSMF from $M_* \approx 10^7$ to $\approx 10^{12}\ M_\odot$.

Figure 12 presents our combined GSMF (solid line) and some GSMFs reported in the literature: the two used by us (see above), and those from Wright et al. (2017), Papastergis et al. (2012), and Baldry et al. (2008) in small but deep volumes, and D'Souza et al. (2015) in a large volume. We plot both the original data from Bernardi et al. (2013) (pink symbols) and after decreasing M_* by 0.12 dex (blue symbols) to homogenize the stellar masses to the BC03 population synthesis model (see Appendix F). There is very good agreement between our combined GSMF and the recent GSMF reported in Wright et al. (2017) for the GAMA data.

Since the GSMF will be used as an interface for constructing the HI and H_2 mass functions, the assumption that each galaxy with a given stellar mass has its respective HI and H_2 content is implicit. Hence, the gas mass functions presented below exclude the possibility of galaxies with gas but no stars, and are equivalent to gas mass functions constructed from optically-selected samples (as in e.g., Baldry

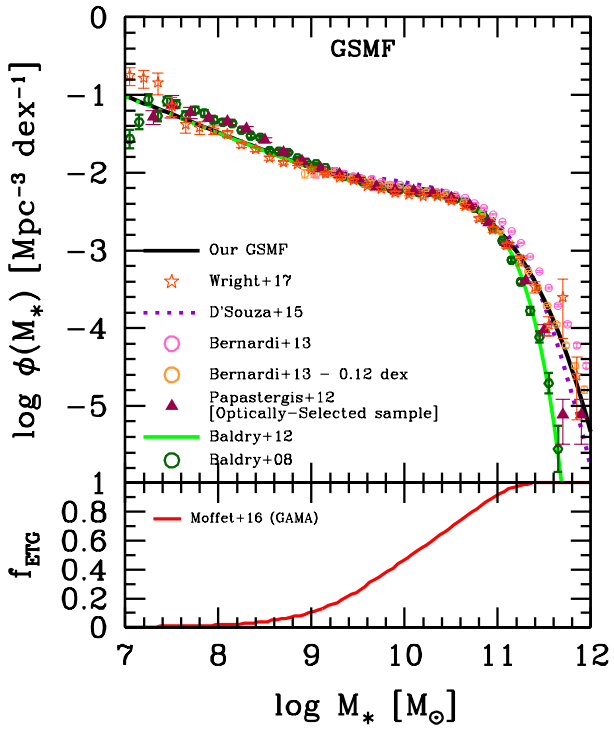

Fig. 12. Our GSMF obtained from the combination of three observational GSMFs following Kravtsov et al. (2014) (thick solid line): first from the large SDSS DR7 volume but complete only down to $\approx 10^9\ M_\odot$ (Bernardi et al. 2013, pink open circles with error bars; the orange open circles with error bars were obtained after correcting M_* by 0.12 dex, see text); and then, two complete GSMFs down to smaller masses but drawn from a very local volume (Wright et al. (2017), Papastergis et al. 2012 and Baldry et al. 2012). We also plot for comparison, the GSMFs reported in Baldry et al. (2008) and D'Souza et al. (2015). The lower panel shows the fraction of ETGs as a function of mass inferred by Moffett et al. (2016), using GAMA galaxies and their visual morphological classification. The color figure can be viewed online.

et al. 2008; Papastergis et al. 2012). In any case, it seems that the probability of finding only-gas galaxies is very low (Haynes et al. 2011).

We generate a volume complete mock galaxy catalog that samples the empirical GSMF presented above, and that takes into account the empirical volume-complete fraction of ETGs, $f_{\rm early}$, as a function of stellar mass (the complement is the fraction of LTGs, $f_{\rm late} = 1 - f_{\rm early}$). The catalog is constructed as follows:

1. A minimum galaxy stellar mass $M_{*,\rm min}$ is set ($= 10^7\ M_\odot$). From this minimum we generate a population of 5×10^6 galaxies that samples the GSMF presented above.

2. Each mock galaxy is assumed to be either an LTG or an ETG. For this, we use the results reported in Moffett et al. (2016), who visually classified galaxies from the GAMA survey. They considered ETGs those classified as Ellipticals and S0-Sa galaxies. The $f_{\rm early}$ fraction as a function of M_* is calculated as $\Phi_{\rm early}(M_*)/\Phi_{\rm all}(M_*)$, with $\Phi_{\rm early}(M_*) = \Phi_{\rm Ell}(M_*) + \Phi_{\rm S0-Sa}(M_*)$, using the fits to the respective GSMFs reported in Moffett et al. (2016).[12]

3. To each galaxy $R_{\rm HI}$ is assigned randomly from the conditional probability distribution $P_j(R_{\rm HI}|M_*)$ that a galaxy of mass M_* and type j =LTG or ETG lies in the $R_{\rm HI} \pm dR_{\rm HI}/2$ bin. Then, M_{HI}=$R_{\rm HI}$×M_*. The probability distributions for LTGs and ETGs are given by the mass-dependent PDFs presented in equations (3) and (7), respectively (their parameters are given in Table 7).

4. The same procedure as in the previous item is applied to assign M_{H_2}=$R_{\rm H_2}$×M_*, by using for the $P_j(M_{H_2}|M_*)$ probability distributions the corresponding mass-dependent PDFs for LTGs and ETGs presented in equations (3) and (7), respectively (their parameters are given in Table 7).

Our mock galaxy catalog is a volume-complete sample of $5{\times}10^6$ galaxies above $M_* = 10^7\ M_\odot$, corresponding to a co-moving volume of 5.08×10^7 Mpc^3. Since the HI and H_2 mass functions are constructed from the GSMF, the mass limit $M_{*,\rm min}$ will propagate in a different way from that of these mass functions. The co-moving volume in our mock galaxy catalog is big enough to avoid significant effects from Poisson noise. This noise affects specially the counts of massive galaxies, which are the less abundant objects.

6.1. *The Mock Galaxy Mass Functions*

6.1.1. *Stellar Mass Function*

The mock GSMF is plotted in panel (a) of Figure 13 along with the Poisson errors given by the thickness of the gray line; except for the highest masses, the Poisson errors are actually thinner than the line. The mock GSMF is an excellent realization of the empirical GSMF (compare it with Figure 12). We also plot the corresponding contributions to the mock GSMF of the LTG and ETG populations (blue and red dashed lines). As expected, LTGs dominate at small stellar masses and ETGs dominate at large

[12]Note that Sa galaxies are not included in our definition of ETGs, so that $f_{\rm early}$ is probably overestimated at masses where Sa galaxies are abundant, making that $f_{\rm early} = 0.5$ at masses lower than the break mass, M^* (see Figure 13).

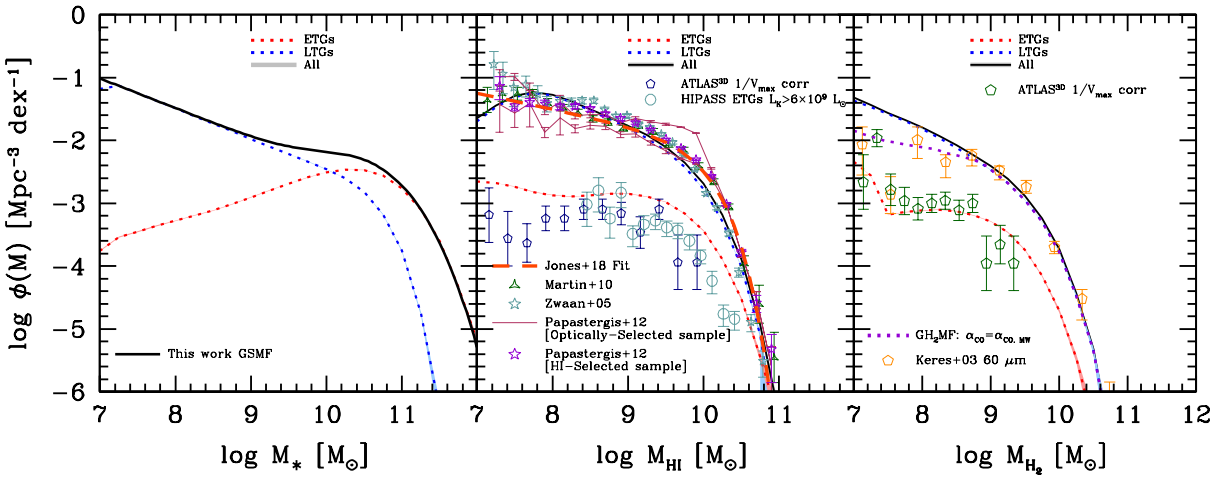

Fig. 13. *Panel (a):* Total GSMF from the mock catalog that reproduces the empirical GSMF of Figure 12 (solid line). The gray shadow represents the Poisson errors (except for large masses, these errors are thinner than the line thickness). The GSMF from the mock catalog samples very well the empirical GSMF used as input. The blue/red dotted lines and shadows correspond to the LTG/ETG mass function components, using the empirical ETG fraction as a function of M_* shown in Figure 12. *Panel (b):* Same as Panel (a) but for atomic gas, using the mean $R_{\rm HI}$–M_* relation and its scatter distribution as given in § 5. Several observational GHIMF's from blind HI samples, and the ETG GHIMF from ATLAS^{3D} and HIPASS surveys are reproduced (see labels inside the panel). *Panel (c):* Same as Panel (a) but for molecular gas. The GH_2MF calculated from the Keres et al. (2003) L_{CO} function is reproduced. The dotted purple line is the total GH_2MF from the mock catalog using a R_{H_2}–M_* correlation obtained from our compilation, but assuming that α_{CO}=$\alpha_{CO,MW}$=const., as done in Keres et al. (2003). The color figure can be viewed online.

stellar masses. The contribution of both populations is equal ($f_{\rm early} = f_{\rm late} = 0.5$) at $M_*^{\rm cross} = 10^{10.20}$ $M_\odot$ (recall that the fraction $f_{\rm early}$ used here comes from Moffett et al. (2016), who included Sa galaxies as ETGs; if consider Sa galaxies as LTGs, then $M_*^{\rm cross}$ would likely be higher). In order to predict accurate gas and baryonic mass functions, the present analysis will be further refined in Rodriguez-Puebla et al. (in prep.), where several sources of systematic uncertainty in the GSMF measurement and in the definition of the LTG/ETG fractions will be taken into account. Our aim here is only to test whether the empirical correlations derived in § 4 are roughly consistent with the total HI and H_2 empirical mass functions.

6.1.2. HI *Mass Function*

In panel (b) of Figure 13, we plot the predicted GHIMF for our mock galaxy catalog using the mean (LTG+ETG) $R_{\rm HI}$–M_* relations and their scatter distributions as given in § 5 (black line, the gray shadow shows the Poisson errors). For comparison, we plot also the HI mass functions estimated from the blind HI surveys ALFALFA (Martin et al. 2010; Papastergis et al. 2012, for both their HI- and optically-selected samples; and the latest results from Jones et al. 2018) and HIPASS (Zwaan et al. 2005). At masses larger than $M_{HI} \approx 3 \times 10^{10}$ $M_\odot$, our GHIMF is in vey good agreement with those from the ALFALFA survey but significantly above that from HIPASS. Martin et al. (2010) argue that the larger volume of the ALFALFA survey compared to HIPASS, makes ALFALFA more likely to sample the mass function at the largest masses, where objects are very rare. The volume of our mock catalog is even larger than the ALFALFA one. At intermediate masses, $9 \lesssim \log(M_{HI}/M_\odot) \lesssim 10.5$, our GHIMF is in reasonable agreement with the observed mass functions but it has in general a slightly less curved shape. At small masses, $\log(M_{HI}/M_\odot) \lesssim 8$, the observed GHIMF's flatten more than our predicted mass function. It could be that the blind surveys start to be incomplete due to sensitivity limits in the radio observations. Note that Papastergis et al. (2012) imposed additional optical requirements to their HI blind sample (see their § 2.1), which flatten the low-mass slope. Regarding the optically-selected sample of Papastergis et al. (2012), since it is constructed from a GSMF that starts to be incomplete below $\log(M_*/M_\odot) \approx 8$ (see Figure 12), one expects incompleteness in the GHIMF starting at a larger HI mass. Since our GHIMF is mapped from a volume-complete GSMF from $M_{*,\rm min} \approx 10^7$ $M_\odot$, "incompleteness" in M_{HI} is expected to start from the HI

masses corresponding to $M_{*,\rm min} \times P(R_{\rm HI}|M_{*,\rm min})$, where the latter is the scatter around the $R_{\rm HI}$–M_* relation. This shows that our GHIMF can be considered complete from $\log(M_{HI}/M_\odot) \approx 8$. The slope of the GHIMF around this mass is -1.52, steeper than the slope at the low-mass end of the corresponding GSMF ($\alpha = -1.47$).

In Figure 13 are also plotted the LTG and ETG components of the GHIMF as obtained from our mock catalog. The GHIMF is totally dominated by the contribution of LTGs. Our ETG GHIMF is compared with the ones obtained from observations using the ATLAS^{3D} and HIPASS surveys as reported in Lagos et al. (2014).

