UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE POSGRADO EN ASTRONOMÍA

INSTITUTO DE ASTRONOMÍA

CONEXIÓN HALO - GALAXIA: FRACCIÓN BARIÓNICA DE GALAXIAS ROJAS Y AZULES

TESIS

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS (ASTRONOMÍA)

PRESENTA:
ÁNGEL RUBÉN CALETTE MORÍN

TUTOR:
VLADIMIR ANTÓN ÁVILA REESE
INSTITUTO DE ASTRONOMÍA

MÉXICO D.F., OCTUBRE 2014

Universidad Nacional Autónoma de México.

Abstract

Instituto de Astronomía
Department of Extragalactic Astronomy and Cosmology.

Master's Degree in Astrophysics.

Conexión Halo-Galaxia: Fracción bariónica de galaxias rojas y azules.

by Ángel Rubén Calette Morín

Con base en una exhaustiva compilación de la literatura y su homogeneización (considerando límites superiores), determinamos las correlaciones empíricas entre las masas de gas atómico y molecular, $M_{\rm H_I}$ y $M_{\rm H_2}$, con la masa estelar, M_*, para galaxias locales separadas en dos poblaciones, azules/tardías y rojas/tempranas (ATs y RTs). Para el cálculo de $M_{\rm H_2}$ introducimos un factor de conversión CO-to-H_2 dependiente de M_*. El cociente $M_{\rm H_2}/M_{\rm H_I}$ incrementa lentamente (considerablemente) con M_* para las galaxias ATs (RTs). Con las correlaciones empíricas que determinamos hasta $M_* \approx 10^7$ $M_\odot$ y usando como "interface" las funciones de masa estelar de las galaxias ATs y RTs, calculamos las correspondientes funciones de masa (FM) de H_I, H_2, gas y bariones para ambas poblaciones locales y el promedio pesado en densidad de ambas. Nuestras FM promedio de H_I y H_2 están de acuerdo con los resultados de catastros ciegos; en estas funciones de masa la población dominante por mucho son las galaxias ATs. La pendiente de la FM bariónica promedio (total) a bajas masas es -1.52, ligeramente más empinada que la pendiente de la FM estelar (-1.49) pero menos empinada que las funciones de masa de H_I, H_2, y gas.

Generamos un catálogo sintético de 3 millones de galaxias que muestrea estadísticamente las FMs estelar de las galaxias ATs y RTs. Con un método estadístico logramos la conexión entre M_* y masa de halo M_h y usando nuevamente las correlaciones empíricas $M_{\rm H_I}$–M_* y $M_{\rm H_2}$–M_* que encontramos para galaxias ATs y RTs, asignamos masas de H_I y H_2. Nuestro catálogo reproduce por ende las FMs de H_I, H_2 y bariónicas mencionadas antes. Encontramos que las relaciones de M_*, $M_{\rm H_I}$, $M_{\rm H_2}$, $M_{\rm gas}$ y $M_{\rm bar}$ con M_h se segregan por el color/tipo de las galaxias. En el caso de la relación $M_{\rm gas}$–M_h de las ATs, crece rápido hasta $\log(M_h/{\rm M}_\odot)\approx 11.4$ y a masas mayores tienden a aplanarse, mientras que en el caso de las RTs, en halos más masivos que $\log(M_h/{\rm M}_\odot)\approx 11.2$, la relación decrece con M_h. La segregación entre ATs y RTs en la relación $M_{\rm bar}$–M_h es menor que en el caso del gas pero mayor que el estelar; a paridad de M_h, las galaxias ATs tienen masas bariónicas mayores que las RTs. Esta diferencia se incrementa hacia bajas masas debido a que las galaxias ATs enanas son mucho más dominadas por gas que las RTs enanas. Las relaciones $M_{\rm gas}$–M_h y $M_{\rm bar}$–M_h de galaxias ATs y RTs locales contienen información clave sobre la evolución de galaxias y pueden ser usadas para restringir modelos y simulaciones numéricas.

Finalmente, hemos extendido la recopilación de la literatura de abundancias de gas hasta $z \sim 2-3$ y hemos inferido la evolución de la relación masa de gas a masa estelar. El punto cero de esta relación crece fuertemente con el corrimiento al rojo z, debido principalmente al incremento de la fracción de gas molecular.

Galaxy-Halo connection: Baryonic fraction of red and blue galaxies.

Abstract

Based on an exhaustive literature compilation and homogenization (taking into account reported upper limits), we have obtained the empirical correlations of the atomic and molecular gas masses, $M_{\rm H_I}$ and $M_{\rm H_2}$, with the stellar mass, M_*, for local galaxies divided into two populations, late/blue and early/red galaxies (LTGs and ETGs). To calculate $M_{\rm H_2}$, we introduced an M_* dependent CO-to-H_2 conversion factor. The $M_{\rm H_2}/M_{\rm H_I}$ ratio slowly (strongly) increases with M_* for local LTGs (ETGs). Once determined the empirical correlations up to $M_* \approx 10^7 {\rm M}_\odot$ and using as an "interface" the galaxy stellar mass functions of LTGs and ETGs, we calculate the corresponding H_I, H_2, gas, and baryonic mass functions (MFs) for both populations and the weighted density mean of both. Our H_I and H_2 mean MFs agree with the results of blind surveys; these MFs are by much dominated by the LTG population. The slope of the mean (total) baryonic mass function at the low-mass end is -1.52, slightly steeper than the stellar MF slope (-1.49) but less steep than the H_I, H_2, and gas MFs.

We have generated a mock catalog of 3 million galaxies that statistically samples the stellar MFs of LTGs and ETGs. Using an statistical method, we obtained the connection between M_* and the halo mass M_h, and using once again the empirical correlations $M_{\rm H_I}$-M_* and $M_{\rm H_2}$-M_* found for LTGs and ETGs, we assign the H_I and H_2 masses. Thus, our synthetic catalog reproduces the H_I, H_2, and baryonic MFs mentioned above. We find that the M_*-, $M_{\rm H_I}$-, $M_{\rm H_2}$-, $M_{\rm gas}$- and $M_{\rm bar}$-M_h relations are segregated by color/type. The $M_{\rm gas}$-M_h relation of LTGs grows fast up to $\log(M_h/{\rm M}_\odot) \approx 11.4$ and then it tends to flatten at higher masses, whereas for ETGs, in halos more massive than $\log(M_h/{\rm M}_\odot) \approx 11.2$, the relation decreases with M_h. The segregation among LTGs and ETGs in the $M_{\rm bar}$-M_h relation is smaller than in the gas cases but higher than the stellar one. At the same M_h, LTGs have larger baryonic masses than ETGs. This difference increases towards lower masses due to the fact that dwarf LTGs are much more gas dominated than dwarf ETGs. The $M_{\rm gas}$-M_h and $M_{\rm bar}$-M_h relations of local LTGs and ETGs contain key information about galaxy evolution and they can be used to constrain models and numerical simulations.

Finally, we have extended the literature compilation of gas abundances up to $z \sim 2-3$ and inferred the evolution of the gas-to-stellar mass relation. The zero point of this relation strongly increases with redshift z, mainly due to increasing of the molecular gas fraction.

Conexión Halo-Galaxia: Fracción bariónica de galaxias rojas y azules.

Contents

Abbreviations

AMT	**A**bundance **M**atching **T**echnique
DLA	**D**umped **L**yman-alpha **A**bsorption system
ETG	**E**arly **T**ype **G**alaxy
LTG	**L**ate **T**ype **G**alaxy
GBT	**G**reen **B**ank **T**elescope
GSMF	**G**alaxy **S**tellar **M**ass **F**unction
GH_IMF	**G**alaxy H_I **M**ass **F**unction
GH_2MF	**G**alaxy H_2 **M**ass **F**unction
GGMF	**G**alaxy **G**as **M**ass **F**unction
GBMF	**G**alaxy **B**aryonic **M**ass **F**unction
IMF	**I**nitial **M**ass **F**unction
MATCH	**M**ulti-**A**bundance **M**atching **T**echnique **C**onstraints in **H**alos
NGAs	**N**úcleos **G**alácticos **A**ctivos
NFGS	**N**ear **F**ield **G**alaxy **S**ample
NYU-VAGC	**N**ew **Y**ork **V**alue **A**dded **G**alaxy **C**atalogue
SFR	**S**tar **F**ormation **R**ate
sSFR	**s**pecific **S**tar **F**ormation **R**ate

Chapter 1

Introduction

Nuestro avance en el entendimiento de la formación y evolución de las galaxias se ha dado vertiginosamente en la última década, gracias principalmente a la obtención de enormes catastros homogéneos, tanto de galaxias locales como a altos corrimientos al rojo, así como a la consolidación de un modelo cosmológico, mismo que provee las condiciones iniciales y a la frontera claves para entender la evolución de las galaxias. El modelo cosmológico mejor constreñido por las observaciones es el de Materia Oscura Fría con Constante Cosmológica (ΛCDM por su abreviatura en inglés). La formación de estructuras en el contexto de este modelo (ver por ejemplo reseñas en Avila-Reese, 2006; Frenk and White, 2012) se da a partir de la evolución gravitacional de las perturbaciones de materia oscura, mismas que con su potencial gravitacional capturan el gas bariónico. Este se enfría por procesos radiativos, cae al centro de las estructuras de materia oscura ya virializadas (halos) y se dan entonces los procesos de formación y evolución estelar, retroalimentación, etc. Dichos procesos son muy complejos y se ha avanzado en su entendimiento a través de modelos semi-analíticos y simulaciones numéricas de N cuerpos + Hidrodinámica, quedando aún mucho camino por recorrer para consolidar una teoría de formación y evolución de galaxias (ver por ej. Baugh, 2006; Benson, 2010; Mo et al., 2010, etc).

En los últimos años han entrado en auge enfoques semi-empíricos cuyo principal objetivo es el de lograr la conexión halo-galaxia, tanto localmente como a altos corrimientos al rojo, al combinar la información estadística de grandes catastros de galaxias con la de simulaciones cosmológicas de N cuerpos. De esta manera se logra inferir cómo dependen ciertas propiedades globales de las galaxias (principalmente la M_* y luminosidad) con la masa de los halos donde supuestamente se han formado. Con los resultados obtenidos

con este enfoque se introducen o se obtienen importantes cotas a los procesos físicos y evolutivos en el interior de los halos de materia oscura en crecimiento, los cuales son los responsables de imprimir las propiedades globales de las galaxias (ver por ej. Firmani et al., 2010; Behroozi et al., 2010, 2013; Moster et al., 2010, 2013; Leauthaud et al., 2012; Rodríguez-Puebla et al., 2013, 2015).

En los trabajos arriba mencionados, y muchos otros, se ha logrado la conexión halo-galaxia a nivel de la masa estelar, tanto localmente como a altos corrimientos al rojo. Esto se reduce principalmente a determinar la relación entre masa estelar y de halo, M_*–$M_{\rm h}$. Se logró también hacer inferencias de esta relación para galaxies centrales y satélites por separado, con sus halos y subhalos, respectivamente. Aunque se encontró que hay diferencias entre ambas relaciones (principalmente debido a que cuando un halo se convierte en subhalo pierde masa por desnudamiento gravitacional), la relación M_*–$M_{\rm h}$ promedia se parece siempre a la de las galaxias centrales. Finalmente en Rodríguez-Puebla et al. (2015) se logró constreñir la relación M_*–$M_{\rm h}$ para galaxias azules y rojas por separado, aunque sólo para el caso de las centrales. Estos autores encontraron que en realidad hay una diferencia significativa estadísticamente entre las relaciones M_*–$M_{\rm h}$ de ambas poblaciones.

Las galaxias además de estrellas, están compuestas de medio interestelar, en el cual predomina el gas frío, compuesto de gas atómico ($\rm H_I$) y molecular ($\rm H_2$). Inicialmente, las galaxias empezaron siendo gaseosas y como producto de su evolución, han ido transformando el gas en estrellas, acretando posteriormente más gas en ciertos casos, o calentando y perdiendo su gas por procesos de retroalimentación positiva, por ejemplo, por vientos galácticos en el caso de halos de baja masa o eyecciones de los Núcleos Galácticos Activos (NGAs) en halos masivos. Como producto de los complejos procesos astrofísicos, dependientes de la escala (masa de halo), la época, el medio ambiente, etc., es que las galaxias al día de hoy terminan con las masas estelares y gaseosas que tienen en función de la masa de sus halos. Entonces, una propiedad importante de las galaxias, además de sus masas estelares y sus colores, es la masa de gas frío que se compone de $\rm H_I$, $\rm H_2$, He y metales:

$$M_{\rm gas} = M_{H_{\rm I}} + M_{H_2} + M_{\rm He} + M_Z = 1.4(M_{H_{\rm I}} + M_{H_2}). \tag{1.1}$$

Las abundancias de helio y metales (Z) al día de hoy son del orden del 40%, por eso el

factor 1.4 en la ecuación (1.1). La masa bariónica total de una galaxia se define entonces como $M_{bar} = M_* + M_{\rm gas}$.[1] La fracción de gas frío de una galaxia es $f_{gas} = M_{\rm gas}/M_{bar}$.

En el espíritu de la conexión halo-galaxia mencionado arriba, un siguiente paso de relevancia es el determinar la relación M_{bar}–$M_{\rm h}$. Debido a que las fracciones de gas f_{gas} de las galaxias en función de M_* son diferentes para galaxias azules/tardías y rojas/tempranas, además de que las relación M_*–$M_{\rm h}$ se segrega por color (Rodríguez-Puebla et al., 2015), en realidad es importante determinar la relación M_{bar}–$M_{\rm h}$ *por separado* para estas dos poblaciones. El promedio pesado por la densidad numérica de ambas poblaciones será entonces la relación M_{bar}–$M_{\rm h}$ global. La fracción bariónica de las galaxias, $f_{\rm bar}$= $M_{bar}/M_{\rm h}$, en función de $M_{\rm h}$, refleja la capacidad de captura de bariones (gas y estrellas) que tiene la galaxia dado el potencial de su halo, así como de expulsión del gas bariónico por los procesos de retroalimentación positiva.

Mientras que la determinación de la masa estelar es rutinaria en grandes catastros de galaxias como el "Sloan Digital Sky Survey" (SDSS), el "Galaxy Mass Assembly" (GAMA) y muchos otros a altos corrimientos al rojo, la observación del gas atómico y, peor aún del molecular o sus trazadores como el CO, es una tarea mucho más difícil. Por esta razón la información de masa de HI y H_2 a nivel de catastros es limitada. Los catastros más extensos en HI fueron hechos con el radiotelescopio CSIRO's 64-m Parkes (Zwaan et al., 2005), el radiotelescopio australiano Westerbork (WSRT) (Serra et al., 2012) y con el radiotelescopio de Arecibo en Puerto Rico (e.g., Martin et al., 2010; Papastergis et al., 2012). En base a estos estudios se construyeron funciones de masa en HI. No obstante, no ha sido trivial contar con muestras completas que contengan la información tanto de $M_{H_{\rm I}}$ como de M_*. Mucho menos, que contengan información de M_{H_2}. Además en el caso del H_2, su determinación es casi siempre a través de la molécula trazadora de CO y ahí surge la incertidumbre del factor de conversión de masa en CO a masa en H_2, mismo que parece no es constante sino que depende por ejemplo de la metalicidad (ver por ej. Schruba et al., 2012; Boselli et al., 2014; Bolatto et al., 2013, etc).

En esta Tesis se propone realizar una extensa recopilación y uniformización de la literatura de muestras de galaxias locales que contengan determinaciones de M_* y $M_{H_{\rm I}}$,

[1] En las galaxias también hay gas caliente en estado ionizado; no obstante la fracción de esta componente del medio interestelar es generalmente muy pequeña. Su detección no es fácil pues este gas es muy difuso y posiblemente se extiende mucho más allá de lo que es la galaxia visible, formando un halo caliente y tenue alrededor de la misma (ver por ejemplo, Sharma et al., 2012; Jeltema et al., 2008, etc.).

así como de M_* y M_{H_2}, en un amplio intervalo de masas y separadas en azules/tardías y rojas/tempranas. De esta manera, se pretende inferir las correlaciones M_{H_I}–M_* y M_{H_2}–M_* tanto para galaxias azules/tardías como para rojas/tempranas. Conjuntando estas correlaciones, se infieren entonces las correspondientes correlaciones M_{gas}–M_* y, añadiendo la información de M_*, finalmente se obtienen las correlaciones M_{bar}–M_* (Cap. 2).

Una vez constreñidas las correlaciones mencionadas, se puede hacer uso de las funciones de masa estelar de galaxias azules y rojas para "proyectar" estas correlaciones a funciones de distribución de M_{H_I}, M_{H_2}, M_{gas}, y finalmente, M_{bar}, tanto para galaxias azules como rojas, así como para el caso promedio pesado por densidad (total). Este es entonces otro objetivo propuesto en esta Tesis (Cap. 3).

Finalmente, con las funciones de masa encontradas, aplicando el enfoque estadístico de conexión halo-galaxia anteriormente mencionado, se puede determinar la relación M_{bar}–M_{h} separada en galaxias azules y rojas así como la promedio (total). En realidad, se pueden obtener también las correspondientes relaciones M_{H_I}–M_{h}, M_{H_2}–M_{h}, y M_{gas}–M_{h}. Todas estas relaciones, pero en especial la de M_{bar} vs M_{h}, reflejan, como ya se dijo, aspectos claves de cómo se ensamblaron las galaxias en función de las masas de sus halos y permiten sondear el modelo cosmológico subyacente, el ΛCDM (Cap. 4).

1.1 Restricciones en diferentes épocas

Como ya se mencionó, en base a los catastros observacionales a diferentes corrimientos al rojo, con los cuales se determinan las GSMFs y funciones de correlación de las galaxias a diferentes épocas, los métodos semi-empíricos han permitido inferir la conexión halo-galaxia hasta épocas muy en el pasado, principalmente la relación M_*–M_{h}.

Para lograr inferencias de la la relación M_{bar}–M_{h} hasta $z \sim 3$, el primer paso que se presenta en esta Tesis, es lograr un cuadro empírico de la correlación M_{gas}–M_* a diferentes z's, para luego, como se procedió a $z \sim 0$, proyectar esta correlación con la ayuda de las GSMFs observadas a funciones de distribución de M_{gas} y finalmente de M_{bar}. Con estas funciones de masa bariónicas y el método estadístico de correspondencia de abundancias, se pueden entonces inferir las relaciones M_{bar}–M_{h} a diferentes épocas (Cap. 5).

La información observacional relacionada al contenido gaseoso de las galaxias lejanas ($z > 0.1$) proviene prácticamente sólo de estudios en CO que se traducen a masas de H_2. En la Tesis hemos compilado varias muestras de galaxias con información de M_{H_2} y M_* hasta $z \sim 3$. Se introduce también un estimado de cómo evoluciona la masa de gas atómico, M_{H_I}. Con esta información se plantea en un futuro inferir la evolución de la relación M_{bar}–M_h siguiendo la estrategia mencionada arriba.

1.1.1 Contenido

El plan de la presente Tesis es como sigue. En el Capítulo 2 se presenta la recopilación y caracterización del tipo de galaxia en tardías/azules o tempranas/rojas ya sea por la morfología, color o tasa de formación estelar de datos observacionales de la literatura para galaxias del Universo local con determinaciones de M_* y M_{H_I} al igual que para M_* y M_{H_2}. Posteriormente se homogeneizan estos datos y se toma en cuenta aquellas galaxias observadas en radio, pero que no son detectadas dado el límite de sensibilidad de el telescopio o el detector utilzado. Por último, se obtienen relaciones M_{H_I}-M_* M_{H_2}-M_*, M_{gas}-M_* y M_b-M_*.

En el Capítulo 3, utilizando la fracción de galaxias rojas, la función de masa estelar promedio en el universo local como interfaz para asignar masas de gas y bariónica con las relaciones obtenidas en el capitulo 2, se infieren las funciones de masa de M_{H_I}, M_{H_2}, M_{gas}y M_b para galaxias tardías/azules y tempranas/rojas.

En el Capítulo 4 se infiere las relaciones M_h-M_*, M_h-M_{H_I}, M_h-M_{H_2}, M_h-M_{gas} y $M_h$mbar, utilizando el método **MATCH** (Multi-Abundance Matching Technique Constraints in Halos) el cual consiste en determinar la relación masa stelar- masa halo (SHMR por sus siglas en inglés) para galaxias tardías/azules y tempranas/rojas a través del métododod de empate de abundancias. Posteriormente asignamos masas de M_{H_I}, M_{H_2}, M_{gas} y M_b utilizando las relaciones de masa de gas y bariónica con la masa estelar, determinadas en el capitulo 2.

El Capítulo 5, incluye la dependencia de las correlaciones masa de gas-masa estelar con el corrimiento al rojo. Esta dependencia está basada en resultados observacionales de la evolución de la fracción de masa de hidrógeno molecular y de la densidad de masa de H_I a diferentes corrimientos al rojo de sistemas DLA (Dumped Lyman-alpha absorption systems). Para el caso del hidrógeno molecular presentamos el contraste con resutlados

observacionales, corrigiendo por un factor de conversión dependiente del corrimiento al rojo.

El Capítulo 6 presenta las principales conclusiones de esta Tesis.

Chapter 2

The observational data and inferred correlations

We are in the era of large homogenous surveys of galaxies, mainly build up in the optical and infrared bands. However, in order to attain information about the cold gas content of galaxies, both atomic and molecular, observations in the radio are necessary. Unfortunately, the extent of samples that contain both optical/infrared information and radio information on H_I and H_2 contents is very limited. In view of this, our goals in this Chapter are:

- To compile from the literature catalogs and samples of galaxies that contain as primary information stellar masses, M_* (inferred from photometry and/or from spectral energy distributions), a characterization of the galaxy type, either morphological or by color/star formation activity AND integrated fluxes in H_I from which the H_I gas mass, M_{H_I}, can be inferred AND/OR CO observations from which the H_2 gas mass, M_{H_2}, can be inferred;

- To homogenize as much as possible the data, and take into account those galaxies observed in radio but non detected due to the sensitivity limit of the telescope/detector.

- To infer the M_{H_I}–M_*, M_{H_2}–M_*, M_{gas}–M_* and M_{bar}–M_* correlations in the stellar mass range $7 < \log_{10}(M_*/M_\odot) < 12$. Recall that we have defined the total galaxy cold gas mass as M_{gas}=1.4(M_{H_I}+ M_{H_2}), and the galaxy baryonic mass as M_{bar}= M_*+ M_{gas}.

2.1 Characterization of the samples and definitions

The separation of the galaxy population at least into two broad groups of "late-type/blue" and "early-type/red" is important because the gas contents of these two populations seems to be very different at a given M_*. In order to characterize these two broad galaxy types we use one or another among the following criteria:

• **Criterion 1.-** Morphological type based on images: for those samples, where the morphological type is given, we characterize as our "early-type/red" those galaxies having morphological types of Ellipticals (E) and lenticulars (S0) or $T \leq 0$, and as our "late-type/blue" those having morphological types of Spirals (S) and Irregulars (Irr) or $T > 0$.

• **Criterion 2.-** Photometrical integral colors, concentrations, and/or SFR's: The separation into two populations for those samples that do not report galaxy morphology based on images can be done based on some global photometric or spectro-photometric information as the integral color, the concentration index[1], and/or the specific star formation rate (sSFR=SFR/M_*). The separation into blue and red galaxies by colors has been done as a function of stellar mass. It is well known that galaxies follow a bivariate distribution in color, distribution that changes with M_*. For a given M_*, the bivariate distribution can be fitted for instance by two Gaussians and the color at which the Gaussians intersect at the 1σ level, can be defined as the transition color: galaxies bluer than this color are classified as blue (the blue cloud) and those redder are classified as red (the red sequence). This was done for example in Li et al. (2006) by using $g - r$ colors or in Wyder et al. (2007) by using $NUV - r$ color. Other authors (see e.g., Deng, 2013), by making an statistical study using the Galaxy Zoo, with the definition of contamination and completeness from Shimasaku et al. (2001), find that the concentration index in the r-band can characterize well late-type galaxies (those having $c \leq 2.85$). Nevertheless, for those with $c > 2.85$ it is not straightforward to say that they are of early type due to the high contamination by late-type galaxies (see e.g., Bernardi et al., 2010). Therefore, in order to characterize galaxies as early type by concentration, an extra parameter such as the $NUV-$r color or the sSFR are needed.

[1]The concentration index is defined commonly as $c \equiv \frac{R_{90}}{R_{50}}$, where R_{90} and R_{50} are the radii enclosing 90 and 50 percent of the r-band Petrossian Flux respectively

Based on these criteria and depending on the information provided by the samples and catalogs compiled here, we separate the galaxies into the two mentioned above broad groups. Hereafter, we will refer to these two broad groups just as *late-type and early-type galaxies (LTGs and ETGs).* Note that color is not attempted to substitute the morphological type. Although it is well known that color correlates with the galaxy morphology and star formation activity, it is also known that there is a fraction of galaxies that do not follow these correlations and, according to the criteria used to divide them into two groups, they can be disk-dominated but red and passive (e.g., Maller et al., 2009; Masters et al., 2010, the fraction of these galaxies is not larger than $\sim 10\%$ of the late-type ones and it is more common among massive galaxies), or bulge-dominated but blue and star forming (e.g., Schawinski et al., 2009, these galaxies are less than 6% of early-type ones and are not found among massive galaxies). The correlations and mass functions that we attempt to infer here are not for a fine and continuous separation of galaxies by color or morphological type, but only for two rough galaxy populations. Therefore, *the "contaminations" produced by the above mentioned deviations from the general correlations are not so relevant as to change these correlations relations and mass functions when separated just into two groups.*

Regarding the gas masses, in several of the samples taken from the literature there are observed galaxies that, at the sensitivity level of the radiotelescope, could not be detected. In some cases the fraction of such non-detections in the sample is small but in other ones it is large (in particular for ETGs). The non-detection at the level of sensitivity (reported then as an upper limit) gives information about the galaxy that we can not obviate (this would introduce a bias in the data).

To use observational data containing upper-limits, we resort to the Kaplan-Meier product limit estimator, which is a survival analysis method commonly used in Astronomy for combining censored (detections) and uncensored (upper limits for non-detections) data. Thus, mean, median, percentiles, etc. can be calculated for a given sample of censored and uncensored data; see the Appendix B for more information. In order to take into account the non-detections and calculate mean, standard deviations, median, and percentiles of the $M_{\rm H_I}$- and $M_{\rm H_2}$-to-M_* ratios in M_* bins, we use the ASURV (Astronomy SURVival analysis) package developed by Takashi Isobe, Michael LaValley and Eric Feigelson in 1992, and implemented in the **stsdas** package (Space Telescope Science Science Data Analysis) in IRAF. In particular, we make use of the *kmestimate*

(Kaplan-Meier estimator) routine.

In the next Section, we will plot and infer from the compiled data the $M_{\rm H_I}$ and $M_{\rm H_2}$ to M_* ratios, that we define as:

$$R_{\rm H_I} \equiv \frac{M_{\rm H_I}}{M_*} \tag{2.1}$$

$$R_{\rm H_2} \equiv \frac{M_{\rm H_2}}{M_*} \tag{2.2}$$

For most of the samples and catalogs taken from the literature, the authors multiply by 1.4 the $\rm H_I$ mass in order to get the total gas mass (i.e., including He and metals, but in many cases, $\rm H_2$ is not included). We prefer first define the masses of $\rm H_I$ and $\rm H_2$ (or the $M_{\rm H_I}/M_*$ and $M_{\rm H_2}/M_*$ ratios) as a function of M_* and then sum up them taking into account He and metals in order to derive the total cold gas masses (see Eq. 1.1). Thus, when necessary, we (de)correct the reported data by the factor related to He and metals, which is taken into account when the total cold gas mass is calculated:

$$M_{\rm gas} = 1.4 \times (M_{\rm H_I} + M_{\rm H_2}) \tag{2.3}$$

Thus, we can define the cold gas to stellar mass ratio,

$$R_{\rm gas} \equiv \frac{M_{\rm gas}}{M_*}, \tag{2.4}$$

and the galaxy baryonic mass,

$$M_{\rm bar} = M_{\rm gas} + M_*. \tag{2.5}$$

It is common in the literature to use the galaxy gas fraction instead of the galaxy gas-to-stellar mass ratio. The former is defined as

$$f_{\rm gas} \equiv \frac{M_{\rm gas}}{M_{\rm gas} + M_*} = \frac{M_{\rm gas}}{M_{\rm bar}}. \tag{2.6}$$

The gas fraction is related to the gas-to-stellar mass ratio by:

$$f_{\rm gas} = \frac{R_{\rm gas}}{R_{\rm gas} + 1} \tag{2.7}$$

The complement to f_{gas} is the galaxy stellar mass fraction,

$$f_* = 1 - f_{\rm gas} \tag{2.8}$$

2.2 The compilation

In the following, we present our compilation and homogenization of samples of galaxies for which the authors report M_*, the information we require to classify the galaxy as LTG or ETG, and either measurements of $R_{\rm H_I}$ (Sect. 2.2.1) or $R_{\rm H_2}$ (Sect. 2.2.2).

2.2.1 Hydrogen atomic gas

Measurements of the amount of atomic hydrogen in galaxies are derived from the H$_{\rm I}$ 21 cm emission intensity. Thus, if the integrated flux of the 21 cm line and the redshift of a galaxy are known, the H$_{\rm I}$ mass is given by:

$$M_{\rm H_I} \approx 2.356 \times 10^5 (1+z)^{-1} \left(\frac{D_L(z)}{\rm Mpc}\right)^2 \left(\frac{\int F_\nu dv}{\rm Jy\ Km\ s^{-1}}\right) M_\odot, \tag{2.9}$$

where $D_L(z)$ is the luminosity distance and $\int F_\nu dv$ is the observed integrated H$_{\rm I}$-line flux density. Following, we present the observational data sets compiled of atomic hydrogen to stellar mass ratio ($R_{\rm H_I}$) for LTG and ETGs, including dwarf galaxies.

- **The GALEX Arecibo SDSS Survey** (GASS Catinella et al., 2013): It consists of 760 galaxies (473 detections and 287 non detections) with stellar masses greater than 10^{10} M$_\odot$ and redshifts $0.025 < z < 0.05$; the $R_{\rm H_I}$ limit of the sample is 0.015 for $\log(M_*/{\rm M}_\odot) > 10.5$ and up to 0.05 for smaller masses. The concentration parameter R_{90}/R_{50} was used to characterize the galaxy type, with a value of $c = 2.85$ used to separate LTGs from ETGs. The stellar masses are calculated from the SDSS spectra using the method described in Salim et al. (2007) and assuming a Chabrier (2003) IMF. The GASS galaxies are located in the intersection of the footprints of the SDSS primary spectroscopic survey, the projected GALEX Medium Imaging Survey (MIS) and ALFALFA. Thus, if there is a H$_{\rm I}$ detection in ALFALFA or objects in the Cornell H$_{\rm I}$ digital archive, they use them to obtain H$_{\rm I}$ masses. When there is no H$_{\rm I}$ information, they use the Arecibo radio telescope to measure the integrated H$_{\rm I}$-line intensity and thus $M_{\rm H_I}$. The authors do not apply

corrections for inclination or to account for helium. Not correcting by inclination disks similar to the Milky Way observed close to edge-on or higly inclined, leads to underestimate the H_I column density and the inferred H_I masses (see Appendix B of Haynes and Giovanelli, 1984), this effect seems to be larger for intermediate-type spirals. On the other hand if galaxies are low inclined, the column density is underestimated by less than 5%. The fraction of spiral galaxies in this work affected by this effect is small.

- **The Nearby Field Galaxy Survey** (NFGS Jansen et al., 2000; Wei et al., 2010; Kannappan et al., 2013, see more references therein): A broadly representative sample of 198 galaxies spanning stellar masses $M_* \sim 10^8 - 10^{12}$ $M_\odot$ and all morphologies. Stellar mass were estimated using a variant of the code described in Kannappan and Gawiser (2007) and improved in Kannappan et al. (2009), which fits the spectral energy distribution (SED) and integrated spectrum of a galaxy if exists with a suite of stellar populations models. Both the diet Salpeter IMF and the one adopted here, the Chabrier (2003) IMF, were used. The single-dish H_I fluxes for most galaxies were taken from the HyperLeda database (Paturel et al., 2003) or were obtained by the authors with the Green Bank Telescope (GBT) Spectrometer. They obtain the 21 cm line fluxes and derive H_I masses using the expression from Haynes and Giovanelli (1984) combined with a 1.4 correction factor for helium. The sample provides strong upper limits up to $R_{H_I} \sim 0.1$; all galaxies with ratios larger than this are detected. At difference of the GASS sample, for the NFGS, there is not a clear R_{H_I} limit as a function of M_*.

- **The Stark et al. (2013) compilation**[2]: It is a compilation of galaxies with available CO, H_I, and multi-band imaging data. The primary sample comes from NFGS galaxies with CO data. The compilation from the literature includes galaxies from three large surveys: the Spitzer Infrared Near Galaxy Survey (SINGS Kennicutt et al., 2003), ATLAS-3D (Young et al., 2011), and CO Legacy Database for the GASS (COLD GASS Saintonge et al., 2011). For low-mass galaxies, objects from Barone et al. (2000), Garland et al. (2004), Leroy et al. (2005), Taylor et al. (1998), and Kannappan et al. (2009) were added. Most of these references are themselves the sources of the CO data, and the H_I data often come from the same source as

[2]Dr. David Stark kindly sent us in electronic form the data reported in their paper

the CO data, or else from alternate sources in the literature or HyperLeda. Morphology classification is done by eye, and the stellar masses are calculated as in Kannappan et al. (2013) for the diet Salpeter IMF. Except for a subset outside the SDSS footprint but that has $BVRI$ photometry (see Muñoz-Mateos et al. (2009) for details) from SINGS sample, all the optical data come from the SDSS DR8. NIR imaging is available for all galaxies from 2MASS. Note that a significant fraction of the galaxies from the Stark et al. (2013) compilation are in common with those of Kannappan et al., 2013. Most of the massive ETGs in the Stark et al. compilation are from Serra et al. (2012); we handle these data separately and refer to them as the Serra et al. (2012) sample. This sample is from the ATLAS-3D survey, a Integral-Field Unit Spectroscopy survey of 260 early-type galaxies within the 42 Mpc local volume many of which are from the Coma cluster. Only a fraction of galaxies from the ATLAS-3D survey have HI information. The HI masses were obtained from radio observations in the WRST.

- **The HERACLES/THINGS survey** (Leroy et al., 2008): From this work we select from this work 23 nearby, star-forming galaxies, which we associate to LTG objects; 11 are dwarf, H_I-dominated galaxies and 12 are large spirals. The former are defined as those galaxies with rotational velocities $v_{\rm rot} \leq 125$ Km s^{-1}, stellar masses $M_* \leq 10^{10} M_\odot$ and absolute magnitude in the blue band $M_B \geq -20$ mag. The latter have $v_{\rm rot} \geq 125$ Km s^{-1}, $M_* \geq 10^{10} M_\odot$ and $M_B \leq -20$ mag. The authors derived atomic hydrogen surface density from 21 cm line integrated intensity maps from Walter et al. (2008) (THINGS survey), correcting for inclination and considering the 1.36 factor to account for helium (He). Then, from this surface mass density, the atomic hydrogen mass is obtained. The galaxy stellar mass is inferred from 3.6 μm maps from Spitzer, mostly of these from SINGS (Kennicutt et al., 2003). Leroy et al. (2008) constructed radial profiles with the 3.6 μm intensity in each tilted ring to avoid contamination by hot dust and foreground stars. To convert the 3.6 μm intensity to surface stellar mass density (Σ_*), they use a K-to-3.6 μm calibration and adopt a fixed K-band mass-to-light ratio, $\Upsilon_*^{\rm K} = 0.5 M_\odot/L_\odot$, assuming a Kroupa (2001) IMF. From this surface stellar mass density they obtain the total galaxy stellar mass.

- **Dwarf LTGs** (Geha et al., 2006): In this work, 101 dwarf galaxies with H_I measurements are presented. Galaxies are selected if they have absolute magnitudes

in the r-band, $M_r - 5\log_{10}(h_{70}) > -16$ from the low-luminosity catalog of Blanton et al. (2005). Stellar masses are based in the optical SDSS i-band magnitude and $g-r$ colors using the mass-to-light ratios of Bell et al. (2003) and considering a Kroupa et al. (1993) IMF. The sample is composed by a 10% of early type galaxies (dwarf spheroidals) and the rest of the sample are late types (dwarf irregulars). $M_{\rm H_I}$ observations are obtained using the Arecibo 305 m telescope and the Green Bank 100 m telescope, by measuring the HI integrated flux.

- **Stewart et al. (2009) & Papastergis et al. (2012) fits:** In both cases, the authors provide linear fits in the logarithm to the $M_{\rm H_I}$–M_* (or R_{HI}–M_*) relation from small samples compiled by them. Stewart et al. (2009) used mainly the observational data presented in McGaugh (2005) for disk-dominated galaxies. Papastergis et al. (2012) have actually $M_{\rm H_I}$ determinations for thousands of galaxies from ALFALFA and stellar masses from the SDSS in common with the ALFALFA survey, but they do not make public these data and do not use them for inferring the $R_{\rm H_I}$–M_* relation due to sample selection effects. Instead, they present a fit to R_{HI}–M_* relation by using data sets from Swaters and Balcells (2002), Garnett (2002), Noordermeer et al. (2005) and Zhang et al. (2009).

The left panels of Fig. 2.1 show the data from the above mentioned samples and fits in the R_{HI} vs M_* plane (see the legends indicating each sample/fit inside the panels). The upper and lower panels are for the LTGs and ETGs, respectively. The down arrows indicate the cases of non-detection, with the values plotted corresponding to an upper limit as reported in the original works. The data show that the two populations of galaxies have different loci in the R_{HI} vs M_* plane, and this is why *it is important to separate them.* In these plots we can appreciate that the GASS survey has a defined $R_{\rm H_I}$ and $R_{\rm H_2}$ limit, dependent on M_*, as mentioned above. Unfortunately, this is not the case for the other samples. Given the "controlled" completeness of the GASS sample in the $R_{\rm H_I}$ and $R_{\rm H_2}$ distributions, this sample is useful for exploring the nature of these distributions and for determining the intrinsic scatter around the $M_{\rm H_I}$–M_* and $M_{\rm H_2}$–M_* relations, at least for massive galaxies (see Cortese et al., 2011).

In the right panels of Fig. 2.1, we reproduce the compiled data showed in the left ones but not making differences among the different sources; only detections and upper limits are differentiated. Note that we did not take into account the galaxies from Serra et al.