6.1.3. H_2 *Mass Function*

In Panel (c) of Figure 13, we plot the predicted GH_2MF from our mock galaxy catalog using the mean (LTG+ETG) $R_{\rm H_2}$–M_* relations and their scatter distributions as given in § 5 (black line, the gray shadow shows the Poisson errors). We compute the H_2 mass function from the CO luminosity function derived by Keres et al. (2003), who used the small and incomplete FCRAO CO survey (Young et al. 1995) and combined it with the volume-complete FIR survey. We adopt the MW H_2-to-CO conversion factor and correct their h parameter to 0.7. Unfortunately, this derivation is highly uncertain since is based on a empirical correlation between the 60μm and CO luminosities, and the selection effects in both surveys introduce several biases. The GH_2MF obtained is plotted in Figure 13. Our GH_2MF decreases faster than the one by Keres et al. (2003) at large masses, roughly agrees with it at intermediate masses, and for masses below $\log(M_{H_2}/M_\odot) \approx 8.5$, our mass function is steeper. The reason for this latter difference seems to be the mass-dependent CO-to-H_2 conversion factor introduced by us (see Appendix C). This factor increases as M_* decreases while in the case of Keres et al. (2003) it is constant. We recalculate the GH_2MF by using in the conversion from $L_{\rm CO}$ to M_{H_2} a constant CO-to-H_2 factor equal to the MW value, and plot it as the purple dotted line; the mass function at the small-mass side is now in good agreement with that of Keres et al. (2003). In Figure 13 are also plotted the LTG and ETG components of the GH_2MF as obtained from our mock catalog. The GH_2MF is totally dominated by the contribution of LTGs. Our ETG GH_2MF is compared with the one obtained from observations by using the ATLAS^{3D} survey as reported in Lagos et al. (2014).

7. DISCUSSION

7.1. *The* H_2*-to-*HI *Mass Ratio*

The global H_2-to-HI mass ratio of a galaxy characterizes its global efficiency of converting atomic into molecular hydrogen. This efficiency is tightly related to the efficiency of large-scale SF in the galaxy (see e.g., Leroy et al. 2008). From the empirical correlations inferred in § 4, we can calculate M_{H_2}/M_{HI} as a function of M_* for both the LTG and ETG populations. We do this by using our double power-law fits to the data. The left panel of Figure 14 presents the obtained M_{H_2}/M_{HI}–M_* relations and their 1σ scatter calculated by propagating the dispersions in the assumption of null covariance. In this sense, the plotted scatters are upper limits, since there is evidence of some (weak) correlation between the HI and H_2 content of galaxies, in particular among those deficient in HI and H_2 (Boselli et al. 2014c). We can plot the same correlations for the mock catalog presented in § 6, which samples the observed GSMF, the LTG and ETG fractions as a function of M_*, and the empirical correlations inferred by us. The middle panels of Figure 14 present our measures from the mock catalog for LTG (blue), ETG (red), and all galaxies (gray). The lines are the logarithmic means in small mass bins and the shaded regions are the corresponding standard deviations. At small masses, LTGs dominate, so the correlation of all galaxies is practically the one of LTGs. At large masses, ETGs become more important.

According to Figure 14, the molecular-to-atomic mass ratio of LTGs increases with M_*, albeit with a large scatter. On average, M_{H_2}/M_{HI} increases from ≈ 0.1 to ≈ 0.8 for masses ranging from $M_* = 10^8\ M_\odot$ to $3 \times 10^{11}\ M_\odot$. Given that the surface density of LTGs correlates significantly with M_*, one can expect this dependence of M_{H_2}/M_{HI} on M_* at least from two arguments: (1) Disk instabilities, which drive the formation of molecular clouds (e.g., the Toomre criterion Toomre 1964), are more probable to occur as the disk surface density is higher. (2) The H_2-to-HI mass ratio in galaxies has been shown to be directly related to the hydrostatic gas pressure (Blitz & Rosolowsky 2006; Krumholz et al. 2009), and this pressure depends on the (gas and stellar) surface density (Elmegreen 1989). In fact, the physics of H_2 condensation from HI is very complex and it is expected to be driven by local parameters of the ISM (see e.g., Blitz & Rosolowsky 2006; Krumholz et al. 2009; Obreschkow & Rawlings 2009). Therefore, the dependence of the H_2-to-HI mass ratio on M_* should be understood as consequence of the correlations of these parameters (their mean values along

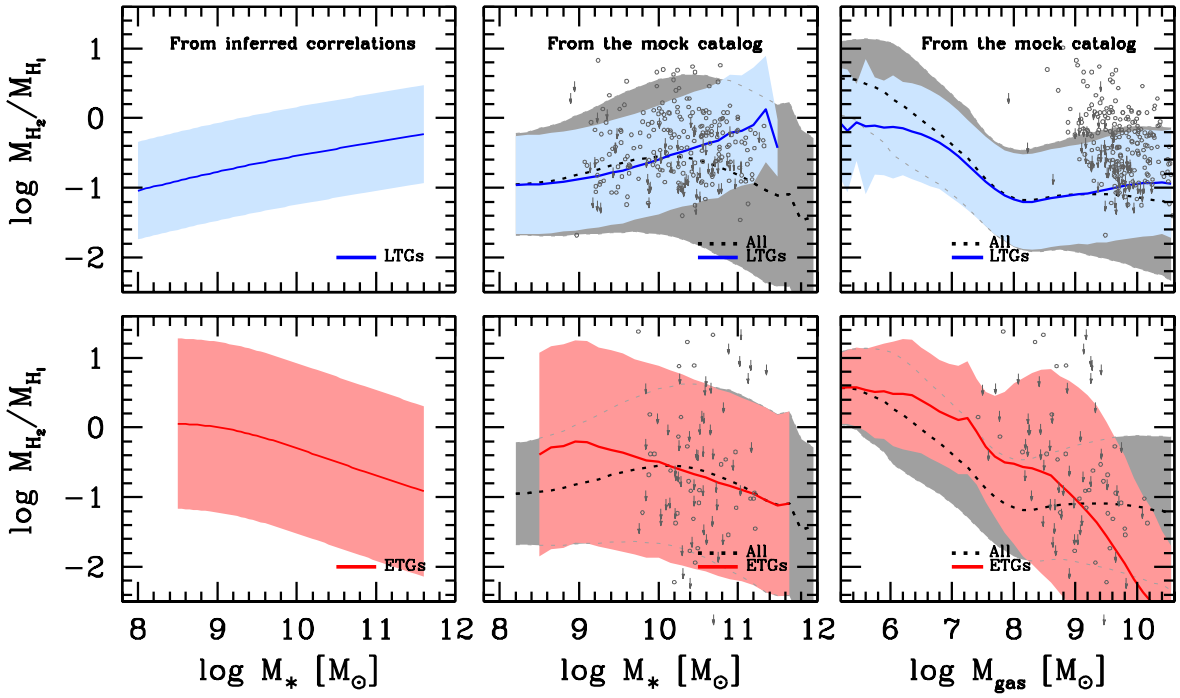

Fig. 14. *Left panels:* Molecular-to-atomic mass ratio, M_{H_2}/M_{HI}, for LTGs (upper panel) and ETGs (lower panel) inferred from our double power-law fits to the $R_{\rm HI}$-M_* and $R_{\rm H_2}$-M_* correlations. The shaded areas are the 1σ scatter obtained by error propagation of the scatter around the $R_{\rm HI}$-M_* relations. *Middle panels:* Same as left panels, but from our mock catalog generated to sample the empirical GSMF, volume-complete ETG/LTG fractions as a function of mass, and $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* correlations. The dotted line surrounded by the gray area is the total M_{H_2}/M_{HI} ratio and the 1σ scatter as a function of stellar mass. *Right panels:* Molecular-to-atomic mass ratio as a function of the cold gas mass, $M_{\rm gas}$, from the mock catalog for LTGs (upper panel) and ETGs (lower panel). We plot available detected and undetected cold gas observational data as gray unfilled circles and downward arrows respectively. The color figure can be viewed online.

the galaxy) with M_*, introducing this actually large scatter in the dependence of M_{H_2}/M_{HI} on M_*. Indeed, several authors have shown that M_{H_2}/M_{HI} correlates better with the mean gas-phase metallicity or mean stellar surface density than with M_* (e.g., Saintonge et al. 2011; Boselli et al. 2014a).

For ETGs, the trend of the H_2-to-HI mass ratio is inverse to the one of LTGs and shows a very large scatter. The ETGs more massive than $\approx 10^{11}$ $M_\odot$ have mean ratios around 0.15 and a 1-σ scatter of $\approx \pm 1$ dex; for intermediate masses, this ratio increases on average, and for ETGs with masses $M_* \approx 10^9$ $M_\odot$, which are actually very rare, their mean H_2-to-HI mass ratios are ≈ 1 with the same scatter of $\approx \pm 1$ dex. Even though the gas fraction in ETGs is much smaller than in LTGs at all masses (see Figure 5), the former are also typically more compact than the latter, resulting probably in similar or higher gas pressures on average, and consequently similar or even higher M_{H_2}/M_{HI} ratios, especially at masses smaller than $M_* \approx 10^{10}$ $M_\odot$. In fact, given the large scatter in M_{H_2}/M_{HI} for ETGs, this ratio depends likely on many other internal and external factors (mergers, environment, etc.) that do not correlate significantly with M_*.

Regarding M_{H_2}/M_{HI} vs. $M_{\rm gas}$, for LTGs, which for $M_* > 10^7$ $M_\odot$ have mostly gas masses $> 10^8$ $M_\odot$, there is no significant dependence, while for ETGs, which are almost inexistent with $M_{\rm gas} \gtrsim 10^9$ $M_\odot$, M_{H_2}/M_{HI} is larger on average for lower values of $M_{\rm gas}$. This can be seen in the right panel of Figure 14, where the mock catalog has been used. Basically, for a given $M_{\rm gas}$, in the mass range $M_{\rm gas} \approx 10^7 - 10^9 M_\odot$, ETGs have typically larger H_2-to-HI mass ratios than LTGs. In combination, the H_2-to-HI ratio appears to be larger for lower values of $M_{\rm gas}$. Such a dependence has been reported by Obreschkow & Rawlings (2009) for their compiled sample of galaxies, and predicted by these authors for a physical model.

The dependences of the H_2-to-HI mass ratio on M_*, M_{gas}, and morphological type discussed above are in qualitative agreement with several previous observational works, which actually are part of our compilation (Leroy et al. 2008; Obreschkow & Rawlings 2009; Saintonge et al. 2011; Boselli et al. 2014a; Bothwell et al. 2014). However, our results extend to a larger mass range and separate explicitly the two main populations of galaxies.

7.2. *The Role of Environment*

There are several pieces of evidence that the atomic gas fraction of galaxies is lower in higher-density environments (e.g., Haynes & Giovanelli 1984; Gavazzi et al. 2005; Cortese et al. 2011; Catinella et al. 2013; Boselli et al. 2014c). The fact that the ETG population has lower HI gas fractions than the LTG one (§ 4), the former being commonly found in higher-density environments, agrees with the mentioned environment trends. Thus, due to the morphology-density relation, our determinations of the R_{HI}–M_* (as well as R_{H_2}–M_* and R_{gas}–M_*) correlations for the LTG and ETG populations, partially account for the dependence of these correlations on environment. Moreover, for the very isolated LTGs and for the subsample of LTGs in the Virgo cluster central regions, we confirm higher and lower HI-to-stellar mass ratios than the average of the overall LTG sample, respectively (see § 3.1). However, this systematic difference according to the environment is within the 1σ scatter of the R_{HI}–M_* correlation of LTGs (see Figure 3). Instead, in the case of ETGs, the isolated galaxies have much larger R_{HI} values than the mean of all ETGs, above the 1σ scatter; isolated ETGs are almost as HI gas rich as the mean of LTGs.

Regarding the molecular gas fraction, the observational results are controversial in the literature. Recent studies seem to favor the the fact that galaxies in clusters are actually H_2–deficient compared to similar galaxies in the field. However, the deficiencies are smaller than in the case of HI (Boselli et al. 2014c, and references therein). Here, for isolated and Virgo-center LTGs, we do not see any systematical segregation of R_{H_2} from the rest of our compiled LTGs (Figure 4), but in the case of ETGs, the isolated galaxies have on average larger values of R_{H_2}.

In summary, the results from our compilation point out that the HI content of LTGs has a (weak) dependence on environment, mainly due to the fact that at high densities LTGs are HI deficient. However, the H_2 content of LTGs seems not to change on average with the environment. In the case of ETGs, those very isolated are significantly more gas rich (both in HI and H_2) than the average at a given mass.

An important aspect related to the environment is whether a galaxy is central or satellite. The local environmental effects once a galaxy becomes a satellite inside a halo. Ram pressure, viscous stripping, starvation, harassment and tidal interactions work in the direction of lowering the gas content of the galaxy, likely more the more massive the halo is (Boselli & Gavazzi 2006; Brown et al. 2017). Part of the scatter in the gas-to-stellar mass correlations is probably due to the external processes produced by these local-environment mechanisms. A result in this direction has been recently shown for the R_{HI}–M_* correlation by Brown et al. (2017). These authors have found that the HI content of satellite galaxies in more massive halos has, on average, lower HI-to-stellar mass ratios at fixed stellar mass and specific SFR. According to their analysis, the systematic environmental suppression of the HI content at both fixed stellar mass and fixed specific SFR in satellite galaxies begins at halo masses typical of the group regime ($> 10^{13}$ $M_\odot$), and fast-acting mechanisms such as ram-pressure stripping are suggested to explain their results. In a future study, we will attempt to characterize the central/satellite nature of our compiled galaxies, as well as calculate a proxy to their halo masses, in order to clarify this question.

7.3. *Comparisons with Previous Works*

In Figure 15 we compare our results with those of previous works. When necessary, the data are corrected to a Chabrier (2003) IMF. Most of the previous determinations of the HI- and H_2-to-stellar mass correlations are not explicitly separated for the two main galaxy populations as done here, and in several cases non-detections are assumed to have the values of the upper limits, or are not taken into account.

In the upper panel, our empirical R_{HI}–M_* correlations for LTGs and ETGs are plotted along with the linear relations given by Stewart et al. (2009) (cyan line, the dashed lines show the 1σ scatter) and Papastergis et al. (2012) (gray line). The former authors used mainly the observational data presented in McGaugh (2005) for disk-dominated galaxies, and the latter authors used samples from Swaters & Balcells (2002), Garnett (2002), Noordermeer et al. (2005), and Zhang et al. (2009), which refer mostly to late-type galaxies. Their fits are slightly above the mean of our LTG R_{HI}–M_* correlation. This is likely due to their ignoring non-

detections. We also plot the logarithmic average values in mass bins reported by Catinella et al. (2013) for GASS (green open circles). Since ETGs progressively dominate in number as the mass increases, our total (density-weighted) $R_{\rm HI}$–M_* correlation would fall below the one by Catinella et al. (2013), especially at the largest masses. Note that for the data plotted from Catinella et al. (2013), the HI masses of non-detections were set equal to their upper limits. Therefore, the plotted averages are biased to high values of $R_{\rm HI}$, specially for ETGs which are dominated by non-detections. Furthermore, recall that we have corrected the upper limits of GASS for distance to make them compatible with those of the nearer ATLAS^{3D} survey.

More recently, Brown et al. (2015) have used the HI spectral stacking technique for a volume-limited, stellar mass selected sample from the intersection of SDSS DR7, ALFALFA, and *GALEX* surveys. With this technique the stacked signal of co-added raw spectra of detected and non-detected galaxies (about 80% of the ALFALFA selected sample) is converted into a (lineal) average HI mass. The authors have excluded from their analysis HI-deficient galaxies –typically found within clusters– because of their significant offset to lower gas content. The black dots connected by a dotted line show the logarithm of the average $R_{\rm HI}$ values reported at different stellar mass bins in Brown et al. (2015). Since HI-deficient galaxies –which typically are ETGs– were excluded, the Brown et al. (2015) correlation should be compared with our correlation for LTGs. Note that with the stacking technique it is not possible to obtain the population scatter in $R_{\rm HI}$ because the reported mean values come from stacked spectra instead of from averaging individual values of detections and non-detections. However, the stacking can be applied to subsets of galaxies, for example, selected by color. Brown et al. (2015) have divided their sample into three groups by their NUV$-r$ colors: [1,3), [3,5), and [5,8]. The average $R_{\rm HI}$ values at different masses corresponding to the bluest and reddest groups are reproduced in Figure 15 with the blue and red symbols, respectively. Note that the logarithmic mean is lower than the logarithm of the mean. For a lognormal distribution, $\langle \log x \rangle = \log\langle x \rangle - 0.5 \times \sigma^2_{\log x} \ln 10$ (see e.g., Rodríguez-Puebla et al. 2017). Then, for the typical scatter of 0.44 dex corresponding to LTGs, the logarithm of the stacked values of $R_{\rm HI}$ should be decreased by ≈ 0.2 dex to compare formally with our reported values of logarithmic means; this is shown with a black arrow in Figure 15. If the reddest galaxies in the Brown et al. (2015) stacked sample are associated with ETGs (which is true only partially), then for them the correction to a logarithmic mean is of ≈ 1 dex, shown with a red arrow.

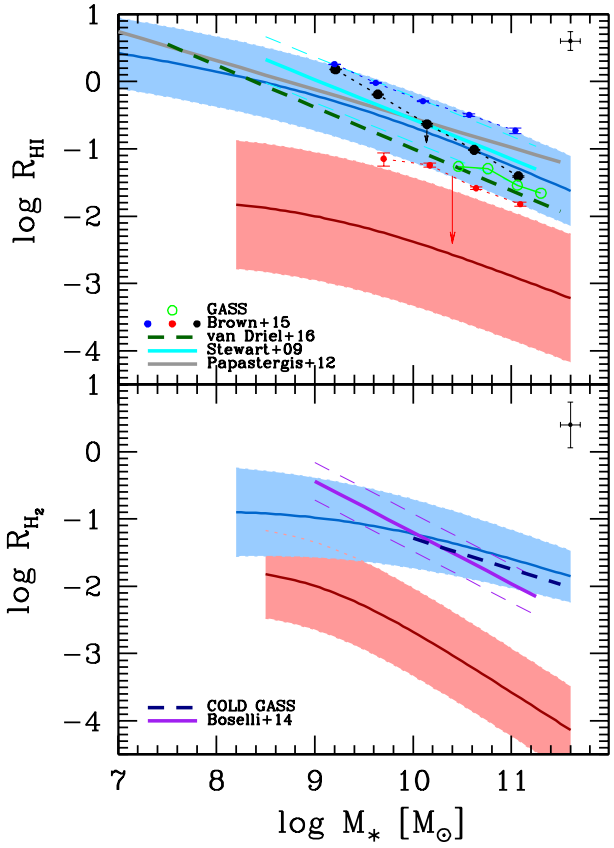

Fig. 15. *Upper panel:* Our empirical HI-to-stellar mass correlations for LTGs and ETGs (blue and red shaded areas, respectively) compared with some previous determinations (see labels inside the panel and details of each determination in the text). Previous determinations are based on compilations typically biased to late-type, blue galaxies, and/or do not take into account non-detections. The blue and red arrows correspond to estimates of the difference between the logarithm of the mean (the stacking technique provides the equivalent of the mean value) and the logarithmic mean (our determinations are for this case) for standard deviations of 0.52 and 0.99 dex, respectively (see text for more details). *Lower panel:* Our empirical molecular H_2-to-stellar mass correlations for LTGs and ETGs (blue and red shaded areas, respectively) compared with very rough previous determinations not separated into LTGs and ETGs (see labels inside the panel and details of each determination in the text). The color figure can be viewed online.

Finally, recently van Driel et al. (2016) reported the results for HI observations at the Nancay Radio Telescope (NRT) of 2839 galaxies selected evenly from SDSS. The authors present a Buckley-James

linear regression to their data (long-dashed green line in Figure 15), taking into account this way upper limits for non-detections (though their upper limits are quite high given the low sensitivity of NRT). They fitted the entire sample, that is, they do not separate into LTG/ETG or blue/red groups. In a subsequent paper (Butcher et al. 2016), the authors obtained ≈ 4 times more sensitive follow-up HI observations at Arecibo for a fraction of the galaxies that were either not detected or marginally detected; 80% of them were detected with HI masses ≈ 0.5 dex smaller than the upper limits in van Driel et al. (2016), and the rest, mostly luminous red galaxies, were not detected. If this trend is representative of the rest of the NRT undetected galaxies, Butcher et al. (2016) expect the fit plotted in Figure 15 to be offset toward lower $R_{\rm HI}$ values by about 0.17 dex and even more at the highest masses. This fit lies between a density-weighted fit to our two correlations when taking into account that at large masses the fraction of ETG/red galaxies increases and at small masses LTG/blue galaxies dominate.

The lower panel of Figure 15 is similar to the upper panel, but for the $R_{\rm H_2}$–M_* correlations. Regarding the molecular gas content, there are only a few attempts in the literature to determine the relation between M_{H_2} and M_*. In fact, those works that report approximate correlations are included in our compilation: Saintonge et al. (2011) for COLD GASS, and Boselli et al. (2014a) for HRS. The former authors report a linear regression to their binned data assuming H_2 masses for non-detection set equal to their upper limits. The latter authors present a bisector fit using only detected, late-type gas-rich galaxies. Therefore, in both cases the reported relations are clearly biased to LTGs and toward high $R_{\rm H_2}$ values.

The differences we find between our correlations and those plotted in Figure 15, as discussed above, can be understood on the basis of the different limitations that are present in each of the previous works. Bearing in mind these limitations we can conclude that the correlations presented here are in rough agreement with previous ones, but compared to them (i) they extend the correlations to a larger mass range, (ii) they separate explicitly galaxies into their two main populations, and (iii) they take into account adequately the non-detections.

8. SUMMARY AND CONCLUSIONS

The fraction of stars and atomic and molecular gas in local galaxies is the result of complex astrophysical processes along their evolution. Thus, the observational determination of how these fractions vary as a function of mass provides key information on galaxy evolution at different scales. Before the new generation of radio telescopes becomes available, which will bring extragalactic gas studies more in line with optical surveys, the main way to get this kind of information are studies based on radio follow-up observations of (small) optically-selected galaxy samples. In this work, we have compiled and homogenized samples from the literature with information on M_* and M_{HI} and/or M_{H_2} for galaxies that can be identified to belong to two main (in a statistical sense) groups: the LTG and ETG populations. To estimate M_{H_2} from CO observations, we have introduced a mass-dependent CO-to-H_2 conversion factor in agreement with studies that show that this factor is not constant and depends on metallicity (hence, statistically, on mass). Results using a constant CO-to-H_2 factor were also presented. Figures 1 and 2 summarize our compilation in the $R_{\rm HI}$ vs. M_* and $R_{\rm H_2}$ vs. M_* logarithmic diagrams.

Before inferring the correlations, we have tested how much each one of the compiled samples deviates from the rest and have classified them into three categories: (1) samples complete in limited volumes (or selected from them) without selection effects that could affect the calibration of the correlations (Golden); (2) samples that are not complete but are representative of the average galaxy population, without obvious selection effects (Silver); and (3) samples selected by their environment (Bronze). We showed that most of the samples, after our homogenization, are suitable to infer the $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* correlations, except those from the Bronze category in the case of ETGs. These galaxies in extreme environments show significant deviations from the mean trends, and are not taken into account in our determinations. From the combination of all the chosen samples, we have calculated the mean, standard deviation, and percentiles of the logarithms of the $R_{\rm HI}$ and $R_{\rm H_2}$ mass ratios in several stellar mass bins, taking into account non-detected galaxies and their reported upper limits, which are a non-negligible fraction of the data, especially for the ETG population. Taking into account non-detected galaxies and their homogenization among different samples is relevant for determining the gas-to-stellar mass correlations of ETGs.

The mean logarithmic values in mass bins, $\langle \log R_{\rm HI} \rangle$ and $\langle \log R_{\rm H_2} \rangle$, with the corresponding (intrinsic) standard deviations calculated by means of the Kaplan-Meier estimator were fitted to the logarithm of single and double power-law functions

(equation 1). The parameters of the best fits to these functions, both for LTGs and ETGs, are reported in Tables 5 and 6, respectively. We highlight the following results from our analysis:

• The $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* correlations for the LTG and ETG populations can be described roughly equally well by a single or a double power law at masses larger than $\log(M_*/M_\odot)\gtrsim 9$. For smaller masses, we see some hints of a flattening in these correlations. LTGs have significantly higher HI and H_2 gas fractions than ETGs, the differences increasing at the high- and low-stellar mass ends. For the ETG population, the scatters of the $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* correlations are much larger than for the LTG population.

• Combining the $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* correlations and propagating errors, we calculated the cold gas ($M_{\rm gas}$=1.4[$M_{HI}+ M_{H_2}$])-to-stellar mass correlations of the LTG and ETG populations. For the former, $R_{\rm gas}$ is around 4 on average at $M_* = 10^7\ M_\odot$ and ≈ 1 at $M_* = 1.60\times10^9\ M_\odot$. At larger masses, $R_{\rm gas}$ continues to decrease significantly. For the ETG population, $R_{\rm gas}$ on average is smaller than 1 even for the smallest galaxies. Galaxies as massive as $M_* = 10^{11}\ M_\odot$ have on average $R_{\rm gas}$ ratios smaller than 2.5×10^{-3}. The intrinsic standard deviation of the $R_{\rm gas}$–M_* correlation of the LTG population is $\approx$ 0.44 dex while for ETG it is larger, around 0.68 dex.

• The H_2-to-HI mass ratio implied by our correlations is such that for LTGs it increases on average with M_*, from ≈ 0.1 to 0.8 for masses ranging from $M_* = 10^8\ M_\odot$ to $3\times10^{11}\ M_\odot$. For ETGs, the trend is the opposite but with a large scatter (standard deviation of $\approx \pm1$ dex). While ETGs have much less gas content than LTGs, the H_2-to-HI mass ratio at intermediate and small masses is higher on average in the former than in the latter, and lower at large masses.