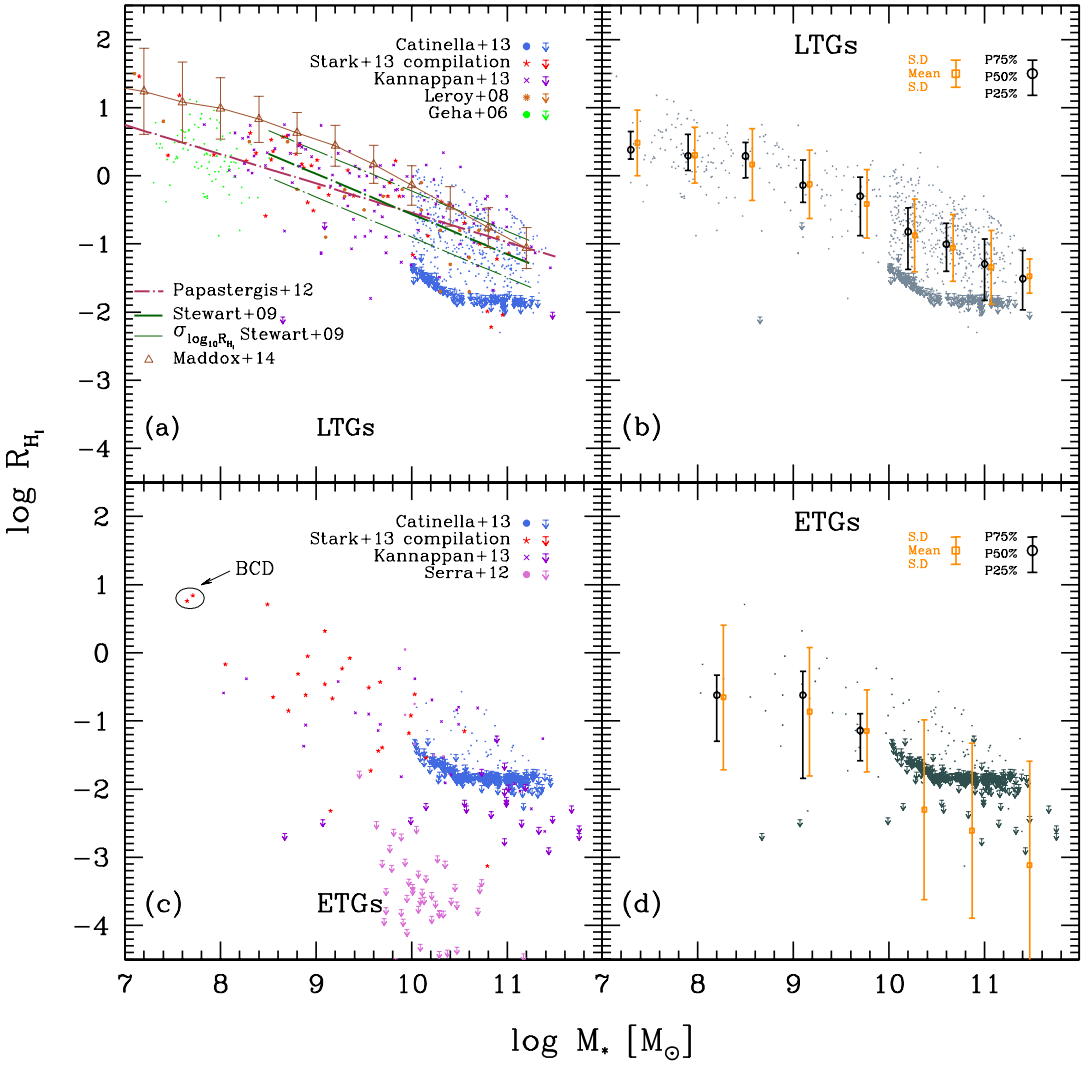

Figure 2.1: *Upper panels:* Compiled data with information on $R_{\rm H_I}$ and M_* for LTGs (see inside the panels for the different sources). We include the reported upper limits for non detections (arrows), as well as the fits to previous compilations by Stewart et al. (2009) and Papastergis et al. (2012). Right panel is the same as left one, but with the data plotted without distinguishing the different sources. These data has been grouped in different stellar mass bins in order to calculate the mean, standard deviations, median and 25-75 percentiles in each bin by means of the Kaplan-Meier estimator for censored + uncensored data (see text for more details). The brown 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 to high values of $R_{\rm H_I}$. *Lower panels:* The same as in upper panels but for ETGs. The galaxies from Serra et al. (2012) are not included in the right panel due to the high bias to a high-density environment in this sample (see text). The two lowest mass galaxies from Stark et al. (2013) compilation were also excluded because they actually are blue compact star forming dwarfs (BCD). For the bins where more than 50% of the data are upper limits, the median and percentiles are not calculated.

(2012) since a significant fraction of them are actually from the Coma cluster and other clusters (these galaxies are from the ATLAS3D catalog); hence, this sample is strongly biased to a particular dense environment, where galaxies have less gas contents than in

the field.

As a preliminary step for inferring correlations (see Section 2.3), we calculate the mean, median and 25-75 percentiles of the data in different stellar mass bins. Since in the data there are non detections reported as upper limits, we use the Kaplan-Meier estimator to obtain these quantities in stellar mass bins, as described at the beginning of this Chapter (see also Appendix B). When the data are numerous, we use bins in $\log M_*$ of 0.4-0.6 dex. Otherwise, the width of the bins is made as large as to have at least ~ 10 data points and/or a fraction of detections larger than $\sim 20\%$ in the bin.

The empty black circles with (asymmetric) error bars in Fig. 2.1 show the median and 25-75 percentiles calculated with the Kaplan-Meier estimator in different stellar mass bins as described above. The orange squares and error bars show the corresponding mean and standard deviation in each mass bin (the points are slightly shifted to avoid superposition). The IRAF package provides actually the standard error of the mean; we estimate the standard deviation of the population from it, see the Appendix B for details. As mentioned also in the Appendix B, the median and percentiles are very uncertain or impossible to be calculated for detection fractions smaller than 50%, while the mean yet can be estimated for fractions as small as $\sim 20\%$, though with a larger uncertainty. In the case of the massive ETGs, the fraction of detections are smaller than 50%, therefore, the median and percentiles are not plotted in the corresponding mass bins. However, the mean and standard deviations are plotted, though it should be taken into account that the determinations are uncertain in these cases.

2.2.2 Molecular hydrogen gas

Direct observational measurements of molecular hydrogen fluxes in galaxies able to give us the H_2 masses are not possible. H_2 is a diatomic molecule with identical nuclei, therefore it does not possess a permanent dipole moment neither dipolar rotational transitions. Purely rotational quadrupole transitions are the lowest energy transitions of H_2 in the far infrared (FIR) and these are weak due to their long spontaneous decay lifetimes ($\tau_{\rm decay} \sim 100$ years). H_2 transitions such as the ortho and para or the lowest vibrational energy transitions, require temperatures $T > 100$ K. Thus, these conditions make impossible the emission of cold H_2 in the molecular ISM (Bolatto et al., 2013).

In order to solve this problem, given that molecular gas is not pure H_2, a tracer such as CO molecule can be used. This molecule consists of Oxigen and Carbon, which are the most abundant ($\sim 10^{-4}/H_2$) heavy elements in the ISM, and under the physical conditions in molecular clouds, it can be formed. Also, it has a weak permanent dipole moment and a ground rotational transition with low excitation energy ($J \rightarrow 1-0$) that even in molecular clouds can be easily excited, and this emission lies in a transparent atmospheric window (see e.g., Bolatto et al., 2013; Narayanan et al., 2012). Thus, the way to relate the H_2 column density (N_{H_2}), with the observed CO intensity (W_{CO}) is:

$$N_{H_2} = X_{CO} W_{CO} \tag{2.10}$$

If equation (2.10) is integrated over the emitting area, a relation among the H_2 mass (M_{H_2}) and the CO line luminosity (L_{CO}) is obtained:

$$M_{H_2} = \alpha_{CO} L_{CO} \tag{2.11}$$

X_{CO} and α_{CO} are known as the CO-to-H_2 conversion factor, which can be transformed by $X_{CO} = 6.3 \times 10^{19} \alpha_{CO}$ (Narayanan et al., 2012).

The determination of M_{H_2} in galaxies is commonly done by considering a constant CO-to-H_2 conversion factor (see e.g., Leroy et al., 2008; Saintonge et al., 2011; Bauermeister et al., 2013; Stark et al., 2013). This factor could have been determined with reasonable accuracy in molecular clouds in the solar neighborhood ($X_{CO} = 2{\times}10^{20} \mathrm{cm}^{-2}$ K km s^{-1} or $\alpha_{CO} = 3.2(\mathrm{K\ km\ s^{-1}\ pc^{-1}})^{-1}$ with a systematic uncertainty of $\pm 30\%$). For Milky-Way like disks, the conversion factor is similar to the one found for our Milky-Way, though the scatter and systematic uncertainty increase compared to the MW uncertainty. Traditionally authors assumed a constant value for the CO-to-H_2 conversion factor, equal to the one of the Milky Way. This is the case for almost all of the observational samples that we compiled (see below).

In the last years, the consensus increased in favor of considering that the CO-to-H_2 conversion factor is not constant. There are many observational and theoretical works that strongly suggest that the CO-to-H_2 conversion factor is sensitive, for instance, to the gas phase (nebular) metallicity Z_{gas}: it increases with decreasing Z_{gas} (e.g., Boselli et al., 2002; Schruba et al., 2012; Narayanan et al., 2012; Bolatto et al., 2013). We consider

that a CO-to-H_2 conversion factor varying with $Z_{\rm gas}$ must be used in order to get more realistic molecular hydrogen masses. In the Appendix A we discuss works that propose the dependence of $\alpha_{\rm CO}$ on $Z_{\rm gas}$ (Schruba et al., 2012). Since $Z_{\rm gas}$ on its own depends on M_* (see e.g. Andrews and Martini, 2013; Sánchez et al., 2013) , in the Appendix A we infer an approximate dependence of the CO-to-H_2 factor on M_*:

$$\log_{10}(\alpha_{\rm CO}) = 0.42 + 2 \times \log_{10}\left[1 + 0.1\left(\frac{3\times10^{10}{\rm M}_\odot}{M_*}\right)^{0.64}\right] \quad (2.12)$$

for LTGs less massive than 3×10^{10} $M_\odot$, and $\alpha_{\rm CO} = \alpha_{{\rm CO},MW} = 3.2$ for more massive galaxies. For ETGs we leave $\alpha_{\rm CO} = \alpha_{{\rm CO},MW}$ at all masses[3]. These galaxies formed on average their stars early and quickly and exhausted much of the gas in that process. The little amount of gas left is enriched by stellar mass loss. The $Z_{\rm gas}$ of these galaxies is expected then to be similar or larger than in the Milky Way; for such metallicities, $\alpha_{\rm CO}$ tends to be close to the one the Milky Way (see Appendix A).

As mentioned above, almost all the galaxy samples with information on $M_{\rm H_2}$ assumed $\alpha_{\rm CO}$ constant and equal to the Milky-Way value; only in a few cases some metallicty-dependent factors were used. For one or another case, in order to homogenize the results, we reestimate the H_2 masses of LTGs from the compiled samples by dividing them by the value the respective authors used, and multiplying then by our mass-depending $\alpha_{\rm CO}$. This implies that the H_2 masses of LTGs will be larger as smaller they are with respect to those reported in the original studies[4]. We have also uncorrected by helium and metals if this correction has been applied for a given sample; our aim here is to have estimates of the $R_{\rm H_2}$ mass only.

The compiled H_2 data

- **The CO Legacy Legacy Database for GASS** (COLD GASS Saintonge et al., 2011): This is a large program aimed at observing CO(1-0) lines fluxes at the IRAM 30 m telescope for galaxies from the GASS survey (Catinella et al., 2013; see section 2.2.1). The stellar masses and morphologies are from this survey. From

[3]Eq. (2.12) is normalized in such a way that $\alpha_{\rm CO}$=3.2 for Milky Way like galaxies, in other words, those galaxies having stellar masses $M_* \gtrsim 3\times10^{10}{\rm M}_\odot$ (see e.g., Schruba et al., 2012). This is consequence of the dependence on metallicity of $\alpha_{\rm CO}$ conversion factor (see figure A.3 in appendix A), for high mass LTGs the mass-metallicity relation flattens at high stellar masses, thus $\alpha_{\rm CO}$ flattens.

[4]Once we correct by our $\alpha_{\rm CO}$ conversion factor, H_2 masses are larger by a factor ~ 30 at $M_* \sim 1\times10^7{\rm M}_\odot$ and larger by a factor ~ 8 at $M_* \sim 1\times10^9{\rm M}_\odot$

the CO fluxes, the total CO luminosities and hence the H_2 masses are calculated; the Milky-Way CO-to-H_2 conversion factor was used.

- **The Stark et al. (2013) compilation**: This is the same data compiled sample in 2.2.1 for H_I. The IRAM 30 m and the ARO 12 m telescopes were used to measure the CO ($J \rightarrow 2-1$) (IRAM) and ($J \rightarrow 1-0$) (IRAM & ARO) lines. Thus, molecular hydrogen masses are obtained from the CO flux, using a constant CO-to-H_2 conversion factor $X_{\rm CO} = 2 \times 10^{20} {\rm cm}^{-2} ({\rm K\ km\ s}^{-1})^{-1}$ and beam correction. In addition, Stark et al. (2013) make a compilation from the literature for the H_2 sample, including the following works: Kennicutt et al. (2003), Young et al. (2011), Barone et al. (2000), Garland et al. (2005), Leroy et al. (2005), Taylor et al. (1998), and Kannappan et al. (2009).

- **The HERACLES/THINGS survey** (Leroy et al., 2008): This is the same sample presented in Section 2.2.1 for H_I. The H_2 surface mass density (Σ_{H_2}) is estimated by using the IRAM 30 m telescope to map CO $J \rightarrow 2-1$ emission for the full optical disk of the 18 THINGS galaxies (HERACLES sample). Then, to relate CO $J \rightarrow 2-1$ to CO $J \rightarrow 1-0$ intensity, it is assumed the ratio $I_{\rm CO}(2 \rightarrow 1) = 0.8 I_{\rm CO}(1 \rightarrow 0)$, finding this as a typical value in their sample, based on direct comparison with HERACLES and previous surveys. There are two galaxies which are not included in the HERACLES sample. For these, they use CO $J \rightarrow 1-0$ maps from BIMA SONG to estimate Σ_{H_2}. The CO-to-H_2 conversion factor is constant for this work, adopting the value for the solar neighborhood, $X_{\rm CO} = 2 \times 10^{20} {\rm cm}^{-2} ({\rm K\ km\ s}^{-1})^{-1}$. A correction for inclination is considered.

- **Bauermeister et al. (2013) compilation:** From this study, we use 8 of their normal star-forming galaxies. These are in the redshift range $0.005 \leq z \leq 0.1$ and the stellar mass range $4 \times 10^{10} M_{\odot} \leq M_* \leq 16 \times 10^{10} M_{\odot}$. Stellar masses are obtained by fitting SDSS *ugriz* photometry to a grid of models spanning a wide range of star formation histories. They measure CO $J \rightarrow 1-0$ intensity with CARMA, and thus obtain H_2 masses, using a constant CO-to-H_2 Milky-Way conversion factor.

- **Boselli et al (2014) compilation:** In this work are used CO observations obtained using the NRAO (National Radio Astronomy Observatory) Kitt Peak 12 m telescope and compiling from literature CO observations of the HRS (Hershel

Reference Survey) galaxies. These observations consist of 225 galaxies, with 143 detections and 82 non-detections, ranging from LTG to ETG type systems using the NED classification of morphology or their own characterization if not available. They use a constant and a variable $X_{\rm CO}$ factor. The constant value of the CO-to-H_2 conversion factor is $X_{\rm CO} = 2.3 \times 10^{20} {\rm cm}^{-2} ({\rm K\ km\ s}^{-1})^{-1}$ and the variable $X_{\rm CO}$, is a H-band luminosity-dependent, which Boselli et al. (2002) found to be:

$$\log_{10}(X_{\rm CO}) = -\,0.38 \times \log_{10}(L_H) + 24.23\ [{\rm cm}^{-2}/({\rm K\ km\ s}^{-1}]) \tag{2.13}$$

Stellar masses are derived from i-band luminosities and using the $g-i$ colour-dependent stellar mass-to-light ratio from Zibetti et al. (2009), with an assumed Chabrier (2003) IMF.

- **Bothwell et al (2014) compilation:** Using the APEX telescope, the $CO(2 \rightarrow 1)$ emission line was measured to trace the molecular hydrogen of 42 late-type galaxies in the stellar mass rage $8.5 < \log_{10}(M_*) < 10$, resdshift range $0.01 < z < 0.03$ and metallicities $12 + \log_{10}(O/H) > 8.5$. The stellar masses are derived based on SED fitting (Kauffmann et al., 2003) using the SDSS DR7. In order to obtain the $CO(1 \rightarrow 0)$ line luminosities, they assume that the $CO(2 \rightarrow 1)$ emission line is fully thermalized, this is, $T_b(2 \rightarrow 1)/T_b(1 \rightarrow 0) = 1$ (T_b is the equivalent Rayleigh-Jeans brightness temperature in excess of that of the microwave background). A constant and a variable CO-to-H_2 conversion factor are used to infer H_2 masses. The former case assumes a Milky Way like value $\alpha_{\rm CO} = 4.5({\rm K\ kms}^{-1} pc^{-2})^{-1}$ that includes a correction of 1.36 to account for interestellar Helium. The latter case considers an $\alpha_{\rm CO}$ varying with metallicity according to the Wolfire et al. (2010) models.

The left panels of Fig. 2.2 show the data from the above mentioned samples in the $R_{\rm H_2}$–M_* plane (see the legends indicating each sample inside the panels). The upper and lower panels are for the LTGs and ETGs, respectively. The down arrows indicate the cases of non-detection, with the values plotted corresponding to an upper limit as reported in the original works. The data show that the two populations of galaxies have different loci in the R_{H_I} vs M_* plane, and this is why *it is important to separate them.*

In the right panels of Fig. 2.2, we reproduce the compiled data showed in the left ones but not making differences among the different sources; only detections and upper limits are differentiated. Note that we did not take into account the galaxies from Young et al. (2011) since a significant fraction of them are actually from the Coma cluster and other clusters (these galaxies are from the ATLAS3D catalog); hence, this sample is strongly biased to a particular dense environment, where galaxies have less gas contents than in the field.

As in the case of $R_{\rm H_I}$ vs. M_* (see above), we calculate the mean, standard deviation, median and 25-75 percentiles of the censored + uncensored data in different stellar mass bins by using the Kaplan-Meier estimator. The width of the bins is made as large as to have at least $\sim$ 10 data points and/or a fraction of detections larger than $\sim$ 20% in the bin. For the smallest masses, only upper limits are available, so that it is not possible to calculate a mean or median in these cases. Similar to Fig. 2.1, the black empty circles with error bars show the median and 25-75 percentiles in each mass bin, while the orange squares with error bars show the corresponding mean and standard deviation (the points are slightly shifted to avoid superposition). In the case of the massive ETGs, the fraction of detections are smaller than 50%, therefore, the median and percentiles are not plotted in the corresponding mass bins. However, the mean and standard deviations are plotted, though it should be taken into account that the determinations are uncertain in these cases due to the very low fraction of detections.

2.3 The stellar-gas mass correlations.

In the previous Section, we presented our literature compilation and homogenization of several samples and catalogues of galaxies with information on stellar mass, galaxy color/type, and atomic and/or molecular hydrogen mass. From the compiled data it is clear that the LTG and ETG populations should be studied by separate in that regards their gas contents. In the left panels of Figs. 2.1 and 2.2, we have shown estimates of the mean, standard deviations, median and 25-75 percentiles of the $R_{\rm H_I}$ and $R_{\rm H_2}$ ratios in several stellar mass bins, taking into account the upper limits of the non detections by using the Kaplan-Meier estimator. These estimates will help us below to find the best correlations of $R_{\rm H_I}$ and $R_{\rm H_2}$ with M_* for both LTGs and ETGs.

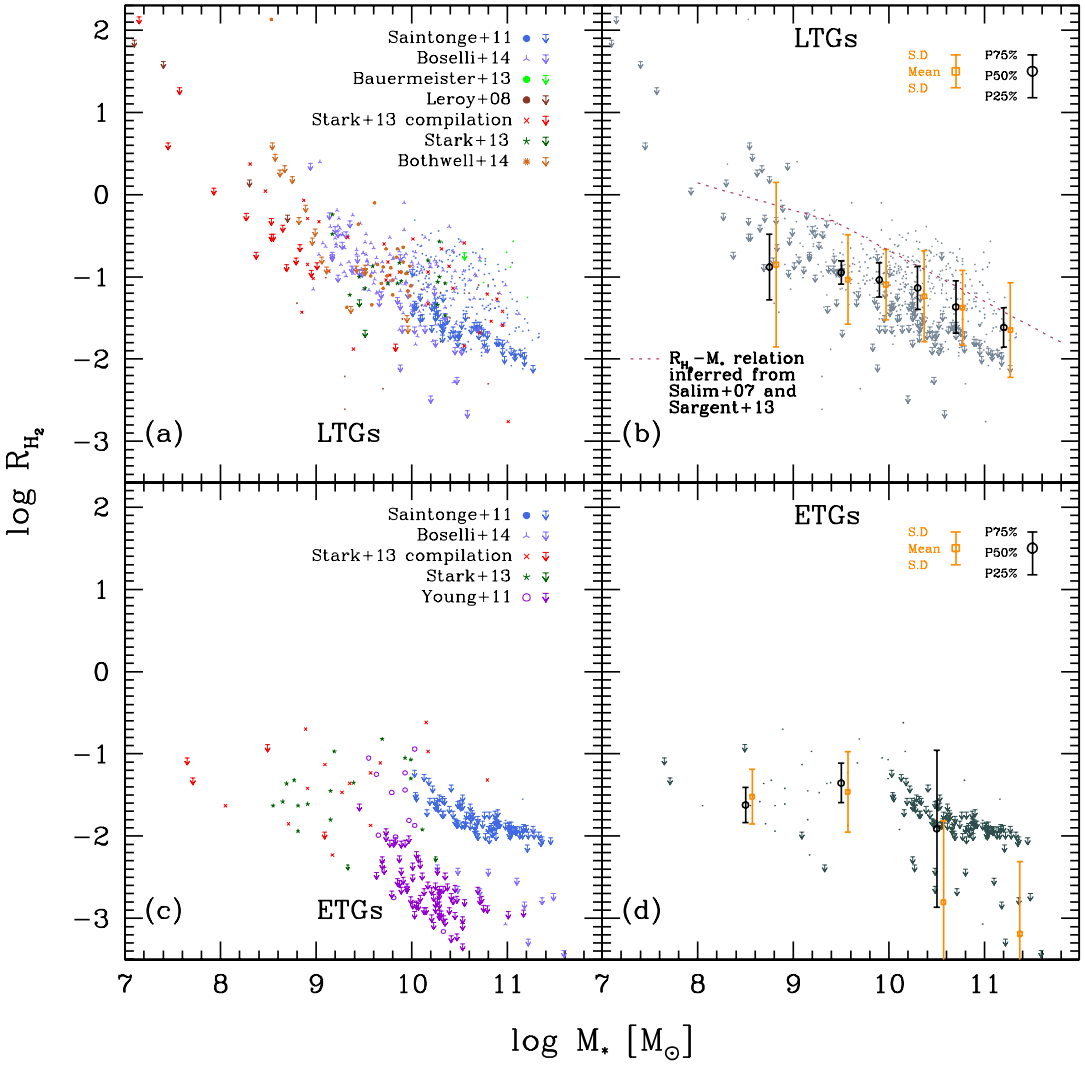

Figure 2.2: *Upper panels:* Compiled data with information on $R_{\rm H_2}$ and M_* for LTGs (see inside the panels for the different sources). We include the reported upper limits for non detections (arrows). Right panel is the same as left one, but with the data plotted without distinguishing the different sources. These data has been grouped in different stellar mass bins in order to calculate the mean, standard deviations, median and 25-75 percentiles in each bin by means of the Kaplan-Meier estimator for censored + uncensored data (see text for more details). For the bins where more than 50% of the data are upper limits, the median and percentiles are not calculated. For the lowest masses there are only upper limits so that mean and median were not calculated. The dashed line shows an estimate for the $R_{\rm H_2}$–M_* relation inferred by combining the empirical SFR–$M_{\rm H_2}$ and SFR–M_* correlations for blue/star-forming galaxies (see text for more details). *Lower panels:* The same as in upper panels but for ETGs. The galaxies from Young et al. (2011) are not included in the right panel due to the high bias to a high-density environment in this sample (see text).

It is important to mention that until the coming of telescopes as the Square Kilometre Array (SKA Carilli, 2014; Blyth et al., 2015), which will bring extragalactic gas studies more in line with optical surveys, the gas-to-stellar ratios of galaxies as a function of mass should be constrained from the current heterogeneous and limited studies based on

radio follow-up observations of small optically-selected galaxy samples as we intend to do here. For H_I, the cross match of the large $\alpha.40$ ALFALFA survey (Giovanelli et al., 2005; Haynes et al., 2011) with the SDSS could be a good option, however, since this survey is H_I-selected, it is biased to high values of R_{H_I} (see e.g., Papastergis et al., 2012; Maddox et al., 2015) we do not this data. In Fig.2.1, we plot the mean and standard deviations from the ALFALFA/SDSS sample as reported by the latter authors; their data are clearly above the data from the optically-selected samples.

According to the mentioned above, the many selection effects and biases of the different samples introduce uncertainties much larger than the accuracy required for formal fittings to the data. In view of this, it is not possible to demand strong statistical rigor in the empirical inferences of the correlations of the gas-to-stellar mass ratios with stellar mass.

In this Section, we push as much as possible a rigorous inference of the R_{H_I}–M_* and R_{H_2}–M_* correlations but it is unavoidable not to resort to extra criteria and physically-motivated assumptions in order to obtain them. Besides, bear in mind that our aim is to get consistency of these correlations with the observed galaxy stellar, H_I and H_2 mass functions.

2.3.1 $R_{H_I} - M_*$ correlations

The mean and median shown in panels (b) and (d) of Fig. 2.1, suggest that the R_{H_I}–M_* correlations for both LTGs and ETGs are not linear in the log-log plane. In fact, if we extrapolate the power law that describes the high-mass end of the R_{H_I}–M_* correlations, at low masses, the values of R_{H_I} would be much larger than the few observational data at these masses. Therefore, the low-mass end should flattens. Besides, for the values of R_{H_I} at these masses being so large, simple models show that this implies too low metallicities as compared to observations. For instance, Baldry et al. (2008) have applied the well-established relation between stellar mass and metallicity coupled with a metallicity to gas mass fraction relation, which can be determined from a simple chemical evolution model. They obtained a gas-to-stellar mass ratio that flattens at low masses, this ratio not being on average larger than ~ 10 at masses as low as $M_* = 10^7$ $M_\odot$.

Another evidence that low mass galaxies (as low as $M_* \sim 10^6 - 10^7$ $M_\odot$) do not have R_{H_I} ratios on average larger than ~ 10 –as would imply the extrapolation of the high-mass R_{H_I}–M_* relation– comes from Maddox et al. (2015). By cross matching the $\alpha.40$ ALFALFA survey with SDSS, they obtain the H_I and stellar masses for thousands of local galaxies. While this sample is not useful to infer the average M_{H_I}–M_* (or R_{H_I}–M_*) relation of galaxies due to its bias towards high values of R_{H_I} (see above), an upper envelope of this relation can be actually constrained; the high-R_{H_I} side does not suffer of selection limit effects. This envelope shows a lack of low mass galaxies with too high R_{H_I} ratios; at $M_* \approx 10^7$ $M_\odot$, the mean value of R_{H_I} is lower than 10. Besides, the envelope shows a clear change of slope at $M_* \sim 10^9$ $M_\odot$, in the sense that at lower masses, the dependence of R_{H_I} on M_* flattens.

Based on the arguments mentioned above, we propose that the R_{H_I}–M_* relation of local LTGs and ETGs down to $M_* \sim 10^7$ $M_\odot$ is better described by a double power law than by a simple power law as usually has been done (e.g., Stewart et al., 2009; Papastergis et al., 2012). We assume also that the break or transition in the R_{H_I}–M_* relation happens at $M_* = 3 \times 10^9$ $M_\odot$, a value close to the one seen in Maddox et al. (2015) or in the result from the mass–metalicity based model of Baldry et al. (2008). The double power law function that we use for fitting the data in the plane spanned by R_{H_I} and M_* is:

$$R_{H_I} = \frac{B}{\left(\frac{M_*}{M^s}\right)^{\xi} + \left(\frac{M_*}{M^s}\right)^{\rho}} \tag{2.14}$$

where B is the normalization, ξ and ρ are the slopes of the function and $M^s = 3 \times 10^9$ $M_\odot$ is the break mass. This function is continuous and derivable. Regarding the *intrinsic* scatter around the R_{H_I}–M_* relations, we assume that it is log-normally distributed (or normally distributed in $\log R_{H_I}$). In a detailed study, Cortese et al. (2011) conclude that the distribution of R_{H_I} is close to a log-normal function. More recently, Lemonias et al. (2013) used the sample by Catinella et al. (2012) to study the H_I mass functions in fixed stellar mass bins (bivariate H_I mass function); by fitting log-normal functions to these H_I mass distributions in fixed bins of M_*, they found widths (scatters) of $\sim 0.6 - 0.7$ dex. The Catinella et al. (2012) sample include both LTGs and ETGs, but the former dominate in number. Note that standard deviations of $\log R_{H_I}$ from our compiled data

for LTGs (where the Catinella et al. sample is also included) are of the order of these values. For ETGs, which are a minor fraction, the scatter is larger.

Since we are interested in the *intrinsic* scatter, it should be taken into account that the measured scatter in the observed samples contains also the observational errors, which may account for up to 0.2-0.25 dex; hence, the obtained standard deviations are actually upper limits to the intrinsic scatters. The available data are too heterogeneous as to attempt to estimate the *intrinsic* scatter. As an educated guess, we will consider that the normally distributed intrinsic scatter around the $R_{\rm H_I}$–M_* relation of LTGs and ETGs are $\sigma_{\log R_{\rm H_I}} = 0.45$ and 0.70 dex, respectively, independent of mass. We use also these scatters for characterizing the intrinsic standard deviation around the mean values of $R_{\rm H_I}$ in each mass bin.

We fit the function given by eq. (2.14) to the mean (orange squares) and intrinsic standard deviations in the $R_{\rm H_I}$–M_* planes showed in Fig. 2.1, under the mentioned above assumptions. The Levenberg-Marquardt method is used for the fitting (Press et al., 1996). For ETGs, the information is quite limited, in special at low masses where there are almost no data. The two lowest mass galaxies are actually blue compact galaxies. While they were classified as spheroidal galaxies, they are actually blue, actively forming stars and could be classified also in the group of LTGs. Because of the lack of information and large uncertainties for ETGs at low masses, we prefer to fix the low-mass end slope ξ to 0. In fact, there are almost no low-mass ($M_* \lesssim 10^9$ $\rm M_\odot$) field galaxies in the local volume which are red/quiescent (Geha et al., 2012).

In Table 2.1, we present the best-fit parameters of the double-power law function for both the LTG and ETG samples. The reduced χ^2 are 0.86 and 0.80, respectively. These χ^2 values lower than 1 are likely due to the large error bars and relatively low number of data (the mean in the different M_* bins) used in the fits.

Galaxy population	**Parameters**			
ETG	$\log_{10}(B)$	ξ	ρ	$\log_{10}(M^s)$
	-0.50	0.00	1.70	9.50
LTG	$\log_{10}(B)$	ξ	ρ	$\log_{10}(M^s)$
	-0.01	0.30	0.85	9.50

Table 2.1: Parameters of the $R_{\rm H_I}$–M_* two-power law relations

Our fits to the $R_{\rm H_I}$–M_* correlations and their (assumed) 1σ scatters for the LTG (left panel) and ETG (right panel) galaxies are shown in Fig. 2.3 (solid lines and shaded areas, respectively). For completeness, we plot also the corresponding fits directly in the $M_{\rm H_I}$–M_* plane, Fig. 2.4. For LTGs, the $M_{\rm H_I}$–M_* correlation at the low-mass and high-mass ends goes as ${M_*}^{0.70}$ and ${M_*}^{0.15}$, respectively (the significant change in slope occurs at $M_* \approx 3 \times 10^9$ $\rm M_\odot$). For ETGs, these dependences go as ${M_*}^{1}$ and ${M_*}^{-0.70}$, respectively; $M_{\rm H_I}$ increases with M_* up to $M_* \approx 3 \times 10^9$ $\rm M_\odot$ and then decreases with M_*.

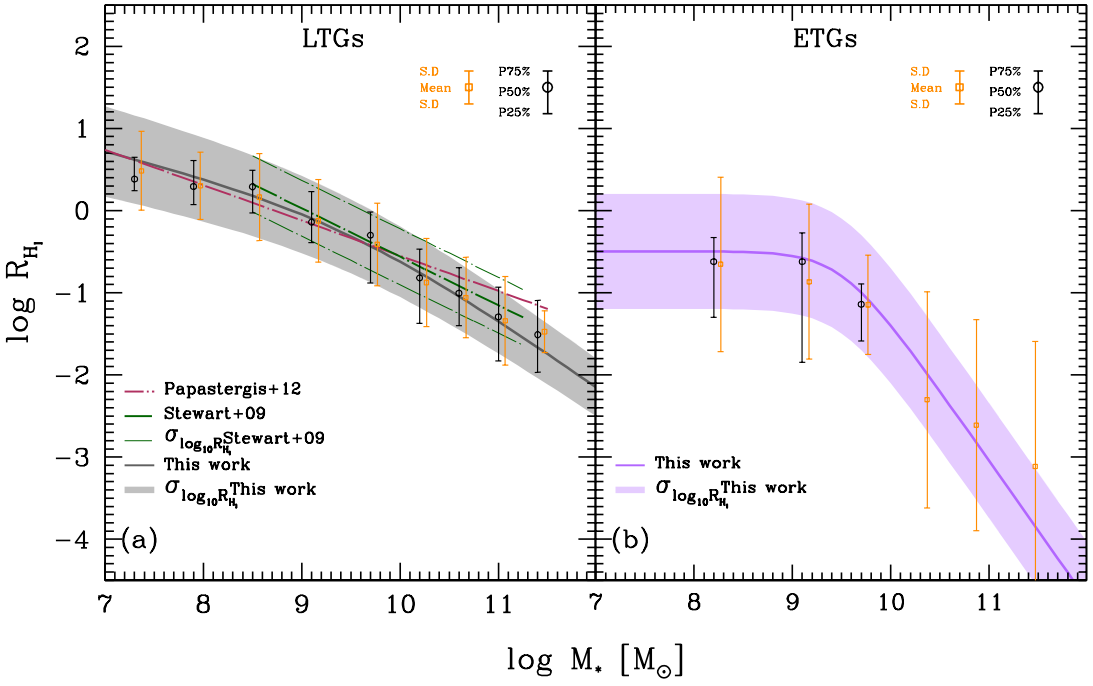

Figure 2.3: Compiled data for $R_{\rm H_I}$ vs. M_* as shown in right panels of Fig. 2.1 and the best fits (gray solid lines) to the mean and standard deviations also showed in Fig. 2.1. The color shaded area are our estimates for the 1σ intrinsic scatter around the relations; this scatter is assumed to be normally distributed in $\log R_{\rm H_I}$. Panels a and b are for the LTG and ETG populations, respectively.

2.3.2 $R_{H_2} - M_*$ correlations

The process of star formation (SF) in galaxies happens inside molecular clouds. Therefore, the SF rate (SFR) in galaxies is expected to correlate with their H_2 content. This is actually the molecular gas version of the global Kennicutt-Schmidt law. Empirical inferences of the SFR–$M_{\rm H_2}$ relations have been obtained for galaxies at different redshifts for a large range of masses and SFR's. For instance, Sargent et al. (2013) infer this relation from several previous observations by using a metallicity-dependent $\alpha_{\rm CO}$ factor.

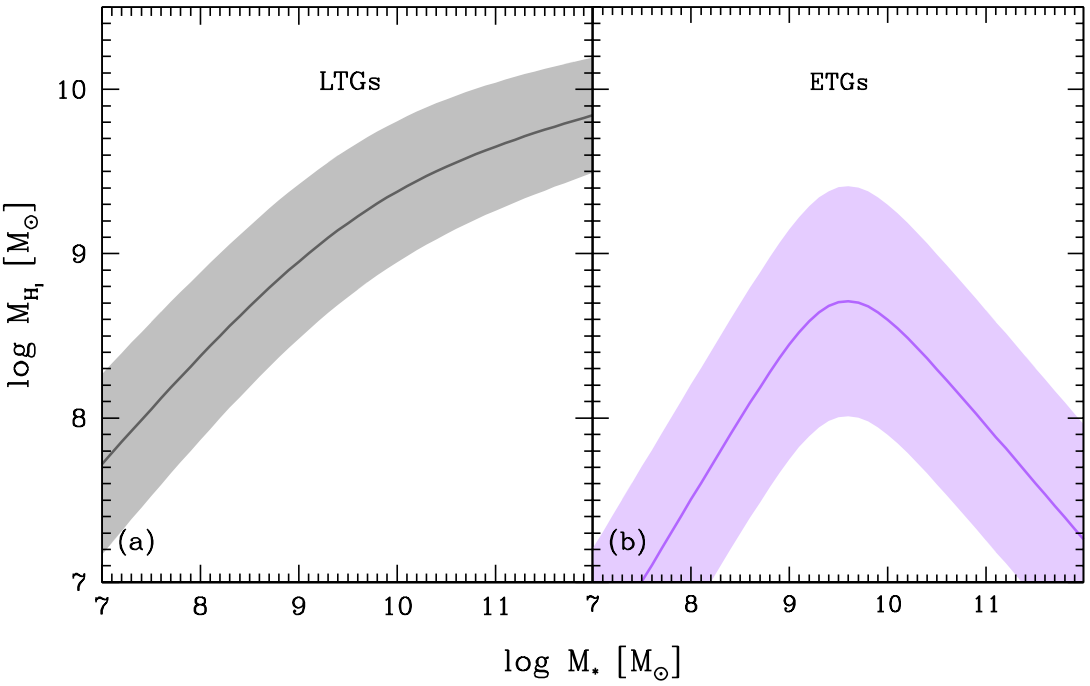

Figure 2.4: Our fits to the $M_{\rm H_I}$-M_* correlation. Panels a and b are for the LTG and ETG populations, respectively. For LTGs (ETGs) $M_{\rm H_I} \propto {M_*}^{0.70}$ and $M_{\rm H_I} \propto {M_*}^{0.15}$ ($M_{\rm H_I} \propto {M_*}^{1}$ and $M_{\rm H_I} \propto {M_*}^{-0.70}$) at the low and high mass end respectively. The color shaded area are the 1σ intrinsic scatters around the fitted $M_{\rm H_I}$-M_* relations.

The SFR–$M_{\rm H_2}$ correlation of "normal" star-forming galaxies and that one of star burst, ULIRG, and sub-milimmetre galaxies have roughly a similar slope but the latter have a normalization (higher SFR) a factor of approximately ten than the former. In the local Universe the fraction of star burst, ULIRG, and sub-milimmetre galaxies is negligible. The Sargent et al. (2013) relation for star-forming galaxies is:

$$\log(M_{\rm H_2}/{\rm M}_\odot) = 9.18 \pm 0.02 + (0.83 \pm 0.03)\log(SFR/{\rm M}_\odot yr^{-1}). \tag{2.15}$$

By using this relation of star-forming galaxies and the SFR–M_* relation for these galaxies (the so-called main sequence), one can find an approximate relationship between $M_{\rm H_2}$ and M_*. We use the SFR/M_*-M_* relation found from SDSS galaxies by Salim et al. (2007) for blue/star-forming local galaxies (their eq. 11):

$$\log(SFR/M_*) = \begin{cases} -0.17(\log(M_*) - 10) - 9.65 & \text{if } \log(M_*) \leq 9.4 \\ -0.53(\log(M_*) - 10) - 9.87 & \text{if } \log(M_*) > 9.4 \end{cases} \tag{2.16}$$

The resulting $M_{\rm H_2}/M_*$–M_* relation is plotted with a dashed line in panel (b) of Fig. 2.2 (for LTGs). It clearly follows a double power law with a break at $M_* \approx 3 \times 10^9$ $\rm M_\odot$.

In panels (b) and (d) of Fig. 2.2, we have presented our compilations of $R_{\rm H_2}$ vs. M_* for LTGs and ETGs, along with the mean and median of the censored + uncensored data in different M_* bins. Again, it seems that a simple power law is not enough to describe the data. Besides, as discussed above, a simple empirical inference based on the observed SFR–$M_{\rm H_2}$ and specific SFR–M_* relations suggests a double power law function, at least for the LTGs. We propose the same double power law function eq. (2.14) to fit the $R_{\rm H_2}$–M_* correlations, as in the case of the $R_{\rm H_I}$–M_* correlations. Following the result obtained above, we assume that the break or transition in the $R_{\rm H_2}$–M_* relations happens at $M_* = 3 \times 10^9$ $\rm M_\odot$. For LTGs, we impose the condition that the slope of the correlation in the mass range $8.5 \lesssim \log(M_*/\rm M_\odot) \lesssim 9.4$ should be close to the one of our empirical inference, which extends down to this mass range; the slope at this mass range is ≈ 0.3 (see the dashed line in Fig. 2.2). Note that at low masses the information is quite limited, in particular for the ETGs. The data in this case, show some evidence that the $R_{\rm H_2}$ ratio not only does not continue increasing at lower masses, but it decreases. In fact, red dwarf spheroidals are not expected to have available molecular gas.

Regarding the intrinsic scatter around the $R_{\rm H_2}$–M_* relations, again, we will assume that it is log-normally distributed. The width of this distribution, in the case of LTGs, which are mostly star-forming objects, is expected to be close to the width of the specific star formation rate (sSFR) vs. M_* relation (the main sequence), because of the connection discussed above between SFR and molecular gas content. The observations of local galaxies show that the scatter around this relation is $0.3 - 0.5$ dex, increasing slightly to lower masses (e.g., Salim et al., 2007). Based on this reasoning as well as on the scatter that we see in our largest compiled samples, we find that a good description for the intrinsic scatter of LTGs is:

$$\sigma_{\log_{10} R_{\rm H_2}} = A' + \varphi' \log_{10} \left(\frac{M_*}{M^s} \right) . \tag{2.17}$$

with $A' = -0.45$ and $\varphi' = -0.04$. In the case of ETGs, the intrinsic scatter seems to be larger than for LTGs. The former are in general passive, devoided of gas reservoirs, but probably a fraction of them can acquire some gas and trigger star formation during interactions and mergers. Then, the amount of $\rm H_2$ depends on the kind of merger and on

the conditions to transform the atomic gas to molecular one. The range of possibilities is large, hence, the scatter around the $R_{\rm H_2}$–M_* relation should be large. We assume the width of the log-normal distribution to be $\sigma_{\log_{10} R_{\rm H_2}} = 0.7$ dex and constant with mass.

As in the case of the H_2 data, we use the Levenberg-Marquardt method to fit the binned data in the $R_{\rm H_2}$–M_* plane for both LTGs and ETGs. The fit is performed for the mean with an intrinsic standard deviation assumed as above explained, and under the conditions also mentioned above. In Table 2.2, we present the best-fit parameters of the double-power law function for both the LTG and ETG samples. The reduced χ^2 are 0.92 and 0.72, respectively. The χ^2 values lower than 1 are likely due to the relatively large standard deviations and low number of data (the mean in the different M_* bins) used in the fits.

Galaxy population	**Parameters**			
ETG.	$\log_{10}(B)$	ξ	ρ	$\log_{10}(M^s)$
	-1.05	-0.45	1.35	9.50
LTG.	$\log_{10}(B)$	ξ	ρ	$\log_{10}(M^s)$
	-0.50	0.20	0.65	9.50

Table 2.2: Parameters of the $R_{\rm H_2}$–M_* double power-law correlations

Our fits to the $R_{\rm H_2}$–M_* correlations and their scatters for the LTG (left panel) and ETG (right panel) galaxies are shown in Fig. 2.5 (solid lines and shaded areas). For completeness, we plot also our fits directly in the $M_{\rm H_2}$–M_* plane, Fig. 2.6. For LTGs, the $M_{\rm H_2}$–M_* correlation at the low-mass and high-mass ends goes as ${M_*}^{0.80}$ and ${M_*}^{0.35}$, respectively (the significant change in slope occurs at $M_* \approx 3 \times 10^9$ $\rm M_\odot$). For ETGs, these dependences go as ${M_*}^{1.45}$ and ${M_*}^{-0.35}$, respectively; $M_{\rm H_I}$ increases with M_* up to $M_* \approx 3 \times 10^9$ $\rm M_\odot$ and then decreases with M_*.

2.3.3 $R_{\rm gas} - M_*$ correlations.

Once determined the $R_{\rm H_I}$–M_* and $R_{\rm H_2}$–M_* correlations, we can obtain the $R_{\rm gas}$–M_* correlation, where $R_{\rm gas} = M_{\rm gas}/M_* = 1.4(R_{\rm H_I} + R_{\rm H_2})$, or in logarithmic form:

$$\log_{10}(R_{\rm gas}) = \log_{10}(1.4) + \log_{10}(R_{\rm H_I} + R_{\rm H_2}) \tag{2.18}$$

By introducing the $R_{\rm H_I}$–M_* and $R_{\rm H_2}$–M_* (eq. 2.14 with parameters from tables 2.1 and 2.2) relations in eq. (2.4), the $R_{\rm gas}$–M_* relations for both galaxy populations are

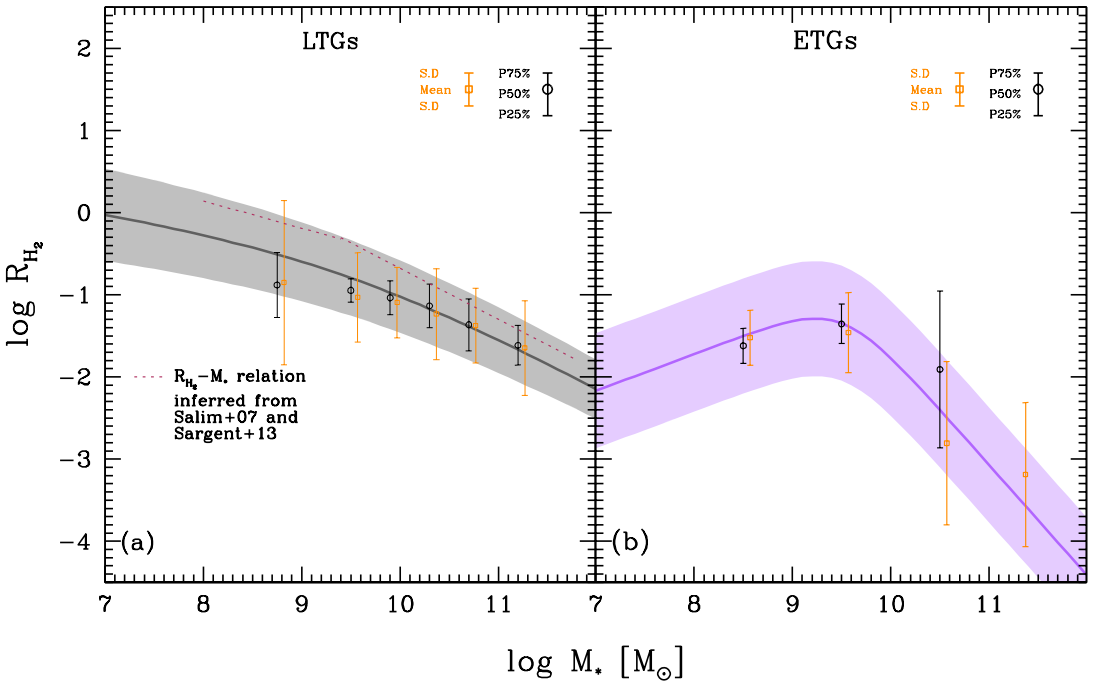

Figure 2.5: Compiled data for $R_{\rm H_2}$ vs M_* as shown in right panels of Fig. 2.2 and the best fits (gray solid lines) to the mean and standard deviations also showed in Fig. 2.2. The color shaded area are our estimates for the 1σ intrinsic scatter around the relations; this scatter is assumed to be normally distributed in $\log R_{\rm H_I}$. Panels a and b are for the LTG and ETG populations, respectively.

obtained. To get the scatters around these relations we propagate the scatters determined for the $R_{\rm H_I}$–M_* and $R_{\rm H_2}$–M_* relations separately for both galaxy populations. Propagating errors, in the assumption of null covariances, we obtain the logarithmic scatter around the $R_{\rm gas}$–M_* relation[5],

$$\sigma_{\log_{10} R_{\rm gas}} = \frac{1}{R_{\rm H_I} + R_{\rm H_2}} \left(R^2_{\rm H_I} \sigma^2_{\log_{10} R_{\rm H_I}} + R^2_{\rm H_2} \sigma^2_{\log_{10} R_{\rm H_2}} \right)^{\frac{1}{2}} \tag{2.19}$$

In Fig. 2.7, we plot the obtained $R_{\rm gas}$–M_* correlations for LTG (left panel) and ETG (right panel) populations. The color regions are the corresponding propagated scatters. For completeness, we plot in this figure also those galaxies from our compilation that have determinations for both the $\rm H_I$ and $\rm H_2$ masses. Note that a large fraction of our compilation have not determinations for both quantities at the same time.

According to Fig. 2.7, the $R_{\rm gas}$–M_* correlations of LTG and ETG galaxies are significantly different. The gas content in the former is at all masses larger than in the latter,

[5]If the value of $R_{\rm H_I}$ $R_{\rm H_2}$ or both cases are an upper limit, then Eq. (2.19) is an upper limit of the intrinsic scatter of $M_{\rm gas}$–M_* relation.

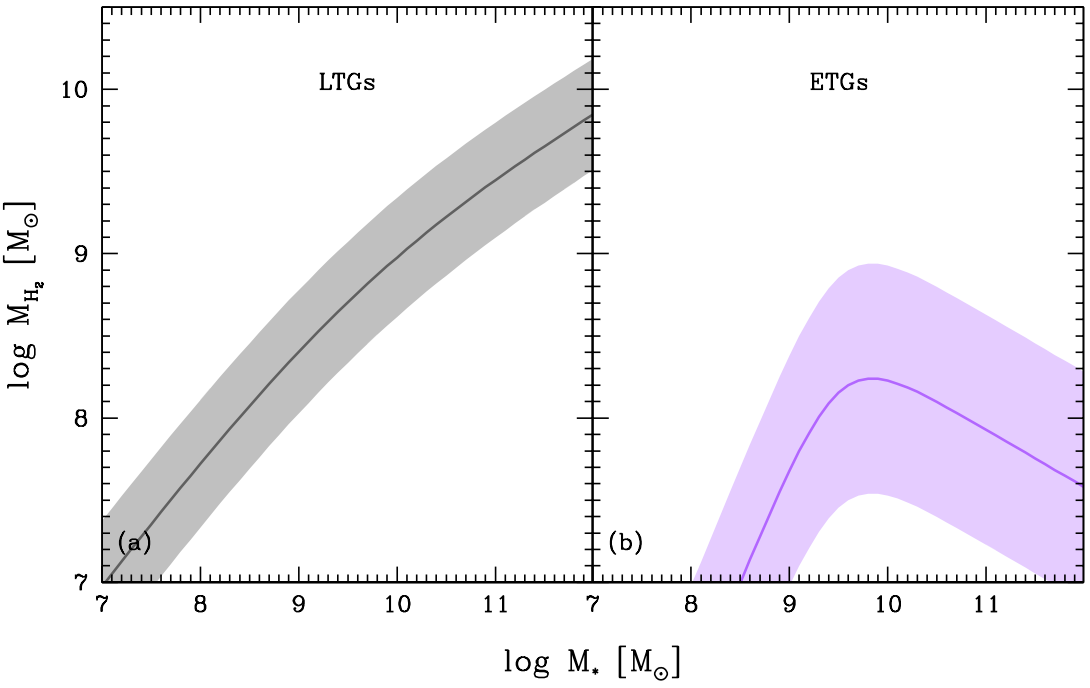

Figure 2.6: Our fits to the $M_{\rm H_2}$-M_* correlation. Panels a and b are for the LTG and ETG populations, respectively. For LTGs (ETGs) $M_{\rm H_2} \propto {M_*}^{0.80}$ and $M_{\rm H_2} \propto {M_*}^{0.35}$ (${M_{\rm H_2}} \propto {M_*}^{1.45}$ and $M_{\rm H_2} \propto {M_*}^{-0.35}$) at the low and high mass end respectively. The color shaded area are the 1σ scatters around the fitted $M_{\rm H_2}$-M_* relations.

specially at larger masses. In general, in both cases, the smaller the galaxies, the more dominated by cold gas they are. For LTGs, cold gas can be on average ten times more massive than stellar mass for galaxies of stellar masses around $10^7 {\rm M}_\odot$.

2.3.4 Gas fractions

In the literature it is common to define the cold gas fraction, $f_{\rm gas}$=$M_{\rm gas}$/($M_{\rm gas}$+ M_*), instead of the cold gas-to-stellar mass ratio, $R_{\rm gas}$. We can easily get $f_{\rm gas}$ as a function of M_* by using eq. (2.7) and the $R_{\rm gas}$–M_* relation above obtained. The scatter around this relation can be obtained by error propagation from the one around the $R_{\rm gas}$–M_* relation eq. (2.19) and by using eq. (2.7):

$$\sigma_{f_{\rm gas}} = \frac{R_{\rm gas}}{\log_e(10)\,[R_{\rm gas}+1]} \left(1 - \frac{R_{\rm gas}}{R_{\rm gas}+1}\right) \sigma_{\log_{10} R_{\rm gas}} \tag{2.20}$$

In Fig. 2.8, we present the $f_{\rm gas}$-M_* correlations and their scatters for the LTG (left panel) and ETG (right panel) populations. As in Fig. 2.7, we plot, for completeness, the data from our compilation that has determinations for both $\rm H_I$ and $\rm H_2$ masses. As

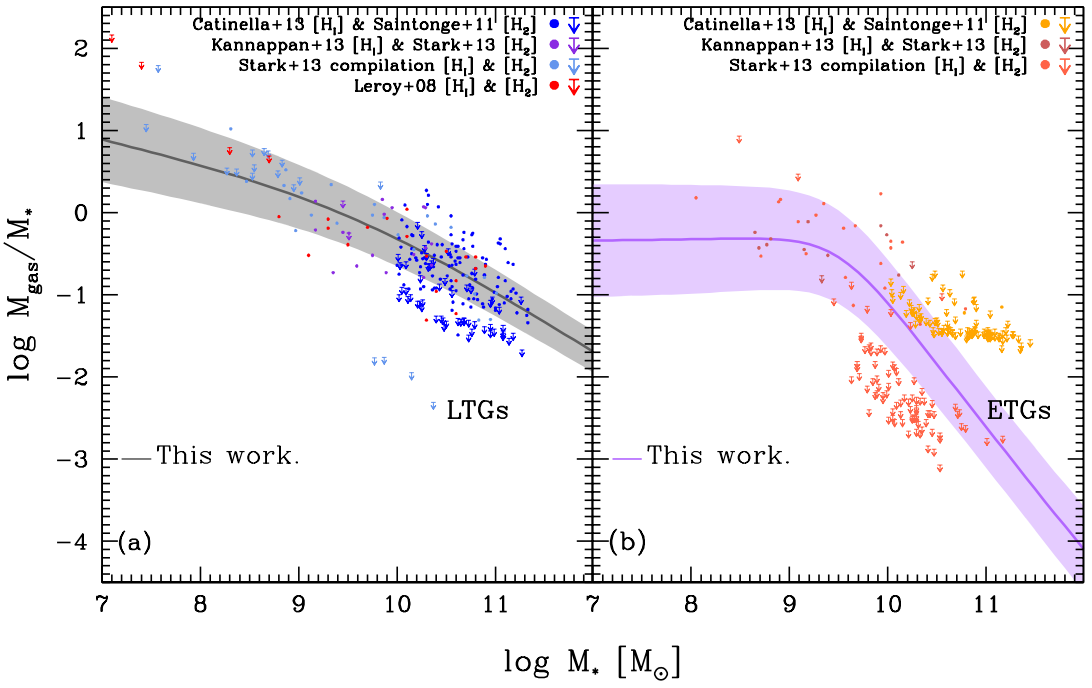

Figure 2.7: Cold gas-to-stellar mass ratio, $R_{\rm gas}$, vs. M_* for LTG (left panel) and ETG (right panel) populations inferred from our fits to the $R_{\rm H_I}$–M_* and $R_{\rm H_2}$–M_* correlations. The shaded area are the 1σ scatters obtained by error propagation of the scatters around the $R_{\rm H_I}$–M_* and $R_{\rm H_2}$–M_* relations. For completeness, also are plotted the data from our compilation that have determinations of both $\rm H_I$ and $\rm H_2$ masses. Note that the plotted $f_{\rm gas}$–M_* relations are not fits to these data.

expected, the gas fraction strongly increases as less massive are the galaxies, tending to 90% at $M_* \approx 10^7$ $\rm M_\odot$. For a Milky-Way sized galaxy, in the case of LTGs, $f_{\rm gas} \approx 12\%$, while for ETGs, $f_{\rm gas} < 0.5\%$. Our Milky Way has $f_{\rm gas} = 10 - 15\%$.

2.4 $M_{\rm H_2}/M_{\rm H_I}$-M_* correlation.

Understanding the process that makes possible the conversion of atomic into molecular hydrogen, would give us insights on how do galaxies form stars. A quantity that characterizes the global efficiency of this process is the $M_{\rm H_2}/M_{\rm H_I}$ ratio. We study then how this ratio correlates with M_* for the LTG and ETG populations by using the observational data compilation and fits presented in Sections 2.2.1 and 2.2.2.

The $M_{\rm H_I}$-M_*, $M_{\rm H_2}$-M_* correlations and their scatters as fitted to the observations were given in sections 2.3.1 and 2.3.2. From them, we obtain the $M_{\rm H_2}/M_{\rm H_I}$ ratios at each M_* as well as the scatters. The results are plotted in Fig. 2.9, where the galaxies for which there are information on both $M_{\rm H_I}$ and $M_{\rm H_2}$ are also included.

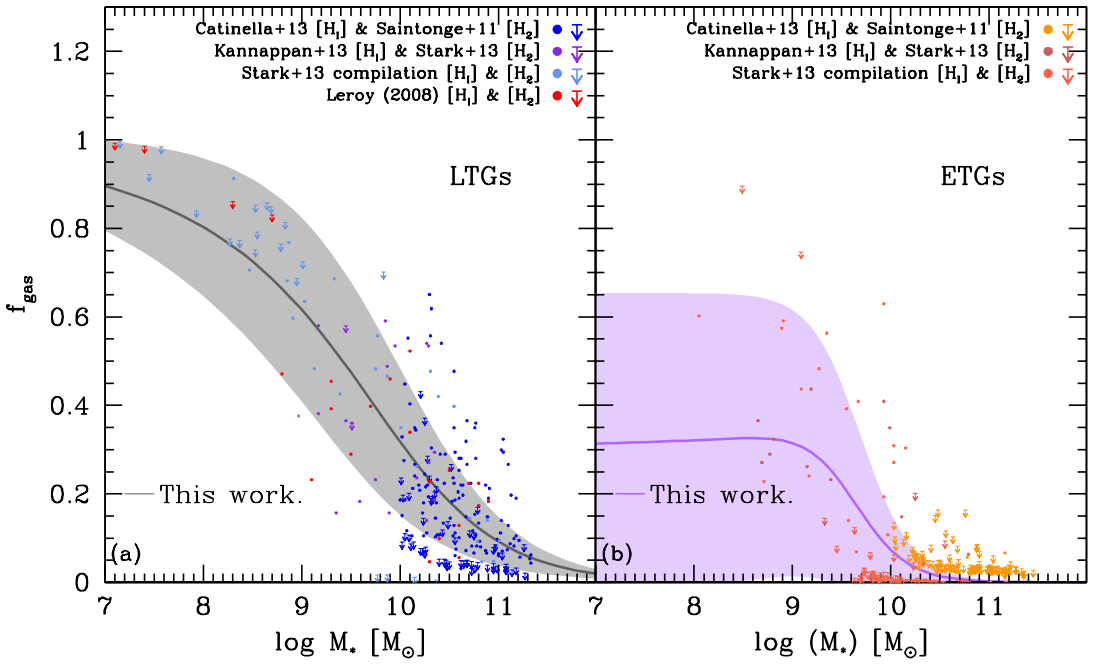

Figure 2.8: Cold gas fraction, $f_{\rm gas}$, vs. M_* for LTG (left panel) and ETG (right panel) populations inferred from our fits to the $R_{\rm H_I}$–M_* and $R_{\rm H_2}$–M_* correlations. The shaded area are the 1σ scatters obtained by error propagations of the scatters around the $R_{\rm H_I}$–M_* and $R_{\rm H_2}$–M_* relations. For completeness, also are plotted the data from our compilation that have determinations of both $\rm H_I$ and $\rm H_2$ masses. Note that the plotted $f_{\rm gas}$–M_* relations are not fits to these data.

According to figure 2.9, the molecular-to-atomic mass ratio of LGTs is on average 0.2 for the lowest mass galaxies, meanwhile for the largest masses the ratio is on average 1. Our results are similar at $M_* \gtrsim 3\times10^{10}\rm M_\odot$ when we compare with the observational $\rm H_2$-to-$\rm H_I$ metallicity dependent mass ratio of Bothwell et al. (2014) and the semi-analytical inferences of Fu et al. (2012) and Lagos et al. (2011), but we differ at low stellar masses, our results are higher. Thus, LTGs are expected to have higher global SF efficiencies at high masses, i.e., the ability of local LTGs to transform atomic gas into molecular gas (and then into stars) slowly increases with stellar mass. For ETGs, $M_{\rm H_2}/M_{\rm H_I}$ strongly decreases with decreasing M_*, so that the less massive the galaxy, the less efficient they are. ETGs have total cold gas contents much smaller than LTGs. For the lowest ETGs, the molecular-to-atomic mass ratio is on average 0.02, while for the largest masses the ratio is on average 2, however, there is a huge scatter around the main trend of increasing the ratio with M_*.

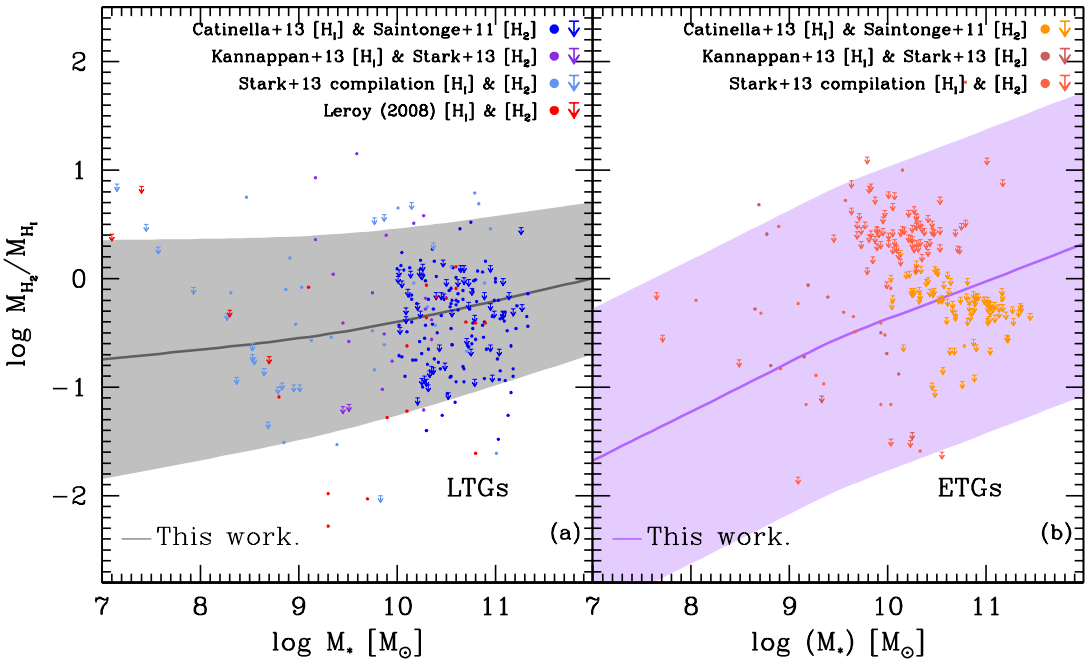

Figure 2.9: Molecular-to-atomic mass ratio, $M_{\rm H_2}/M_{\rm H_I}$ for LTG (left panel) and ETG (right panel) galaxies inferred from our fits to the $R_{\rm H_I}$-M_* and $R_{\rm H_2}$-M_* correlations. The shades areas are the 1σ scatters obtained by error propagation of the scatters around the $R_{\rm H_I}$-M_* relations. For completness, also are plotted the data from our compilation that have determinations for both H_I and H_2 masses. Note that the plotted $M_{\rm H_I}/M_{\rm H_2}$-M_*relations are not fits to this data.

Chapter 3

The stellar, gas, and baryonic mass function of local late and early type galaxies

In the previous Chapter, we have obtained from a compilation of observational data the M_{H_I}– and M_{H_2}–M_* correlations, for ETGs and LTGs. In this Chapter we use these correlations in order to *predict* the galaxy mass distribution functions for M_{H_I}, M_{H_2}, $M_{\rm gas}$ and $M_{\rm bar}$ by using as an "interface" the observed galaxy stellar mass function, GSMF. In what follows, we describe the method employed for constructing these galaxy mass functions, and probe this way how the mass correlations do project into distribution functions and viceversa.

3.1 The method

3.1.1 Observed Galaxy Stellar Mass Functions

The different local M_{H_I}–M_* and M_{H_2}–M_* correlations, parameterized in Chapter 2, in combination with the local observed GSMF, allow us to determine galaxy mass distribution functions for M_{H_I}, M_{H_2}, $M_{\rm gas}$ and $M_{\rm bar}$. In this subsection, we explain first how we construct the local (total) GSMF separated into blue and red galaxies (LTGs and ETGs). For this, we combine the overall SDSS GSMF determined by Baldry et al. (2008) down to $M_* \approx 10^7$ $M_\odot$ and the one determined by Rodríguez-Puebla et al. (2015) for galaxies above $M_* \approx 10^9$ $M_\odot$; the latter is also based on SDSS galaxies but it takes into account the separation into blue/red populations and introduces an "aperture" correction to the stellar mass estimates. Basically, what we do is to extend the Rodríguez-Puebla et al. (2015) GSMF to smaller masses by "matching" the Baldry et al. (2008) GSMF. The latter authors have introduced careful corrections relevant for the completeness. On the

other hand, the data release of the SDSS catalog used by the former authors is posterior to the one used by Baldry et al. (2008) in such a way that more galaxies and a larger volume is considered; this improves the issue with cosmic variance that affects the high–mas end of the GSMF. The composed GSMF is presented in Fig 3.1 as filled black circles with error bars.

Baldry et al. (2008) derived the GSMF based on the NYU-VAGC catalog from the SDSS DR4 for galaxies at $z < 0.05$. By combining several methods for stellar mass determinations, Baldry et al. (2008) estimated the GSMF between $M_* = 10^7$ $M_\odot$ and $M_* = 10^{11.8}$ $M_\odot$. The authors perform a careful analysis to take into account incompleteness at low masses. In spite of that, their determinations may be yet incomplete due to missing faint low-surface brightness galaxies at $M_* \lesssim 10^8 M_\odot$. We do not attempt to introduce uncertain corrections to this possible incompleteness and use just the GSMF reported by them down to $M_* \sim 10^7$ $M_\odot$. Regarding the shape of this GSMF, the authors find hints for an upturn at the low mass-end when the GSMF is fitted to a double Schechter.

In a recent analysis of the SDSS DR7, Rodríguez-Puebla et al. (2015) constructed the GSMF between $M_* = 10^9$ and $M_* = 10^{12.3}$ $M_\odot$ based on the NYU-VAGC catalog for all, *blue and red* galaxies. The stellar masses were taken from the Spectral Energy Distribution inferences reported in the MPA-JHU DR7 database.[1] Colors and absolute magnitudes were K-corrected to $z = 0.1$ and separated into blue and red galaxy populations according to the magnitude-color criteria given in Li et al. (2006). In addition, the GSMF constructed according to Rodríguez-Puebla et al. (2015) takes into account that galaxy surface brightness profiles extends further away than the commonly used aperture limits in SDSS. These authors introduced a correction following Mendel et al. (2014), who compared the MPA-JHU DR7 stellar masses with the ones obtained by computing accurate bulge+disk and Sérsic profile photometric decompositions in several bands. In the case of a Sérsic profile, they find masses larger by up to ≈ 0.08 dex at the smallest masses and by ≈ 0.23 dex at the largest masses. Following these results, Rodríguez-Puebla et al. (2015) corrected conservatively the masses by ≈ 0.05 dex for masses up to $\log(M_*/M_\odot) \sim 10.7$ and then increase smoothly the correction ending with 0.23 dex at $\log(M_*/M_\odot) \sim 12$. The correction is particularly important for (massive) central galaxies in clusters (see e.g., Bernardi et al., 2013; He et al., 2013; Kravtsov

[1] Available at http://www.mpa-garching.mpg.de/SDSS/DR7.

et al., 2014; Mendel et al., 2014, and reference therein). We assume that this aperture correction is the same for blue and red galaxies.

For our purposes of using the observed GSMF as an interface to determine gas and baryonic mass functions, it is more convenient to have an analytical version of the GSMF. Following Baldry et al. (2008) and Rodríguez-Puebla et al. (2015), we fit our combined GSMF by a function that is composed of a single Schechter plus a Schechter function with a sub-exponential decay at the high-mass end,

$$\phi_g(X)dX = \phi_1^* X_1^{\alpha_1} e^{-X_1} dX_1 + \phi_2^* X_2^{\alpha_2} e^{-X_2^{\beta}} dX_2, \tag{3.1}$$

where $X_i = M_*/\mathcal{M}_i^*$ with $i = 1, 2$. The corresponding best-fit parameters are listed in Table 3.1.

Figure 3.1 shows the best fit function (solid line) to the our combined GSMF (filled circles with error bars). In the same figure, the short-dashed and long-dashed lines show the Schechter and the Schechter with a sub-exponential decay fitted components, respectively. The bottom panel shows the residuals between the model fit and observations. Our fit reproduce accurately observations at the level of less than 5% for most of the mass range.

For comparison, we also reproduce in Fig. 3.1 previous GSMFs published in Li and White (2009, green squares with error bars) and Bernardi et al. (2013, magenta triangles with error bars). Note that Li and White (2009) used the standard aperture limits from the SDSS. In contrast, Bernardi et al. (2013) used more sophisticated models fits for the surface brightness profile than the standard fits from SDSS, similar to the correction that Rodríguez-Puebla et al. (2015) have introduced.

In order to separate the GSMF into blue and red galaxies, we approximate the fraction of red galaxies as:

$$f_r = \frac{1}{1 + (M_*/M_{\rm char})^{\gamma}}. \tag{3.2}$$

The fraction of blue galaxies is the complement, $f_b = 1 - f_r$. When fitting this function to the observed SDSS-DR7 NYU-VAGC red fraction from Rodríguez-Puebla et al. (2015), we obtain the following best fit parameters: $\log(M_{\rm char}/{\rm M}_{\odot}) = 10.1$ and $\gamma = -0.40$. We

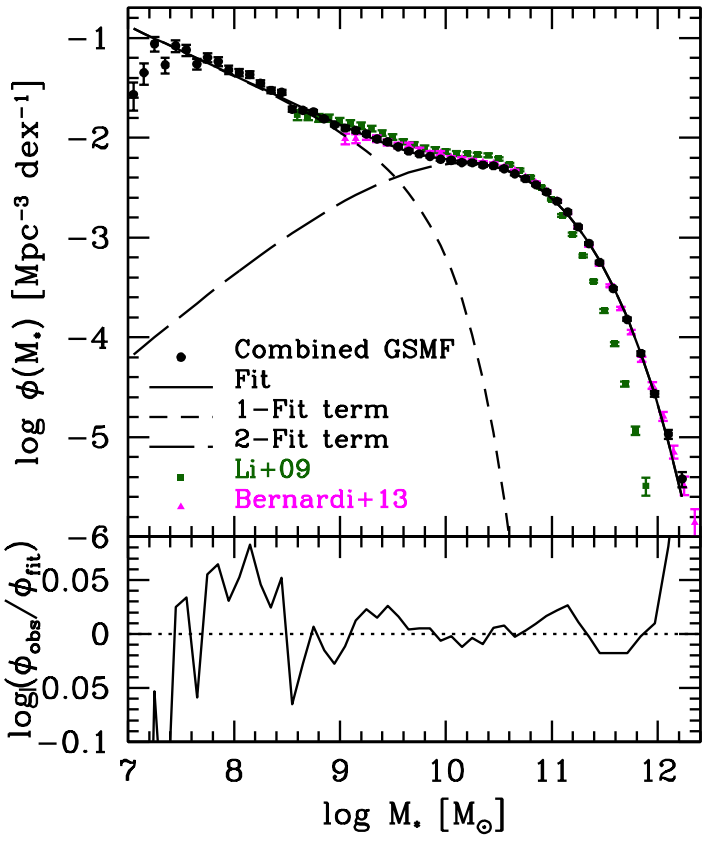

Figure 3.1: Total GSMF constructed here from matching the SDSS GSMFs by Baldry et al. (2008) and Rodríguez-Puebla et al. (2015, black solid circles with error bars). The short and long dashed lines correspond to the Schechter and Schechter with sub exponential decay functions that in combination (solid line) provide the best fit to our GSMF. The lower panel shows the fractional deviations of the best fit from the observed GSMF. The GSMFs reported in Li and White (2009) and Bernardi et al. (2013) are also plotted for comparison.

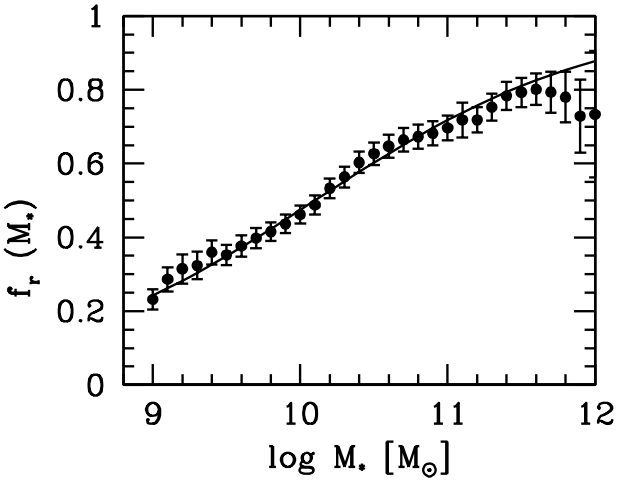

Figure 3.2: Fraction of red (ETG) galaxies as a function of M_* in the NYU-VAGC DR7 sample processed in Rodríguez-Puebla et al. (2015) and used here to construct our GSMF. The solid line is the best-fit function given in eq. (3.2). This fraction extrapolated down to 10^7 $M_\odot$ is the one we use to calculate the red (ETG) GSMF. The blue (LTG) GSMF is calculated using the complement, i.e., $f_b = 1 - f_r$.

show this in Figure 3.2. Solid circles with error bars present the observed fraction of red galaxies as a function of mass, while the solid lines indicates our best fitting model.

We use these fractions in order to compute the GSMFs of blue and red galaxies as: $\phi_{g,i} = f_i \times \phi_g$, where subscript '$i$' refers either to blue or red galaxies. Note that for masses smaller than $\sim 5 \times 10^8$ $M_\odot$, the red/blue fractions are just extrapolations of the determinations mentioned above. In any case, observations suggest that most of dwarfs are indeed blue, star-forming galaxies, unless they are satellites (see e.g., Geha et al., 2012). However, the ratio of satellite-to-centrals at a given mass is small, specially at low masses.

3.1.2 A Simple Model For Predicting Galaxy Mass Functions.

As discussed in Chapter 2, the M_{H_I} and M_{H_2} empirical correlations with M_* strongly differ among LTGs and ETGs. This is why we decided namely to separate both populations and obtain separate correlations. Fortunately, we are able to obtain the GSMF also separated into LTGs and ETGs. Actually, the separation is only by color, into blue and red galaxies. However, as mentioned in the previous Chapter, we use color and/or type in order to separate all galaxies into two rough populations. In this way, we are now able to estimate the mass distribution functions for M_{H_I}, M_{H_2}, $M_{\rm gas}$ and $M_{\rm bar}$ separated into LTG and ETG populations, as well as the overall functions. The correlations parameterized as a function of M_* mentioned in section 2.3 are then "projected" into the different mass distributions by using the GSMFs as an interface. For this, we start by generating a volume complete mock galaxy population based on our GSMF separated into blue (LTG) and red (ETG) galaxies. Each galaxy of mass M_* in this mock catalog has a chance of being either blue or red according to the fraction f_r (see eq. 3.2). By using the $\rm H_I$ and H_2-to-stellar mass empirical relations and their scatters (for both galaxy populations), the masses M_{H_I} and M_{H_2}(and therefore $M_{\rm gas}$ and $M_{\rm bar}$) are assigned to each M_* according to its color/type. Therefore, a mock catalog of LTGs/ETGs with M_*, M_{H_I}, M_{H_2}, $M_{\rm gas}$, and $M_{\rm bar}$ is generated, in such a way that the total, blue (LTG), and red (ETG) GSMFs of this catalog reproduce well the corresponding observed GSMFs presented above.