• In an attempt to describe the full distributions of $R_{\rm HI}$ and $R_{\rm H_2}$ as a function of M_* for both the LTG and ETG populations, the respective PDFs from the censored+uncensored data in different mass bins provided by the Kaplan-Meier estimator were used. For LTGs, we found that a Schechter function with their parameters depending on M_* offers a good description of the $R_{\rm HI}$ and $R_{\rm H_2}$ distributions as a function of M_* (equation 3). For ETGs, these distributions look bimodal, with a (broken) Schechter function and a uniform distribution at the low-end side providing an approximate description of them (equation 7). These mass-dependent PDFs offer a full description of the $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* relations and their scatter distributions for both LTGs and ETGs. Their first and second moments agree very well with our previously determined double power-law correlations (Figures 10 and 11).

• The mass-dependent distribution functions of $R_{\rm HI}$ and $R_{\rm H_2}$ were used to map the GSMF into the corresponding HI and H_2 mass functions, both for LTGs and ETGs. We use an empirical GSMF from the combination of GSMFs from a low-z survey and from the overall DR7 sample, following Kravtsov et al. (2014). The fractions of LTGs/ETGs as a function of M_* are calculated from the fitted mass functions of ETGs obtained by Moffett et al. (2016) using the GAMA survey. The predicted total HI and H_2 mass functions agree with those obtained from empirical determinations in the mass ranges where these determinations are reliable.

Our (marginal) finding of a flattening in the HI- and H_2-to-stellar mass correlations at small masses has been suggested in some previous works (see § 4 for references). For our double power-law fits (equation 1), we find that the transition mass M_*^{tr} is around $1-2\times10^9\ M_\odot$ for both the $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* correlations and for both the LTG and ETG populations. Interestingly enough, this is the mass that roughly separates normal and dwarf galaxies.

We are aware that our determination of the gas-to-stellar mass relations is based on an heterogeneous mix of samples. However, we have shown that there are no significant differences in the $R_{\rm HI}$ and $R_{\rm H_2}$ values as a function of M_* for volume-limited complete and incomplete samples. Significant differences are observed only for samples selected by environment in the case of ETGs. On the other hand, our correlations for ETGs (and LTGs in the case of molecular gas), are very limited at small masses. They are actually just extrapolations for stellar masses below several $10^8\ M_\odot$, but we have checked them to be consistent with the very few available determinations (mostly non-detections) below these masses.

In spite of these shortcomings, it is encouraging that the correlations (in fact, the full mass-dependent distributions), when mapped onto the HI and H_2 mass functions using the observed GSMF as an interface, are consistent with the mass functions determined from observational radio surveys, at least in the mass ranges where these surveys do not suffer strong selection, volume, and cosmic variance effects. Such a self-consistency between the gas-to-stellar correlations and mass functions supports the reliability of our results, which help to pave the way for the next generation of radio telescopes.

The empirical gas-to-stellar mass correlations and the approximate scatter distributions presented in this paper for the two main populations of galaxies are useful for understanding global aspects of galaxy evolution as a function of mass. We encourage the use of these correlations (or the full mass-dependent PDFs) for comparisons with predictions of models and simulations of galaxy formation and evolution.

Finally, we provide upon request to A. R. Calette a Python-based code that allows to generate plots and electronic tables for both LTGs and ETGs of (1) the $R_{\rm HI}$-M_* and $R_{\rm H_2}$-M_* double power-law relations and their 1σ intrinsic scatters as presented in Figure 5 and Table 6; and (2) the mass-dependent full $R_{\rm HI}$ and $R_{\rm H_2}$ PDFs as constrained in § 5, including the first and second moments (mean and standard deviation) of these PDFs.

We thank Dr. David Stark for kindly making available to us his compilation of data in electronic form, and Dr. Claudia Lagos for providing us the ETG data plotted in Figure 13. We thank the anonymous referee for useful comments and suggestions, which improved the quality of the manuscript. The authors acknowledge CONACyT grant (Ciencia Básica) 285721 for partial funding. ARC acknowledges a PhD Fellowship provided by CONACyT. ARP has been suported by a UC-MEXUS Fellowship.

APPENDIX

A. THE COMPILED GALAXY SAMPLES WITH HI INFORMATION

A.1. *Golden Category*

Updated Nearby Galaxy Catalog (UNGC; Karachentsev et al. 2013, 2014): It is the most representative and homogeneous sample of galaxies (869, most of them of small masses) in the Local Volume, located within 11 Mpc or with corrected radial velocities $V_{LG} < 600$ Km s^{-1}. The authors mention that the sample is complete to $M_B \approx -11$ mag, spanning all morphologies. However, we take a more conservative limit, bearing in mind that at low luminosities the fraction of hard-to-detect low surface brightness (LSB) galaxies strongly increases. Karachentsev et al. (2013) report the mean $B-$band surface brightness (SB) within the Holmberg isophote, $\bar{\mu}_{B,26}$ for the UNGC galaxies. The SB decreases on average with decreasing luminosity. For LTGs, the distribution of SBs appears to be incomplete from $M_B \approx -13.5$ mag, in such a way that most of the galaxies could be lost at lower luminosities. This is in agreement with the completeness limit suggested by Klypin et al. (2015) for UNGC, based on the turnover of the luminosity function constructed by them. In view of these arguments, we consider complete the UNGC sample for LTGs only to $M_B \approx -13.5$ mag ($M_* \approx 10^{7.2-7.4}$ $M_\odot$); the few LTGs below this limit are of high SB and are expected to contain less gas than the average. Since ETGs are of higher SBs than LTGs, the SB distribution for the small fraction of them seems not to be affected even at the faintest observed luminosities, $M_B \approx -11$ mag. There are 561 galaxies with available HI data (for details regarding the data sources on HI fluxes, see Table 3 from Karachentsev et al. 2013); 90 of them do not satisfy our completeness limit. We estimate stellar masses from the reported K-band luminosities and $B-K$ colors as in Avila-Reese et al. (2008), who calculated the mass-to-light ratios for HSB and LSB galaxies following Bell et al. (2003) and Verheijen (1997), respectively. The obtained masses (assuming a diet Salpeter IMF) were corrected to the Chabrier IMF. To separate HSB and LSB galaxies we use the reported $\bar{\mu}_{B,26}$, and transform it to a central surface brightness, $\mu_{0,B}$ assuming an exponential disk. Thus, the criterion $\mu_{B,0} > 22.5$ mag/arcsec^2 for selecting LSB galaxies corresponds to $\bar{\mu}_{B,26} > 24.6$ mag/arcsec^2. Karachentsev et al. (2013) apply corrections for peculiar motions in the determination of the distances of all the galaxies.

GALEX Arecibo SDSS Survey (GASS; Catinella et al. 2013): It is an optically-selected sub-sample of 760 galaxies more massive than 10^{10} $M_\odot$ taken from a parent SDSS DR6 sample volume limited in the redshift range $0.025 < z < 0.05$ and cross-matched with the ALFALFA and GALEX surveys. The HI information comes from follow-up observations carried out with the Arecibo 305 m telescope and detections taken from the ALFALFA survey or the Cornell HI digital archive. The $R_{\rm HI}$ limit of the sample is well controlled: 0.015 for $\log(M_*/M_\odot) > 10.5$ and up to 0.05 for smaller masses. There are 473 detections and 287 non-detections; for the latter, upper limits are provided. For the morphological type, we use the Huertas-Company et al. (2011) automatic classification applied to the SDSS DR7. These authors provide for each galaxy the probability of being of early type, PE, i.e., E or S0. We have tested this probability in a catalog of galaxies with careful visual morphological classification (UNAM-KIAS, see below; Hernández-Toledo et al. 2010) and found that galaxies of types $T \leq 1$ are mostly those with $PE > 0.65$,

and those with $PE \leq 0.65$ correspond mostly to $T > 1$.[13] Thus, we consider here as ETGs those with $PE > 0.65$, and the complement are LTGs. We find a good correlation between the ETGs and LTGs in this way defined with those defined using the concentration parameter $c = R_{90}/R_{50}$ to characterize the galaxy type, with the value of $c = 2.85$ for separating the LTGs from the ETGs (for the latter, it is also required to obey the color criterion $NUV - r > 5$, Deng 2013). The stellar masses in Catinella et al. (2013) were calculated from the spectral energy distribution (SED) of the SDSS galaxies (Salim et al. 2007) assuming a Chabrier (2003) IMF.

Herschel Reference Survey – field galaxies (HRS; Boselli et al. 2010, 2014a,b,c): It is a $K-$band volume limited ($15 \leq D/\mathrm{Mpc} \leq 25$) sample of 323 galaxies complete to $K_s = -12$ and -8.7 mag for LTGs and ETGs, respectively. The authors collected from the literature and homogenized HI data for 315 galaxies, and CO data for most of them. The morphological type was taken from NED or, if unavailable, from their own classification. Stellar masses were derived from i-band luminosities and $g-i$ colors (from Cortese et al. 2012) using stellar mass-to-light ratios as given in Zibetti et al. (2009), and assuming a Chabrier IMF. The distances were corrected for peculiar motions and presence of clusters. The sample includes objects in environments of different density, from the core of the Virgo cluster, to loose groups and fairly isolated systems. To match the Golden category, we excluded the numerous galaxies from the Virgo Cluster center (regions A and B), which bias the sample toward high densities.

ATLAS^{3D} HI sample – field ETGs (Serra et al. 2012): ATLAS^{3D} is a sample of 166 local ETGs observed in detail with integral field unities (IFUs; Cappellari et al. 2011). The distance range of the sample is between 10 and 47 Mpc; the sample includes 39 galaxies (24% of the galaxies) from the Virgo Cluster. For the Golden category, we exclude those ETGs in the Virgo core. The sample is not complete, but after excluding the large number of Virgo core galaxies, it is expected to be representative of the local population of ETGs since the galaxies were selected from a complete volume-limited parent sample. The masses range from $\approx 10^{9.8}$ to $10^{11.3}$ $M_\odot$; more massive galaxies are tipically not found in small volumes. We estimate stellar masses using $\log(M_*) = \log(0.5) + \log(L_K)$, where L_K is the K-band luminosity inferred from the K-band absolute magnitude. The HI observations were carried out in the Westerbork Synthesis Radio Telescope (Serra et al. 2012). They used ALFALFA spectra to determine M_{HI} upper limits using one resolution element and found that M_{HI} limit is a factor ≈ 2 over the HI mass limit obtained with their data. The $R_{\rm HI}$ detection limit increases with mass on average by more than 1.5 orders of magnitude, attaining values as low as $\approx 10^{-4}$ for the most massive systems. Because the ATLAS^{3D} galaxies are nearby, the upper limits are much lower than in the case of the GASS galaxies in the same mass range.

A.2. *Silver Category*

Nearby Field Galaxy Survey (NFGS; Jansen et al. 2000a,b; Wei et al. 2010; Kannappan et al. 2013, and references therein): It is a broadly representative sample of 198 local galaxies spanning stellar masses $M_* \approx 10^8 - 10^{12}$ $M_\odot$ and all morphological types. Morphological classification was obtained from Jansen et al. (2000b). The sample is not volume-complete; galaxies span distances from 2 to 306 Mpc. Distances were derived from the Virgo centric flow corrected velocities with respect to the centroid of the Local Group. Stellar masses were estimated using a variant of the code described in Kannappan & Gawiser (2007) and improved in Kannappan et al. (2009), which fits the SED and the integrated spectrum of a galaxy with a suite of stellar populations models. Both the diet Salpeter and the Chabrier (2003) IMFs were used. The single-dish HI fluxes for most of the galaxies were taken from the HyperLeda database (Paturel et al. 2003) or were obtained by the authors with the Green Bank Telescope (GBT) Spectrometer. The sample provides strong upper limits up to $R_{\rm HI} \approx 0.1$; all galaxies with larger ratios are detected (139), and for the rest only upper limits are provided.

Stark et al. (2013) compilation: These authors compiled from the literature and homogenized 323 galaxies with available HI, CO, and multi-band imaging data. Most of the compiled galaxies are from the GASS, NFGS and ATLAS^{3D} surveys described above. We use here only those galaxies that are not in these surveys (67 galaxies). The authors use morphological type to separate galaxies into two groups, coincident with our morphology criterion for ETGs and LTGs. In their compilation are included some blue compact dwarfs (BCDs). We exclude those BCDs classified as early types. The stellar masses were calculated following Kannappan et al. (2013). The optical and NIR information required

[13] Huertas-Company et al. (2011) define as ETGs those with $T \leq 1$, but the T index in their case is from the Fukugita et al. (2007) notation, which assigns $T = 1$ to lenticulars instead of $T = 0$ as in the usual de Vacouleours notation.

for this calculation was taken from SDSS DR8 (for those galaxies outside the SDSS footprint, the *BVRI* photometry from the SINGS sample was used) and 2MASS, respectively.

Leroy et al. (2008) THINGS sample: It is a sample of 23 nearby, star-forming galaxies, which we associate with LTGs; 11 are dwarf, HI-dominated galaxies and 12 are large well-defined spiral galaxies. The HI information of the galaxies comes from "The HI Nearby Galaxy Survey" (THINGS, Walter et al. 2008) and was obtained with the NRAO Very Large Array (VLA). The stellar masses were calculated from 3.6 μm information taken from the Spitzer Infrared Nearby Galaxies Survey (SINGS Kennicutt et al. 2003). To convert the 3.6 μm intensity to surface stellar mass density, they used a K-to-3.6 μm calibration and adopted a fixed $K-$band mass-to-light ratio, $\Upsilon_*^{\rm K} = 0.5 M_\odot/L_\odot$, assuming a Kroupa (2001) IMF; M_* was calculated from integrating the surface stellar mass density.

Dwarf LTGs (Geha et al. 2006): It is a sample of 101 dwarf galaxies, 88 out of them with HI measurements and being of late type. Galaxies with absolute magnitudes $M_r - 5\log_{10}(h_{70}) > -16$ were selected from the low-luminosity spectroscopy catalog of Blanton et al. (2005b), based on the SDDS. Distances were estimated based on a model of the local velocity field (Willick et al. 1997). Possible selection effects related to the Blanton et al. (2005b) catalog include the fact that it does not span the full range of environments (there are no clusters), and LSB dwarfs are missed. Stellar masses were based on the optical SDSS i-band magnitude and $g-r$ colors using the mass-to-light ratios of Bell et al. (2003). The M_{HI} masses were obtained by Geha et al. (2006) from the H I integrated fluxes measured with the Arecibo 305 m telescope and the GBT.

ALFALFA dwarf sample (Huang et al. 2012b): It consists of 176 low HI mass dwarf galaxies from the ALFALFA survey. The galaxies were selected to have $M_{HI} < 10^{7.7}\ M_\odot$ and HI line widths < 80 km s^{-1} (*s-com* sample). This sample is not complete in a volume-limited sense but it probes the extreme low HI mass tail of the ALFALFA survey. Stellar masses were obtained through SED fitting following Salim et al. (2007), assuming a Chabrier (2003) IMF. Only 57 out of the 176 galaxies have a stellar mass determination. These galaxies have HI detections and high gas fractions, they are dwarf irregulars.

A.3. *Bronze Category*

UNAM-KIAS catalog of isolated galaxies (Hernández-Toledo et al. 2010): It is a magnitude-limited sample ($m_r > 15.2$ mag) of galaxies from the SDSS DR5 that fulfill strict isolation criteria; it is composed of 1520 galaxies spanning all morphological types. The morphological classification was carried out by the authors. We searched HI information for these galaxies in HyperLeda (the 21-cm line magnitudes corrected for self-absorption, m_{21}^c). The HI masses were calculated as $M_{H_I}[M_\odot] = 2.356 \times 10^5 \cdot d_L^2 \cdot F_{21}$, where $F_{21}[\text{Jy Kms}^{-1}] = 10^{0.4(17.40-m_{21}^c)}$ and d_L is the luminosity distance to the galaxy in Mpc. For the HI non-detections, we searched rms noise limits in the Digital archive of HI 21 centimeter line spectra of optically selected galaxies (Springob et al. 2005), finding data only for 7 galaxies. Non-detected HI upper mass limits were estimated as $M_{H_I}^{\rm lim}[M_\odot] = 1.5 \cdot \text{rms} \cdot \delta W$, where δW is the full width of the HI line obtained from the Tully-Fisher relation of Avila-Reese et al. (2008) ($\delta W = 2V_m$ is assumed). For LTGs (ETGs), we found 272 (24) detections and 7 (0) non-detections. Stellar masses were taken from the group catalog of Yang et al. (2007), where the Bell et al. (2003) mass-to-light ratios for a Kroupa (2001) IMF were used.

Analysis of the interstellar Medium of Isolated GAlaxies (AMIGA; Lisenfeld et al. 2011): It is a redshift-limited sample ($1500 \leq v_{rec}$ [km s^{-1}] ≤ 5000) consisting of 273 isolated galaxies with reported multi-band imaging and CO data. We perform the same procedure described above for the UNAM-KIAS sample to estimate detected and non-detected HI masses. For LTGs (ETGs) galaxies, we found 203 (11) detections. Only 4 non-detections were found, all for ETGs. The stellar masses were calculated as described above for the UNGC sample. Morphologies were obtained using higher resolution images from SDSS or their own images.

Low-mass Isolated galaxies (Bradford et al. 2015): It is a sample of 148 isolated low-mass galaxies ($7 \leq \log(M_*/M_\odot) \leq 9.5$) drawn from the SDSS NSA catalog (see Geha et al. 2012). Isolated galaxies were defined as those without massive hosts (at least 0.5 dex more massive than the given galaxy) at projected distances less than 1.5 Mpc. HI measurements were obtained using the 305 m Arecibo and the 100 m Greenbank telescopes. Stellar masses were calculated in the NSA catalog with the correct software of Blanton & Roweis (2007), using the SDSS and GALEX photometric bands and assuming a Chabrier 2003 IMF. For the morphology, we used the Huertas-Company et al. (2011) automatic clas-

sification, following the same procedure described above for the GASS survey; a classification was obtained for 128 out of the 148 galaxies; all of them are of late type. Indeed, according to Geha et al. (2012) all the isolated low-mass galaxies in the local Universe are star forming (late-type) objects.

Herschel Reference Survey – Virgo galaxies: This is the same HRS sample described above but including only galaxies from the Virgo Cluster central regions A and B (59). Therefore, this sample is biased toward galaxies in a very high density environment.

ATLAS^{3D} HI sample – Virgo core ETGs: This is the same ATLAS^{3D} sample described above but taking into account account only the Virgo core ETGs (15). Therefore, this sample is biased in favorm of ETGs in a very high density environment.

B. THE COMPILED GALAXY SAMPLES WITH CO (H_2) INFORMATION

B.1. *Golden Category*

Herschel Reference Survey (HRS)– field galaxies: It is the same sample described in § A.1 (excluding Virgo Cluster core), with 155 galaxies with available CO information (101 detections and 54 non-detections). The authors used either compiled CO observations from the literature or they carried out their own observations with the National Radio Astronomy Observatory (NRAO) Kitt Peak 12 m telescope (Boselli et al. 2014a). A MW constant or H-band luminosity-dependent (Boselli et al. 2002) CO-to-H_2 conversion factor was applied to calculate M_{H_2}.

CO Legacy Legacy Database for GASS (COLD GASS; Saintonge et al. 2011): This is a program aimed at observing CO(1-0) line fluxes with the IRAM 30 m telescope for galaxies from the GASS survey described in § A.1. From the CO fluxes, the total CO luminosities, (and hence the H_2 masses) were calculated for 349 galaxies. The authors applied the MW constant CO-to-H_2 conversion factor.

ATLAS^{3D} H_2 sample – field ETGs (Young et al. 2011): This is the same sample described in § A.1 (excluding the Virgo Cluster core) but with observations in CO using the IRAM 30 m Radio Telescope. The sample amounts for 243 ETGs with CO observations. The authors use the constant MW CO-to-H_2 conversion factor.

B.2. *Silver Category*

Stark et al. (2013) compilation: It corresponds to the same compiled galaxy sample described in § A.2. The authors observed 35 galaxies of the NFGS with the IRAM 30 m and the ARO 12 m telescopes to measure the CO ($J \rightarrow 2-1$) (IRAM) and ($J \rightarrow 1-0$) (IRAM & ARO) lines. For the other galaxies, the H_2 information from previous works was used. Stark et al. (2013) used the MW constant CO-to-H_2 factor for estimating M_{H_2}.

Leroy et al. (2008) HERACLES sample: It is the same sample described in § A.2. The H_2 information for the 23 galaxies (LTGs) comes from the CO $J \rightarrow 2-1$ maps from the HERA CO-Line Extragalactic Survey HERACLES (Leroy et al. 2008). CO $J \rightarrow 2-1$ is related to CO $J \rightarrow 1-0$ by assuming the ratio $I_{\rm CO}(2 \rightarrow 1)/I_{\rm CO}(1 \rightarrow 0) = 0.8$, and CO $J \rightarrow 1-0$ maps from the Berkeley-Illinois-Maryland Association (BIMA) Survey of Nearby Galaxies (BIMA SONG Helfer et al. 2003). The MW constant CO-to-H_2 conversion factor was used.

APEX Low-redshift Legacy Survey for MOlecular Gas: (ALLSMOG; Bothwell et al. 2014) Using the APEX telescope, the $CO(2 \rightarrow 1)$ emission line was measured to trace H_2 in 42 late-type galaxies with masses $8.5 < \log(M_*/M_\odot) < 10$, in the redshift range $0.01 < z < 0.03$ and with metallicities $12 + \log(O/H) > 8.5$. Morphological classification was taken from NED. The stellar masses were derived by SED fitting (Kauffmann et al. 2003) using the SDSS DR7 optical data. To obtain the $CO(1 \rightarrow 0)$ line luminosities, the $CO(2 \rightarrow 1)$ emission line was assumed to be fully thermalized. A MW constant, or metallicity-dependent (Wolfire et al. 2010), CO-to-H_2 conversion factor was applied to infer the H_2 masses.

Bauermeister et al. (2013) compilation: We took from this literature compilation 8 galaxies in the low-redshift range $0.05 \leq z \leq 0.1$. All of them are star forming and we associated them to LTGs. Their stellar masses are in the range $4 \times 10^{10} M_\odot \leq M_* \leq 1.6 \times 10^{11} M_\odot$ and they were calculated by fitting SDSS *ugriz* photometry to a grid of models spanning a wide range of star formation histories. The H_2 masses were obtained by the authors from CO $J \rightarrow 1-0$ intensity maps with CARMA, using a MW constant CO-to-H_2 conversion factor.

B.2. *Bronze Category*

Analysis of the interstellar Medium of Isolated GAlaxies (AMIGA; Lisenfeld et al. 2011): This is the same sample described in § A.3. The authors carried out their own observations of CO($J \rightarrow 1-0$) with the IRAM 30 m or the 14 m FCRAO telescopes for 189 galaxies; 87 more were

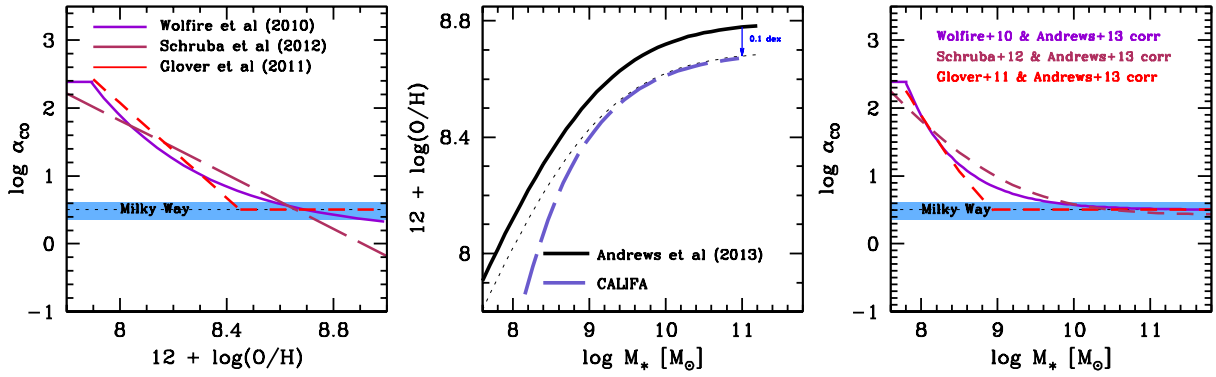

Fig. 16. *Left panel:* Dependence of the CO-to-H_2 factor on gas-phase metallicity as given by physical models (Wolfire et al. 2010; Glover & Mac Low 2011) calibrated by observations and by a purely empirical approach (Schruba et al. 2012). Observations do not allow to constrain these relations for metallicities lower than $12 + \log_{10}(\mathrm{O/H}) \approx 7.9$ *Middle panel:* Dependence of metallicity on mass according to the CALIFA (Sánchez et al. 2013) and SDSS (Andrews & Martini 2013) surveys. We use an updated relation for CALIFA that includes more galaxies, especially at small masses (S. Sanchez, priv. communication); the masses were corrected from the Salpeter to the Chabrier IMF. The dotted line shows the SDSS relation lowered by 0.1 dex to correct for the aperture effect; notice how well it agrees with the CALIFA relation, but it extends to smaller masses, so this is the relation we use. *Right panel:* Dependence of the CO-to-H_2 factor on mass inferred from the α_{CO}–M_* and $Z - M_*$ dependences plotted in the other panels. The color figure can be viewed online.

compiled from the literature. An aperture correction was applied to the CO data. A MW constant CO-to-H_2 conversion factor was used to compute M_{H_2}.

Herschel Reference Survey – Virgo core: This is the same HRS sample described above but including only the Virgo Cluster core regions A and B galaxies (62). Therefore, this sample is biased toward galaxies in a very high density environment.