For our propose, it is necessary to introduce the conditional probability distribution functions $P_b(M_{H_I}|M_*)$ and $P_r(M_{H_I}|M_*)$ that a galaxy of mass M_* hosts either a blue or

red galaxy in the HI mass bin $M_{H_I} \pm dM_{H_I}/2$, respectively. Each probability distribution function $P_j(M_{H_I}|M_*)$ is assumed to be a lognormal function with mean $\langle \log M_{H_I}(M_*) \rangle_j$ and width (standard deviation) $\sigma_{H_I,j}(M_*)$, where the subscript 'j' denotes either blue (LTG) or red (ETG); that is, we are using the empirical M_{H_I}–M_* correlation determined in Chapter 2, Section 2.3.1. Similarly, we introduce the distribution functions $P_b(M_{H_2}|M_*)$ and $P_r(M_{H_2}|M_*)$ for assigning H_2 masses, assuming that they are lognormal functions with mean $\langle \log M_{H_2}(M_*) \rangle_j$ and standard deviation $\sigma_{H_2,j}(M_*)$; again here the subscript 'j' denotes either blue (LTG) or red (ETG), and we are using the empirical M_{H_2}–M_* correlation determined in Chapter 2, Section 2.3.2.

In more detail, our mock catalog is generated as follows:

1. A minimum galaxy stellar mass $M_{*,\rm min}$ is set, for example = 10^7 $M_\odot$. From this minimum we generate a population of 3×10^6 galaxies sampled from the empirical (combined) GSMF presented above. Each galaxy is randomly picked from this function by generating a random number $\mathcal{U}$ uniformly distributed within the interval $[0, 1]$ and finding the value for M_* that solves the equation $n_g(> M_*)/n_g(> M_{*,\rm min})$.[2] We denote the cumulative GSMF as n_g. With this procedure we sample the given empirical GSMF distribution.

2. Each mock galaxy is assigned either as blue (LTG) or red (ETG), according to the empirical probability given by Eq. (3.2).

3. Once each galaxy has assigned an M_* and marked either as blue or red, the HI gas mass M_{H_I} is assigned randomly according to the probability distribution function $P_j(M_{H_I}|M_*)$. The subscript 'j' refers either to blue (LTG) or red (ETG).

4. Similarly, we repeat the procedure in the previous item but this time for assigning the H_2 gas mass, M_{H_2}, according to the probability distribution function $P_j(M_{H_I}|M_*)$.

5. Finally, total gas mass, $M_{\rm gas}$, and baryon mass $M_{\rm bar}$, are assigned as $M_{\rm gas} = 1.4(M_{H_I} + M_{H_2})$ and $M_{\rm bar} = M_{\rm gas} + M_*$, respectively.

[2] We make use of the fact that the integrals of two probability distributions are equal, $\int_0^x h(x')dx' = \int_0^y g(y')dy'$. In the case of a uniform distribution between 0 and 1, $\int_0^x U(x')dx' = x$. Therefore, a random number y from the distribution $g(y)$ can be obtained by resolving the equation $x = \int_0^y g(y')dy'$, where x is a random number from the [0,1] interval.

Our mock galaxy catalog is a volume-complete sample of 3×10^6 members corresponding to a co-moving volume of 2.28×10^7 Mpc^3, which is comparable with one of the largest and most resolved dissipationless cosmological ΛCDM simulation, Bolshoi($V_{Bolshoi} \sim 4.5 \times 10^7$Mpc^3 (see, Klypin et al., 2011)). Note that our mock galaxy catalog is mass-limited in the GSMF, with $M_{*,\rm min} = 1 \times 10^7$ M$_\odot$. Since the H$_{\rm I}$, H_2, $M_{\rm gas}$, and $M_{\rm bar}$ distribution functions are constructed from the GSMF, this mass limit will propagates in different ways to these mass functions. The co-moving volume of 2.28×10^7 Mpc^3 in our mock galaxy catalog is big enough in order to avoid significant effects from Poisson noise. This noise affects specially the counts of massive galaxies, which are the less abundant objects. Nevertheless, real observations are subject to several source of uncertainties. For the GSMF, these uncertainties (black error bars in Fig. 3.1) are mostly from cosmic variance plus statistical counts. In order to take into account these source of uncertainties in our mock galaxy catalog, *which is intended to be as the observed one*, we model the error bars from the observational GSMF and include them in the mock catalog. Figure 3.3 shows with dots the magnitude of these error bars as a function of M_*. Note that the error bars (in dex) can be slightly asymmetrical. At $M_* \approx 10^9$ M$_\odot$ there is a discontinuity; recall that our GSMF was obtained as the match of two GSMFs (see above). We fit a three-step function to the upper/lower error bars (black line). The function is given by:

$$\begin{aligned} \sigma_{\log_{10}\phi(M_*)} &= \varphi_1 \left(\log_{10}(M_*) - \lambda_1\right)^3 + \omega_1 && 7 \leq \log_{10}(M_*) \leq 9.5 \\ \sigma_{\log_{10}\phi(M_*)} &= 0.02 && 9.5 \leq \log_{10}(M_*) \leq 11.5 \\ \sigma_{\log_{10}\phi(M_*)} &= \varphi_2 \left(\log_{10}(M_*) - \lambda_2\right)^2 + \omega_2 && 11.5 \leq \log_{10}(M_*) \leq 12.4 \end{aligned} \tag{3.3}$$

The best fit parameters are $\varphi_1 = -0.0064$, $\lambda_1 = 9.5$, $\omega_1 = 0.02$, $\varphi_2 = 0.083$, $\lambda_2 = 11.5$, $\omega_2 = 0.02$. The best fit is shown in Fig. 3.3 with the solid line. The fit to the observational uncertainties in the GSMF, assumed as lognormal, is used at each of the assignment steps described above for introducing an observational-like scatter to these assignments.

Finally, we should say that we ignore any contribution from systematical uncertainties in stellar mass estimates in the GSMF. The systematical uncertainties in stellar mass estimates can be as large as ~ 0.25 dex (e.g., Behroozi et al., 2010) and will result only in a systematic shift in the stellar mass axis of the GSMF.

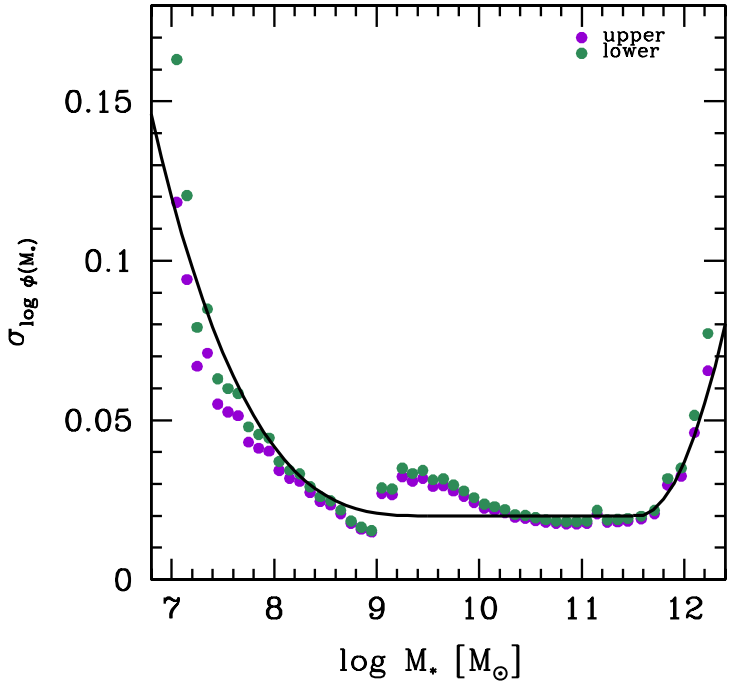

Figure 3.3: Upper and lower errors in the GSMF plotted in Fig. 3.1. The solid line is a three-step function that we use to fit these errors as a function of M_*, assuming they to be lognormally distributed. The discontinuity seen at $M_* \sim 10^9$ $M_\odot$ is due to the differences in the errors of the two observational GSMFs used to construct our one.

3.2 Results

In this Section we present our inferred stellar, H_I, H_2, gas, and baryonic mass functions split in all, late/blue (LTG) and early/red (ETG) galaxies (black, blue, and red solid circles respectively in all the figures presented in this Section). Also, we use the Levenberg–Marquardt method to fit our MF results with a Schechter, Schechter with subexponential decay or a combination of both functions. These are given by,

$$\phi(M_k)_j = \sum_i \phi_i(M_k)_j \tag{3.4}$$

with

$$\phi_i(M_k)_j = \ln(10) \cdot \phi_i^* \cdot \left(\frac{M_k}{M_i^*}\right)^{(\alpha_i+1)} e^{-\left(\frac{M_k}{M_i^*}\right)^{\beta_i}} \tag{3.5}$$

Where $i = 1, 2...$; j = All, LTG or ETG; $k = *$, H_I, H_2, gas or bar; ϕ_i^* is the normalization parameter, α_i the slope towards the low mass end, M_i^* the breakdown mass, and β_i is the parameter that controls how steep is the decay at the high mass end. It is worth

to mention that if $\beta_i = 1$, equation (3.5) takes the form of a simple Schechter function, otherwise, we have a Schechter with subexponential decay. Thus, the sum of mass functions for each galaxy type, of a certain mass component, should be the mass function of all galaxies:

$$\phi(M_k)_{\rm All} = \phi(M_k)_{\rm LTG} + \phi(M_k)_{\rm ETG} \tag{3.6}$$

In fact, we will present both fittings for the total mass functions, that is, the one obtained by directly fitting the given total mass function (Table 3.1), and the one obtained as the sum of its LTG and ETG components (Table 3.2).

3.2.1 Galaxy Stellar Mass Functions

Figure 3.4 shows our GSMF results. In panel (a) we repeat the average GSMF (black dots with error bars) presented in Fig. 3.1, along with the LTG and ETG components (blue and red dots; only the mean are shown, the corresponding error bars are plotted in the other two panels). We compare our average (total) GSMF with those of Bernardi et al. (2013), Li and White (2009), and Papastergis et al. (2012) (for the latter, both their H_I and optically selected samples). Regarding the scatter around our results, we present with solid bars the case where we consider only observational scatter and with dashed lines when we take into account the quadratic sum of observational plus Poissonian errors. Basically there is no contribution from the latter for the average, LTG and ETG GSMF in the whole stellar mass range.

As already shown in Fig. 3.1, at the high-mass end, the GSMF is in agreement with Bernardi et al. (2013). This is because of the correction related with the SDSS aperture limits mentioned in 3.1.1, which is not considered neither in Papastergis et al. (2012) nor Li and White (2009). At intermediate stellar masses, our results agree well with all the observational works, and for the low-mass end, there is a rough agreement with the optically selected sample of Papastergis et al. (2012); the Li and White (2009) and Bernardi et al. (2013) GSMFs are valid only down to $M_* \approx 3 \times 10^8$ and $\approx 10^9$ $M_\odot$respectively. LTGs dominate from low to intermediate stellar masses, while ETGs dominate the high-mass end.

In panels (b) and (c), our GSMFs (including the modeled error bars) for LTGs and ETGs are plotted, respectively. The solid lines are the best fits to these GSMFs. In

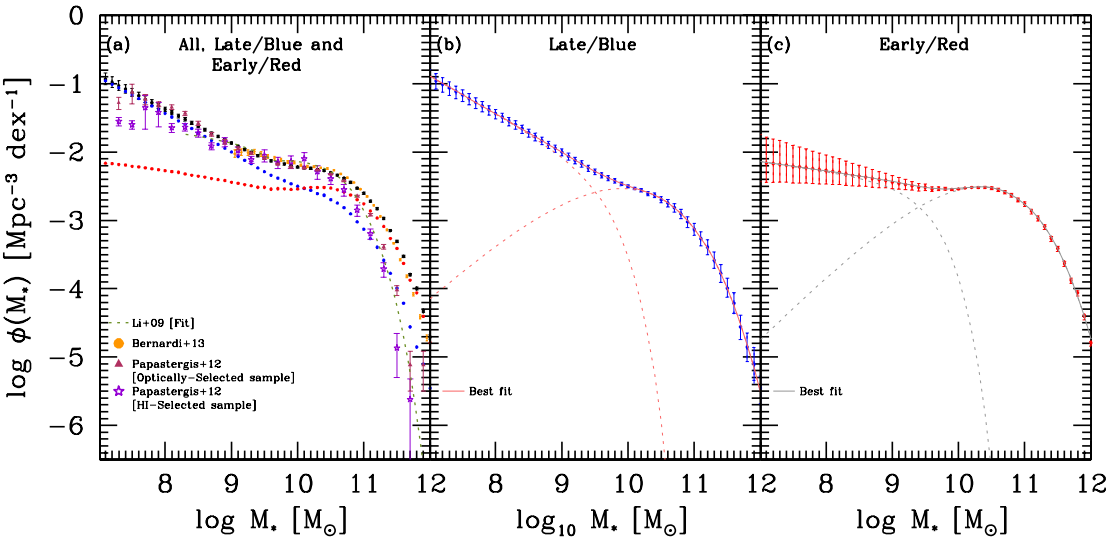

Figure 3.4: *Panel a:* Total (average) GSMF as constructed here from observations (black solid circles and error bars) and the best fit to it (solid black line, see Table 3.1 for the fitting function parameters). The blue and red solid circles show the mean of the LTG and ETG components of this GSMF. Other GSMFs reported in the literature are also plotted; see the labels inside the panel. *Panel (b):* Mean and error bars of the LTG component of the GSMF and the best fit to it (pink solid line, which is composed of a Schechter and Schechter with subexponential decay functions represented by the pink dashed lines, see Table 3.2 for the fitting functions parameters). *Panel (c):* Same as panel (b) but for the ETG component of the GSMF. In all the panels are plotted error bars considering only observational scatter (solid bars) and the quadratic sum of observational plus Poissonian scatters (dashed bars). Actually, the Poissonian noise is negligible in our GSMF and their LTG/ETG components.

both cases, the best fits are given by a composition of a Schechter and a Schechter with subexponential decay functions (short-dashed lines in these panels; see eqs. 3.4 and 3.5). The corresponding best-fit parameters are presented in Table 3.2.

The total GSMF constructed from our mock galaxy catalog has a slope at its low-mass end (down to $M_* \approx 10^7$ $M_\odot$) of -1.49. This is mainly the slope of the blue (LTG) GSMF ($= -1.55$), which by far dominates at low masses. The characteristic mass at which the GSMF falls strongly is $\log(M_*/M_\odot) \approx 10.4$. As already said, the fall is shallower than an exponential function; it is better described by a sub exponential function. At these very large masses, the ETG population dominates.

3.2.2 Galaxy H_I Mass Functions

The inferred galaxy H_I mass functions, GH_IMFs, for LTGs/ETGs and all galaxies are shown in Fig. 3.5. In panel (a) we plot the results for the average (total) GH_IMF (dots with error bars), as well as the mean of the LTG and ETG components (blue and red dots, respectively). The solid line is the best fit given by Eq. (3.6), which is described

by the combination of a Schechter and Schechter with sub exponential decay functions; the parameters of the best fit are given in Table 3.1. LTGs dominate basically all the mass function, except at very high masses, where LTGs contribute more to the total GH_IMF.

The LTG and ETG GH_IMFs (mean and modeled error bars), along with the fits for both, are presented in panels (b) and (c), respectively. The fit for the former case is a combination of a Schechter plus a Schechter with subexponential decay functions, while for the second case the fit is composed of a Schechter with subexponential decay function (dashed lines in figure 3.5. See Eqs. 3.4 and 3.5). The corresponding best fit parameters are given in Table 3.2. In all the cases, the fits were performed to the GH_IMFs down to the mass where they are complete, $\log_{10}(M_{H_I}/M_\odot) \gtrsim 8.5$ (dashed black line).

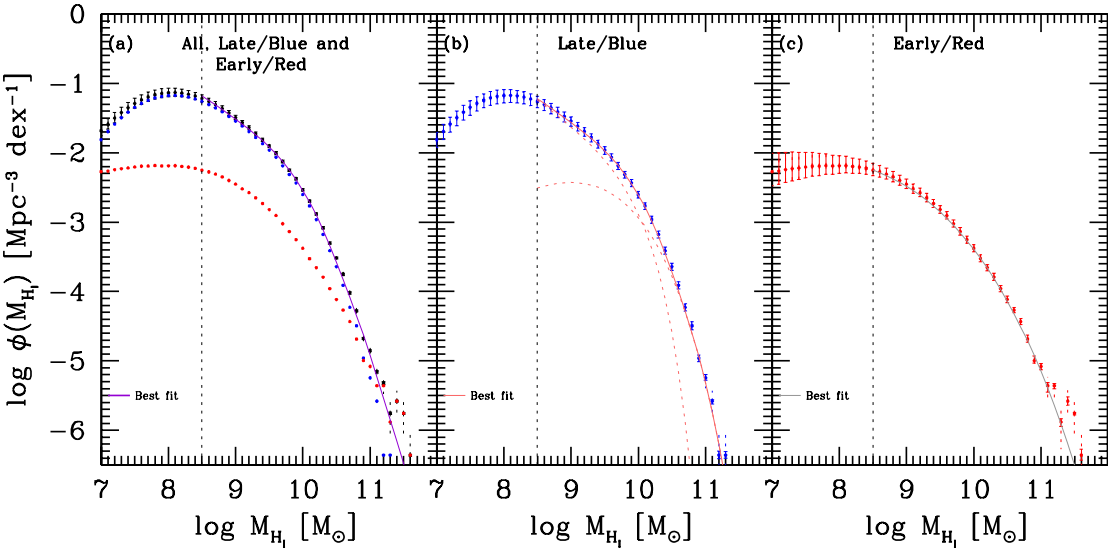

Figure 3.5: *Panel a:* Total (average) GH_IMF as constructed here from our mock catalog (black solid circles and error bars) and the best fit to it (solid black line, see Table 3.1 for the fitting function parameters). The blue and red solid circles show the mean of the LTG and ETG components of this GH_IMF. *Panel (b):* Mean and error bars of the LTG component of the GH_IMF and the best fit to it (pink solid line, which is composed of a Schechter and Schechter with subexponential decay functions represented by the pink dashed lines, see Table 3.2 for the fitting functions parameters). *Panel (c):* Same as panel (b) but for the ETG component of the GH_IMF. The best fit in this case is a single Schechter with subexponential decay function. In all the panels are plotted error bars considering only observational scatter (solid bars) and the quadratic sum of observational plus Poissonian scatters (dashed bars). Dashed lines indicate the mass where our fits to the GH_IMFs are complete, $\log_{10}(M_{H_I}/M_\odot) \gtrsim 8.5$.

In all panels are plotted error bars considering only observational scatter (solid bars) and the quadratic sum of observational plus Poissonian scatters (dashed bars). Unlike the GSMF, the average, LTG and ETG GH_IMFs are sensitive to Poissonian error at

the high-mass end, where the number counts become very small even for our 3 million objects catalog.

The total GH_IMF constructed from our mock galaxy catalog has a slope at its low-mass end (down to $M_{H_I} \approx 3 \times 10^8$ $M_\odot$) of -1.72. This is slightly steeper than the slope of the blue (LTG) GH_IMF, which is -1.69. The characteristic mass at which the GH_IMF falls strongly is $\log(M_*/M_\odot) \approx 9.9$. This fall is shallower than an exponential function; it is better described by a subexponential function. At these very large masses the ETG population is slightly more abundant than the LTG one.

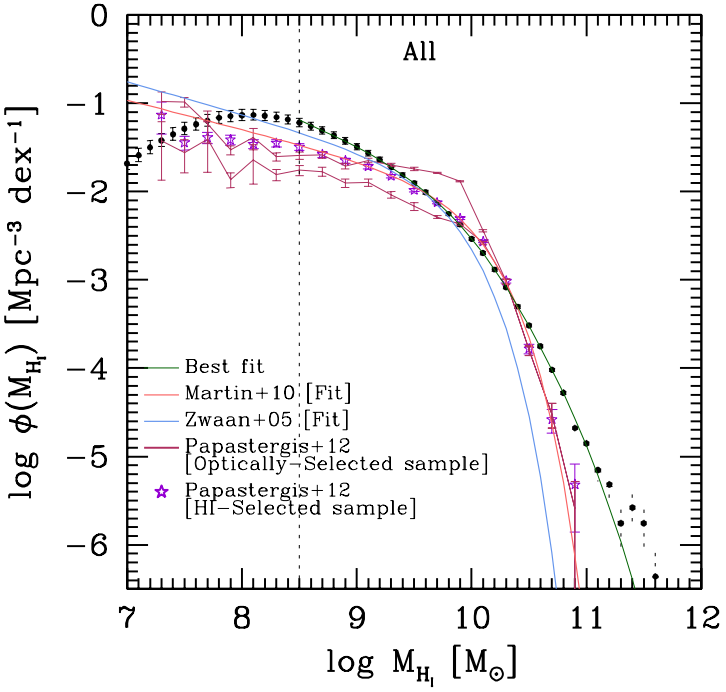

Figure 3.6: Comparison of our average GH_IMF (black solid circles with error bars) with several observational GH_IMFs indicated with labels inside the panels, along with the best fit (solid green line, which consist of a Schechter and Schechter with subexponential decay functions, see Table 3.1 for the fitting parameters). Dashed line indicates the mass where our fits to the GH_IMF is complete, $\log_{10}(M_{H_I}/M_\odot) \gtrsim 8.5$

In Fig. 3.6, we compare our average GH_IMF (solid dots with error bars) and its best fit (green solid line) with the observational total GH_IMFs reported in Martin et al. (2010), Papastergis et al. (2012) (for their H_I and optically selected samples), and Zwaan et al. (2005). At masses larger than $M_{H_I} \sim 10^{10}$ $M_\odot$, our GH_IMF slightly agrees with those of Martin et al. (2010) and Papastergis et al. (2012) but it is higher than the one by Zwaan et al. (2005). Martin et al. (2010) argue that the larger volume of ALFALFA survey compared with the HIPASS survey, makes ALFALFA more likely to be complete at the

large masses. Our result confirms this. At intermediate masses, $10^9 \lesssim M_{H_I}/\mathrm{M}_\odot \lesssim 10^{10}$, our $\mathrm{GH_IMF}$ is in good agreement with all the observed ones. In the case of the optically-selected sample of Papastergis et al. (2012), our results lie actually in between the two limits they present for estimating the $\mathrm{H_I}$ masses of those SDSS galaxies that lack ALFALFA detections (a significant fraction). These limits are a lower and an upper limit to the atomic hydrogen content ($M_{\mathrm{H_I}}^{\mathrm{min}}, M_{\mathrm{H_I}}^{\mathrm{max}}$), where $M_{\mathrm{H_I}}^{\mathrm{min}} = 0$, corresponding to $\mathrm{H_I}$-devoid galaxies. $M_{\mathrm{H_I}}^{\mathrm{max}}$ is determined from the galaxy distance D obtained using the SDSS optical redshift and $S_{\mathrm{int}}^{25\%\mathrm{lim}}$, the flux level at which the completeness of ALFALFA falls to 25% (eq. (3) in Papastergis et al., 2012). In order to obtain $S_{\mathrm{int}}^{25\%\mathrm{lim}}$, the M_*-v_{rot} relation is used for estimating W_{50} for the ALFALFA $\alpha.40$ catalog; this is because non-detected galaxies lack measurements of W_{50}. The assignment of these mass limits leads to a flattening of the $\mathrm{GH_IMF}$ slope, specially for the $M_{\mathrm{H_I}}^{\mathrm{min}}$ case. Therefore, there is a bias in the galaxy count and this have a direct effect in the $\mathrm{GH_IMF}$ determination.

At low masses, $M_{H_I} \lesssim 10^9$ $\mathrm{M}_\odot$, the $\mathrm{GH_IMF}$ starts to be incomplete due to sensitivity limits in the radio observations, so that it is expected a flattening as seen in the case of the $\mathrm{H_I}$ selected sample of Papastergis et al. (2012). Martin et al. (2010) considered sensibility corrections through the distribution of observed profile widths. Papastergis et al. (2012) impose additional optical requirements to their $\mathrm{H_I}$-selected sample (see their Section 2.1), which makes even less steep the $\mathrm{GH_IMF}$ low-mass slope than the one obtained by Martin et al. (2010) and by us. At the low masses, our $\mathrm{GH_IMF}$ agrees well with the Schechter fit to the results given in Zwaan et al. (2005, blue solid line). These authors construct the $\mathrm{GH_IMF}$ using the $\mathrm{H_I}$ Parks All-Sky Survey (HIPASS) and do not consider the requirements made by Papastergis et al. (2012) and Martin et al. (2010). In any case, their fit at the lowest masses shown in Fig. 3.6 is already an extrapolation.

Since our $\mathrm{GH_IMF}$ is constructed from a mass-complete GSMF down to $\approx 10^7$ $\mathrm{M}_\odot$ (this is analogous to the optically-selected case of Papastergis et al.), the flattening and posterior decreasing starts from the $\mathrm{H_I}$ masses corresponding to this stellar mass multiplied by the upper-scatter value of the M_{H_I}-to-M_* ratio at this mass. So, our $\mathrm{GH_IMF}$ can be considered complete down to these $\mathrm{H_I}$ masses, $M_{H_I} \approx 10^8$ $\mathrm{M}_\odot$. In the case of the optically-selected sample of Papastergis et al. (2012), their GSMF becomes to be incomplete at a larger mass, $M_* \sim 3 \times 10^7$ $\mathrm{M}_\odot$.

In general, the $\mathrm{GH_IMF}$ constructed by us from the observed GSMF and the empirical

M_{H_I}–M_* correlation is quite robust, and it seems to be a description of different observed GH_IMFs, at least in the mass ranges where the different samples used to construct these GH_IMF are reliable. In this sense, our GH_IMF is the most reliable in all the mass range down to $M_{H_I} \sim 10^8$ $M_\odot$, where the mass completeness becomes an issue. Moreover, our GH_IMF is actually the composition of the MFs corresponding to the LTG and ETG populations. The method used here, can help to estimate the completeness in other surveys.

3.2.3 Galaxy H_2 Mass Functions

In Fig. 3.7, we plot the inferred galaxy H_2 mass functions, GH_2MFs, for LTGs/ETGs and all galaxies. In panel (a) we plot the results for the average (total) GH_2MF (dots with error bars), as well as the mean of the LTG and ETG components (blue and red dots, respectively). The solid line is the best fit to the average GH_2MF given by Eq. (3.6), which is described by the combination of a Schechter and Schechter with sub exponential decay functions; the parameters of the best fit are presented in Table 3.1. As seen in panel (a), LTGs dominate basically at all masses, the contribution of ETGs being negligible. Only at the largest H_2 masses, the abundance of ETGs becomes comparable to the one of LTGs, however, these high-M_{H_2} ETGs are not the most massive galaxies in stellar mass.

The LTG and ETG GH_2MFs (mean and modeled error bars), along with the fits for both, are presented in panels (b) and (c), respectively. The fit for the former case is composed of a Schechter with subexponential decay function, while for the second case the fit is a combination of two Schechter with subexponential decay functions (dashed lines in figure 3.7. See eqs. 3.4 and 3.5). The best-fit parameters are reported in Table 3.2. In all the cases, the fits were performed to the GH_2MFs down to the mass where they are complete, $\log_{10}(M_{\rm H_2}/M_\odot) \gtrsim 8.5$ (dashed black line).

In all panels are plotted error bars considering only observational scatter (solid bars) and the quadratic sum of observational plus Poissonian scatters (dashed bars). Unlike the GSMF, the average, LTG and ETG GH_2 MFs are sensitive to Poissonian error at the high-mass ends.

In Fig. 3.8 we compare our average GH_2MF with observations from Keres et al. (2003). There is a good agreement at high H_2 masses, but for $\log_{10}(M_{\rm H_2}/M_\odot) \lesssim 9.5$, our MF

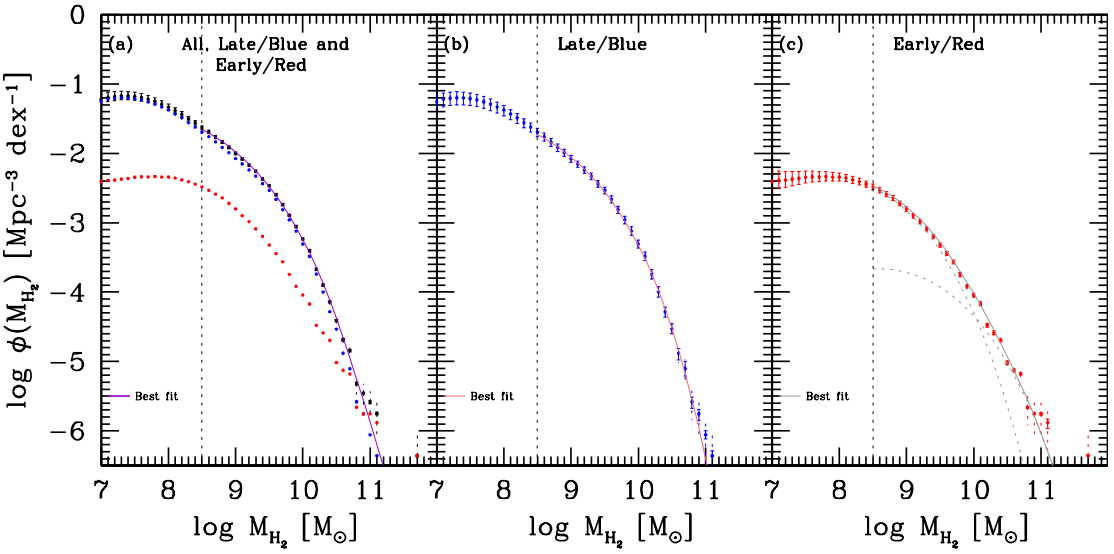

Figure 3.7: *Panel a:* Total (average) GH_2 MF as constructed here from our mock catalog (black solid circles and error bars) and the best fit to it (solid black line, see Table 3.1 for the fitting function parameters). The blue and red solid circles show the mean of the LTG and ETG components of this GH_2 MF. *Panel (b):* Mean and error bars of the LTG component of the GH_I2 MF and the best fit to it (pink solid line, which is a single Schechter with subexponential decay function represented by the pink dashed lines, see Table 3.2 for the fitting functions parameters). *Panel (c):* Same as panel (b) but for the ETG component of the GH_2 MF. The best fit in this case is composed by two Schechter with subexponential decay functions. In all the panels are plotted error bars considering only observational scatter (solid bars) and the quadratic sum of observational plus Poissonian scatters (dashed bars). Dashed lines indicate the mass where our fits to the GH_2MFs are complete, $\log_{10}(M_{H_2}/\mathrm{M}_\odot) \gtrsim 8.5$.

is steeper. One reason is probably the incompleteness that the observed GH_2MF starts to suffer due to the instrument flux limit. The other reason is the mass-dependent H_2-to-CO conversion factor introduced by us; see Section 2.2.2. This factor increases as M_* is smaller while in the case of Keres et al. (2003) it is constant. We have widely discussed in Section 2.2.2 the need of considering a more realistic H_2-to-CO conversion factor than previously done.

The total GH_2MF constructed from our mock galaxy catalog has a slope at its low-mass end (down to $M_{H_2} \approx 3 \times 10^8$ $\mathrm{M}_\odot$) of ≈ -1.79, which is steeper than the slope of the dominant LTG component, -1.09. The characteristic mass at which the GH_2MF falls strongly is $\log(M_*/\mathrm{M}_\odot) \approx 9.1$. This fall is shallower than an exponential function; it is better described by a subexponential function.

3.2.4 Galaxy Gas Mass Functions

Figure 3.9 shows our resulting galaxy cold gas mass functions, GGMFs. In panel (a), we plot the results for the average (total) GGMF (dots with error bars), as well as the

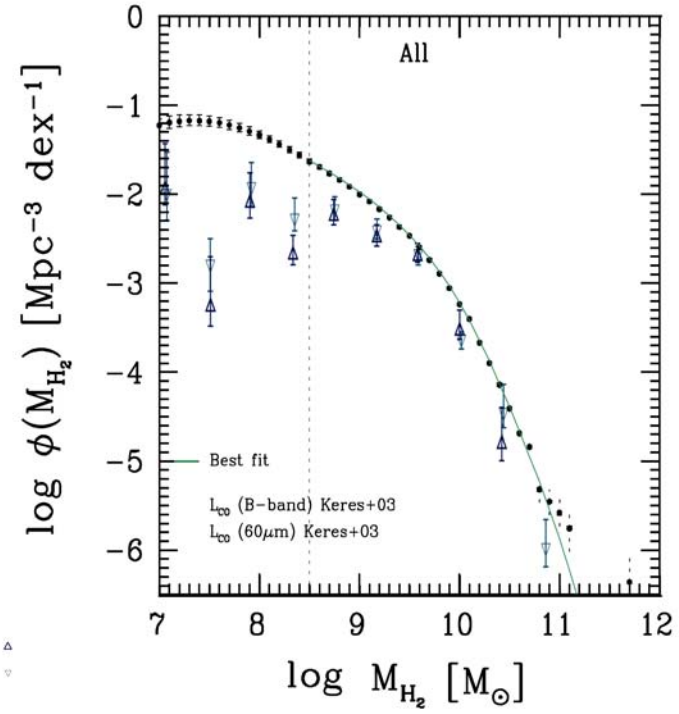

Figure 3.8: Comparison of our average GH_2MF (black solid circles with error bars) with the observations reported in Keres et al. (2003). Note that these authors use the Milky Way H_2-to-CO conversion factor while we have introduced a mass-dependent conversion factor that increases as M_* is smaller. The best fit (solid green line, which consist of a Schechter and Schechter with subexponential decay functions, see table 3.1 for the fitting parameters) is plotted also. Dashed line indicates the mass where our fits to the GH_2MF is complete, $\log_{10}(M_{H_2}/\mathrm{M}_\odot) \gtrsim 8.5$

mean of the LTG and ETG components (blue and red dots, respectively). The solid line is the best fit to the average GGMF given by Eq. (3.6), which is described by the combination of a Schechter and Schechter with sub exponential decay functions. As a consequence of the abundance dominion of LTGs in the GH_IMF and GH_2 MF, they also dominate in the GGMF. Only in the high-mass end, the abundance of ETGs is of the order of the abundance of LTGs.

The LTG and ETG GGMFs (mean and modeled error bars), along with the fits for both, are presented in panels (b) and (c), respectively. The fit for the former case is a combination of a Schechter function and a Schechter with subexponential decay function, while for the second case, the fit is composed of two Schechter with subexponential decay functions (dashed lines in figure 3.9). See eqs.(3.4 and 3.5). The best-fit parameters are reported in Table 3.2. In all the cases, the fits were performed to the GGMFs down to the mass where they are complete, $\log_{10}(M_{\rm gas}/\mathrm{M}_\odot) \gtrsim 8.5$ (dashed black line).

The total GGMF constructed from our mock galaxy catalog has a slope at its low-mass end (down to $M_{H_2} \approx 3 \times 10^8$ $M_\odot$) of -1.75. The low-mass end slope of the GGMF is somewhat steeper than the one of the GSMF. The characteristic mass at which the GGMF falls strongly is $\log(M_*/M_\odot) \approx 9.6$. This fall is shallower than an exponential function; it is better described by a subexponential function, and it is produced mainly by the ETG component, which at the largest masses has a contribution similar to the LTG one. The characteristic mass of the GGMF is much lower than the one of the GSMF.

In all panels are plotted error bars considering only observational scatter (solid bars) and the quadratic sum of observational plus Poissonian scatters (dashed bars). The scatters follow the same trend as in the H_IMF and H_2 MF. At the high gas masses, the Poissonian error contributes considerably to the total scatter. This is because the number counts at the high-mass ends of these functions are very small in spite that our catalog contain millions of objects.

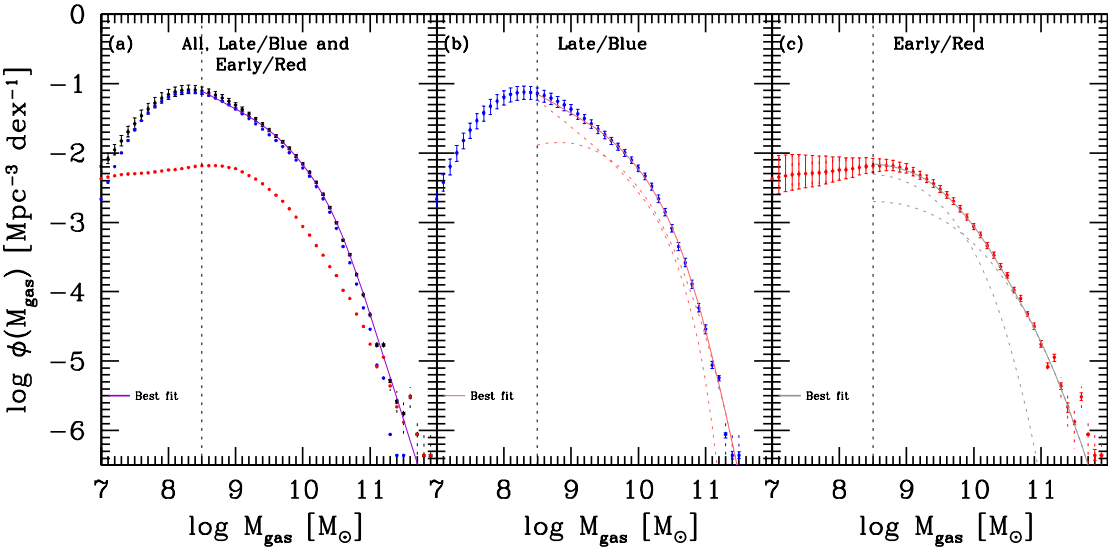

Figure 3.9: *Panel a:* Total (average) GGMF as constructed here from our mock catalog (black solid circles and error bars) and the best fit to it (solid black line). The blue and red solid circles show the mean of the LTG and ETG components of this GGMF. *Panel (b):* Mean and error bars of the LTG component of the GGMF and the best fit to it (pink solid line, which is composed of a Schechter and Schechter with subexponential decay functions represented by the pink dashed lines, see Table 3.2 for the fitting functions parameters). *Panel (c):* Same as panel (b) but for the ETG component of the GGMF. The best fit in this case is composed of two Schechter with subexponential decay functions. In all the panels are plotted error bars considering only observational scatter (solid bars) and the quadratic sum of observational plus Poissonian scatters (dashed bars). Dashed lines indicate the mass where our fits to the GGMFs are complete, $\log_{10}(M_{\rm gas}/M_\odot) \gtrsim 8.5$.