ATLAS^{3D} H_2 sample – Virgo core ETGs: This is the same ATLAS^{3D} sample described above but taking into account account only the Virgo core ETGs (21). Therefore, this sample is biased toward ETGs in a very high density environment.

C. THE CO-to-H_2 CONVERSION FACTOR

Several authors have shown that the CO-to-H_2 conversion factor depends on the gas phase metallicity, Z (see e.g., Boselli et al. 2002; Schruba et al. 2012; Narayanan et al. 2012; Bolatto et al. 2013). In a recent review on the topic, among the several approaches for determining the dependence of α_{CO} on Z in galaxies, Bolatto et al. (2013) recommend a prescription based on a local physical model for the H_2 and CO production and calibrate it with extragalactic observations. In particular, they find that the prescription given in Wolfire et al. (2010), based on photodissociation models with shielding, is the most consistent with the scarce observational data that provide α_{CO} vs. Z in galaxies. According to Wolfire et al. (2010):

$$\alpha_{\mathrm{CO}} = \alpha_{\mathrm{CO,MW}} \exp\left[\frac{+4.0\Delta A_V}{Z' \bar{A}_{V,MW}}\right] \exp\left[\frac{-4.0\Delta A_V}{\bar{A}_{V,MW}}\right] \quad \text{(C9)}$$

where $\alpha_{\mathrm{CO,MW}} = 3.2$ (in units $M_\odot$ pc^{-2}/K km s^{-1}) is the adopted conversion factor for the Milky Way, $Z' = Z/Z_\odot$ where $Z \equiv 12 + \log_{10}(\mathrm{O/H})$, $\Delta A_V \approx 1$, and $\bar{A}_{V,MW}$ is the mean extinction through a giant molecular cloud with Milky Way metallicity $Z_\odot$, with $\bar{A}_{V,MW} \approx 5$ for $\Sigma_{\mathrm{GMC}} \approx 100\ M_\odot \mathrm{pc}^{-2}$. According to equation (C9), $\alpha_{\mathrm{CO}} \approx \alpha_{\mathrm{CO,MW}}$ for $Z \gtrsim Z_\odot$. The left panel of Figure 16 shows the Wolfire et al. (2010) relation along with those of Glover & Mac Low (2011) and Schruba et al. (2012).

To relate α_{CO} to stellar mass, we use the mass-metallicity relation for galaxies in the local Universe. Sánchez et al. (2013) and Andrews & Martini (2013) determined the mass-metallicity relation for galaxies using the CALIFA and SDSS surveys in the stellar mass range $8.4 \leq \log(M_*/M_\odot) \leq 11.2$ and $7.4 \leq \log(M_*/M_\odot) \leq 11.2$, respectively. The work by Sánchez et al. (2013) provides a more reliable estimate of the mass-metallicity relation; recall that the SDDS galaxies are mapped by only one central fiber of fixed aperture, while CALIFA maps the whole galaxies with many integral field units. However, the mass range in the CALIFA sample is limited, while that of Andrews & Martini (2013) extends to very

low masses. We use an updated version of the CALIFA mass-metallicity relation (S. F. Sanchez, priv. communication) and correct M_* to pass from the Salpeter IMF to the Chabrier one used in Andrews & Martini (2013). At the mass range where both studies coincide, they agree modulo a shift in the SDSS relation by $\approx +0.1$ dex in metallicity compared to the CALIFA one (see the middle panel of Figure 16). This is expected given that CALIFA covers the galaxies up to 2-3 effective radii while SDSS in most cases covers only the central regions which are typically more metallic than the outer ones (see for a discussion Sánchez et al. 2013). Thus, we use the relation as reported in Andrews & Martini (2013) but lowering it by 0.1 dex. They found that the function proposed by Moustakas et al. (2011) fits well their observational results:

$$12 + \log_{10}(\mathrm{O/H}) = (12 + \log_{10}(\mathrm{O/H})_{\mathrm{asm}}) - \log_{10}\left(1 + \left(\frac{M_{TO}}{M_*}\right)^{\gamma}\right), \quad \text{(C10)}$$

with $12+\log_{10}(\mathrm{O/H})_{\mathrm{asm}} = 8.798$ (we use 8.698, after subtracting 0.1 dex), $M_{TO} = 8.901$, and $\gamma = 0.640$.

Combining equations (C9) and (C10), we are able to obtain the mean α_{CO}–M_* relation. In fact, the metallicity in any calibration is one of the hardest astronomical quantities to measure with precision. However, for our purpose, given the large uncertainties and scatter the exact calibration is not relevant; the average dependence of the α_{CO} factor with mass is sufficient. Following Bolatto et al. (2013), we actually normalize the α_{CO}–M_* dependence to $\alpha_{\mathrm{CO}} = \alpha_{\mathrm{CO,MW}}$ at $M_* = 3 \times 10^{10} M_\odot$, corresponding to a metallicity slightly lower than $Z_\odot$. For larger masses (metallicities), we assume that this value remains constant, and for smaller masses, we use the mass dependence given by the combination of equations (C9) and (C10):

$$\log(\alpha_{\mathrm{CO}}) = 0.15 + 0.35\left[1 + 0.1\left(\frac{3 \times 10^{10} M_\odot}{M_*}\right)^{0.64}\right]. \quad \text{(C11)}$$

This equation is valid roughly down to $M_* \approx 10^8 M_\odot$, which corresponds to metallicities ≈ 0.8 dex below the solar one (or $12 + \log(\mathrm{O/H}) \approx 7.9$); there are no observational determinations of α_{CO} at lower metallicities. Therefore, for $M_* < 10^8\ M_\odot$, we use the same value of α_{CO} at $10^8 M_\odot$, i.e., $\alpha_{\mathrm{CO}} \approx 250$. Besides, as highlighted in Bolatto et al. (2013), as one moves to increasingly low metallicities, the use of CO emission to quantify the H_2 reservoir becomes increasingly an extrapolation and eventually a practical floor should appear past which CO is not a useful tracer of total H_2 mass; rather, CO will be a tracer of high column density peaks and well-shielded regions.

The above α_{CO}–M_* dependence is applied to LTGs. The right panel of Figure 16 shows this dependence along with those calculated from the $\alpha_{\mathrm{CO}} - Z$ dependences from Glover & Mac Low (2011) and Schruba et al. (2012). For ETGs, which typically have higher metallicities than $Z_\odot$, we assume $\alpha_{\mathrm{CO}} = \alpha_{\mathrm{CO,MW}} =$ const. at all masses.

D. CORRECTIONS TO THE UPPER LIMITS OF ETGs

In § 5, we noted that the upper limits reported for the GASS (HI) and COLD GASS (H_2) samples in the case of ETGs are significantly larger than those reported for the ATLAS^{3D} or HRS samples. Following Serra et al. (2012), we corrected the ATLAS^{3D} upper limit values by a factor of two in order to take into account differences between the different telescopes and signal-to-noise thresholds used in this survey and in GASS (see § 5). However, the main reason for the differences in the upper limits among these samples is a selection effect due to the different volumes covered by them. To illustrate this, in the left panel of Figure 17 we plot the histogram of HI masses for ETGs in the $10.10 - 10.65 \log M_*$ bin for GASS (solid black line) and ATLAS^{3D} (dotted black line). Non-detections are also included, with values of M_{HI} corresponding to their upper limits. The red lines show the histograms of detections only. The number of GASS ETGs increases as M_{HI} decreases and it has a peak at $\log(M_{HI}/M_\odot) \approx 8.4 - 9.0$, contributed mainly by the upper limits and consistent with the sensitivity limit of the ALFALFA survey at the distances of the GASS galaxies in the mentioned stellar mass range. For ATLAS^{3D}, with distances much closer than GASS, some ETGs are detected in HI with masses lower than $\log(M_{HI}/M_\odot)$=8.4, but most of them are actually undetected, having upper limits 1–1.5 orders of magnitude lower than in the case of GASS, which is consistent with the distance differences between both samples. The main difference between the M_{HI} distributions of both samples is their upper limits, and this is clearly due to a selection effect imposed by the different distance ranges of these samples. Basically, if the undetected GASS ETGs were at the distances of ATLAS^{3D} ETGs, then probably most of them would not be detected in HI, having upper limits lower by 1–1.5 orders of magnitude. Thus, the high values of their upper limits

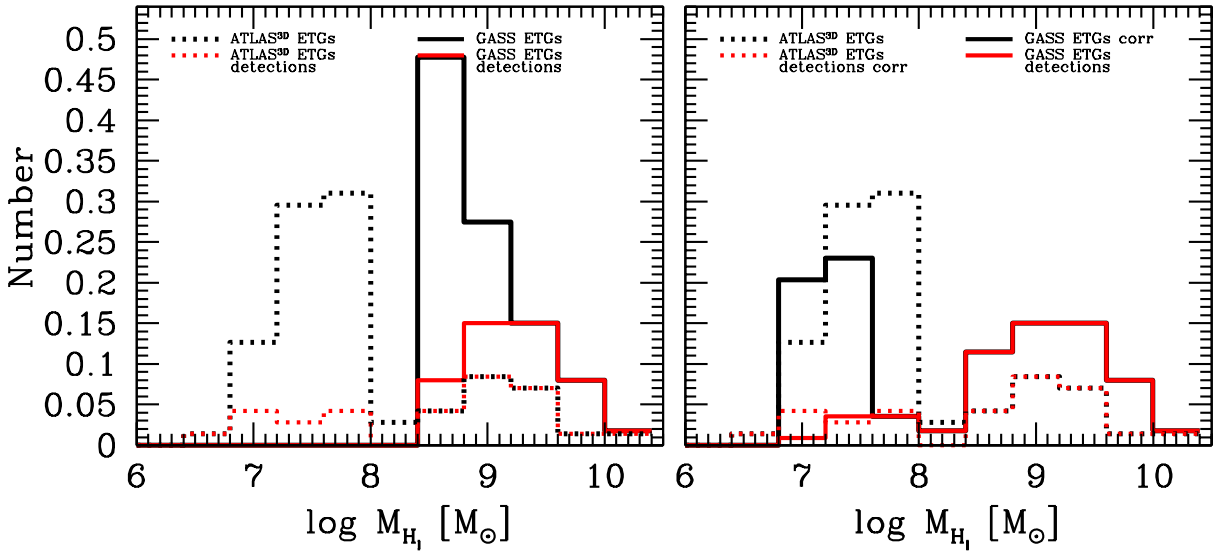

Fig. 17. *Left panel:* Distributions of HI masses for ETGs in the $10.10 - 10.65$ $\log M_*$ bin for GASS (solid black line) and ATLAS^{3D} (dashed black line). Non-detections are also included, with values of M_{HI} corresponding to their upper limits (for ATLAS^{3D}, we use the upper limits increased by a factor of two as explained in § 5). The red lines show the contribution of detected galaxies. The GASS distribution is clearly limited to much higher upper limits than ATLAS^{3D}, and this is mainly due to a distance selection effect. *Right panel:* Same as left panel but after correcting the upper limits of GASS according to the observations of ATLAS^{3D}. The color figure can be viewed online.

imposed by the volume of GASS, are expected to introduce a bias in the determination of the gas-to-stellar mass correlations of ETGs.

In an attempt to correct for this selection effect, we will assume that the ETGs in the GASS and ATLAS^{3D} (and HRS, too) samples are representative of the same local ETG populations. Then, the fact that the upper limits for the ATLAS^{3D} (or HRS) ETGs are significantly lower than those *of similar stellar mass* galaxies from GASS, is mainly due to the distance differences among these samples. If the GASS ETGs were as near as the ATLAS^{3D} ones, then the upper limit region in the plots of HI-to-stellar mass ratio vs. M_* would be on average lower by a factor equal to the squared distance ratio. Thus, to homogenize the upper limits in $R_{\rm HI}$ given by the GASS and ATLAS^{3D} samples, we decrease the upper limits of the galaxies in the volume-limited sample with more distant galaxies (GASS) by $(D_i/\bar{D}_{\rm ATLAS^{3D}})^2$, where D_i is the distance of each GASS ETG and $\bar{D}_{\rm ATLAS^{3D}} = 25$ Mpc is the average distance of the ATLAS^{3D} ETGs. In fact, according to the ATLAS^{3D} observations, 25% of ETGs below the upper limit region of GASS were detected (see for an example Figure 17). Therefore, we decrease the GASS upper limits as mentioned above for 75% of the galaxies, and for the remaining ones we assign randomly an $R_{\rm HI}$ value between its upper limit and the average upper limit of the ATLAS^{3D} galaxies at the corresponding stellar mass. The same procedure is applied to the COLD GASS ETGs for the $R_{\rm H_2}$ upper limits, where the corresponding $\bar{D}_{\rm ATLAS^{3D}}$ for COLD GASS is 26 Mpc.