3.2.5 Galaxy Baryonic Mass Functions

Finally, the GBMFs constructed from our mock catalog are plotted in Fig. 3.10. In panel (a), we plot the results for the average (total) GBMF (dots with error bars), as well as the mean of the LTG and ETG components (blue and red dots, respectively). The solid line is the best fit to the average GBMF given by Eq. (3.6), which is described by the combination of a Schechter and Schechter with sub exponential decay functions; the parameters of the best fit are given in Table 3.1. The GBMF follows closely the GGMF in the low-mass end (which is limited by completeness at gas masses much larger than in the case of stellar masses). At larger masses, as larger is the mass, the closer is the GBMF to the GSMF. According to the panel (a), the LTG population dominates in the GBMF at low masses, $\log(M_{\rm bar}/{\rm M}_\odot) \lesssim 10.5$, while the ETG population dominates at larger masses.

The LTG and ETG GBMFs (mean and modeled error bars), along with the fits for both, are presented in panels (b) and (c), respectively. The fit for the former case is a combination of a Schechter function and a Schechter with subexponential decay function, while for the second case, the fit is composed of two Schechter with subexponential decay functions (see eqs. 3.4 and 3.5). The best-fit parameters are reported in Table 3.2. In all the cases, the fits were performed to the GBMFs down to the mass where they are complete, $\log_{10}(M_{\rm bar}/{\rm M}_\odot) \gtrsim 8.5$ (dashed black line).

In all panels are plotted error bars considering only observational scatter (solid bars) and the quadratic sum of observational plus Poissonian scatters (dashed bars). It is evident that Poissonian noise does not contributes to the total scatter of the average, LTG or ETG GBMFs, as in the case of the GSMFs.

The total GBMF constructed from our mock galaxy catalog has a slope at its low-mass end (down to $M_{\rm bar} \approx 3 \times 10^8$ ${\rm M}_\odot$) of -1.52, the same slope of the blue (LTG) GBMF. As expected from the GGMF contribution, the low-mass end slope of the GBMF is somewhat steeper than the one of the GSMF. The characteristic mass at which the GBMF falls strongly is $\log(M_{\rm bar}/{\rm M}_\odot) \approx 10$. This fall is shallower than an exponential function; it is better described by a subexponential function, and it is produced mainly by the ETG component, which at the largest masses is the dominant one in the total GBMF.

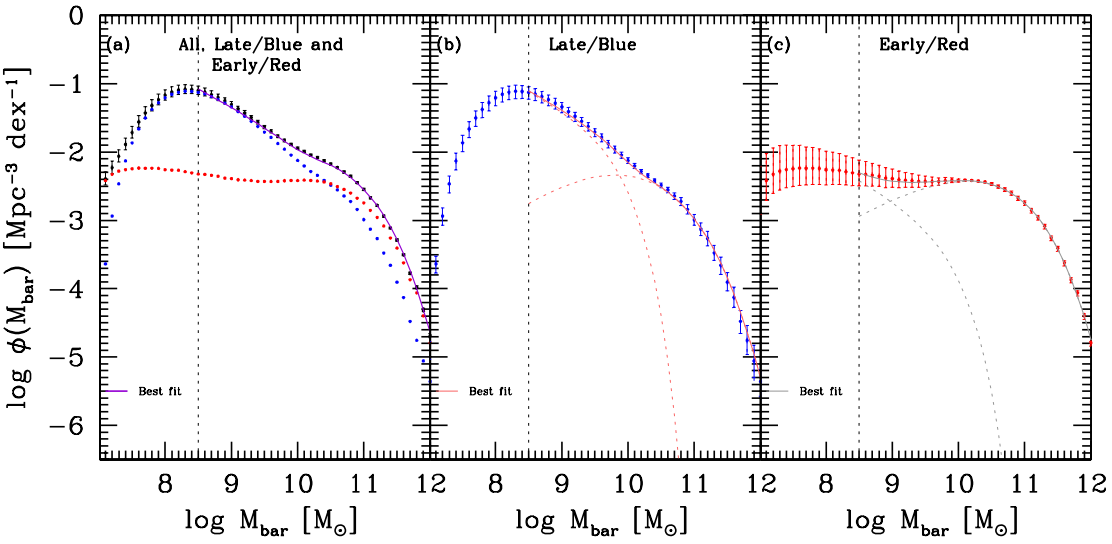

Figure 3.10: *Panel a:* Total (average) GBMF as constructed here from our mock catalog (black solid circles and error bars) and the best fit to it (solid black line, see Table 3.1 for the fitting function parameters). The blue and red solid circles show the mean of the LTG and ETG components of this GBMF. *Panel (b):* Mean and error bars of the LTG component of the GBMF and the best fit to it (pink solid line, which is composed of a Schechter and Schechter with subexponential decay functions represented by the pink dashed lines, see Table 3.2 for the fitting functions parameters). *Panel (c):* Same as panel (b) but for the ETG component of the GBMF. In all panels are plotted error bars considering only observational scatter (solid bars) and the quadratic sum of observational plus Poissonian scatters (dashed bars). It is evident that Poissonian noise does not contributes to the total scatter of the average, LTG or ETG GBMFs. Dashed lines indicates the mass where our fits to the GBMFs is complete, $\log_{10}(M_{\rm bar}/{\rm M}_{\odot}) \gtrsim 8.5$.

In Fig. 3.11, we compare our average GBMF with the Papastergis et al. (2012) (both, their H$_{\rm I}$– and optically–selected samples) and Baldry et al. (2008) observational works. At the $10 \lesssim \log(M_{\rm bar}/{\rm M}_{\odot}) < 11$ mass range, the three mass functions roughly agree. At lower masses, our GBMF agrees with the one obtained by Baldry et al. (2008)[3] and has a steeper increase than in Papastergis et al. (2012). This is because (1) at these low masses the H$_{\rm I}$ contribution becomes important, and as discussed above, our GH$_{\rm I}$MF is steeper than the one of these authors; (2) at difference of these authors, in our GBMF is taken into account the H$_2$ mass contribution, which at low masses becomes important and the GH$_2$MF increases steeply as lower is M_{H_2}; (3) the flattening of the observed GBMF starts probably at a larger mass than in our case due to the flux limits in radio as well as observational selection effects.

[3]Baldry et al. (2008) constructed the baryonic mass function from their stellar masses and the gas-to-stellar mass ratio, derived from well- established relation between stellar mass and metallicity coupled with a metallicity to gas mass fraction relation, which can be determined from a simple chemical evolution model. Thus, their baryonic mass is defined as Eq. (2.5) and it is expected their baryonic mass function to be similar than ours, unlike Papastergis et al. (2012).

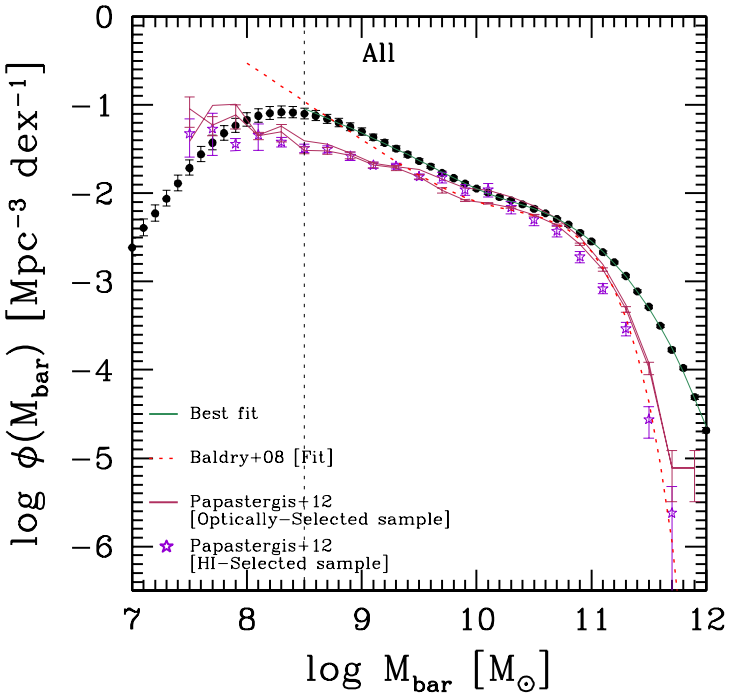

Figure 3.11: Comparison of our average GBMF (black solid circles with error bars) with those ones reported in Papastergis et al. (2012), both for their H_I- and optically-selected samples. The latter authors did not take into account H_2 in their baryonic masses. We also compare with Baldry et al. (2008) (red dashed line). Black dashed line indicates the mass where our fits to the GBMF are complete, $\log_{10}(M_{\rm bar}/M_\odot) \gtrsim 8.5$.

At masses around and larger than $\log(M_{\rm bar}/M_\odot)= 11$, our GBMF falls shallower than those of Papastergis et al. (2012) and Baldry et al. (2008). At these masses, the contribution of cold gas to the baryonic mass becomes negligible. Therefore, the difference is mainly due to the difference in the GSMFs (see Fig. 3.4). Recall that we take into account a correction in the stellar masses due to "aperture" effects (see Section 3.2.1), which is increasingly important as the larger is the galaxy.

We recall that our galaxy H_I and H_2 mass functions were constructed from the GSMF. In this sense, these mass functions, as well as the baryonic one, can be considered as obtained from an "optically"-selected galaxy sample. This procedure has been followed also by Papastergis et al. (2012), who used a galaxy sample from the SDSS and cross-matched it with the ALFALFA galaxies to assign them H_I masses; for those that were not observed by ALFALFA, these authors introduced two approximations to estimate statistically their H_I masses. In order to explore the effect of the (completeness) mass limit in the GSMF ("optical" sample) on the GBMF, we extrapolate our GSMF at low

masses down to $M_\star^{\rm lim} = 10^6$ or 10^8 $M_\odot$. The constructed GBMFs and their decomposition into LTG and ETG components by using these M_* limits in the GSMF are shown in Fig. 3.12, along with the real case of $M_\star^{\rm lim} = 10^7$ $M_\odot$ discussed above.

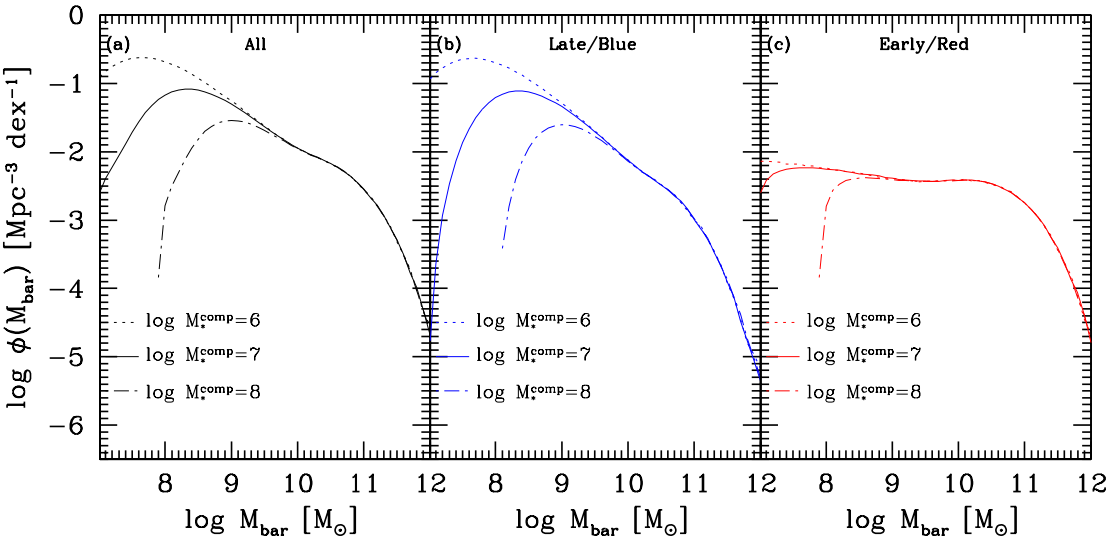

Figure 3.12: GBMFs constructed from our mock galaxy catalog by applying different "completeness" stellar mass limits, $M_\star^{\rm lim}$, in the GSMF (indicated inside the panels). Panels a, b, and c are for the average (all), LTG component, and ETG component, respectively. The completeness of the GBMFs at their low-mass ends are clearly affected by the GSMF low-mass limit. The GBMF start to flatten at a baryonic mass approximately ten times $M_\star^{\rm lim}$.

We find a clear trend: the M_* limit in the GSMF produces a mass limit in the GBMF at a baryonic mass that is $\sim$ 10 times larger than $M_\star^{\rm lim}$. In fact, this factor slightly increases as $M_\star^{\rm lim}$ is smaller. This trend is the same for the LTG and ETG components of the GBMF. Based on this result, care should be taken when using an "optically"-based sample to construct the gas and baryonic mass functions. A flattening of these mass functions is expected at the low-mass end due to the completeness M_* limit, flattening that starts at gas or baryonic masses at least a factor of ten larger than $M_\star^{\rm lim}$.

Figure 3.13 presents a plot that resumes the main results of this Chapter. Here, we present all the average (total) mass functions constructed from our mock catalog. Recall that this catalog of 3 million of galaxies was constructed to sample the observed GSMF (and its decomposition into LTG and ETG components) down to $M_\star^{\rm lim} = 10^7$ $M_\odot$, as well as the empirical M_{H_I}–M_* and M_{H_2}–M_* correlations. In order to avoid overcrowding, we plot only the mean of each mass function, excepting the GBMF for which the error bars are also plotted.

Figure 3.13 shows that the high-mass end of the GBMF is completely dominated by the GSMF, which on its own is dominated by the ETG population (see Fig. 3.4; the most massive galaxies are ETGs with negligible cold gas mass fractions).

The low-mass end of the average GBMF is steep; down to $M_{\rm bar} \approx 3 \times 10^8$ M$_\odot$) the slope is -1.52. At these masses the major contribution to $M_{\rm bar}$ comes from the cold gas mass. As seen in Fig. 3.13, the low-mass end slopes of the GH_IMF and GH_2MF are indeed very steep; the combination of both, gives a slope of ≈ -1.75 a these masses, which is steeper than the one the GSMF. Low-mass galaxies are gas dominated; a significant fraction of the baryons in these low-mass systems is locked in the cold gas. This has important consequences for understanding the formation and evolution of low-mass galaxies. What are dark the matter halo masses of these galaxies? This is part of the questions to be addressed in the next Chapter.

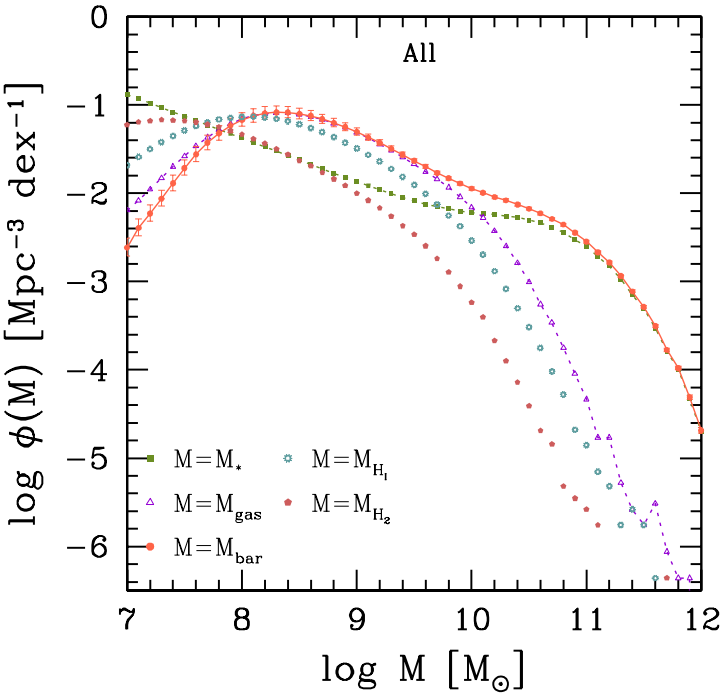

Figure 3.13: Average stellar, H_I, H_2, gas and baryonic mass functions. The mean for each MF are plotted with different symbols. The error bars are shown only for the GBMF.

Mass component	**Functions**	**Best-fit parameters**				$\boldsymbol{\chi^2_{red}}$
Stellar	Sc	$\boldsymbol{\phi_1^*}$	$\boldsymbol{\log_{10}(M_1^*)}$	$\boldsymbol{\alpha_1}$	$\boldsymbol{\beta_1}$	1.25
		0.0027	9.71	-1.49	–	
	+	$\boldsymbol{\phi_2^*}$	$\boldsymbol{\log_{10}(M_2^*)}$	$\boldsymbol{\alpha_2}$	$\boldsymbol{\beta_2}$	
	ScSexp	0.0028	9.07	-0.05	0.37	
H_I	Sc	$\boldsymbol{\phi_1^*}$	$\boldsymbol{\log_{10}(M_1^*)}$	$\boldsymbol{\alpha_1}$	$\boldsymbol{\beta_1}$	1.14
		0.0017	9.89	-1.72	–	
	+	$\boldsymbol{\phi_2^*}$	$\boldsymbol{\log_{10}(M_2^*)}$	$\boldsymbol{\alpha_2}$	$\boldsymbol{\beta_2}$	
	ScSexp	0.0419	7.76	-0.74	0.32	
H_2	Sc	$\boldsymbol{\phi_1^*}$	$\boldsymbol{\log_{10}(M_1^*)}$	$\boldsymbol{\alpha_1}$	$\boldsymbol{\beta_1}$	1.12
		0.0007	9.80	-1.79	–	
	+	$\boldsymbol{\phi_2^*}$	$\boldsymbol{\log_{10}(M_2^*)}$	$\boldsymbol{\alpha_2}$	$\boldsymbol{\beta_2}$	
	ScSexp	0.0016	5.59	-0.18	0.23	
Baryonic	Sc	$\boldsymbol{\phi_1^*}$	$\boldsymbol{\log_{10}(M_1^*)}$	$\boldsymbol{\alpha_1}$	$\boldsymbol{\beta_1}$	1.35
		0.0089	9.77	-1.52	–	
	+	$\boldsymbol{\phi_2^*}$	$\boldsymbol{\log_{10}(M_2^*)}$	$\boldsymbol{\alpha_2}$	$\boldsymbol{\beta_2}$	
	ScSexp	0.0054	8.95	-0.18	0.36	

Table 3.1: Best-fit parameters of the functions that describe our average (total) stellar, H_I, H_2 and baryonic mass functions. The second column indicates which function(s) we use for the corresponding fitting: Sc=Schechter and ScSexp=Schechter with subexponential decay. The third column shows the best-fit parameters of each function. The fourth column reports the χ^2_{red} for each fit.

Mass component	**Galaxy Pop.**	**Functions**	**Best-fit parameters**				$\chi^2_{\rm red}$
Stellar	ETG	Sc	ϕ_1^*	$\log_{10}(M_1^*)$	α_1	β_1	1.09
		+	0.0014	9.44	-1.15	–	
			ϕ_2^*	$\log_{10}(M_2^*)$	α_2	β_2	
		ScSexp	0.0015	9.22	-0.01	0.38	
	LTG	Sc	ϕ_1^*	$\log_{10}(M_1^*)$	α_1	β_1	1.05
		+	0.0020	9.64	-1.55	–	
			ϕ_2^*	$\log_{10}(M_2^*)$	α_2	β_2	
		ScSexp	0.0014	8.71	-0.09	0.35	
H_I	ETG	ScSexp	ϕ_1^*	$\log_{10}(M_1^*)$	α_1	β_1	1.17
			0.0019	4.80	-0.37	0.19	
	LTG	Sc	ϕ_1^*	$\log_{10}(M_1^*)$	α_1	β_1	1.26
		+	0.0031	9.84	-1.69	–	
			ϕ_2^*	$\log_{10}(M_2^*)$	α_2	β_2	
		ScSexp	0.0025	8.08	-0.15	0.38	
H_2	ETG	ScSexp	ϕ_1^*	$\log_{10}(M_1^*)$	α_1	β_1	1.35
		+	0.0006	5.40	-0.10	0.24	
			ϕ_2^*	$\log_{10}(M_2^*)$	α_2	β_2	
		ScSexp	1.7×10^{-5}	5.91	-0.08	0.23	
	LTG	ScSexp	ϕ_1^*	$\log_{10}(M_1^*)$	α_1	β_1	1.05
			0.0250	8.40	-1.09	0.41	
Gas	ETG	ScSexp	ϕ_1^*	$\log_{10}(M_1^*)$	α_1	β_1	1.49
		+	0.0057	8.1316	-0.49	–	
			ϕ_2^*	$\log_{10}(M_2^*)$	α_2	β_2	
		ScSexp	6×10^{-5}	5.09	-0.06	0.20	
	LTG	Sc	ϕ_1^*	$\log_{10}(M_1^*)$	α_1	β_1	1.16
		+	0.0012	10.29	-1.74	–	
			ϕ_2^*	$\log_{10}(M_2^*)$	α_2	β_2	
		ScSexp	0.0005	6.24	0.21	0.26	
Baryonic	ETG	Sc	ϕ_1^*	$\log_{10}(M_1^*)$	α_1	β_1	1.30
		+	0.0003	9.84	-1.57	–	
			ϕ_2^*	$\log_{10}(M_2^*)$	α_2	β_2	
		ScSexp	0.0042	9.86	-0.40	0.45	
	LTG	Sc	ϕ_1^*	$\log_{10}(M_1^*)$	α_1	β_1	1.14
		+	0.0076	9.78	-1.52	–	
			ϕ_2^*	$\log_{10}(M_2^*)$	α_2	β_2	
		ScSexp	0.0023	8.58	-0.15	0.33	

Table 3.2: Best-fit parameters of the functions that describe our (total) stellar, H_I, H_2 and baryonic mass functions for the ETG and LTG populations. The second column indicates the galaxy population. The third column indicates which function(s) we use for the corresponding fitting: Sc=Schechter and ScSexp=Schechter with subexponential decay. The fourth column shows the best-fit parameters of each function. The fifth column reports the $\chi^2_{\rm red}$ for each fit. Parameters to stellar, H_I, H_2, gas and baryonic mass functions fit for late/blue and early/red populations. In the first column we indicate which mass component we are fitting. The second column is tells which galaxy populations of the mass component is the case we fit. The third column indicates which function(s) we use, where Sc=Schechter and ScSexp=Schechter with subexponential decay. The fourth column shows the parameters for the function(s) fitted and in fifth column we present the $\chi^2_{\rm red}$ for each fit.

Chapter 4

The galaxy-halo connection: Gas and baryon contents of local late and early type galaxies

In the last decade, several statistical approaches have appeared for connecting galaxy properties (mainly the stellar mass, M_*) to their host dark matter halos. These approaches range from those that probe halo mass ($M_{\rm h}$) directly, such as weak-lensing (Mandelbaum et al., 2006) and satellite kinematics (More et al., 2011), to those based in statistical semi-empirical inferences, such as halo occupation model, the conditional stellar/luminosity function, the abundance matching technique (see Rodríguez-Puebla et al., 2013, for a review and discussion of them). Due to their relative simplicity and practicality, the semi-emiprical approaches are widely used. They allowed to explore large mass ranges and different galaxy populations, at difference of the direct methods. The simplest among the semi-empirical approaches is the abundance matching technique (AMT). This approach is based in that the observed cumulative galaxy number at a given property is matched against the theoretical halo plus subhalo cumulative number density. The resulting output from this match is a relation between a galaxy property (e.g., M_*) and the halo mass, $M_{\rm h}$. The power of the AMT relies on the fact that the galaxy-halo connection is obtained without prior knowledge of the complex and uncertain underlying physics of galaxy formation. Therefore, the AMT is a useful tool that, while it encodes the relevant physics of galaxy formation, it is not subject to the uncertain modeling of this physics.

In this Chapter, we will apply a variant of the AMT to the M_{H_I}, M_{H_2}, $M_{\rm gas}$, and $M_{\rm bar}$ distribution functions that were inferred in the previous Chapter separated into LTG and ETG galaxies (blue/late-type and red/early-type, respectively). This with the

aim of obtaining the galaxy-halo connection for M_{H_I}, M_{H_2}, $M_{\rm gas}$, and $M_{\rm bar}$, i.e. the correlation of these properties with the halo mass $M_{\rm h}$, for the two populations of LTGs and ETGs separately and for all the galaxies.

The main code used here was kindly provided by Dr. Aldo Rodrǵuez-Puebla (see for more details Rodríguez-Puebla et al., 2015).

4.1 The Model

When connecting galaxies to halos it is important to have in mind that a halo may have subhalos inside its virial radius. The main halo, called commonly *distinct* (i.e., a halo that is not contained inside a larger halo but that can have subhalos inside it), has associated to it a *central* galaxy, and each subhalo has also associated a galaxy that is *satellite* around the central. Therefore, halos are occupied generally by a central galaxy and satellites. The statistical models that describe how halos are occupied by central and satellite galaxies as a function of mass are called Halo Occupation Distribution (HOD) models, and in order to be constrained require as observational input not only the GSMF but also the spatial clustering of galaxies.

Galaxies are not only characterized by mass. In the era of the large galaxy surveys, along with luminosity (or M_*), properties like integral colors, star formation rates, photometric concentrations, etc. can be determined homogeneously. The most general and well studied property after stellar mass is probably the color. Here we will consider that galaxies are separated roughly into two populations: blue and red galaxies. In general, we associate these two populations to what we called LTGs and ETGs.

Within the context of the AMT, the connection between the overall GSMF, $\phi_{\rm g}(M_*)$, and the distinct halo plus subhalo mass functions, $\phi_{\rm h}(M_{\rm h}) + \phi_{\rm sub}(M_{\rm h})$, arises naturally if one assumes a probability distribution function, denoted by $P(M_*|M_{\rm h})$, that a (sub)halo of mass $M_{\rm h}$ hosts a galaxy in the stellar mass bin $M_* \pm dM_*/2$. Mathematically, this is written as

$$\phi_{\rm g}(M_*) = \int P(M_*|M_{\rm h})\phi_{\rm DM}(M_{\rm h})dM_{\rm h}, \tag{4.1}$$

where $\phi_{\rm DM}(M_{\rm h}) = \phi_{\rm h}(M_{\rm h}) + \phi_{\rm sub}(M_{\rm h})$. Thus, once the left hand side of this equation is known, it can be used to determine the average stellar-to-halo mass relation (SHMR hereafter) of galaxies, $M_*(M_{\rm h})$. The overall GSMF is actually composed of the GSMFs

of central and satellite galaxies, $\phi_{\rm g}(M_*) = \phi_{\rm g,c}(M_*) + \phi_{\rm g,s}(M_*)$. In order to obtain this relation separately for central/halos and satellites/subhalos, the HOD and conditional luminosity models need to be introduced (Rodríguez-Puebla et al., 2012, 2013).

The above approach can be extended even further in order to connect different galaxy populations, –for example blue and red galaxies– to their host dark matter halos. This has been done for *central* galaxies in Rodríguez-Puebla et al. (2015). Unfortunately, we can not repeat this approach for our galaxy mock samples (separated into LTGs and ETGs) since we do not account for information regarding their spatial clustering. Besides, in our case the inferred mock galaxy mass functions refer to all galaxies, and we do not have information about their separation into centrals and satellites, i.e., about their conditional mass functions.

Thus, our approach here will be based:

1. on the simple AMT, for which only the total galaxy and halo mass functions of blue (LTG) and red (ETG) galaxies are necessary,
2. on the halo mass function decomposition into those that host *central* blue and red galaxies from the accurate results obtained in Rodríguez-Puebla et al. (2015),
3. on the assumption that the SHMRs of central and satellite galaxies are the same, and
4. on a simple empirically-motivated model for the fraction of subhalos hosting blue and red satellites.

Under these assumptions, we will obtain then the average SHMRs for LTGs and ETGs by separate using the corresponding GSMFs presented in the previous Chapter. Further, based on a mock galaxy catalog that samples these GSMFs and follows the mentioned SHMRs, the corresponding $\rm H_I$-, $\rm H_2$-, gas- and baryonic-to-halo mass relations will be obtained by using the empirical M_{H_I}– and M_{H_2}–M_* correlations inferred in Chapter 2. Following, we describe in detail our full approach.

4.1.1 Central galaxies

For *central* blue and red galaxies, one can introduce the conditional probability distribution functions $P_{\rm c,b}(M_*|M_{\rm h})$ and $P_{\rm c,r}(M_*|M_{\rm h})$ to establish the statistical connection

between the "blue", $\phi_{\rm h,b}$, and "red", $\phi_{\rm h,r}$, distinct halo mass functions and the GSMFs of blue and red centrals, $\phi_{\rm g_c,b}$ and $\phi_{\rm g_c,r}$, respectively. As above, the mean relations $M_{*,b}(M_{\rm h})$ and $M_{*,r}(M_{\rm h})$ are the result of this connection. More formally, this connection is given by,

$$\phi_{{\rm g_c},j}(M_*) = \int P_{{\rm c},j}(M_*|M_{\rm h})\phi_{{\rm h},j}(M_{\rm h})dM_{\rm h}. \tag{4.2}$$

Here the subscript 'j' refers either to blue or red galaxies (b or r).

In order to fully characterize the link between blue and red central galaxies to their host dark matter halos, we need detailed information about the functions $\phi_{\rm h,b}$, and $\phi_{\rm h,r}$, i.e., about the *halo mass functions associated to blue or red galaxies*. Unfortunately these functions are unknown but in practice they can be constrained. As mentioned above, Rodríguez-Puebla et al. (2015) introduced a statistical framework that combines different statistical methods of galaxy-halo connection to constrain the occupational number of blue and red galaxies in dark matter halos of different masses. Using this framework, combined with measurements of the local SDSS galaxy stellar mass functions and galaxy clustering of blue and red galaxies, Rodríguez-Puebla et al. (2015) constrained the fraction of red central galaxies as a function of halo mass: $f_{\rm h,r}(M_{\rm h}) = \phi_{\rm h,r}(M_{\rm h})/\phi_{\rm h}(M_{\rm h})$. The complement, $f_{\rm h,b} = 1 - f_{\rm h,r}$, is the fraction of blue galaxies at each $M_{\rm h}$. Motivated by recent observational results from SDSS galaxy groups in Woo et al. (2013), Rodríguez-Puebla et al. (2015) parameterized this fraction as

$$f_{\rm h,r}(M_{\rm h}) = \frac{1}{b + M_{\rm h}^{\star}/M_{\rm h}}, \tag{4.3}$$

where $M_{\rm h}^{\star} = \beta \times 10^{12}{\rm M}_{\odot}h^{-1}$. By means of their statistical approach, they constrained $\beta = 0.481 \pm 0.065$ and $b = 1.032 \pm 0.014$. This fraction is such a that at $M_{\rm h} = 10^{11}{\rm M}_{\odot}$ only $\sim 13\%$ of dark matter halos host red centrals, while at $2 \times 10^{12}{\rm M}_{\odot}$ this fraction has increased up to $\sim 80\%$, increasing even more at higher masses. At $M_{\rm h} \sim 7 \times 10^{11}{\rm M}_{\odot}$ the fraction of halos hosting blue and red centrals is the same (i.e., $f_{\rm h,r} = f_{\rm h,b} = 0.5$).

Here, we use the best fitting model constrained in Rodríguez-Puebla et al. (2015) for the fraction of halos hosting red central galaxies to derive the central SHMRs by color (or type) necessary in this work. However, the mock galaxy mass function inferred in the previous Chapters are for all the galaxies, not only central ones. Therefore, we need an approach to take into account satellites/subhalos.

Conexión Halo-Galaxia: Fracción bariónica de galaxias rojas y azules.

4.1.2 Satellite galaxies

For connecting satellite galaxies to their host subhalos we use similar arguments as for centrals. We introduce the conditional probability distribution functions $P_{\rm sub,b}(M_*|M_{\rm h})$ and $P_{\rm sub,r}(M_*|M_{\rm h})$ to establish the statistical connection between the "blue", $\phi_{\rm sub,b}$, and "red", $\phi_{\rm sub,r}$, subhalo halo mass functions and the GSMFs of blue and red satellite galaxies, $\phi_{\rm g_s,b}$ and $\phi_{\rm g_s,r}$, respectively. Therefore we write,

$$\phi_{{\rm g_s},j}(M_*) = \int P_{{\rm sub},j}(M_*|M_{\rm h})\phi_{{\rm sub},j}(M_{\rm h})dM_{\rm h}. \tag{4.4}$$

As before, the subscript 'j' refers either to blue or red galaxies.

Analogously to central galaxies, in order to obtain the relation between satellites and their host subhalos of blue and red galaxies we need to specify functions $\phi_{\rm sub,b}$ and $\phi_{\rm sub,r}$. Based in an analysis of the Yang et al. (2007) galaxy group catalog, Peng et al. (2012) studied the fraction of satellite galaxies as a function of stellar mass M_* and environments ρ. By defining ϵ as the fraction of previously central galaxies that are red because they become satellite galaxies, Peng et al. (2012) found that the fraction of red satellite galaxies $f_{\rm sat,r}$ is given by

$$f_{\rm sat,r}(\rho, M_*) = f_{\rm cen,r}(M_*) + \epsilon(\rho) + f_{\rm cen,r}(M_*)\epsilon(\rho). \tag{4.5}$$

Motivated by this phenomenological model we now rewrite this equation to characterize the fraction of subhalos hosting red satellite galaxies,

$$f_{\rm sub,r}(M_{\rm h}) = f_{\rm h,r}(M_{\rm h}) + \epsilon - f_{\rm h,r}(M_{\rm h})\epsilon. \tag{4.6}$$

In the above equation, $f_{\rm h,r}$ is the fraction of halos hosting red central galaxies defined in eq. (4.3) above, while ϵ is a function that, in principle, depends on both subhalo mass and the mass of the host halo mass where it resides. The principal reason for this change is motivated by the group analysis developed in Woo et al. (2013). These authors conclude that the phenomenological model presented in Peng et al. (2012) should be valid if one substitutes $M_{\rm h}$ for M_*. Note that the environmental dependence of $f_{\rm sub,r}$

in Eq. (4.6) enters now through halo mass.[1] Peng et al. (2012) showed that a constant value of $\epsilon \approx 0.4$ reproduces well the mass functions of blue and red satellite galaxies as well as the observed conditional satellite stellar mass functions in halos of different masses.

4.1.3 Model Assumptions

In the standard AMT, it is usual to assume that the SHMRs of central and satellite galaxies are identical. Formally this is not true, but the differences are actually moderate (Rodríguez-Puebla et al., 2013). We assume that both SHMRs are equal and generalize it for the case of central/satellite blue and central/satellite red galaxies by separate. This is equivalent to assume that $P_{\mathrm{c},j}(M_*|M_\mathrm{h}) = P_{\mathrm{sub},j}(M_*|M_\mathrm{h})$ (j =blue or red) if their intrinsic scatters are the same. For simplicity and because they are equal by assumption, we will denote the distributions $P_{\mathrm{c},j}(M_*|M_\mathrm{h})$ and $P_{\mathrm{sub},j}(M_*|M_\mathrm{h})$ as $P_j(M_*|M_\mathrm{h})$. Each probability distribution function $P_j(M_*|M_\mathrm{h})$ was assumed to be lognormal distributed with mean SHMR $\langle \log M_{*,j}(M_\mathrm{h})\rangle$ and standard deviation (intrinsic scatter) σ_j,

$$P_j(M_*|M_\mathrm{h})dM_* = \frac{e}{2\pi\sigma_j^2}\exp\left[-\frac{(\log M_* - \langle \log M_{*,j}(M_\mathrm{h})\rangle)^2}{2\sigma_j^2}\right]\frac{dM_*}{M_*}. \tag{4.7}$$

Each mean $\langle \log M_{*,j}(M_\mathrm{h})\rangle$ is parametrized using the functional form proposed in Behroozi et al. (2013),

$$\langle \log M_{*,j}(x)\rangle = \log(\epsilon_j M_{1,j}) + g(x) - g(0), \tag{4.8}$$

where

$$g(x) = \delta_j \frac{(\log(1+e^x))^{\gamma_j}}{1+e^{10^{-x}}} - \log(10^{\alpha_j x}+1), \tag{4.9}$$

and $x = \log(M_\mathrm{h}/M_{1,j})$. This function behaves as a power law with slope α_j at masses much smaller than $M_{1,j}$, and as a sub-power law with slope γ_j at large masses. Each σ_j were assumed to be constant with halo mass. Based on the results in Rodríguez-Puebla et al. (2015) we adopt $\sigma_b = 0.12$ and $\sigma_r = 0.14$.

[1] Analysis of dark matter halo clustering from both both analytical models (Mo and White, 1996) and $N-$body simulations (Tinker et al., 2010) have shown that halos of different halos masses cluster differently: more massive halos are more clustered.

For the distinct halo mass function we use the fit to large N-body cosmological simulations published in Tinker et al. (2008). We define halo masses at the radius where the spherical over density is $\Delta_{\rm vir}$ times the mean matter density. For the subhalo mass function we use the analytic fit function described in Behroozi et al. (2013) to the Bolshoi N-body high resolution cosmological simulation.

4.1.4 Constraining the stellar-to-halo mass relation

We constrain the total SHMRs separately for blue and red galaxies by finding the best fit parameters of each $\langle \log M_{*,j}(M_{\rm h}) \rangle$ (Eq.4.8) that maximize the likelihood function $L \propto e^{\chi^2/2}$. Here the total χ^2 is defined as,

$$\chi^2 = \chi^2_{\phi_{\rm b}} + \chi^2_{\phi_{\rm r}}. \tag{4.10}$$

Each $\chi^2_{\phi_{\rm j}}$ has been defined as,

$$\chi^2_{\phi_{\rm j}} = \sum \frac{(\phi_{{\rm mock},j} - \phi_{{\rm mod},j})^2}{\sigma^2_{{\rm mock},j}} \tag{4.11}$$

where the subscript 'mock' indicates that we are using the resulting blue/red (LTG/ETG) GSMFs as obtained in Chapter 3. The quantity $\phi_{{\rm mod},j}$ is given by its central and satellite component calculated according to Eqs. (4.2) and (4.4), respectively. As before 'j' refers either to blue or red galaxies. Note that the corresponding SHMRs are implicitly contained in these equations.