The right panel of Figure 17, shows the same histograms as the left panel but now the upper limits of the GASS sample are corrected as explained above. Observe how similar are now the upper limit distributions of GASS and ATLAS^{3D} galaxies after correcting for the distance selection effect. Further, we use a large mock galaxy catalog to test the procedure applied here to the GASS (or COLD GASS) upper limits to homogenize them with those of nearby samples such as ATLAS^{3D}. The mock catalog is a volume-limited sample (up to 313 Mpc) of 5×10^6 galaxies that represents well the observational GSMF and LTG/ETG fractions as a function of M_* (see § 6). We assign HI masses to each LTG/ETG galaxy by using an *input* $R_{\rm HI}$ distribution for a given M_* (a $R_{\rm HI}$–M_* relation and its scatter) for LTGs and ETGs. Distances are assigned assuming an isotropic distribution within a sphere of the same radius of the volume sampled. Note that we ignore any clustering properties of the galaxies. This is a safe assumption as we are only interested in the selection effects introduced by the detection limits of the GASS and ATLAS^{3D} samples. Then, we select ETGs more

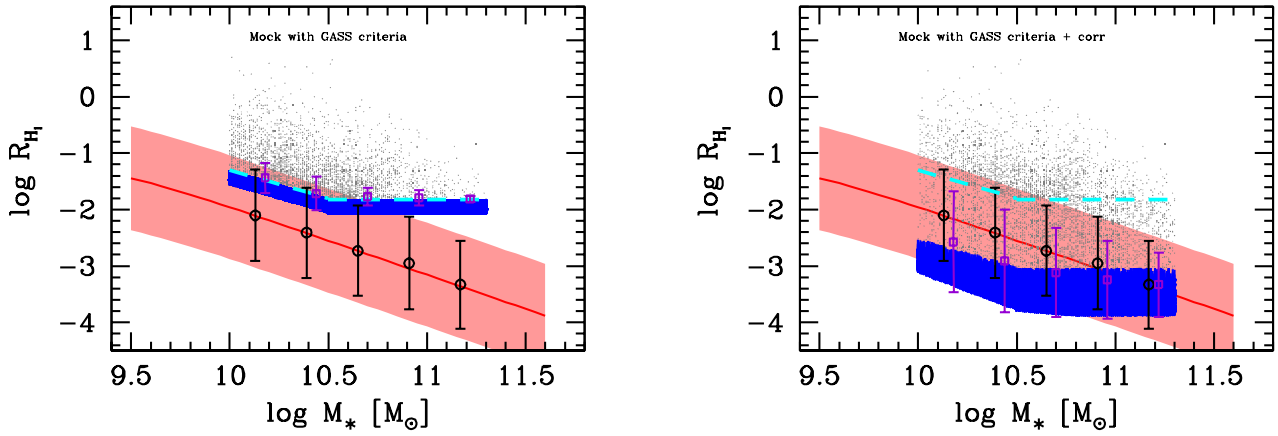

Fig. 18. *Left panel:* ETGs from our $109 < D < 222$ Mpc volume mock catalog in the $R_{\rm HI}$ vs. M_* plane, following the selection and $R_{\rm HI}$ limits of GASS. All mock ETGs below the GASS $R_{\rm HI}$ limits (dashed line) are assumed as undetected and assigned an $R_{\rm HI}$ value equal to the $R_{\rm HI}$ limit (upper limit; blue arrows). The magenta squares with error bars show the mean and standard deviation calculated in different mass bins with the Kaplan-Meier estimator. The $R_{\rm HI}$–M_* correlation for ETGs used in the generation of the mock catalog is plotted with the red solid line and shaded area. The circles with error bars show the mean and standard deviation calculated in different mass bins for all the ETGs from the mock catalog. The mock catalog samples very well the input correlation but this is not the case when the $R_{\rm HI}$ limit of GASS is imposed, even when using the Kaplan-Meier estimator to take into account the upper limits. *Right panel:* Same as left panel but after applying our ATLAS$^{\rm 3D}$-based corrections to the upper limits of GASS (see text). The mean and standard deviation in the different mass bins, taking into account the (corrected) upper limits, now follow closely the input correlation. The color figure can be viewed online.

massive than 10^{10} $M_\odot$ in the $109 < D < 222$ Mpc range (the GASS volume), and impose upper limits to the $R_{\rm HI}$ ratio as a function of mass like the one of GASS (see Catinella et al. 2012). Then, we calculate the mean $R_{\rm HI}$ and its standard deviation taking into account the upper limits in mass bins as we did for the observational sample (using the Kaplan-Meier estimator). The question now is whether we recover the input $R_{\rm HI}$–M_* correlation for ETGs.

In the left panel of Figure 18, we plot our input $R_{\rm HI}$–M_* correlation for ETGs (for this exercise, it is described by a double power-law function with the parameters given in Table 6 and assuming a log normal scatter) along with the values from the mock catalog in the $109 < D < 222$ Mpc volume and imposing the sensitivity limit of the GASS sample (dots). All the dots below this limit are plotted as upper limits (blue arrows); they populate the imposed sensitivity limit in the $R_{\rm HI}$ vs M_* diagram. The open circles with error bars are the mean and standard deviation calculated directly from the catalog in $\log M_*$ bins for ETGs in the $109 < D < 222$ Mpc volume, while the magenta squares and error bars are the same means and standard deviations calculated with the Kaplan-Meier estimator after imposing the GASS sensitivity limit. Thus, after imposing this limit, *the recovered correlation is far from the input one.*

Then, we apply the same corrections we have used for the real GASS data, based on the information from the ATLAS$^{\rm 3D}$ sample, i.e., the GASS-like imposed upper limits to the mock catalog galaxies were decreased by $D_i^2[{\rm Mpc}^2]/25^2{\rm Mpc}^2$ in 75% of the cases, and for the remaining, a random detection value for $R_{\rm HI}$ was assigned as explained above. The right panel of Figure 18 shows the result of these corrections along with the mean and standard deviations calculated with the corrected data in the same three mass bins as in the left panel (magenta squares with error bars). Note that after our corrections, the calculated mean and standard deviation in each mass bin are in better agreement with those corresponding to the mock catalog without any selection, that is, the input $R_{\rm HI}$–M_* correlation is reasonably well recovered; this shows the necessity of applying the mentioned corrections.

The effect of introducing the mentioned correction to the GASS and COLD GASS upper limits for the determination of the HI- and H_2-to-stellar mass correlations of ETGs is, of course, not as significant as in the experiment shown in Figure 18 because these samples are not the only ones used (§ 2.2 and 2.3). In Tables 5 and 6 (cases ETG$^{\rm ndc}$),

we present the fitted HI-to-stellar mass correlation for ETGs for the case where the upper limits of the GASS sample were not corrected by distance. The double power-law correlation, without the correction, changes slightly at the large-mass end: it is shallower but with a much larger scatter than when we made the correction; the latter is expected due to the strong segregation of the upper limits from COLD GASS and from the less distant ATLAS^3D and HRS samples. The single power-law would be shallower. Similarly, in these tables is also shown the fitted H_2-to-stellar correlation for ETGs when the upper limits of the COLD GASS sample were not corrected by distance. The relations are actually almost the same whether or not the correction is taken into account, but the scatter is larger at the large-mass end for the latter case, as expected due to the segregation of the upper limits from COLD GASS and from the less distant ATLAS^{3D} and HRS samples.

E. OBSERVATIONAL ERRORS

To provide a rough estimate of the intrinsic scatter around the $R_{\rm HI}$–M_* and $R_{\rm H_2}$–M_* correlations shown in § 4.2 and 4.3, estimates of the (statistical) observational errors, $\sigma_{\rm err}$, in the determination of $R_{\rm HI}$ and $R_{\rm H_2}$ are necessary. For this, we need to know the respective observational uncertainties in the determination of the stellar, HI, and H_2 masses.

Most of the observational sources included in our compilation do not report the individual errors in the determination of these masses, but they report conservative average estimates for them.

For the stellar mass, the observational errors are typically estimated to be 0.1 dex (see e.g., Conroy 2013). After homogenizing the samples to a fixed IMF (Chabrier 2003) we made the conservative assumption that other sources of systematic errors in the determination of M_* are negligible, see § 2.1.3. For the HI mass, a combination of the statistical errors, distance uncertainties, and errors associated with the absolute 21cm flux scale calibration accounts for a total observational error of ≈ 0.1 dex. Therefore, the average error in $\log R_{\rm HI}$ is ≈ 0.14 dex. For the H_2 mass, most of the works used in our compilation report average observational errors of $0.2-0.25$ dex. The uncertainty in the $\alpha_{\rm CO}$ parameter has been taken into account; however, it was probably significantly underestimated. In a recent review on the subject, Boselli et al. (2014a) suggest that this uncertainty is actually of the order of 0.3 dex. Thus, considering that the observational errors in the CO flux account for 30% (0.11 dex; e.g., Boselli et al. 2014a), and the uncertainty in the $\alpha_{\rm CO}$ parameter is 0.3 dex, an estimate of the typical error in $\log M_{H_2}$ is 0.32 dex. The estimated error in $\log R_{\rm H_2}$ is then ≈ 0.34 dex, using an error of 0.1 dex in $\log M_*$.

TABLE 8

GSMF PARAMETERS

α_1	$\log(M_1^*)$ ($M_\odot$)	$\log(\phi_1^*)$ (Mpc^{-3} dex^{-1})	α_2	$\log(M_2^*)$ ($M_\odot$)	$\log(\phi_2^*)$ (Mpc^{-3} dex^{-1})	β
-1.47	9.74	-2.66	0.07	8.84	-2.66	0.37

F. CALCULATION OF THE GSMF

Here we outline how we construct our GSMF in a large mass range following Kravtsov et al. (2014). For large masses, the SDSS-based GMSF presented in Bernardi et al. (2013) is used. These authors reanalyzed the photometry of the SDSS DR7, taking special care of the background estimate of extended luminous galaxies (see also Simard et al. 2011; He et al. 2013; Mendel et al. 2014; D'Souza et al. 2015; Meert et al. 2016); after this reanalysis, the high-end of the luminosity (mass) function becomes shallower. Their GSMF is well fitted by a Schechter + sub exponential Schechter function. For small masses, the GSMFs determined by Baldry et al. (2012, from the GAMA survey) were used. These authors analyzed low redshift samples that contain low luminosity galaxies, but a correction for surface brightness incompleteness was not applied. So, their determinations at $M_* \lesssim 10^8\ M_\odot$ are actually lower limits. This GSMF is well fitted by double Schechter function. Both high and low mass GSMFs assume the Chabrier (2003) IMF to estimate M_*. However, the masses in Bernardi et al. (2013) were calculated using the Bell et al. (2003) mass-to-luminosity ratios, who employed the PEGASE stellar population synthesis models (Fioc & Rocca-Volmerange 1997). In Baldry et al. (2012) the masses were calculated using the Bruzual & Charlot (2003, BC03) models. Conroy (2013) has shown that the former are systematically larger than the latter by $\approx 0.10 - 0.14$ dex. Therefore, for the Bernardi et al. (2013) GSMF, we decrease uniformly M_* by 0.12 dex to homogenize the masses to the BC03 population synthesis model.

Thus, we use a fit to the Baldry et al. (2012) combined with the Bernardi et al. (2013) GSMF fit corrected by 0.12 dex in mass to obtain a GSMF that spans from $M_* \approx 10^7$ to $10^{12}\ M_\odot$. The match of both fits (at the mass where the latter becomes higher than the former) takes place at $M_* \approx 10^{9.3}\ M_\odot$. The obtained GSMF is well fitted by the combination of a Schechter function and a sub exponen-

tial Schechter function. The respective parameters are given in Table 8. See Figure 12 for the plotted GSMF and its comparison to other GSMFs from the literature.