Finally, for the process of maximization of L we use the Powell's directions set method in multi-dimension (Press et al., 1992).

4.2 Multi-Abundance Matching Technique Constrains in Halos: MATCH

In Chapter 3, we generated a population of 3×10^6 galaxies, which trace in a complete volume the empirical LTG/ETG GSMFs as well as the corresponding M_{H_I}–M_* and M_{H_2}–M_* correlations. As a result, each member in the sample is characterized by M_*, M_{H_I}, M_{H_2}, $M_{\rm gas}$, and $M_{\rm bar}$. Our mock catalog is a stellar mass-complete sample above $M_* = 1 \times 10^7 {\rm M}_\odot$. In this Section we describe how we assign halo masses to each member in the sample.

In the previous Section we described how to constrain separate SHMRs for blue and red (LTG and ETG) galaxies. By using this result, at a fixed halo mass $M_{\rm h}$, its stellar mass could be randomly assigned via Eq. (4.7). Nevertheless, our sample has been constructed based on the abundance of galaxies according to the observed GSMF. This means that we cannot use directly Eq. (4.7) but its inverse, i.e., the distribution of halo masses $M_{\rm h}$ at a fixed stellar mass M_*, $P_j(M_{\rm h}|M_*)$. We compute $P_j(M_{\rm h}|M_*)$ via the Bayes theorem,

$$P_j(M_{\rm h}|M_*) = \frac{P_j(M_*|M_{\rm h})P_j(M_{\rm h})}{P_j(M_*)}, \tag{4.12}$$

where $P_j(M_{\rm h})$ is the subhalo + halo mass function, $\phi_{{\rm DM}j}(M_{\rm h})$, hosting galaxies of color or type 'j', and $P_j(M_*)$ are the satellite+central GSMFs for these two galaxy populations, $\phi_{{\rm g}_j}(M_*)$. Recall that the subscript 'j' refers either to blue or red (LTG or ETG, respectively).

We obtain halo masses $M_{\rm h}$ for each galaxy in the sample as follows. Once a galaxy has been marked as either blue or red, we generate a random number $\mathcal{U}_{\rm h}$ uniformly distributed within the interval [0,1] and then find the value for $M_{\rm h}$ that solves the equation $P_j(> M_{\rm h}|M_*) = \mathcal{U}_{\rm h}$.[2] Here,

$$P_j(> M_{\rm h}|M_*) = \int_{M_{\rm h}}^{\infty} P_j(M'_{\rm h}|M_*)dM'_{\rm h}, \tag{4.13}$$

that is, the cumulative distribution of $P_j(M_{\rm h}|M_*)$. In this way, we sample statistically this distribution. Note that the halo masses generated in this way reproduce the theoretical subhalo + halo mass functions by construction because in Eq. (4.12) the SHMRs for blue and red galaxies (LTG and ETG) are encoded.

Once we have generated the halo mass for each galaxy using the SHMRs, we can now predict the M_{HI}-$M_{\rm h}$, M_{H_2}-$M_{\rm h}$, $M_{\rm gas}$-$M_{\rm h}$, and $M_{\rm bar}$-$M_{\rm h}$ relations for LTG and ETG galaxies. Assigning halo masses in this way is equivalent to apply the AMT to each of the mass distributions determined in the previous Chapter in order to obtain the corresponding mass relations. Thus, finally each member in our stellar mass-complete catalog, separated into LTG and ETG populations, has the following properties: M_{HI}, M_{H_2}, $M_{\rm gas}$, M_*, $M_{\rm bar}$ and $M_{\rm h}$.

[2]We make use of the fact that the integrals of two probability distributions are equal, $\int_0^x h(x')dx' = \int_0^y g(y')dy'$. In the case of a uniform distribution between 0 and 1, $\int_0^x U(x')dx' = x$. Therefore, a random number y from the distribution $g(y)$ can be obtained by resolving the equation $x = \int_0^y g(y')dy'$, where x is a random number from the [0,1] interval.

4.3 The link between halo mass and different galaxy masses

In the previous Section we have constructed a mock catalog of millions of galaxies, which follow the empirical galaxy stellar, H_I, H_2, gas and baryonic mass functions for both LTGs and ETGs, as well as for the averages of both of them. Here, by means of our MATCH formalism we will establish the connection of these mass functions to the halo mass function in order to ultimately obtain the M_{H_I}-, M_{H_2}-, $M_{\rm gas}$-, and $M_{\rm bar}$-to-halo mass relations for both galaxy populations.

In Fig. 4.1, we plot the results of the AMT for blue and red galaxies (LTGs and ETGs) explained in Section 4.1. The blue and red lines are the M_*–$M_{\rm h}$ relations for blue and red galaxies, while the black line is the density-weighted average relation. The 1σ scatter in dex around each one of the relations is shown in the lower panel. Recall that the scatter is assumed lognormal. The right panel shows the $M_*/M_{\rm h}$ ratios as a function of $M_{\rm h}$.

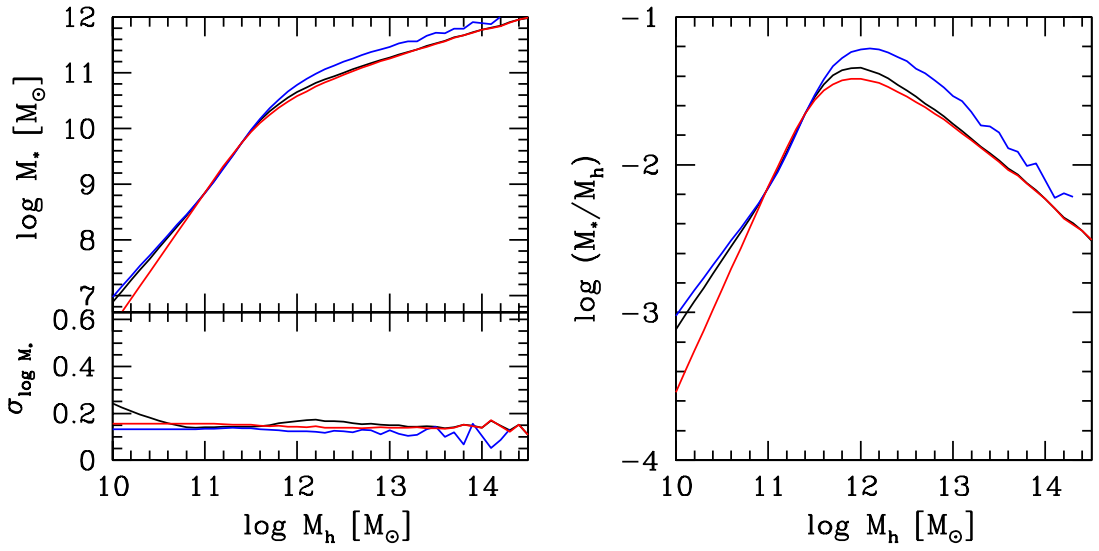

Figure 4.1: Stellar-to-halo mass relations for LTGs (blue lines) and ETGs (red lines), and for the density-weighted average of both populations (black lines). The lower panel shows the intrinsic scatter around each one of these mass relations in dex. The right panel is the same as the left one but for the $M_*/M_{\rm h}$ ratio. The dotted line reproduces the average SHMR reported in Rodríguez-Puebla et al. (2015) for a GSMF similar to the one used here.

As shown in Rodríguez-Puebla et al. (2015), the blue and red M_*-$M_{\rm h}$ relations are different. For a given $M_{\rm h}$, the stellar mass is equal or higher for blue galaxies than for red ones. At high masses, this could be because red (ETG) galaxies, quenched earlier their M_* growth, while their $M_{\rm h}$ continued growing; massive blue (LTG) galaxies, which are a small fraction, were not quenched and continued growing their M_* (see R-P2014). At low masses, the halos typically assembled early but the galaxies delay they star

formation (downsizing in SFR) being today dwarf LTGs; however, a small fraction of the dwarf galaxies might have not suffered the delay processes and formed stars early in bursting processes after which most of gas is ejected by SN-driven outflows. As the result, their stellar masses today are smaller than those of LTGs living in halos of the same mass. Regarding the average M_*-$M_{\rm h}$ relation, at low masse, LTGs are more abundant than ETGs, so that they dominate in the average; the opposite happens at high masses.

4.3.1 Galaxy gas-to-halo mass relations

The obtained M_{H_I}–$M_{\rm h}$ and M_{H_2}–$M_{\rm h}$ relations and their scatters are presented in Figs. 4.2 and 4.3, following the same structure and line coding as in Fig. 4.1. As seen in the lower panels, the scatters around the mass relations are large.

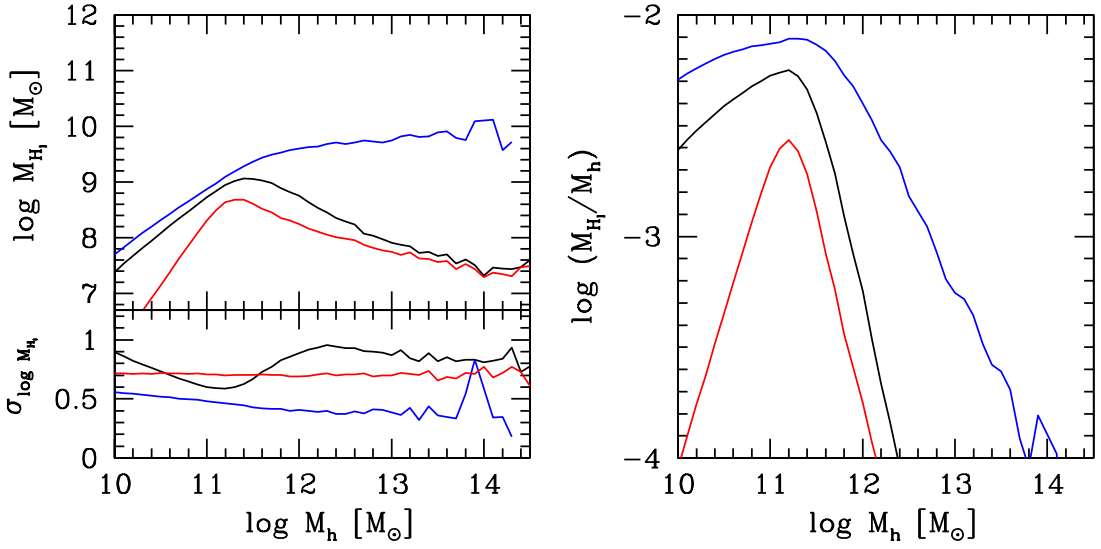

Figure 4.2: H_I-to-halo mass relations for LTG (blue lines) and ETG (red lines) galaxies, and for the density-weighted average of both populations (black lines). The lower panel shows the intrinsic scatter around each one of these mass relations in dex. The right panel is the same as the left one but for the $M_{H_I}/M_{\rm h}$ ratio.

In fact, this scatter is consistent with the results obtained in the M_{H_I}-W_{50} relation, see e.g. figure 5 in Papastergis et al. (2013). For LTGs, the H_I mass contained in low mass halos is larger than the mass contained in stars, as expected. For this population, in both cases, M_{H_I} and M_{H_2}, there is a monotonic increasing relation with $M_{\rm h}$, though very flat already at large masses. However, for the ETG population, the monotonicity is broken; the M_{H_I}–$M_{\rm h}$ and M_{H_2}–$M_{\rm h}$ relations attain a maximum at $M_{\rm h} \approx 10^{11.2}$ $M_\odot$ and decrease at lower and higher masses. Since ETGs dominate at high masses, the density-weighted

average also results non-monotonic. The reason of the non-monotonicity comes from the empirical M_{H_I}–M_* and M_{H_2}–M_* relations for ETGs. Both, M_{H_I} and M_{H_2} increase as M_* increases but at masses around 10^{10} $M_\odot$, these gas masses decrease as M_* increases. However, the large scatter in the M_{H_I}–$M_{\rm h}$ and M_{H_2}–$M_{\rm h}$ relations of ETGs also produces the non-monotonicity.

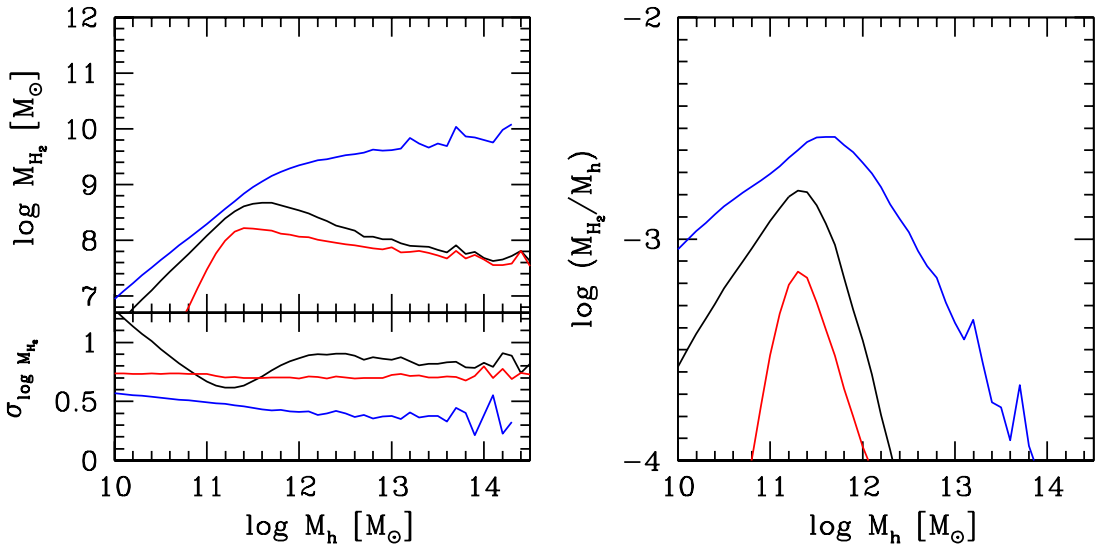

Figure 4.3: H_2-to-halo mass relations for LTG (blue lines) and ETG (red lines) galaxies, and for the density-weighted average of both populations (black lines). The lower panel shows the intrinsic scatter around each one of these mass relations in dex. The right panel is the same as the left one but for the $M_{H_2}/M_{\rm h}$ ratio.

In Fig. 4.4, the corresponding $M_{\rm gas}$–$M_{\rm h}$ relations are plotted. The strong differences between these relations for LTG and ETG populations confirms that both populations have been to be treated separately, as we proceeded. The "average" relations are actually close to the LTG ones at low masses and to the ETG ones at large masses, while at intermediate masses they get far from both, due to the large segregation by type/color. Our results show that the differences between the LTG and ETG populations in the M_*-$M_{\rm h}$ relation become larger in the gas mass-$M_{\rm h}$ relations. For a given $M_{\rm h}$, the gas mass is significantly larger for the LTGs than for the ETGs.

The fact that the gas mass of ETGs in halos more massive than $M_{\rm h} \approx 10^{11.2}$ $M_\odot$ is very low and strongly decreases as more massive they are encodes valuable information about the formation of these galaxies. The processes behind the gas exhaustion in these halos were proposed to be the large cooling time of gas in halos of large virial temperatures and the positive feedback of the AGNs, which are expected to be luminous in massive galaxies. Our results show that the present-day $M_{\rm h}$ at which these "quenching" processes

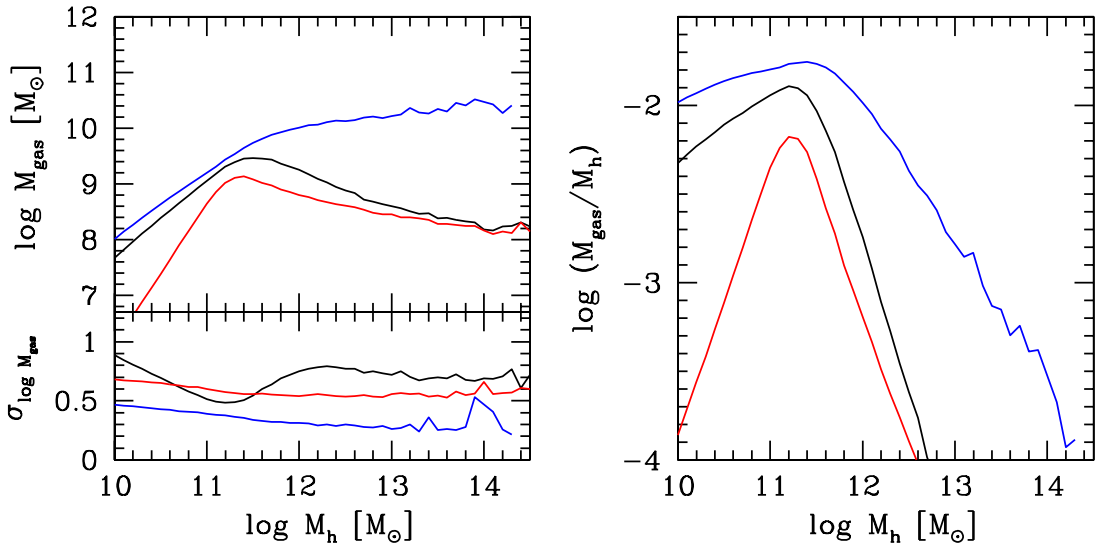

Figure 4.4: Gas-to-halo mass relations for LTG (blue lines) and ETG (red lines) galaxies, and for the density-weighted average of both populations (black lines). The lower panel shows the intrinsic scatter around each one of these mass relations in dex. The right panel is the same as the left one but for the $M_{\rm gas}/M_{\rm h}$ ratio.

start to act is $M_{\rm h} \approx 10^{11.2}$ M$_\odot$. On the other hand, our results show that ETGs can not have gas masses larger than $\sim 10^9$ M$_\odot$. This mass could be associated to the epoch at which the "quenching" processes started to act in halos corresponding to this gas mass at that epoch.

4.3.2 Galaxy baryonic-to-halo mass relations

Our final aim is to constrain the baryonic-to-halo mass function for local galaxies, separated into the LTG and ETG populations, and for the average case. This is plotted in Fig. 4.6, with the same structure and line coding as in Fig. 4.1. The scatters around the relations for the two populations are lower than 0.16 dex for $M_{\rm h} > 3 \times 10^{11}$ M$_\odot$. At lower masses, the relations become noisier. The scatter for the ETG population is ever slightly larger than the one for the LTG one. The scatter around the average relation is larger because of the segregation of the $M_{\rm bar}$–$M_{\rm h}$ relation by type/color.

The segregation by type/color in the $M_{\rm bar}$–$M_{\rm h}$ relation is clearly seen in Fig. 4.6. The LTGs have larger $M_{\rm bar}$ at a given $M_{\rm h}$ than the ETG ones. This is a reminiscence of the M_*–$M_{\rm h}$ relation (see above and Rodríguez-Puebla et al., 2015), but in the case of the $M_{\rm bar}$–$M_{\rm h}$ relation, the segregation increases at low masses. This is because lower mass LTGs are much more dominated by gas than ETG ones. The segregation of the local M_*–$M_{\rm h}$ and $M_{\rm bar}$–$M_{\rm h}$ relations by type/color is the result of important processes

of galaxy evolution. At large masses, the "quenching" processes make that galaxies interrupt or loss the gas and cease star formation; the larger the halos, the earlier this happens on average. Thus, while the halos continue growing hierarchically, the stellar mass remains the same, lowering with time then the $M_*/M_{\rm h}$ ratio (and hence, the $M_{\rm bar}/M_{\rm h}$ ratio). At the low-mass end of the M_*–$M_{\rm h}$ and $M_{\rm bar}$–$M_{\rm h}$ relations, where LTGs dominate, the question is why they keep star forming and gaseous until today, when their halos assembled actually early. Processes that delayed the gas consumption and star formation, the more, the less massive are the halos, should be at play.

As seen, the results obtained here are important to constrain galaxy formation and evolution processes. In Fig. 4.6, we plot the obtained stellar-to-halo (gray color) and baryonic-to-halo (purpure color) mass functions for the average case (upper panel), the LTGs (medium panel), and the ETGs (lower panel). While at masses larger than $M_{\rm h} \approx 10^{12}$ the stellar- and baryonic-to-halo mass relations are similar in all the cases, including the scatters, at lower masses a significant difference reveals for the LTGs a higher baryonic fraction compared with the average and less for the ETGs. Since at low masses dominate by much the former, then the density-weighted average relation is close to that one of them.

The accurate semi-empirical relations shown in Fig. 4.1 offer key constrains to models and simulations of galaxy evolution. They should reproduce these relations and their scatters.

Finally we present the best fit parameters for the $M_{\rm bar}$-$M_{\rm h}$ relation for the average, LTG, and ETG galaxies in Table 4.1. The function used to describe this relation has the form proposed by Behroozi et al. (2013),

$$\log_{10}(M_{\rm bar}) = \log_{10}(\epsilon M_1) + f(x) - f(0) \tag{4.14}$$

$$f(x) = -\log_{10}(10^{-\alpha x}+1) + \frac{\delta[\log_{10}(1+\exp(x))]^{\gamma}}{1+\exp(10^{-x})} \tag{4.15}$$

$$\text{where: } x = \log_{10}(M_{\rm h}/M_1)$$

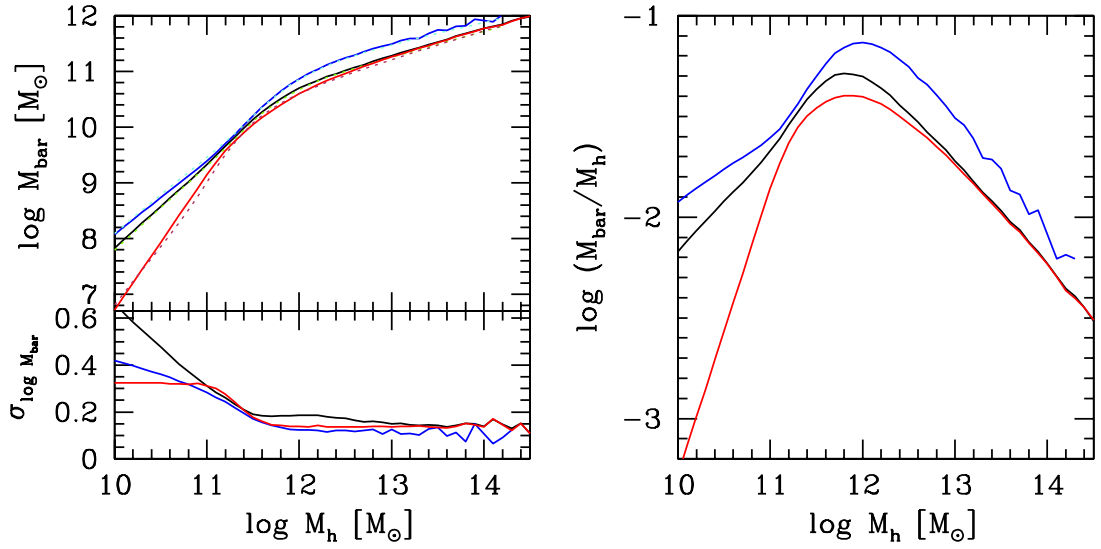

Figure 4.5: Baryonic-to-halo mass relations for LTG (blue lines) and ETG (red lines) galaxies, and for the density-weighted average of both populations (black lines). Also we plot the best fit for each galaxy population with dashed lines as given by eq. (4.14) and parameters in table 4.1. The lower panel shows the intrinsic scatter around each one of these mass relations in dex. The right panel is the same as the left one but for the $M_{\rm bar}/M_{\rm h}$ ratio.

Galaxy population	**Best-fit Parameters**					$\chi^2_{\rm red}$
Average	$\log_{10}(M_1)$	ϵ	α	δ	γ	0.90
	11.56	0.04	1.55	2.80	0.82	
LTG	$\log_{10}(M_1)$	ϵ	α	δ	γ	0.85
	11.68	0.06	1.40	2.78	0.69	
ETG	$\log_{10}(M_1)$	ϵ	α	δ	γ	1.13
	11.29	0.03	2.10	3.89	0.66	

Table 4.1: Parameters to the $M_{\rm bar}$-$M_{\rm h}$ relation given by eq. (4.14) and (4.15) for the the average, LTGs and ETGs.

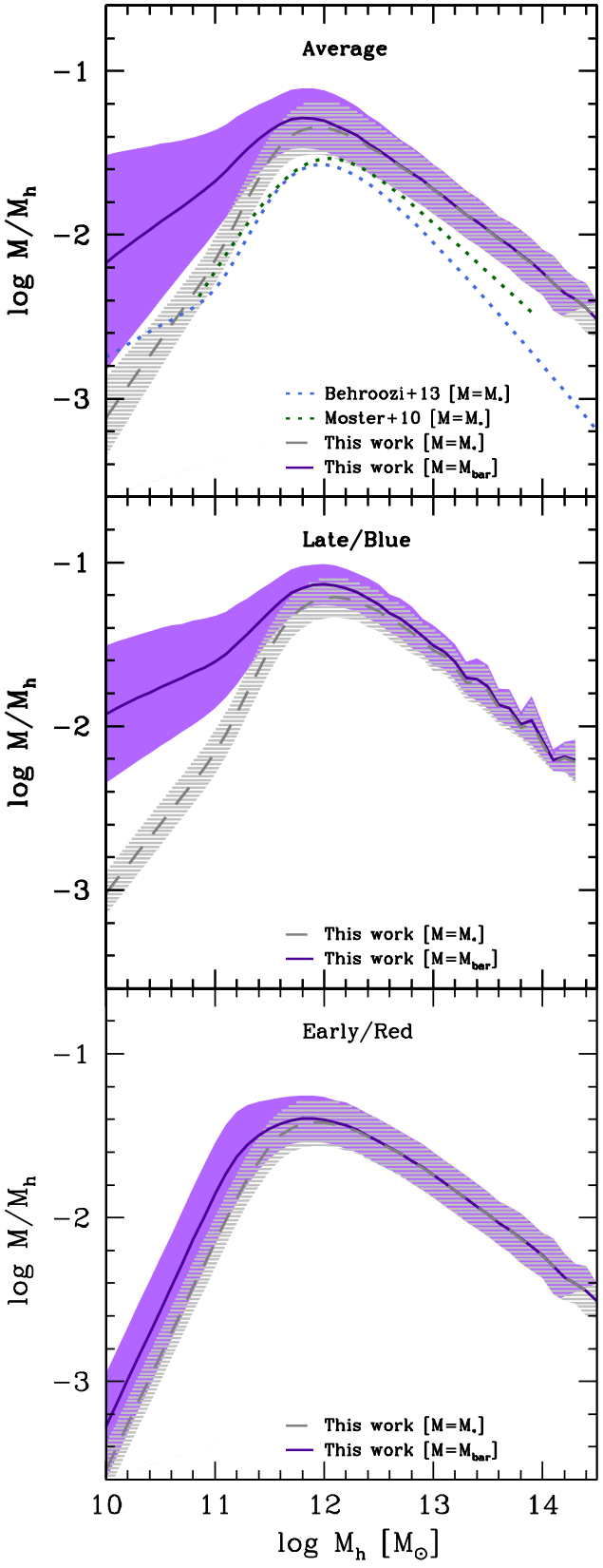

Figure 4.6: Stellar-to-halo (gray line and shaded area) and baryonic-to-halo (purpure line and shaded area) mass relations for the two galaxy populations and the density-weighted average of them.

Conexión Halo-Galaxia: Fracción bariónica de galaxias rojas y azules.

Chapter 5

Redshift evolution of the gas-stellar-mass relations

The results obtained in the previous Sections were for to the local galaxy population. It is expected that galaxies in the past were more gaseous, the more, the higher the redshift. As time goes, the gas is transformed into stars and some fraction can be even ejected by the SN- and AGN-driven outflows. Therefore, a next crucial constrain to galaxy formation and evolution should be given by the stellar-to-gas contents of galaxies as a function of mass at different redshifts.

In this Chapter, we approach the question of the H_2- and H_I-to-stellar mass ratios at different redshifts as inferred empirically. The most robust observational constraints at high redshifts are for the former. We will present a compilation of observational data on H_2 and M_* for galaxies at different redshifts. These determinations refer mostly to star forming, starburst or ultra-luminous infrared (ULIRG) galaxies. In fact, the fraction of active star forming, blue, late-type galaxies, which are basically disk galaxies, increases at higher redshifts; at $z > 1$ the disk galaxy population dominates at all, as observations (e.g., Bruce et al., 2012; Buitrago et al., 2013) and semi-empirical inferences (e.g., Avila-Reese et al., 2014) show. Thus, the gas-to-stellar mass relation at different redshifts to be determined below refers formally to what we called in previous Chapters as LTGs.

The observations in emission at higher redshifts are possible only for the CO lines, which trace H_2. The detection of the H_I 21cm emission line for atomic gas is not possible at higher redshifts due to its inherent weakness. We will constrain the evolution of the H_I content in galaxies by using measurements of atomic gas abundance in absorption systems (mostly by damped Lyα absorption, DLA).

5.1 Evolution of the molecular gas mass content of galaxies

5.1.1 The compilation

We have compiled from the literature several samples of high-redshift galaxies, for which $M_{\rm H_2}$ and M_* are reported. As discussed in Chapter 2, the CO-to-H_2 conversion factor seems to depend on the gas-phase metallicity. We use the same $\alpha_{\rm CO}$-$Z_{\rm gas}$ relation for late-type galaxies adopted there (Moustakas et al., 2010; Schruba et al., 2012, see eq. A.2 in Appendix A); note that for metallicities equal or higher than $Z_{\rm gas} \equiv 12 + \log_{10}({\rm O/H}) = 8.6$, we assume the Milky Way factor, $\alpha_{\rm CO}$=3.2, as in Schruba et al. (2012). However, we consider here a systematical decrease of $\alpha_{\rm CO}$ as z is higher from the argument that the depletion times, $\tau_{\rm dep}$, of higher-redsfhit galaxies are smaller than those of lower redshift galaxies (higher-redhsfit galaxies are on average more gas dominated and unstable and, therefore, with higher star formation efficiencies than the lower-redshift ones). A constant depletion time of 1.8 Gyr was used in Schruba et al. (2012) in order to infer the $\alpha_{\rm CO}$ factor from their local observational sample: $\alpha_{\rm CO}$=SFR$\times\tau_{\rm dep}$/L$_{\rm CO}$ (see Appendix A). We propose that $\log\tau_{\rm dep}$ decreases as:

$$\log\tau_{\rm dep} = 9.25 - 0.42 \times \log(1+z)[{\rm Gyr}], \tag{5.1}$$

in such a way that at $z = 3$, $\tau_{\rm dep} = 1$Gyr.

In the Appendix A, we also used a functionality that fit well the observed mean $Z_{\rm gas}$–M_* relation used to obtain the CO-to-H_2 conversion factor as a function of M_*, see eq. (2.12) in the Appendix A. We adopt the same functionality for the $Z_{\rm gas}$–M_* relation at higher redshifts but with the asymptotic metallicity and turn-over mass parameters changing with z in order to be consistent with the observational determinations reported in Maiolino et al. (2008) for a few high redshifts; see Appendix A for details. These and other authors find that the $Z_{\rm gas}$–M_* relation decreases on average with z (the metallicities are lower) but the decreasing is larger for lower masses.

In the Appendix A, we approximate the turn-over mass, M_{TO}, and the asymptotic metallicity, $12 + \log(\mathrm{O/H})_{\mathrm{asm}}$, dependences on z as:

$$\begin{aligned} \log_{10}(M_{TO}) \rightarrow \text{ increases with } z \text{ as:} \\ \psi(z) = -2.11\,[\log_{10}(1+z)]^2 + 3.36\log_{10}(1+z) \\ 12 + \log(\mathrm{O/H})_{\mathrm{asm}} \rightarrow \text{ decreases with } z \text{ as:} \\ \phi(z) = 0.64\,[\log_{10}(1+z)]^2 \end{aligned} \tag{5.2}$$

This implies that for a given M_*, the α_{CO} factor increases as z is higher (the galaxies have smaller metallicities).

Summarizing, for the compiled samples, we divide the H_2 masses reported therein by the CO-to-H_2 conversion factor used by the corresponding authors and multiply by our factor, which is mass and redshift dependent as mentioned above. As mentioned in the Appendix, we impose the condition that when the α_{CO} factor gets smaller than 3.2, the Milky Way value, then α_{CO} is fixed to this value. For ULIRGs, the CO-to-H_2 conversion factor seems to be systematically smaller than for normal galaxies. For these cases, we do not correct the α_{CO} values assumed by the authors.

Following, we present the different samples compiled by us:

- **Bauermeister et al. (2013):**

  An observational study of molecular gas in 31 star forming galaxies in the redshift range $0.05 \leq z \leq 0.5$ (EGNoG survey). The stellar masses of their galaxies are $4 - 30 \times 10^{10} M_{\odot}$. They observe the $CO(J = 1 \rightarrow 0)$ and $CO(J = 3 \rightarrow 2)$ rotational lines in the Combined Array for Research in Millimeter-wave Astronomy (CARMA), detecting 24 of the 31 galaxies. They adopt a Milky-Way like conversion factor $\alpha_{\mathrm{CO}} = 3.2$.

- **Geach et al. (2011):**

  Four detections and two upper limits of $CO(J = 1 \rightarrow 0)$ emission in the IRAM Plateau de Bure from a 24μm-selected sample of star forming galaxies at $z \sim 0.4$ are presented in this work. Using a Chabrier (2003) IMF, stellar masses were estimated and the galaxies that form this sample have $M_* \sim 10^{11} M_{\odot}$. They adopt

a Milky-Way like conversion factor that multiply by 1.4 to take into account a helium.

- **Combes et al. (2013):**

  CO observations obtained at the IRAM 30-m telescope for 39 galaxies, covering the redshift range $0.2 < z < 1$ and with $L_{FIR} > 10^{12} L_{\odot}$ (they select ULIRGs) are made in this work. They adopt a conversion factor $\alpha_{\rm CO}$= 0.8, which is an appropriate value for ULIRGs.The stellar masses were obtained from observed optical and near infrared magnitudes using the standard relations that exist as a function of colors, derived from stellar population models of Bell et al. (2003).

- **Tacconi et al. (2013):**

  A sample of CO detections in two redshift slices at $z \sim 1.2$ and $z \sim 2.2$, for 52 main-sequence star forming galaxies with stellar masses $M_* > 10^{10.4} M_{\odot}$. The sample is called PHIBSS, the IRAM Plateau de Bure high-z blue sequence $CO(J = 3 \rightarrow 2)$ survey. They adopted the Milky Way value for $\alpha_{\rm CO}$ and consider a Chabrier (2003) IMF in the stellar mass determination.

- **Daddi et al. (2010):**

  Six near-infrared selected galaxies at redshift $z \sim 1.5$ with $CO(J = 2 \rightarrow 1)$ measurements at the IRAM Plateau de Bure Interferometer are presented. They use dynamical models of clumpy disk galaxies in order to derive dynamical masses for their sample and these models are able to reproduce the peculiar spectral line shapes of the CO emission. Then, after accounting for stellar and dark matter masses, they derive molecular hydrogen masses in the range $(0.4 - 1.2) \times 10^{10} M_{\odot}$. They find a value of the conversion factor $\alpha_{\rm CO}$= 3.6 ± 0.8. Stellar masses were derived from synthetic templates assuming a Chabrier (2003) IMF.

- **Magdis et al. (2012a):**

  In this work, robust estimates of $M_{\rm dust}$ are derived from mid-infrared to millimiter data for individual and stacked data. Then, using the correlation of gas-to-dust mass with metallicity, they use their measurements to constraint the gas (H_2) content in galaxies. We use their stacked data at redshift $z \sim 2$.

- **Magdis et al. (2012b):**

  The data presented in this work consist of two massive and Infrared Lyman break galaxies, with measurements of the $CO(J = 3 \rightarrow 2)$ emission, using the IRAM

Plateau de Bure Interferometer. For the molecular gas estimate, they adopt a conversion factor $\alpha_{\rm CO}\sim 3.6$ and stellar masses are calculated assuming a Chabrier (2003) IMF.

5.1.2 H_2 mass fraction vs M_* at different redshfits

In Figs. 5.1 and 5.2, we present our compilation at eight redshift bins. As explained above, we have homogenized the $\alpha_{\rm CO}$ factor to the one introduced here. The H_2 mass fraction, $f_{\rm H_2}$, vs. M_* is plotted, where $f_{\rm H_2}= M_{\rm H_2}/(M_{\rm H_2}+ M_*)$. The different sources are indicated inside each panel. The dashed line in each panel corresponds to the prediction given in Sargent et al. (2013). These authors infer the $f_{\rm H_2}$–M_* relation from the observed specific star formation rate (sSFR)–M_* correlations at different redshifts. They connect sSFR with $f_{\rm H_2}$ by a model based on the empirical Kennicutt-Schmidt law.

Figures 5.1 and 5.2 show that the molecular mass fraction of galaxies systematically increases with z. The data are yet scarce, strongly scattered and for relatively small mass ranges, so that it is not easy to establish a dependence of $f_{\rm H_2}$ on M_* at different z's. We generalize the same $z\sim 0$ dependence of $R_{\rm H_2}\equiv M_{\rm H_2}/M_*$ on M_* (or $f_{\rm H_2}$ on M_*) found in Chapter 2 to the other redshifts, and by eye we "normalize" the mean $f_{\rm H_2}$–M_* relation to the observations, taking care to be also in rough agreement with the model-dependent empirical inferences of Sargent et al. (2013), which are based on the observed sSFR–M_* correlations. Our proposed $f_{\rm H_2}$–M_* mean relations at different redshifts are plotted in Figs. 5.1 and 5.2 with the solid blue lines. The shaded area around them are our estimate for the 1σ scatter. This scatter is the same considered for local galaxies (eq. 2.17), but changes because of the variation of the $R_{\rm H_2}$ ratio with redshift and it seems to agree well with observational data.

The law with which the $R_{\rm H_2}$–M_* zero point increases with z is similar to the one proposed in Sargent et al. (2013) for the sSFR–M_* relation:

$$\gamma(z) = 0.43\left(\frac{A\cdot z}{1+Bz^C}\right) \tag{5.3}$$

The parameters we find better fit the observational data in Figs. 5.1 and 5.2 are $A = 14, B = 3.54, C = 1.50$. The evolution of our mean $R_{\rm H_2}$–M_* relation is plotted in Fig. 5.3. As said above, we assume this relation to be the same that at $z\sim 0$ for LTGs (see

Chapter 2) but with its zero point changing with z, that is:

$$\log_{10} R_{\rm H_2}(z) = \log_{10}(B) - \log_{10}\left[\left(\frac{M_*}{M^s}\right)^{\xi} + \left(\frac{M_*}{M^s}\right)^{\rho}\right] + \gamma(z) \tag{5.4}$$

The evolution of the zero-point value of the $R_{\rm H_2}$–M_* relation agrees very well with the one presented in a recent work by Genzel et al. (2014).