REFERENCES

Accurso, G., Saintonge, A., Catinella, B., et al. 2017, MNRAS, 470, 4750
Andrews, B. H. & Martini, P. 2013, ApJ, 765, 140
Avila-Reese, V., Zavala, J., Firmani, C., & Hernández-Toledo, H. M. 2008, AJ, 136, 1340
Baldry, I. K., Driver, S. P., Loveday, J., et al. 2012, MNRAS, 421, 621
Baldry, I. K., Glazebrook, K., & Driver, S. P. 2008, MNRAS, 388, 945
Barnes, D. G., Staveley-Smith, L., de Blok, W. J. G., et al. 2001, MNRAS, 322, 486
Bauermeister, A., Blitz, L., Bolatto, A., et al. 2013, ApJ, 768, 132
Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57
Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289
Bernardi, M., Meert, A., Sheth, R. K., et al. 2013, MNRAS, 436, 697
Blanton, M. R., Eisenstein, D., Hogg, D. W., Schlegel, D. J., & Brinkmann, J. 2005a, ApJ, 629, 143
Blanton, M. R., Lupton, R. H., Schlegel, D. J., et al. 2005b, ApJ, 631, 208
Blanton, M. R. & Moustakas, J. 2009, ARA&A, 47, 159
Blanton, M. R. & Roweis, S. 2007, AJ, 133, 734
Blitz, L. & Rosolowsky, E. 2006, ApJ, 650, 933
Blyth, S., van der Hulst, T. M., Verheijen, M. A. W., et al. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 128
Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207
Boselli, A., Cortese, L., & Boquien, M. 2014a, A&A, 564, A65
Boselli, A., Cortese, L., Boquien, M., et al. 2014b, A&A, 564, A67
Boselli, A., Cortese, L., Boquien, M., et al. 2014c, A&A, 564, A66
Boselli, A., Eales, S., Cortese, L., et al. 2010, PASP, 122, 261
Boselli, A. & Gavazzi, G. 2006, PASP, 118, 517
Boselli, A., Lequeux, J., & Gavazzi, G. 2002, A&A, 384, 33
Bothwell, M. S., Wagg, J., Cicone, C., et al. 2014, MNRAS, 445, 2599
Bradford, J. D., Geha, M. C., & Blanton, M. R. 2015, ApJ, 809, 146
Brown, T., Catinella, B., Cortese, L., et al. 2015, MNRAS, 452, 2479
Brown, T., Catinella, B., Cortese, L., et al. 2017, MNRAS, 466, 1275
Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000
Buckley, J. & James, I. 1979, Biometrika, 66, 429
Butcher, Z., Schneider, S., van Driel, W., Lehnert, M. D., & Minchin, R. 2016, A&A, 596, A60
Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, MNRAS, 413, 813
Carilli, C. L. & Rawlings, S. 2004, NAR, 48, 979
Catinella, B., Schiminovich, D., Cortese, L., et al. 2013, MNRAS, 436, 34
Catinella, B., Schiminovich, D., Kauffmann, G., et al. 2012, A&A, 544, A65
Chabrier, G. 2003, PASP, 115, 763
Conroy, C. 2013, ARA&A, 51, 393
Conroy, C. & Wechsler, R. H. 2009, ApJ, 696, 620
Cortese, L., Boissier, S., Boselli, A., et al. 2012, A&A, 544, A101
Cortese, L., Catinella, B., Boissier, S., Boselli, A., & Heinis, S. 2011, MNRAS, 415, 1797
Deng, X.-F. 2013, Research in Astronomy and Astrophysics, 13, 651
Dressler, A. 1980, ApJ, 236, 351
D'Souza, R., Vegetti, S., & Kauffmann, G. 2015, MNRAS, 454, 4027
Duffy, A. R., Kay, S. T., Battye, R. A., et al. 2012, MNRAS, 420, 2799
Elmegreen, B. G. 1989, ApJ, 338, 178
Feigelson, E. D. & Babu, G. J. 2012, Modern Statistical Methods for Astronomy: With R Applications (Cambridge)
Feigelson, E. D. & Nelson, P. I. 1985, ApJ, 293, 192
Fioc, M. & Rocca-Volmerange, B. 1997, A&A, 326, 950
Fu, J., Guo, Q., Kauffmann, G., & Krumholz, M. R. 2010, MNRAS, 409, 515
Fukugita, M., Nakamura, O., Okamura, S., et al. 2007, AJ, 134, 579
Garnett, D. R. 2002, ApJ, 581, 1019
Gavazzi, G., Boselli, A., van Driel, W., & O'Neil, K. 2005, A&A, 429, 439
Geha, M., Blanton, M. R., Masjedi, M., & West, A. A. 2006, ApJ, 653, 240
Geha, M., Blanton, M. R., Yan, R., & Tinker, J. L. 2012, ApJ, 757, 85
Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2005, AJ, 130, 2598
Glover, S. C. O. & Mac Low, M.-M. 2011, MNRAS, 412, 337
Haynes, M. P. & Giovanelli, R. 1984, AJ, 89, 758
Haynes, M. P., Giovanelli, R., Martin, A. M., et al. 2011, AJ, 142, 170
He, Y. Q., Xia, X. Y., Hao, C. N., et al. 2013, ApJ, 773, 37
Helfer, T. T., Thornley, M. D., Regan, M. W., et al. 2003, ApJS, 145, 259
Hernández-Toledo, H. M., Vázquez-Mata, J. A., Martínez-Vázquez, L. A., Choi, Y.-Y., & Park, C. 2010, AJ, 139, 2525
Huang, S., Haynes, M. P., Giovanelli, R., & Brinchmann, J. 2012a, ApJ, 756, 113

Huang, S., Haynes, M. P., Giovanelli, R., et al. 2012b, AJ, 143, 133
Huang, S., Haynes, M. P., Giovanelli, R., et al. 2014, ApJ, 793, 40
Huertas-Company, M., Aguerri, J. A. L., Bernardi, M., Mei, S., & Sánchez Almeida, J. 2011, A&A, 525, A157
Jansen, R. A., Fabricant, D., Franx, M., & Caldwell, N. 2000a, ApJS, 126, 331
Jansen, R. A., Franx, M., Fabricant, D., & Caldwell, N. 2000b, ApJS, 126, 271
Johnston, S., Taylor, R., Bailes, M., et al. 2008, Experimental Astronomy, 22, 151
Jones, M. G., Haynes, M. P., Giovanelli, R., & Moorman, C. 2018, MNRAS
Jones, M. G., Papastergis, E., Haynes, M. P., & Giovanelli, R. 2016, MNRAS, 457, 4393
Kannappan, S. J. & Gawiser, E. 2007, ApJ, 657, L5
Kannappan, S. J., Guie, J. M., & Baker, A. J. 2009, AJ, 138, 579
Kannappan, S. J., Stark, D. V., Eckert, K. D., et al. 2013, ApJ, 777, 42
Kaplan, E. L. & Meier, P. 1958, Journal of the American Statistical Association, 53, 457
Karachentsev, I. D., Kaisina, E. I., & Makarov, D. I. 2014, AJ, 147, 13
Karachentsev, I. D., Makarov, D. I., & Kaisina, E. I. 2013, AJ, 145, 101
Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 54
Kauffmann, G., White, S. D. M., Heckman, T. M., et al. 2004, MNRAS, 353, 713
Kennicutt, Jr., R. C., Armus, L., Bendo, G., et al. 2003, PASP, 115, 928
Keres, D., Yun, M. S., & Young, J. S. 2003, ApJ, 582, 659
Klypin, A., Karachentsev, I., Makarov, D., & Nasonova, O. 2015, MNRAS, 454, 1798
Kravtsov, A., Vikhlinin, A., & Meshscheryakov, A. 2014, ApJ submitted, (arXiv:1401.7329)
Kroupa, P. 2001, MNRAS, 322, 231
Kroupa, P., Tout, C. A., & Gilmore, G. 1993, MNRAS, 262, 545
Krumholz, M. R., McKee, C. F., & Tumlinson, J. 2009, ApJ, 693, 216
Lagos, C. D. P., Baugh, C. M., Lacey, C. G., et al. 2011, MNRAS, 418, 1649
Lagos, C. d. P., Crain, R. A., Schaye, J., et al. 2015, MNRAS, 452, 3815
Lagos, C. d. P., Davis, T. A., Lacey, C. G., et al. 2014, MNRAS, 443, 1002
Lee, E. T. & Wang, J. W. 2003, Statistical Methods for Survival Data Analysis, (New York, NY: Wiley)
Lemonias, J. J., Schiminovich, D., Catinella, B., Heckman, T. M., & Moran, S. M. 2013, ApJ, 776, 74
Leroy, A. K., Walter, F., Brinks, E., et al. 2008, AJ, 136, 2782
Lisenfeld, U., Espada, D., Verdes-Montenegro, L., et al. 2011, A&A, 534, A102
Maddox, N., Hess, K. M., Obreschkow, D., Jarvis, M. J., & Blyth, S.-L. 2015, MNRAS, 447, 1610
Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006, MNRAS, 368, 715
Mandelbaum, R., Wang, W., Zu, Y., et al. 2016, MNRAS, 457, 3200
Martin, A. M., Papastergis, E., Giovanelli, R., et al. 2010, ApJ, 723, 1359
McGaugh, S. S. 2005, ApJ, 632, 859
Meert, A., Vikram, V., & Bernardi, M. 2016, MNRAS, 455, 2440
Mendel, J. T., Simard, L., Palmer, M., Ellison, S. L., & Patton, D. R. 2014, ApJS, 210, 3
Meyer, M. J., Zwaan, M. A., Webster, R. L., Staveley-Smith, L., Ryan-Weber, E., Drinkwater, M. J., Barnes, D. G., et al. 2004, MNRAS, 350, 1195
Moffett, A. J., Ingarfield, S. A., Driver, S. P., et al. 2016, MNRAS, 457, 1308
More, S., van den Bosch, F. C., Cacciato, M., et al. 2011, MNRAS, 410, 210
Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121
Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, ApJ, 710, 903
Moustakas, J., Coil, A. L., Aird, J., et al. 2013, ApJ, 767, 50
Moustakas, J., Zaritsky, D., Brown, M., et al. 2011, arXiv:1112.3300
Narayanan, D., Krumholz, M. R., Ostriker, E. C., & Hernquist, L. 2012, MNRAS, 421, 3127
Noordermeer, E., van der Hulst, J. M., Sancisi, R., Swaters, R. A., & van Albada, T. S. 2005, A&A, 442, 137
Obreschkow, D. & Rawlings, S. 2009, MNRAS, 394, 1857
Papastergis, E., Cattaneo, A., Huang, S., Giovanelli, R., & Haynes, M. P. 2012, ApJ, 759, 138
Paturel, G., Petit, C., Prugniel, P., et al. 2003, VizieR Online Data Catalog, 7237, 0
Pforr, J., Maraston, C., & Tonini, C. 2012, MNRAS, 422, 3285
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1996, Numerical Recipes in Fortran 90: The Art of Parallel Scientific Computing (New York, NY: CUP)
Rodríguez-Puebla, A., Avila-Reese, V., & Drory, N. 2013, ApJ, 767, 92
Rodríguez-Puebla, A., Avila-Reese, V., Yang, X., et al. 2015, ApJ, 799, 130
Rodríguez-Puebla, A., Primack, J. R., Avila-Reese, V., & Faber, S. M. 2017, MNRAS, 470, 651
Saintonge, A., Kauffmann, G., Kramer, C., et al. 2011, MNRAS, 415, 32
Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267
Sánchez, S. F., Rosales-Ortega, F. F., Jungwiert, B., et al. 2013, A&A, 554, A58
Schruba, A., Leroy, A. K., Walter, F., et al. 2012, AJ, 143, 138

Serra, P., Oosterloo, T., Morganti, R., et al. 2012, MNRAS, 422, 1835
Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, ApJS, 196, 11
Somerville, R. S. & Davé, R. 2015, ARA&A, 53, 51
Springob, C. M., Haynes, M. P., Giovanelli, R., & Kent, B. R. 2005, ApJS, 160, 149
Stark, D. V., Kannappan, S. J., Wei, L. H., et al. 2013, ApJ, 769, 82
Stewart, K. R., Bullock, J. S., Wechsler, R. H., & Maller, A. H. 2009, ApJ, 702, 307
Swaters, R. A. & Balcells, M. 2002, A&A, 390, 863
Toomre, A. 1964, ApJ, 139, 1217
van Driel, W., Butcher, Z., Schneider, S., et al. 2016, A&A, 595, A118
Verheijen, M. A. W. 1997, PhD thesis, Univ. Groningen, The Netherlands
Walter, F., Brinks, E., de Blok, W. J. G., et al. 2008, AJ, 136, 2563
Wei, L. H., Kannappan, S. J., Vogel, S. N., & Baker, A. J. 2010, ApJ, 708, 841
Welch, G. A., Sage, L. J., & Young, L. M. 2010, ApJ, 725, 100
Willick, J. A., Strauss, M. A., Dekel, A., & Kolatt, T. 1997, ApJ, 486, 629
Wolfire, M. G., Hollenbach, D., & McKee, C. F. 2010, ApJ, 716, 1191
Wright, A. H., Robotham, A. S. G., Driver, S. P., et al. 2017, MNRAS, 470, 283
Yang, X., Mo, H. J., van den Bosch, F. C., et al. 2007, ApJ, 671, 153
Young, J. S., Xie, S., Tacconi, L., et al. 1995, ApJS, 98, 219
Young, L. M., Bureau, M., Davis, T. A., et al. 2011, MNRAS, 414, 940
Zhang, W., Li, C., Kauffmann, G., et al. 2009, MNRAS, 397, 1243
Zibetti, S., Charlot, S., & Rix, H.-W. 2009, MNRAS, 400, 1181
Zu, Y. & Mandelbaum, R. 2015, MNRAS, 454, 1161
Zwaan, M. A., Meyer, M. J., Staveley-Smith, L., & Webster, R. L. 2005, MNRAS, 359, L30

Vladimir Avila-Reese, Angel R. Calette, Héctor Hernández-Toledo, and Aldo Rodríguez-Puebla: Instituto de Astronomía, Universidad Nacional Autónoma de México, A. P. 70-264, 04510, Ciudad de México, México (avila, acalette, hector, apuebla@astro.unam.mx).
Emmanouil Papastergis: Kapteyn Astronomical Institute, University of Groningen, Landleven 12, Groningen NL-9747AD, The Netherlands, and Credit Risk Modeling Department, Coöperative Rabobank U. A., Croeselaan 18, Utrecht NL-3521CB, The Netherlands (papastergis@astro.rug.nl).