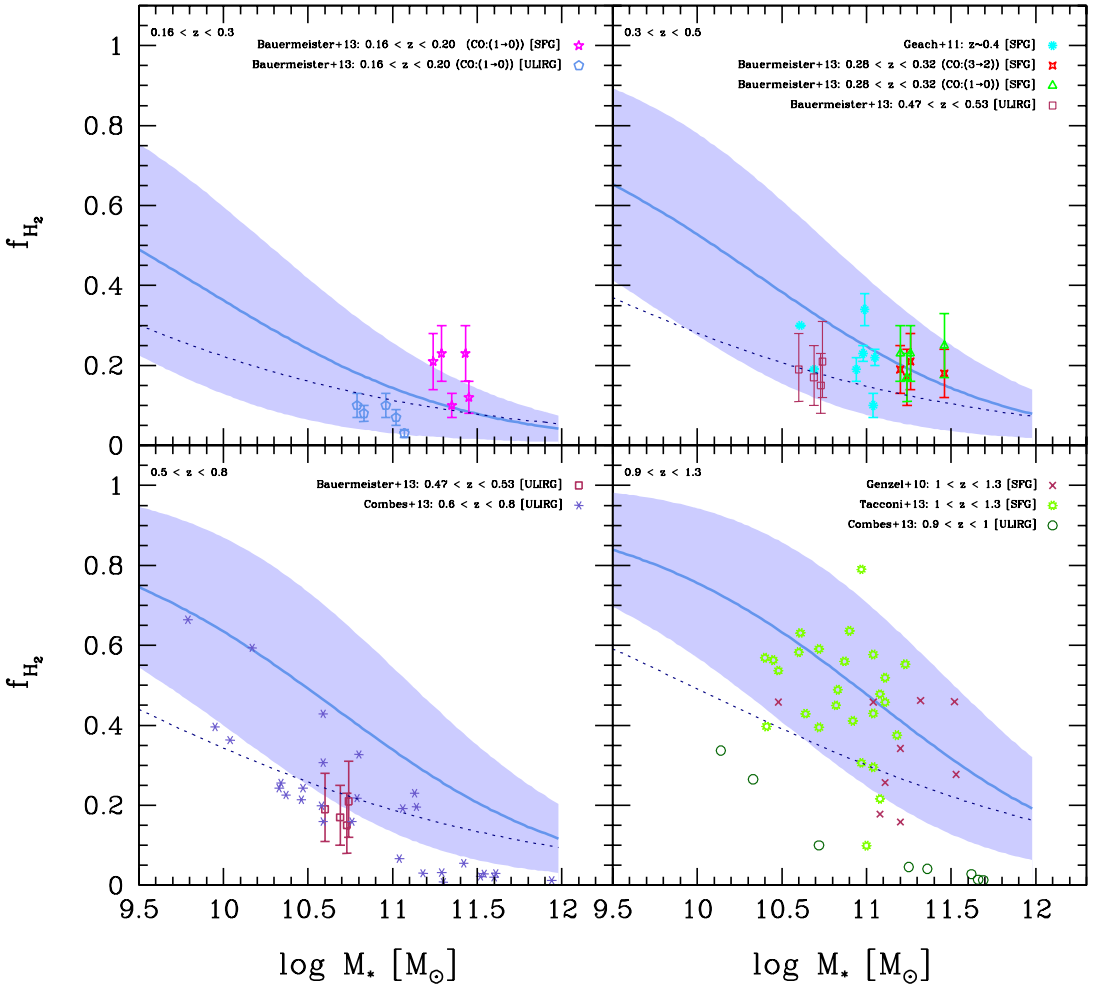

Figure 5.1: Redshift evolution of the $f_{\rm H_2}$–M_* correlation in different redshift ranges up to $z \approx 1.3$ as we infer to be consistent with observations (blue lines for the means and shaded areas for the 1σ scatter). The compiled and corrected observational data are shown with different symbols. Inside each panel it is indicated the redshift range, the observational sources, and the inference from the Sargent et al. (2013) work (dark blue dashed line).

5.2 Evolution of the neutral atomic gas mass content of galaxies

As mentioned above, there are not measurements of the H I 21-cm line flux for galaxies at high redshifts due to the weakness of this line and the sensitivity limit of current

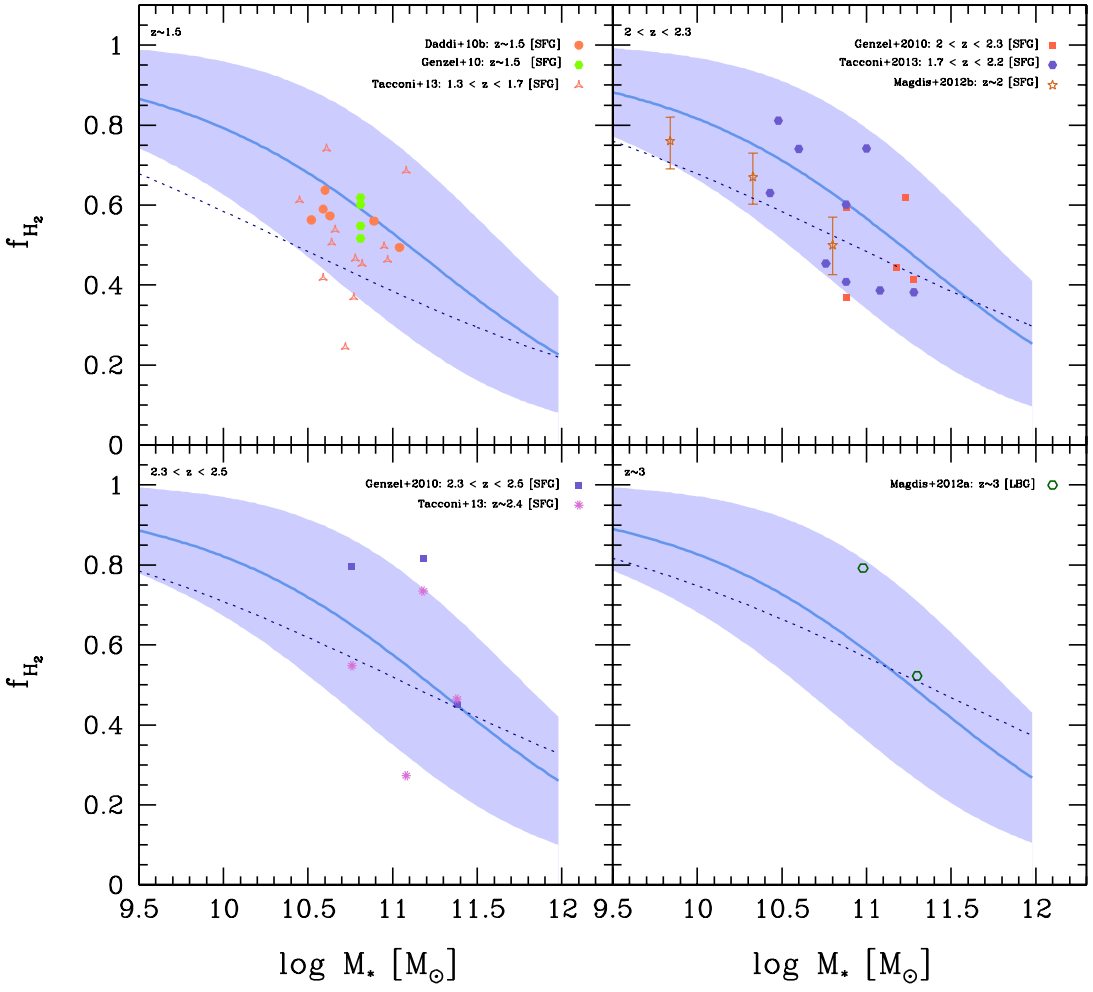

Figure 5.2: As Fig. 5.1 but for redshifts from $z \sim 1.5$ to ~ 3.

detectors. Some determinations were obtained at intermediate redshifts ($z < 0.4$) by applying stacking techniques to attain the necessary signal-to-noise ratios (e.g., Lah et al., 2007; Rhee et al., 2013, and see more references about other determinations therein). For higher redshifts, the only information about neutral atomic gas content we have is the one provided by absorption systems like the DLAs. The DLAs are the highest column density Lyα absorption features seen in the spectrum of quasars, and they helped to constrain the average H_I mass density at different redshifts (see e.g., Rhee et al., 2013).

In Rhee et al. (2013), a compilation of the H_I mass density determinations obtained with DLAs at high redshifts (up to $z \sim 3$) and with the stacking techniques at low redshifts is presented, including their own results with the latter technique. They present this density normalized to the critical density of the Universe, that is, the Ω_{H_I} parameter

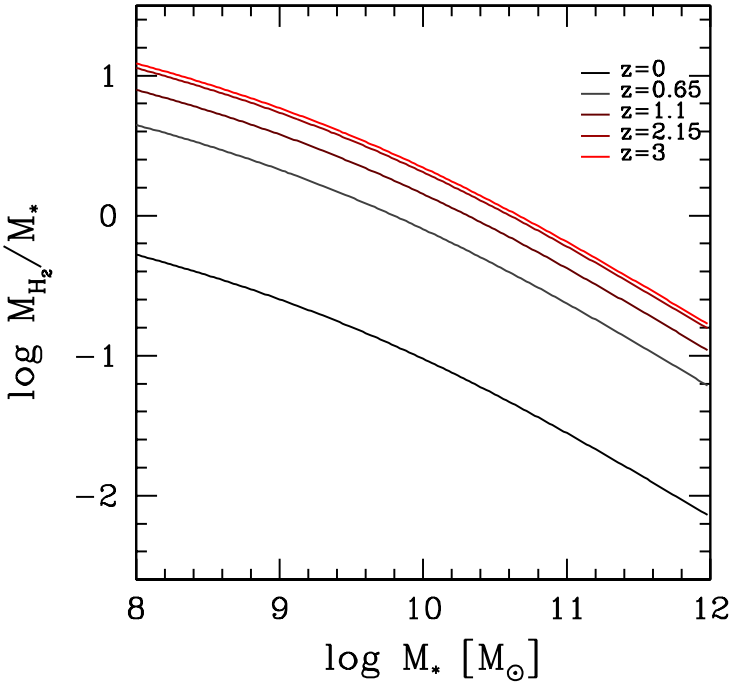

Figure 5.3: Redshift evolution of the mean $R_{\rm H_2}$–M_* relation as inferred here.

at different z's. The data are actually scarce, with large error bars, and not following a clearly monotonic dependence with redshift. Very roughly, we fit to these data a line that grows as $\log_{10}(\Omega_{0,\rm H_I}) + \eta(z)$, with $\eta(z) = 0.1z$. This dependence on z is used to determine how the zero point of the mean $R_{\rm H_I}$–M_* relation increases with z. As in the case of the $R_{\rm H_2}$–M_* relation, we assume that the $R_{\rm H_I}$–M_* relation shape at any redshift is the same that at $z \sim 0$ (for LTGs; see Chapter 2) and the zero point of the relation changes with z, that is:

$$\log_{10} R_{\rm H_I}(z) = \log_{10}(B) - \log_{10}\left[\left(\frac{M_*}{M^s}\right)^{\xi} + \left(\frac{M_*}{M^s}\right)^{\rho}\right] + \eta(z) \qquad (5.5)$$

The values of the parameters related to the M_* dependence were presented in Chapter 2. In Fig. 5.4, the obtained evolution of our mean $R_{\rm H_I}$–M_* relation is plotted.

5.3 Evolution of the total cold gas mass content of galaxies

Once we have determined the evolution of the $\rm H_I$ and $\rm H_2$ gas mass contents, we can finally obtain the evolution of the total cold gas mass content. The cold gas mass is just $M_{\rm gas}$= 1.4($M_{\rm H_I}$+ $M_{\rm H_2}$), and $R_{\rm gas}$=$M_{\rm gas}/M_*$ is the gas-to-stellar mass ratio. In Fig. 5.5 we present the evolution of the mean $R_{\rm gas}$–M_* relation obtained from the $R_{\rm H_I}$–M_* and

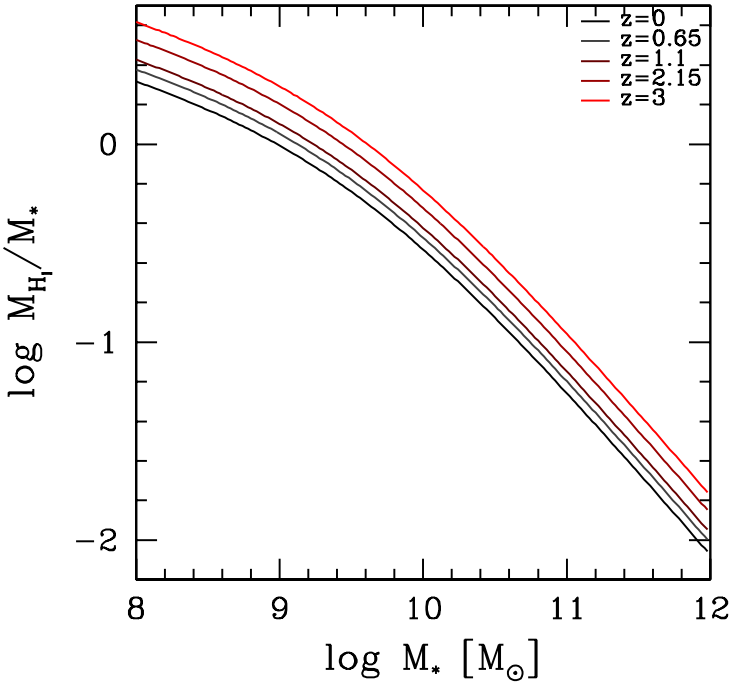

Figure 5.4: Redshift evolution of the mean $R_{\rm H_I}$-M_* relation as inferred here.

$R_{\rm H_2}$–M_* relations and their changes with z presented in the previous sections. The main contribution to the increasing in the $R_{\rm gas}$ ratio with z is due to the strong increasing of the molecular-to-stellar mass ratio in galaxies.

The next step from our analysis is to infer the evolution of the $M_{\rm bar}$–M_* relation. For this, we need the evolution of the GSMF. A compilation of the GSMF evolution is in progress. Once we have determined the $M_{\rm bar}$–M_* relations at different epochs, they will be used to generate mock catalogs of galaxies at different redshifts from where we can obtain the evolution of the galaxy baryonic mass function. By connecting these mass functions with the halo mass functions, as it was done in Chapter 4, we will be able to obtain the evolution of the $M_{\rm bar}$–M_h relation.

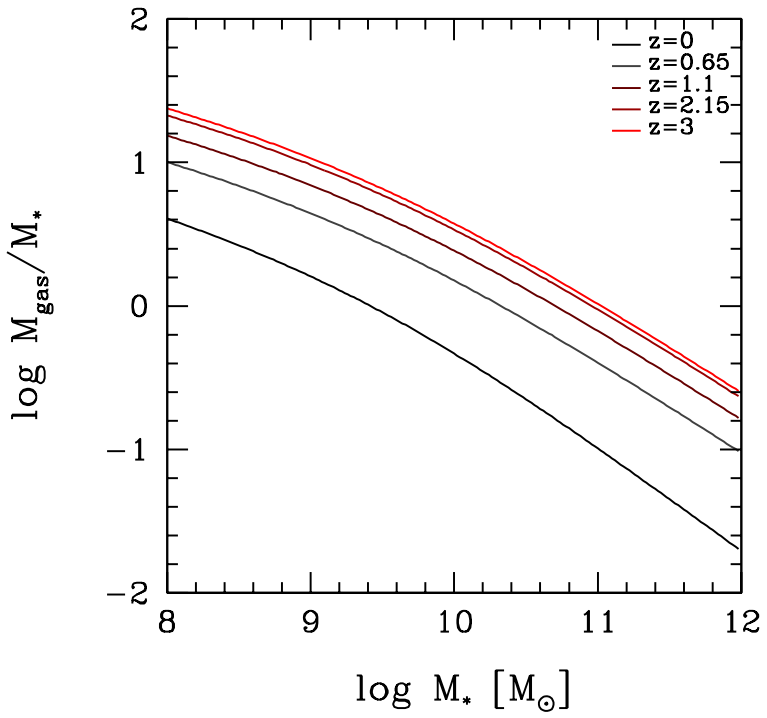

Figure 5.5: Redshift evolution of the mean $R_{\rm gas}$–M_* relation as inferred here.

Chapter 6

Conclusions

La conexión halo-galaxia a través de métodos estadísticos de tipo semi-empírico se ha tornado una herramienta poderosa para entender la formación y evolución de las galaxias pero, sobre todo, para ofrecer importantes restricciones a los procesos evolutivos y astrofísicos que se invocan en los modelos y simulaciones numéricas de evolución de galaxias. En última instancia, estas restricciones podrían también ser importantes para sondear el modelo cosmológico subyacente en el enfoque semi-empírico, el ΛCDM.

En esta Tesis nuestro objetivo fue lograr la conexión halo-galaxia a nivel de contenidos de gas y bariones (gas + estrellas) para las galaxias locales, separadas en dos grandes poblaciones por su tipo/color, así como para el promedio de ambas poblaciones. La determinación de dicha conexión a este nivel, sería la primera en la literatura. Hemos usado la terminología de LTG (late-type galaxies) para las galaxias tardías/azules y ETG (early-type galaxies) para las tempranas/rojas.

A fin de lograr nuestro objetivo fue necesario empezar con una exhaustiva compilación en la literatura de galaxias locales que contengan información de sus masas estelares, tipos/colores y de sus masas en H_I y H_2. En base a esta compilación y su correspondientes homogeneización pudimos entonces establecer las correlaciones entre M_* y M_{H_I} y M_{H_2}. Como resultado de esta primera parte presentada en el Capítulo 2, remarcamos los siguientes resultados

- Existen fuertes diferencias por tipo/color en las relaciones M_{H_I}–M_* y M_{H_2}–M_*, razón por la que no es conveniente considerar un promedio por poblaciones para estas relaciones como se ha hecho en la literatura, más aún si las muestras no son

completas en volumen, es decir cuando las galaxias usadas para obtener estas relaciones no están pesadas por su abundancia. El cociente M_{H_2}/M_{H_I} de las galaxias LTG aumenta lentamente con la masa estelar, mientras que para las galaxias ETG crece significativamente con M_*. Estos resultados dependen en parte del factor de conversión CO-a-H_2, $\alpha_{\rm CO}$, usado para calcular M_{H_2}. Siguiendo las sugerencias más recientes de la literatura sobre la variación de $\alpha_{\rm CO}$ con la metalicidad, hemos inferido aquí una dependencia de este factor con M_* (ver Apéndice A), misma que se usó para calcular M_{H_2} a partir de los datos compilados.

- La relación M_{H_I}–M_* de galaxias LTG crece significativamente (como ${M_*}^{0.7}$) hasta $\log(M_*/{\rm M}_\odot) \approx 9.5$ para luego aplanarse considerablemente y crecer como ${M_*}^{0.15}$; el cociente $R_{\rm H_I} \equiv M_{H_I}/M_*$ es en promedio igual a 1 para galaxias de $\log(M_*/{\rm M}_\odot) \approx 9$. La dispersión estimada alrededor de la relación disminuye en función de la masa estelar desde $\sigma \approx 0.6$ hasta 0.4 dex. La relación M_{H_I}–M_* para las galaxias ETG tiene un comportamiento distinto, apareciendo un pico a $\log(M_*/{\rm M}_\odot) \approx 9.5$. Para masas estelares menores, $M_{H_I} \propto {M_*}^1$ y para masas mayores, $M_{H_I} \propto {M_*}^{-0.7}$; los cocientes $R_{\rm H_I}$ son menores a 1 incluso para las galaxias de muy bajas masas. La dispersión alrededor de la relación es muy grande, en parte porque hay pocas observaciones de HI en estas galaxias y en muchos casos son no detecciones, pero también porque intrínsecamente es así. Hemos estimado un valor de $\sigma \sim 0.7$ dex como aproximado de la dispersión intrínseca.

- La relación M_{H_2}–M_* para galaxias LTGs es relativamente estrecha (dispersión intrínseca estimada que decrece con la masa desde ≈ 0.5 a 0.35 dex) tiene una forma funcional similar al caso de M_{H_I}–M_*, creciendo como ${M_*}^{0.8}$ hasta $\log(M_*/{\rm M}_\odot) \approx 9.5$ y luego como ${M_*}^{0.35}$ a masas mucho mayores. Sin embargo la normalización es un factor 2–3 menor que en el caso de la relación M_{H_I}–M_*; el cociente $R_{\rm H_2} \equiv M_{H_2}/M_*$ alcanza el valor de ~ 1 sólo en galaxias tan pequeñas como $\approx 1 \times 10^7$ ${\rm M}_\odot$. Las masas en H_2 en galaxias de baja masa podrían ser mucho menores si no hubiéramos introducido un factor de conversión $\alpha_{\rm CO}$ que crece cuando M_* es menor. Para las galaxias ETG, la relación M_{H_2}–M_* es mucho más dispersa y la dependencia va como ${M_*}^{1}.45$ hasta $\log(M_*/{\rm M}_\odot) \approx 9.5$ y luego decrece como ${M_*}^{-0.35}$. Para galaxias ETG usamos el factor $\alpha_{\rm CO}$ constante.

- Considerando que la masa de gas frío es $M_{\rm gas}$=1.4(M_{H_I}+ M_{H_2}), de nuestra correlaciones obtuvimos las correlaciones $M_{\rm gas}$–M_* para galaxias LTG y ETG, mismas

que siguen los patrones principalmente de las correlaciones M_{H_I}–M_*. El cociente $R_{\rm gas} \equiv M_{\rm gas}/M_*$ de las LTGs de baja masa crece mientras M_* es más pequeño como ${M_*}^{-0.3}$; para $M_* \approx 10^7$ $M_\odot$, este cociente es en promedio 8 y se hace igual a ≈ 1 para galaxias de 3×10^9 $M_\odot$. A masas mayores, $R_{\rm gas}$ decrece rápido, aproximadamente como ${M_*}^{-0.5}$. Las galaxias ETGs son mucho menos gaseosas, siendo $R_{\rm gas}$ menor a 1 en promedio incluso para las enanas de $M_* \approx 10^7$ $M_\odot$. Para masas mayores as $\approx 3 \times 10^9$ $M_\odot$, $R_{\rm gas}$ cae fuertemente con M_*, de tal manera que galaxias tan masivas como $M_* \sim 10^{11}$ $M_\odot$, tienen cocientes $R_{\rm gas}$ por debajo de 2.5×10^{-3} en promedio; son extremadamente pobres en gas frío.

Habiendo determinado las correlaciones empíricas de M_{H_I}–M_* y M_{H_2}–M_* para las dos poblaciones de galaxias, hicimos uso de las funciones de masa estelar (GSMFs) de galaxias azules y rojas que estimamos aquí hasta $M_* \approx 10^7$ $M_\odot$ en base al análisis de Rodríguez-Puebla et al. (2015) a fin mapear dichas correlaciones a las correspondientes funciones de masa de H_I y H_2, y posteriormente de gas y de bariones. Para esto, generamos un catálogo sintético de 3 millones de galaxias que muestrean estadísticamente la GSMF inferida de las observaciones así como las fracciones de galaxias azules (LTG) y rojas (ETG). A las galaxias de este catálogo se le asignaron masas de H_I y H_2 a partir de su tipo y masa estelar usando las correlaciones empíricas presentadas en el Capítulo 2. Se procedió entonces a construir, a partir del catálogo, las funciones de masa de H_I y H_2 así como de $M_{\rm gas}$ y $M_{\rm bar}$ (abreviadas como GSMF, GH_IMF, GH_2MF y GBMF respectivamente). Algunos resultados que remarcamos de este Capítulo 3 son:

- La GSMF se construyó empatando la GSMF de Baldry et al. (2008), quienes corrigieron meticulosamente por completitud a bajas masas, con la GSMF de Rodríguez-Puebla et al. (2015), quienes usaron una muestra más grande (DR7 vs DR4, ambas del SDSS) y tomaron en cuenta una corrección por "apertura" en la estimación de las masas estelares, misma que afecta principalmente a las altas masas. La GSMF resultante es completa hasta $M_* \sim 10^7$ $M_\odot$; la pendiente en el lado de bajas masas es -1.49. La masa característica donde la GSMF empieza a caer drásticamente es $\log(M_*/M_\odot) \approx 10.6$; la caída no es exponencial, sino que sub exponencial. A masas mucho menores a $\log(M_*/M_\odot) \approx 10$ dominan las galaxias LTG y a masas mucho mayores dominan las ETGs.

- La GH_IMF total tiene una pendiente empinada de -1.72 hacia las masas tendientes a $M_{H_I}= 10^8$ $M_\odot$. La masa característica a la cual la GH_I MF comienza a decrecer fuertemente es $\log(M_{H_I}/M_\odot)\approx$ 9.9. La GH_IMF es dominada casi a todas las masas por la componente de LTGs. Sólo a las masas más grandes, cercanas a $M_{H_I}\sim 10^{11}$ $M_\odot$, la contribución de la población ETG es dominante; la caída a estas masas, más que una exponencial, es de tipo subexponencial. La GH_IMF inferida de nuestro catálogo sintético describe bien las funciones construidas de muestras observacionales en H_I, en los intervalos de masa en que éstas son confiables. En este sentido, nuestra GH_IMF describe integralmente estas previas determinaciones y ofrece una determinación confiable en un intervalo de masas extenso, desde $M_{H_I}\sim 10^8$ $M_\odot$.

- La GH_2MF total tiene una pendiente empinada de ≈ -1.79 hacia las masas tendientes a $M_{H_2}= 10^8$ $M_\odot$. Este comportamiento en parte se debe al factor α_{CO} dependiente de masa que hemos usado para estimar las masas de M_{H_2}. La masa característica a la cual la GH_2MF comienza a decrecer fuertemente es $\log(M_{H_2}/M_\odot)\approx$ 9.1. La GH_2MF es dominada por completo por la componente de LTGs.

- La GBMF obtenida de nuestro catálogo construido en base a la GSMF observada, es completa hasta $\log(M_{bar}/M_\odot)\approx$ 8.5. La pendiente a estas masas es de -1.52, más empinada que la pendiente de la GSMF. La masa característica a la cual la GBMF comienza a decrecer fuertemente es $\log(M_{bar}/M_\odot)\approx$ 10, la caída siendo más bien subexponencial. La GBMF se ajusta con la combinación de una función Schechter y una Schechter subexponencial. A masas menores que $\log(M_{bar}/M_\odot)\approx$ 10.6, domina la componente de LTGs y a masas mayores domina la componente ETG. Esta es la primera determinación de una GBMF que (1) incluye H_I, H_2, helio y metales y (2) está descompuesta en galaxias LTG y ETG.

- La completitud de la GBMF construída en base a la GSMF, es decir con una selección "óptica", depende fuertemente de la masa límite de completitud de la GSMF. En este caso, la GBMF deja de ser significativamente completa a una masa bariónica más de 10 veces mayor que la masa estelar límite.

Finalmente, contando con las funciones de masa de H_I y H_2 así como de estrellas, para las dos poblaciones de galaxias (LTG y ETG), aplicamos un método estadístico

que denominamos MATCH a fin de conectar estadísticamente dichas funciones de masa con la función de masa de los halos ΛCDM. Más que hacer una correspondencia de abundancias analítica, generamos un enorme catálogo sintético que (1) muestrea las GSMFs de las poblaciones LTG y ETG presentadas en el Capítulo 3, (2) establece la conexión M_*–M_h para ambas poblaciones, y (3) asigna las masas correspondientes M_{H_I} y M_{H_2} muestreando las correlaciones empíricas presentadas en el Capítulo 2. Como resultado, nuestro catálogo sintético, mismo que por construcción reproduce todas las funciones de masa empíricas presentadas en el Capítulo 3 y la funciń de masa de halos+subhalos ΛCDM, cuenta para cada población de galaxias con M_*, M_{H_I}, M_{H_2}, y M_h. Debido a que la masa de gas $M_{\rm gas}$ se calcula de M_{H_I} y M_{H_2} y la bariónica a su vez es $M_{\rm bar}$=$M_{\rm gas}$+ M_*, nuestro catálogo contiene entonces todas las masas a fin de lograr la conexión masa de halo con masas estelares, gaseosas y bariónicas. Los resultados obtenidos en base a nuestro catálogo sintético semi-empírico presentados en el Capítulo 4 se resumen así:

- Las relaciones M_*–M_h de galaxias LTG y ETG son algo diferentes. A paridad de M_h, las galaxias LTG tienen en promedio una M_* mayor que las ETG en especial a masas grandes y chicas. Las diferencias en los promedios para esas masas son del orden o algo mayores a dispersiones alrededor de las relaciones, mismas que son son relativamente bajas.

- Las relaciones M_{H_I}–M_h, M_{H_2}–M_h y $M_{\rm gas}$-M_h para las poblaciones LTG y ETG son muy diferentes, lo cual indica que considerar relaciones promedio no es muy apropiado. Eso sí, a bajas (altas) masas, donde dominan en abundancia las LTGs (ETGs), las relaciones promedios se aproximan a las de las galaxias LTGs (ETGs). Las dispersiones alrededor de estas relaciones son grandes, en especial para las relaciones promedio; esto último debido a la fuerte segregación que hay entre LTGs y ETGs.

- Las relaciones M_{H_I}–M_h, M_{H_2}–M_h y $M_{\rm gas}$-M_h de las LTGs crecen monotónamente, rápido hasta $\log(M_h/{\rm M}_\odot)\approx$ 11.4 pero a masas mayores tienden a aplanarse. El comportamiento en el caso de la población ETG es muy diferente: las relaciones crecen hasta $\log(M_h/{\rm M}_\odot)\approx$ 11.2 pero luego decrecen. Este comportamiento, especulamos, es una huella de los procesos de apagado (quenching) que actúan en las galaxias masivas y que hacen justamente que sean ETGs y pobres en gas.

- Las relaciones $M_{\rm bar}$–M_h de galaxias LTG y ETG son relativamente estrechas; las dispersiones son respectivamente de ≈ 0.12 y 0.15 dex pero para masas menores a $\log(M_h/{\rm M}_\odot)\approx 11.4$ aumentan significativamente. Hay una segregación estadísticamente significativa entre ambas relaciones; a paridad de M_h, las galaxias LTG tienen masas bariónicas mayores que las LTGs. Esta diferencia se incrementa hacia bajas masas debido a que las galaxias LTG enanas son mucho más dominadas por gas que las ETG enanas.

- La relaciónes M_*–M_h y $M_{\rm bar}$–M_h son muy similares a las altas masas ($M_* \approx M_{\rm bar} > 10^{12}$ M$_\odot$), tanto para LTGs y ETGs como para los promedios entre ambas. Para masas menores, debido al incremento de la fracción gaseosa mientras menor es la masas, las relaciones se hacen diferentes, en especial para las LTGs: la relación $M_{\rm bar}$–M_h se hace más plana que la M_*–M_h.

Finalmente, en el Capítulo 5, buscamos determinar cómo evolucionan las correlaciones empíricas de las masas de H_I y H_2 con la masa estelar. En el caso del H_2 se hizo una compilación de observaciones a diferentes corrimientos al rojo, hasta $z \sim 3$; se consideró además una corrección del factor de conversión CO-a-H_2 dependiente de la metalicidad y por ende de la masa estelar y el tiempo cósmico. Para el caso del H_I utilizamos el cambio con el corrimiento al rojo de la densidad cósmica del H_I estimado principalmente de observaciones en absorpción de los sistemas DLA. De esta manera estimamos la evolución del punto cero de la relación M_{H_I}–M_* local. Nuestro principal resultado es que la fracción o cociente de masa del H_2 crece bastante con el corrimiento al rojo; a $z \sim 3$, las galaxias tienen cocientes $M_{H_2}/M_* \sim 30$ veces mayores que a $z = 0$. En el caso del H_I el crecimiento es mucho menor, llegando a ser sólo un factor 2 mayor a $z \sim 3$.

Se planea usar la evolución de las correlaciones M_{H_I}–M_* y M_{H_2}–M_* en combinación con determinaciones observacionales de las GSMFs a diferentes corrimientos al rojo para realizar el análisis hecho aquí con galaxias locales. De esta manera podremos inferir la evolución de la relación $M_{\rm bar}$–M_h.

Los resultados presentados en esta Tesis son de gran relevancia para constreñir modelos y simulaciones numéricas de evolución de galaxias así como para entender aspectos claves de los procesos astrofísicos de formación y evolución de galaxias dentro de los halos oscuros en función de la masa y del tipo/color de la galaxia.

Appendices

Appendix A

The CO-to-H_2 conversion factor

In this Appendix, we discuss how to establish an average dependence of the CO-to-H_2 conversion factor with the stellar mass of galaxies based on observational results, as well as how this dependence evolves with redshift.

Several authors have shown that the CO-to-H_2 conversion factor depends on the gas phase metallicity (see e.g., Boselli et al., 2002; Schruba et al., 2012; Narayanan et al., 2012; Bolatto et al., 2013, ;etc). For example, Schruba et al. (2012) infer the CO-to-H_2 conversion factor from SFR and L_{CO} measures in a sample of galaxies, for which the gas phase metallicities were also measured, and this way find a correlation between α_{CO} and metallicity. Schruba et al. (2012) use a sample of 16 dwarf irregulars from the HERACLES survey. They begin by asuming that a constant H_2 depletion time of $\tau_{dep} = 1.8$ Gyr holds for these type of galaxies, based in results found for a large sample of spiral galaxies. Then, from GALEX and SINGS surveys they use FUV and IR data respectively obtaining SFR densities. The samples of galaxies from HERACLES have also available CO luminosities. Thus, with this information they find the α_{CO} conversion factor:

$$\alpha_{CO} = \frac{\tau_{dep}\ SFR}{L_{CO}} \tag{A.1}$$

Metallicities (gas phase oxygen abundances) are obtained from the work of Moustakas et al. (2010), therefore the α_{CO}-metallicity relation for star-forming galaxies is given by a linear relation of the form,

$$\log(\alpha_{CO}) = \zeta + [N \times (12 + \log_{10}(O/H) - 8.7)] \tag{A.2}$$

where normalization ζ is $\alpha_{\rm CO}$ at $12+\log_{10}({\rm O/H}) = 8.7$ and N is the slope of the relation. Then, for the sample without considering starburst galaxies from HERACLES used by Schruba et al. (2012), the parameters are $\zeta = 0.85$ and $N = -2$.

We remark that in this work, they found that for galaxies with metallicities $12 + \log_{10}({\rm O/H}) \gtrsim 8.6$ the conversion factor is $\alpha_{\rm CO} \sim \alpha_{\rm CO,MW}$, this is, late type galaxies at solar gas phase metallicities or higher tend to have a Milky-Way like CO-to-H_2 conversion factor and we use this result as a constraint to our model.

In order to relate the $\alpha_{\rm CO}$ conversion factor with stellar mass, we use the mass-metallicity relation for galaxies in the local universe. Sánchez et al. (2013) and Andrews and Martini (2013) determined the mass-metallicity relation for galaxies using the CALIFA and the SDSS surveys, respectively, in the stellar mass range $8.4 \leq \log_{10}(M_*) \leq 11.2$ for the former and $7.4 \leq \log_{10}(M_*) \leq 11.2$ for the latter. The work by Sánchez et al. (2013) provides a more accurate estimation 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 IFUs. However, the M_* range in the CALIFA sample is limited. Andrews and Martini (2013) cover galaxies of lower stellar masses. Fortunately, at the mass range where both studies coincide, it seems that they agree (see figure A.1).

Thus, we use the mass–metallicity found in Andrews and Martini (2013). They find that the relation proposed by Moustakas et al. (2011) fits well their observational results,

$$12 + \log_{10}({\rm O/H}) = (12 + \log_{10}({\rm O/H})_{\rm asm}) - \log_{10}\left(1 + \left(\frac{M_{TO}}{M_*}\right)^{\gamma}\right) \tag{A.3}$$

where $12 + \log_{10}({\rm O/H})_{\rm asm}$ is the asymptotic metallicity, M_{TO} is the turnover mass and γ is what controls the slope of the mass-metallicity relation. The values found are $12 + \log_{10}({\rm O/H})_{\rm asm} = 8.798$, $M_{TO} = 8.901$ and $\gamma = 0.640$.

Now, we are able to obtain the average $\alpha_{\rm CO}$-M_* relation by combining equations (A.2) and (A.3), and using the result from Schruba et al. (2012) that galaxies with metallicities higher than than solar one have Milky-Way $\alpha_{\rm CO}$ values. Therefore, for all galaxies more massive than $\sim 3 \times 10^{10}$, we assign $\alpha_{\rm CO} = \alpha_{CO,MW}$, otherwise,

$$\log_{10}(\alpha_{\rm CO}) = 0.42 + 2\log_{10}\left[1 + 0.1\left(\frac{3 \times 10^{10}{\rm M}_\odot}{M_*}\right)^{0.64}\right] \tag{A.4}$$

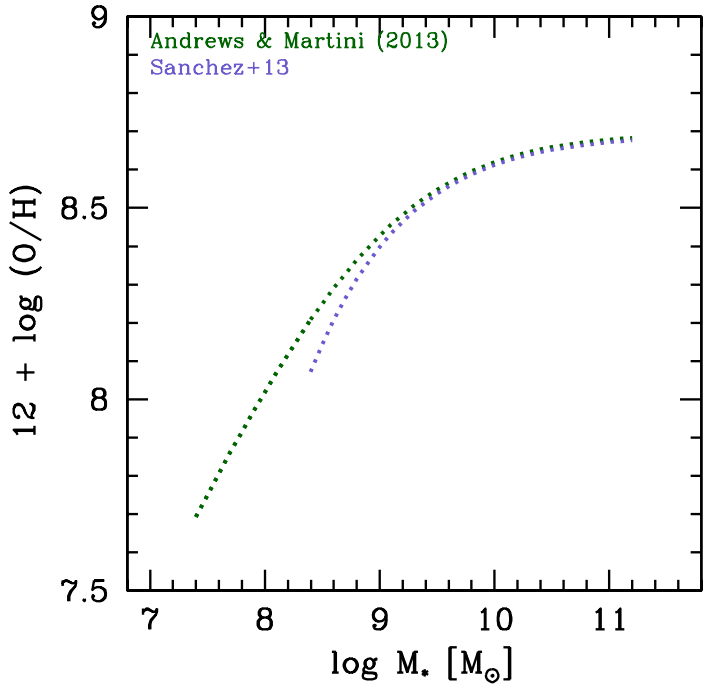

Figure A.1: Comparison of Sánchez et al. (2013) (violet dotted line) and Andrews and Martini (2013) (green dotted line) mass-metalliciy relations. There is a small diferrence at $M_* \approx 10^{8.6} M_\odot$ due to the limited low stellar stellar mass of CALIFA survey.

A.1 Redshift dependence of the $\alpha_{\rm CO}$-M_* correlation.

Schruba et al. (2012) found their $\alpha_{\rm CO}$-gas-metallicity correlation (eq. A.2) is similar for local and galaxies at higher redshifts by comparing their results with those obtained by Genzel et al. (2012), where they consider the same method to estimate $\alpha_{\rm CO}$ but for galaxies at $z \geq 1$. $\alpha_{\rm CO}$ seems to be less steep for galaxies at higher redshift than in the local universe, this by the fact that in the past, galaxies are more efficient in transforming their gas into stars, and this can be seen observationally if one can determine the time of depletion $\tau_{\rm dep}$, this is, the time for a galaxy to transform its gas into stars, which Genzel et al. (2012) found to be $\tau_{\rm dep} \sim 1$ Gyr and Schruba et al. (2012) consider $\tau_{\rm dep} \sim 1.8$ Gyr.

This lead us to perform the same procedure considered for local galaxies in order to relate $\alpha_{\rm CO}$ with stellar masses at different cosmic times, regarding the fact that as redshift increases galaxies have decreasing metallicity and $\alpha_{\rm CO}$ tend to be less steep than value found at the local universe.

Firstly, by considering the different depletion time in local and high redshift galaxies used in Genzel et al. (2012) and Schruba et al. (2012), we infer that the redshift evolution

of depletion time $\tau_{\rm dep}$ for galaxies in the redshift range $0 \leq z \leq 3$ is given by,

$$\log \tau_{\rm dep} = 9.25 - 0.42 \times \log_{10}(1+z)[\rm Gyr], \tag{A.5}$$

in such a way that at $z = 3$, $\tau_{\rm dep} = 1$Gyr.

The work of Maiolino et al. (2008) presents how the mass-metallicity relation decreases with z and it seems that this decreasing is larger for lower stellar masses. From this work we infer a redshift dependence of parameters M_{TO} and $12 + \log_{10}(\rm O/H)_{asm}$ from the mass-metallicity relation for local galaxies given in eq. (A.3) as:

$$\begin{aligned} \log_{10}(M_{TO}(z)) =& \log_{10}(M_{TO}) + \psi(z) \\ 12 + \log_{10}(\rm O/H)_{asm}(z) =& 12 + \log_{10}(\rm O/H)_{asm} - \phi(z) \end{aligned} \tag{A.6}$$

where the functions $\phi(z)$ and $\psi(z)$ are:

$$\begin{aligned} \phi(z) =& 0.64\,[\log_{10}(1+z)]^2 \\ \psi(z) =& -2.11\,[\log_{10}(1+z)]^2 + 3.36\log_{10}(1+z) \end{aligned} \tag{A.7}$$

With this, we have settled a redshift evolution of the mass-metallicity relation from the one used in the local universe by combining (A.3), (A.6) and (A.7).

Then, combining the latter with equation (A.2), considering that $\alpha_{\rm CO}$ is less steep at higher redshifts, varying as given by eq. (A.1) but using the depletion time given by eq. (A.5) and taking into account that if $\alpha_{\rm CO}$ is less than 3.2, the Milky-Way value, we impose $\alpha_{\rm CO}$=$\alpha_{{\rm CO},MW}$ and obtain the redshift evolution of the $\alpha_{\rm CO}$-M_* correlation:

$$\log(\alpha_{\rm CO}(z)) = 0.42 + 2\times\phi(z) + 2\times\log_{10}\left[1 + 10^{0.64\psi(z)-1}\left(\frac{3\times10^{10}}{M_*}\right)^{0.64}\right] - 0.42\times\log_{10}(1+z) \tag{A.8}$$

In figure A.2 we present the evolution of the mass-metallicity relation with redshift. We remark our considerations made by equations (A.6) and (A.7) in the local MZR agree with the redshift dependence found by Maiolino et al. (2008) and the recent work of Genzel et al. (2014), specially at lower stellar masses. On the other hand, at high stellar masses the metallicity does not decrease as strong.

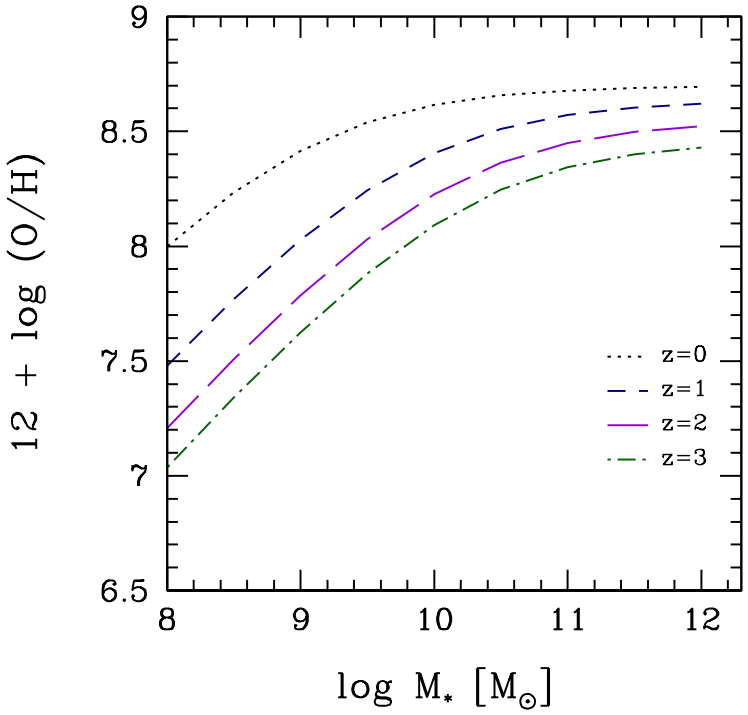

Figure A.2: Redshift evolution of the mass-metallicity relation. It is clear that the deacreasing in metallicity is larger at lower stellar masses, which is a result found by Maiolino et al. (2008) and the recent work of Genzel et al. (2014). In contrast, at high stellar masses, the decreasing in metallicity is not as strong.

In figure A.3, we show the redshift dependence of the $\alpha_{\rm CO}$-M_* correlation as given by eq. (A.8). The $\alpha_{\rm CO}$ varies smoothly in $1 \leq z \leq 3$ at all stellar masses, but a strong redshift dependence is clear in $0 \leq z \leq 1$ at low stellar masses. The dotted line is the Milky-Way value, this is $\alpha_{\rm CO}$= 3.2.

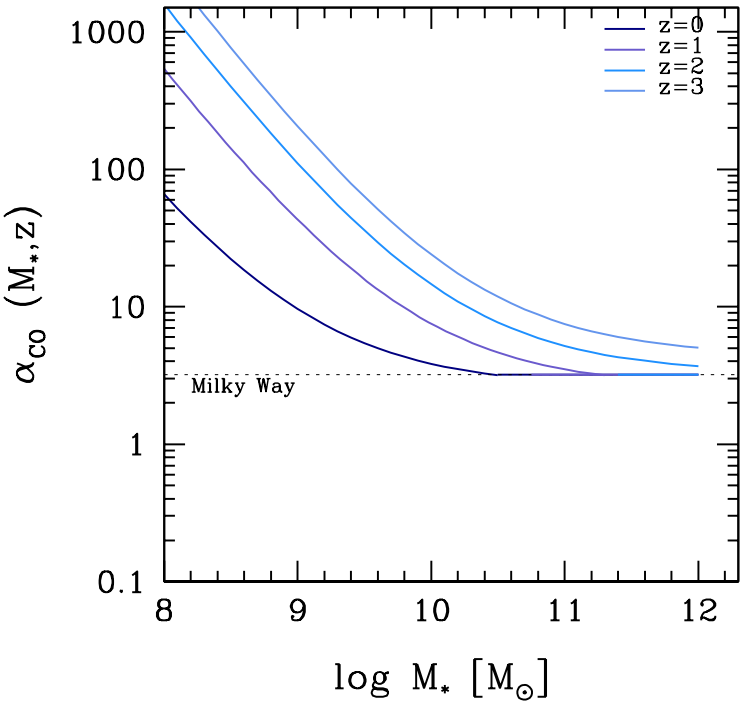

Figure A.3: Redshift evolution of the $\alpha_{\rm CO}$ conversion factor. The dotted line indicates $\alpha_{\rm CO}$= $\alpha_{{\rm CO},MW} = 3.2$.

Appendix B

The Kaplan-Meier estimator

The survival analysis methods can be used to statistically handle data containing along with the detections also non detections, for which upper limits are reported. Survival analysis is an extensive field of statistics that involves modeling of uncensored data (non detections with upper limits reported) taking into account the information of the censored data (detections; see e.g. Feigelson and Nelson, 1985; Isobe et al., 1986; Feigelson and Babu, 2012). A powerful non-parametric statistical estimator is the Kaplan-Meier product limit method (Kaplan and Meier, 1958).

For a random sample of n left-censored data (upper-limits) $\left\{x^L_{(j)}\right\}$ from a single population with distribution function $F^L(t)$, the Kaplan-Meier product limit method, is a non-parametric maximun-likelihood-type estimator of $F^L(t)$ which is stated in terms of the survival function $S(t)$ for right censored data (lower-limits) $\left\{x^R_{(j)}\right\}$ and it is given by:

$$F^L(t) = S(t) \tag{B.1}$$

where,

$$\begin{aligned} S(t) = \prod_{j, x^R(j)<t} (1 - d_j/n_j)^{\delta^L_{(j)}} \text{, when } t > x^R_{(j)} \\ S(t) = 1, \text{ when } t \leq x^R_{(j)} \end{aligned} \tag{B.2}$$

Here $1 \leq j \leq n$ and $\delta^L_{(j)} = 1$ if the element $x^L_{(j)}$ is a detection, otherwise if $x^L_{(j)}$ is an upper limit $\delta^L_{(j)} = 0$. To compute $S(t)$ is necessary to transform the left censored data into right censored data for some constant $M =$ maximum of $x^L_{(j)}$ by:

$$x^R_{(j)} = M - x^L_{(j)} \tag{B.3}$$

Hence, the values of n_j and d_j are:

$$\begin{aligned} n_j &= \text{number of } x_k^R \geq x_{(j)}^R \text{ for } k = 1, .., j \\ d_j &= \text{number of } x_k^R = x_{(j)}^R \text{ for } k = 1, .., j \end{aligned} \tag{B.4}$$

With $F^L(t)$ determined, quantities as the mean, standard deviation and percentiles of the sample of interest can be obtained. There are software packages that perform this kind of calculations, such as ASURV (Astronomy SURVival analysis) developed by Takashi Isobe, Michael LaValley and Eric Feigelson in 1992. The ASURV code is implemented in the **stsdas** package (Space Telescope Science Science Data Analysis) in IRAF. The Kaplan-Meier estimator is in the *kmestimate* routine.

The method becomes very uncertain or not valid in that regards the estimate of the median and percentiles when the non-detections overcome 50% of the data (Lee and Wang, 2003). The mean is more robust and can be yet calculated for samples where the non-detections dominate, down to $\sim 20\%$ of detections.

The IRAF package provides the standard error of the mean, i.e., the standard deviation of the sample mean estimate, $SEM = s/\sqrt{n}$, 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. The sample standard deviation 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 $n - 1.5$ instead of n. In this case $s \to \sigma$. Therefore, a good approximation to the population standard deviation for a normal distribution, given SEM, is $\sigma = SEM\frac{n}{n-1.5}$. This is the expression we use to calculate the reported standard deviations in Figs. 2.1 and 2.2.

According to the tutorial of ASURV Feigelson and Nelson (see also 1985), "the Kaplan-Meier estimator works with any underlying distribution (e.g., Gaussian, power law, bimodal), but only if the censoring is "random." That is, the probability that the measurement of an object is censored can not depend on the value of the censored variable. At first glance, this may seem to be inapplicable to most astronomical problems: we detect the brighter objects in a sample, so the distribution of upper limits always depends on brightness. However, two factors often serve to randomize the censoring distribution.

First, the censored variable may not be correlated with the variable by which the sample was initially identified. Thus, infrared observations of a sample of radio bright objects will be randomly censored if the radio and infrared emission are unrelated. Second, astronomical objects in a sample usually lie at different distances, so that brighter objects are not always the most luminous." *Note that these factors apply also to the samples compiled and analyzed in this Thesis in Chapter 2.*

Bibliography

Andrews, B. H. and Martini, P.: 2013, *ApJ* **765**, 140

Avila-Reese, V.: 2006, *ArXiv Astrophysics e-prints*

Avila-Reese, V., Zavala, J., and Lacerna, I.: 2014, *MNRAS* **441**, 417

Baldry, I. K., Glazebrook, K., and Driver, S. P.: 2008, *MNRAS* **388**, 945

Barone, L. T., Heithausen, A., Hüttemeister, S., Fritz, T., and Klein, U.: 2000, *MNRAS* **317**, 649

Bauermeister, A., Blitz, L., Bolatto, A., Bureau, M., Leroy, A., Ostriker, E., Teuben, P., Wong, T., and Wright, M.: 2013, *ApJ* **768**, 132

Baugh, C. M.: 2006, *Reports on Progress in Physics* **69**, 3101

Behroozi, P. S., Conroy, C., and Wechsler, R. H.: 2010, *ApJ* **717**, 379

Behroozi, P. S., Wechsler, R. H., and Conroy, C.: 2013, *ApJ* **770**, 57

Bell, E. F., McIntosh, D. H., Katz, N., and Weinberg, M. D.: 2003, *ApJS* **149**, 289

Benson, A. J.: 2010, *Phys. Rep.* **495**, 33

Bernardi, M., Meert, A., Sheth, R. K., Vikram, V., Huertas-Company, M., Mei, S., and Shankar, F.: 2013, *MNRAS* **436**, 697

Bernardi, M., Shankar, F., Hyde, J. B., Mei, S., Marulli, F., and Sheth, R. K.: 2010, *MNRAS* **404**, 2087

Blanton, M. R., Lupton, R. H., Schlegel, D. J., Strauss, M. A., Brinkmann, J., Fukugita, M., and Loveday, J.: 2005, *ApJ* **631**, 208

Blyth, S.-L., van der Hulst, J. M., Verheijen, M. A. W., SWG Members, H., Catinella, B., Fraternali, F., Haynes, M. P., Hess, K. M., Koribalski, B. S., Lagos, C., Meyer, M., Obreschkow, D., Popping, A., Power, C., Verdes-Montenegro, L., and Zwaan, M.: 2015, *ArXiv e-prints*

Bolatto, A. D., Wolfire, M., and Leroy, A. K.: 2013, *ARA&A* **51**, 207

Boselli, A., Cortese, L., Boquien, M., Boissier, S., Catinella, B., Lagos, C., and Saintonge, A.: 2014, *ArXiv e-prints*

Boselli, A., Lequeux, J., and Gavazzi, G.: 2002, *A&A* **384**, 33

Bothwell, M. S., Wagg, J., Cicone, C., Maiolino, R., Møller, P., Aravena, M., De Breuck, C., Peng, Y., Espada, D., Hodge, J. A., Impellizzeri, C. M. V., Martín, S., Riechers, D., and Walter, F.: 2014, *MNRAS* **445**, 2599

Bruce, V. A., Dunlop, J. S., Cirasuolo, M., McLure, R. J., Targett, T. A., Bell, E. F., Croton, D. J., Dekel, A., Faber, S. M., Ferguson, H. C., Grogin, N. A., Kocevski, D. D., Koekemoer, A. M., Koo, D. C., Lai, K., Lotz, J. M., McGrath, E. J., Newman, J. A., and van der Wel, A.: 2012, *MNRAS* **427**, 1666

Buitrago, F., Trujillo, I., Conselice, C. J., and Häußler, B.: 2013, *MNRAS* **428**, 1460

Carilli, C. L.: 2014, *ArXiv e-prints*

Catinella, B., Schiminovich, D., Cortese, L., Fabello, S., Hummels, C. B., Moran, S. M., Lemonias, J. J., Cooper, A. P., Wu, R., Heckman, T. M., and Wang, J.: 2013, *MNRAS* **436**, 34

Catinella, B., Schiminovich, D., Kauffmann, G., Fabello, S., Hummels, C., Lemonias, J., Moran, S. M., Wu, R., Cooper, A., and Wang, J.: 2012, *A&A* **544**, A65

Chabrier, G.: 2003, *PASP* **115**, 763

Combes, F., García-Burillo, S., Braine, J., Schinnerer, E., Walter, F., and Colina, L.: 2013, *A&A* **550**, A41

Cortese, L., Catinella, B., Boissier, S., Boselli, A., and Heinis, S.: 2011, *MNRAS* **415**, 1797

Daddi, E., Bournaud, F., Walter, F., Dannerbauer, H., Carilli, C. L., Dickinson, M., Elbaz, D., Morrison, G. E., Riechers, D., Onodera, M., Salmi, F., Krips, M., and Stern, D.: 2010, *ApJ* **713**, 686

Deng, X.-F.: 2013, *Research in Astronomy and Astrophysics* **13**, 651

Feigelson, E. D. and Babu, G. J.: 2012, *Modern Statistical Methods for Astronomy: With R Applications*, Cambridge

Feigelson, E. D. and Nelson, P. I.: 1985, *ApJ* **293**, 192

Firmani, C., Avila-Reese, V., and Rodríguez-Puebla, A.: 2010, *MNRAS* **404**, 1100

Frenk, C. S. and White, S. D. M.: 2012, *Annalen der Physik* **524**, 507

Frieman, J. A.: 1998, *Phys. Rep.* **307**, 61

Fu, J., Kauffmann, G., Li, C., and Guo, Q.: 2012, *MNRAS* **424**, 2701

Garland, C. A., Pisano, D. J., Williams, J. P., Guzmán, R., and Castander, F. J.: 2004, *ApJ* **615**, 689

Garland, C. A., Williams, J. P., Pisano, D. J., Guzmán, R., Castander, F. J., and Brinkmann, J.: 2005, *ApJ* **624**, 714

Garnett, D. R.: 2002, *ApJ* **581**, 1019

Geach, J. E., Smail, I., Moran, S. M., MacArthur, L. A., Lagos, C. d. P., and Edge, A. C.: 2011, *ApJ* **730**, L19

Geha, M., Blanton, M. R., Masjedi, M., and West, A. A.: 2006, *ApJ* **653**, 240

Geha, M., Blanton, M. R., Yan, R., and Tinker, J. L.: 2012, *ApJ* **757**, 85

Genzel, R., Tacconi, L. J., Combes, F., Bolatto, A., Neri, R., Sternberg, A., Cooper, M. C., Bouché, N., Bournaud, F., Burkert, A., Comerford, J., Cox, P., Davis, M., Förster Schreiber, N. M., Garcia-Burillo, S., Gracia-Carpio, J., Lutz, D., Naab, T., Newman, S., Saintonge, A., Shapiro, K., Shapley, A., and Weiner, B.: 2012, *ApJ* **746**, 69

Genzel, R., Tacconi, L. J., Lutz, D., Saintonge, A., Berta, S., Magnelli, B., Combes, F., García-Burillo, S., Neri, R., Bolatto, A., Contini, T., Lilly, S., Boissier, J., Boone, F.,

Bouché, N., Bournaud, F., Burkert, A., Carollo, M., Colina, L., Cooper, M. C., Cox, P., Feruglio, C., Förster Schreiber, N. M., Freundlich, J., Gracia-Carpio, J., Juneau, S., Kovac, K., Lippa, M., Naab, T., Salome, P., Renzini, A., Sternberg, A., Walter, F., Weiner, B., Weiss, A., and Wuyts, S.: 2014, *ArXiv e-prints*

Giovanelli, R., Haynes, M. P., Kent, B. R., Perillat, P., Saintonge, A., Brosch, N., Catinella, B., Hoffman, G. L., Stierwalt, S., Spekkens, K., Lerner, M. S., Masters, K. L., Momjian, E., Rosenberg, J. L., Springob, C. M., Boselli, A., Charmandaris, V., Darling, J. K., Davies, J., Garcia Lambas, D., Gavazzi, G., Giovanardi, C., Hardy, E., Hunt, L. K., Iovino, A., Karachentsev, I. D., Karachentseva, V. E., Koopmann, R. A., Marinoni, C., Minchin, R., Muller, E., Putman, M., Pantoja, C., Salzer, J. J., Scodeggio, M., Skillman, E., Solanes, J. M., Valotto, C., van Driel, W., and van Zee, L.: 2005, *AJ* **130**, 2598

Haynes, M. P. and Giovanelli, R.: 1984, *AJ* **89**, 758

Haynes, M. P., Giovanelli, R., Martin, A. M., Hess, K. M., Saintonge, A., Adams, E. A. K., Hallenbeck, G., Hoffman, G. L., Huang, S., Kent, B. R., Koopmann, R. A., Papastergis, E., Stierwalt, S., Balonek, T. J., Craig, D. W., Higdon, S. J. U., Kornreich, D. A., Miller, J. R., O'Donoghue, A. A., Olowin, R. P., Rosenberg, J. L., Spekkens, K., Troischt, P., and Wilcots, E. M.: 2011, *AJ* **142**, 170

He, Y. Q., Xia, X. Y., Hao, C. N., Jing, Y. P., Mao, S., and Li, C.: 2013, *ApJ* **773**, 37

Isobe, T., Feigelson, E. D., and Nelson, P. I.: 1986, *ApJ* **306**, 490

Jansen, R. A., Franx, M., Fabricant, D., and Caldwell, N.: 2000, *ApJS* **126**, 271

Jeltema, T. E., Binder, B., and Mulchaey, J. S.: 2008, *ApJ* **679**, 1162

Kannappan, S. J. and Gawiser, E.: 2007, *ApJ* **657**, L5

Kannappan, S. J., Guie, J. M., and Baker, A. J.: 2009, *AJ* **138**, 579

Kannappan, S. J., Stark, D. V., Eckert, K. D., Moffett, A. J., Wei, L. H., Pisano, D. J., Baker, A. J., Vogel, S. N., Fabricant, D. G., Laine, S., Norris, M. A., Jogee, S., Lepore, N., Hough, L. E., and Weinberg-Wolf, J.: 2013, *ApJ* **777**, 42

Kaplan, E. L. and Meier, P.: 1958, *Journal of the American Statistical Association* **53(282)**, pp. 457

Kauffmann, G., Heckman, T. M., White, S. D. M., Charlot, S., Tremonti, C., Peng, E. W., Seibert, M., Brinkmann, J., Nichol, R. C., SubbaRao, M., and York, D.: 2003, *MNRAS* **341**, 54

Kennicutt, Jr., R. C., Armus, L., Bendo, G., Calzetti, D., Dale, D. A., Draine, B. T., Engelbracht, C. W., Gordon, K. D., Grauer, A. D., Helou, G., Hollenbach, D. J., Jarrett, T. H., Kewley, L. J., Leitherer, C., Li, A., Malhotra, S., Regan, M. W., Rieke, G. H., Rieke, M. J., Roussel, H., Smith, J.-D. T., Thornley, M. D., and Walter, F.: 2003, *PASP* **115**, 928

Keres, D., Yun, M. S., and Young, J. S.: 2003, *ApJ* **582**, 659

Klypin, A. A., Trujillo-Gomez, S., and Primack, J.: 2011, *ApJ* **740**, 102

Kravtsov, A., Vikhlinin, A., and Meshscheryakov, A.: 2014, *ArXiv e-prints*

Kroupa, P.: 2001, *MNRAS* **322**, 231

Kroupa, P., Tout, C. A., and Gilmore, G.: 1993, *MNRAS* **262**, 545

Lagos, C. D. P., Baugh, C. M., Lacey, C. G., Benson, A. J., Kim, H.-S., and Power, C.: 2011, *MNRAS* **418**, 1649

Lah, P., Chengalur, J. N., Briggs, F. H., Colless, M., de Propris, R., Pracy, M. B., de Blok, W. J. G., Fujita, S. S., Ajiki, M., Shioya, Y., Nagao, T., Murayama, T., Taniguchi, Y., Yagi, M., and Okamura, S.: 2007, *MNRAS* **376**, 1357

Leauthaud, A., Tinker, J., Bundy, K., Behroozi, P. S., Massey, R., Rhodes, J., George, M. R., Kneib, J.-P., Benson, A., Wechsler, R. H., Busha, M. T., Capak, P., Cortês, M., Ilbert, O., Koekemoer, A. M., Le Fèvre, O., Lilly, S., McCracken, H. J., Salvato, M., Schrabback, T., Scoville, N., Smith, T., and Taylor, J. E.: 2012, *ApJ* **744**, 159

Lee, E. T. and Wang, J. W.: 2003, *Statistical Methods for Survival Data Analysis*, Wiley

Lemonias, J. J., Schiminovich, D., Catinella, B., Heckman, T. M., and Moran, S. M.: 2013, *ApJ* **776**, 74

Leroy, A., Bolatto, A. D., Simon, J. D., and Blitz, L.: 2005, *ApJ* **625**, 763

Leroy, A. K., Walter, F., Brinks, E., Bigiel, F., de Blok, W. J. G., Madore, B., and Thornley, M. D.: 2008, *AJ* **136**, 2782

Li, C., Kauffmann, G., Jing, Y. P., White, S. D. M., Börner, G., and Cheng, F. Z.: 2006, *MNRAS* **368**, 21

Li, C. and White, S. D. M.: 2009, *MNRAS* **398**, 2177

Maddox, N., Hess, K. M., Obreschkow, D., Jarvis, M. J., and Blyth, S.-L.: 2015, *MNRAS* **447**, 1610

Magdis, G. E., Daddi, E., Béthermin, M., Sargent, M., Elbaz, D., Pannella, M., Dickinson, M., Dannerbauer, H., da Cunha, E., Walter, F., Rigopoulou, D., Charmandaris, V., Hwang, H. S., and Kartaltepe, J.: 2012a, *ApJ* **760**, 6

Magdis, G. E., Daddi, E., Sargent, M., Elbaz, D., Gobat, R., Dannerbauer, H., Feruglio, C., Tan, Q., Rigopoulou, D., Charmandaris, V., Dickinson, M., Reddy, N., and Aussel, H.: 2012b, *ApJ* **758**, L9

Maiolino, R., Nagao, T., Grazian, A., Cocchia, F., Marconi, A., Mannucci, F., Cimatti, A., Pipino, A., Ballero, S., Calura, F., Chiappini, C., Fontana, A., Granato, G. L., Matteucci, F., Pastorini, G., Pentericci, L., Risaliti, G., Salvati, M., and Silva, L.: 2008, *A&A* **488**, 463

Maller, A. H., Berlind, A. A., Blanton, M. R., and Hogg, D. W.: 2009, *ApJ* **691**, 394

Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., and Brinkmann, J.: 2006, *MNRAS* **368**, 715

Martin, A. M., Papastergis, E., Giovanelli, R., Haynes, M. P., Springob, C. M., and Stierwalt, S.: 2010, *ApJ* **723**, 1359

Masters, K. L., Mosleh, M., Romer, A. K., Nichol, R. C., Bamford, S. P., Schawinski, K., Lintott, C. J., Andreescu, D., Campbell, H. C., Crowcroft, B., Doyle, I., Edmondson, E. M., Murray, P., Raddick, M. J., Slosar, A., Szalay, A. S., and Vandenberg, J.: 2010, *MNRAS* **405**, 783

McGaugh, S. S.: 2005, *ApJ* **632**, 859

Mendel, J. T., Simard, L., Palmer, M., Ellison, S. L., and Patton, D. R.: 2014, *ApJS* **210**, 3

Mo, H., van den Bosch, F. C., and White, S.: 2010, *Galaxy Formation and Evolution*

Mo, H. J. and White, S. D. M.: 1996, *MNRAS* **282**, 347

More, S., van den Bosch, F. C., Cacciato, M., Skibba, R., Mo, H. J., and Yang, X.: 2011, *MNRAS* **410**, 210

Moster, B. P., Naab, T., and White, S. D. M.: 2013, *MNRAS* **428**, 3121

Moster, B. P., Somerville, R. S., Maulbetsch, C., van den Bosch, F. C., Macciò, A. V., Naab, T., and Oser, L.: 2010, *ApJ* **710**, 903

Moustakas, J., Kennicutt, Jr., R. C., Tremonti, C. A., Dale, D. A., Smith, J.-D. T., and Calzetti, D.: 2010, *ApJS* **190**, 233

Moustakas, J., Zaritsky, D., Brown, M., Cool, R., Dey, A., Eisenstein, D. J., Gonzalez, A. H., Jannuzi, B., Jones, C., Kochanek, C. S., Murray, S. S., and Wild, V.: 2011, *ArXiv e-prints*

Muñoz-Mateos, J. C., Gil de Paz, A., Zamorano, J., Boissier, S., Dale, D. A., Pérez-González, P. G., Gallego, J., Madore, B. F., Bendo, G., Boselli, A., Buat, V., Calzetti, D., Moustakas, J., and Kennicutt, Jr., R. C.: 2009, *ApJ* **703**, 1569

Narayanan, D., Krumholz, M. R., Ostriker, E. C., and Hernquist, L.: 2012, *MNRAS* **421**, 3127

Noordermeer, E., van der Hulst, J. M., Sancisi, R., Swaters, R. A., and van Albada, T. S.: 2005, *A&A* **442**, 137

Papastergis, E., Cattaneo, A., Huang, S., Giovanelli, R., and Haynes, M. P.: 2012, *ApJ* **759**, 138

Papastergis, E., Giovanelli, R., Haynes, M. P., Rodríguez-Puebla, A., and Jones, M. G.: 2013, *ApJ* **776**, 43

Paturel, G., Petit, C., Prugniel, P., Theureau, G., Rousseau, J., Brouty, M., Dubois, P., and Cambresy, L.: 2003, *VizieR Online Data Catalog* **7237**, 0

Peng, Y.-j., Lilly, S. J., Renzini, A., and Carollo, M.: 2012, *ApJ* **757**, 4

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: 1992, *Numerical recipes in C. The art of scientific computing*

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: 1996, *Numerical Recipes in Fortran 90 (2Nd Ed.): The Art of Parallel Scientific Computing*, Cambridge University Press, New York, NY, USA

Rhee, J., Zwaan, M. A., Briggs, F. H., Chengalur, J. N., Lah, P., Oosterloo, T., and Hulst, T. v. d.: 2013, *MNRAS* **435**, 2693

Rodríguez-Puebla, A., Avila-Reese, V., and Drory, N.: 2013, *ApJ* **767**, 92

Rodríguez-Puebla, A., Avila-Reese, V., Yang, X., Foucaud, S., Drory, N., and Jing, Y. P.: 2015, *ApJ* **799**, 130

Rodríguez-Puebla, A., Drory, N., and Avila-Reese, V.: 2012, *ApJ* **756**, 2

Saintonge, A., Kauffmann, G., Kramer, C., Tacconi, L. J., Buchbender, C., Catinella, B., Fabello, S., Graciá-Carpio, J., Wang, J., Cortese, L., Fu, J., Genzel, R., Giovanelli, R., Guo, Q., Haynes, M. P., Heckman, T. M., Krumholz, M. R., Lemonias, J., Li, C., Moran, S., Rodriguez-Fernandez, N., Schiminovich, D., Schuster, K., and Sievers, A.: 2011, *MNRAS* **415**, 32

Salim, S., Rich, R. M., Charlot, S., Brinchmann, J., Johnson, B. D., Schiminovich, D., Seibert, M., Mallery, R., Heckman, T. M., Forster, K., Friedman, P. G., Martin, D. C., Morrissey, P., Neff, S. G., Small, T., Wyder, T. K., Bianchi, L., Donas, J., Lee, Y.-W., Madore, B. F., Milliard, B., Szalay, A. S., Welsh, B. Y., and Yi, S. K.: 2007, *ApJS* **173**, 267

Sánchez, S. F., Rosales-Ortega, F. F., Jungwiert, B., Iglesias-Páramo, J., Vílchez, J. M., Marino, R. A., Walcher, C. J., Husemann, B., Mast, D., Monreal-Ibero, A., Cid Fernandes, R., Pérez, E., González Delgado, R., García-Benito, R., Galbany, L., van de Ven, G., Jahnke, K., Flores, H., Bland-Hawthorn, J., López-Sánchez, A. R., Stanishev, V., Miralles-Caballero, D., Díaz, A. I., Sánchez-Blazquez, P., Mollá, M., Gallazzi, A., Papaderos, P., Gomes, J. M., Gruel, N., Pérez, I., Ruiz-Lara, T., Florido, E., de Lorenzo-Cáceres, A., Mendez-Abreu, J., Kehrig, C., Roth, M. M., Ziegler, B., Alves, J., Wisotzki, L., Kupko, D., Quirrenbach, A., Bomans, D., and Califa Collaboration: 2013, *A&A* **554**, A58

Sargent, M. T., Daddi, E., Béthermin, M., Aussel, H., Magdis, G., Hwang, H. S., Juneau, S., Elbaz, D., and da Cunha, E.: 2013, *ArXiv e-prints*

Schawinski, K., Lintott, C., Thomas, D., Sarzi, M., Andreescu, D., Bamford, S. P., Kaviraj, S., Khochfar, S., Land, K., Murray, P., Nichol, R. C., Raddick, M. J., Slosar, A., Szalay, A., Vandenberg, J., and Yi, S. K.: 2009, *MNRAS* **396**, 818

Schruba, A., Leroy, A. K., Walter, F., Bigiel, F., Brinks, E., de Blok, W. J. G., Kramer, C., Rosolowsky, E., Sandstrom, K., Schuster, K., Usero, A., Weiss, A., and Wiesemeyer, H.: 2012, *AJ* **143**, 138

Serra, P., Oosterloo, T., Morganti, R., Alatalo, K., Blitz, L., Bois, M., Bournaud, F., Bureau, M., Cappellari, M., Crocker, A. F., Davies, R. L., Davis, T. A., de Zeeuw, P. T., Duc, P.-A., Emsellem, E., Khochfar, S., Krajnović, D., Kuntschner, H., Lablanche, P.-Y., McDermid, R. M., Naab, T., Sarzi, M., Scott, N., Trager, S. C., Weijmans, A.-M., and Young, L. M.: 2012, *MNRAS* **422**, 1835

Sharma, P., McCourt, M., Parrish, I. J., and Quataert, E.: 2012, *MNRAS* **427**, 1219

Shimasaku, K., Fukugita, M., Doi, M., Hamabe, M., Ichikawa, T., Okamura, S., Sekiguchi, M., Yasuda, N., Brinkmann, J., Csabai, I., Ichikawa, S.-I., Ivezić, Z., Kunszt, P. Z., Schneider, D. P., Szokoly, G. P., Watanabe, M., and York, D. G.: 2001, *AJ* **122**, 1238

Stark, D. V., Kannappan, S. J., Wei, L. H., Baker, A. J., Leroy, A. K., Eckert, K. D., and Vogel, S. N.: 2013, *ApJ* **769**, 82

Stewart, K. R., Bullock, J. S., Wechsler, R. H., and Maller, A. H.: 2009, *ApJ* **702**, 307

Swaters, R. A. and Balcells, M.: 2002, *A&A* **390**, 863

Tacconi, L. J., Neri, R., Genzel, R., Combes, F., Bolatto, A., Cooper, M. C., Wuyts, S., Bournaud, F., Burkert, A., Comerford, J., Cox, P., Davis, M., Förster Schreiber, N. M., García-Burillo, S., Gracia-Carpio, J., Lutz, D., Naab, T., Newman, S., Omont, A., Saintonge, A., Shapiro Griffin, K., Shapley, A., Sternberg, A., and Weiner, B.: 2013, *ApJ* **768**, 74

Taylor, C. L., Kobulnicky, H. A., and Skillman, E. D.: 1998, *AJ* **116**, 2746

Tinker, J., Kravtsov, A. V., Klypin, A., Abazajian, K., Warren, M., Yepes, G., Gottlöber, S., and Holz, D. E.: 2008, *ApJ* **688**, 709

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., Klypin, A., Warren, M. S., Yepes, G., and Gottlöber, S.: 2010, *ApJ* **724**, 878

Walter, F., Brinks, E., de Blok, W. J. G., Bigiel, F., Kennicutt, Jr., R. C., Thornley, M. D., and Leroy, A.: 2008, *AJ* **136**, 2563

Wei, L. H., Kannappan, S. J., Vogel, S. N., and Baker, A. J.: 2010, *ApJ* **708**, 841

Wolfire, M. G., Hollenbach, D., and McKee, C. F.: 2010, *ApJ* **716**, 1191

Woo, J., Dekel, A., Faber, S. M., Noeske, K., Koo, D. C., Gerke, B. F., Cooper, M. C., Salim, S., Dutton, A. A., Newman, J., Weiner, B. J., Bundy, K., Willmer, C. N. A., Davis, M., and Yan, R.: 2013, *MNRAS* **428**, 3306

Wyder, T. K., Martin, D. C., Schiminovich, D., Seibert, M., Budavári, T., Treyer, M. A., Barlow, T. A., Forster, K., Friedman, P. G., Morrissey, P., Neff, S. G., Small, T., Bianchi, L., Donas, J., Heckman, T. M., Lee, Y.-W., Madore, B. F., Milliard, B., Rich, R. M., Szalay, A. S., Welsh, B. Y., and Yi, S. K.: 2007, *ApJS* **173**, 293

Yang, X., Mo, H. J., van den Bosch, F. C., Pasquali, A., Li, C., and Barden, M.: 2007, *ApJ* **671**, 153

Young, L. M., Bureau, M., Davis, T. A., Combes, F., McDermid, R. M., Alatalo, K., Blitz, L., Bois, M., Bournaud, F., Cappellari, M., Davies, R. L., de Zeeuw, P. T., Emsellem, E., Khochfar, S., Krajnović, D., Kuntschner, H., Lablanche, P.-Y., Morganti, R., Naab, T., Oosterloo, T., Sarzi, M., Scott, N., Serra, P., and Weijmans, A.-M.: 2011, *MNRAS* **414**, 940

Zhang, W., Li, C., Kauffmann, G., Zou, H., Catinella, B., Shen, S., Guo, Q., and Chang, R.: 2009, *MNRAS* **397**, 1243

Zibetti, S., Charlot, S., and Rix, H.-W.: 2009, *MNRAS* **400**, 1181

Zwaan, M. A., Meyer, M. J., Staveley-Smith, L., and Webster, R. L.: 2005, *MNRAS* **359**, L30