

The cosmic abundance of cold gas in the local Universe

Thomas J. Fletcher¹, Amélie Saintonge^{1*}, Paula S. Soares^{1,2} and Andrew Pontzen¹

¹ Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

² School of Physics and Astronomy, Queen Mary University of London, Mile End Road, London E1 4NS, UK

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We determine the cosmic abundance of molecular hydrogen (H_2) in the local Universe from the xCOLD GASS survey. To constrain the H_2 mass function at low masses and correct for the effect of the lower stellar mass limit of $10^9 M_\odot$ in the xCOLD GASS survey, we use an empirical approach based on an observed scaling relation between star formation rate and gas mass. We also constrain the HI and HI+ H_2 mass functions using the xGASS survey, and compare them to the HI mass function from the ALFALFA survey. We find the cosmic abundance of molecular gas in the local Universe to be $\Omega_{\text{H}_2} = (5.34 \pm 0.47) \times 10^{-5} h^{-1}$. Molecular gas accounts for $19.6 \pm 3.9\%$ of the total abundance of cold gas, $\Omega_{\text{HI}+\text{H}_2} = (4.66 \pm 0.70) \times 10^{-4} h_{70}^{-1}$. Galaxies with stellar masses in excess of $10^9 M_\odot$ account for 89% of the molecular gas in the local Universe, while in comparison such galaxies only contain 73% of the cold atomic gas as traced by the HI 21cm line. The xCOLD GASS CO, molecular gas and cold gas mass functions and Ω_{H_2} measurements provide constraints for models of galaxy evolution and help to anchor blind molecular gas surveys attempting to determine the abundance of molecular gas at high redshifts.

Key words: galaxies: luminosity function, – galaxies: ISM – ISM: molecules – cosmology: cosmological parameters

1 INTRODUCTION

Cold, dense molecular gas is the fuel for star formation in galaxies. The total molecular gas content of a galaxy and its surface density are the main drivers of its star formation and thus evolution (Kennicutt 1998). The interplay between gas inflows and outflows, star formation and galaxy evolution is one of the remaining open issues in astrophysics. Large surveys in the optical, ultraviolet and infrared have clearly established that the overall star formation rate (SFR) of the Universe was significantly higher in the past (e.g. Lilly et al. 1996; Schiminovich et al. 2005; Le Floch et al. 2005; Madau & Dickinson 2014). The star formation rate per unit volume was at its highest 8-10 Gyr ago, and has since declined by an order of magnitude to the present day. Extensive work is now going into studying the onset of star formation in the Universe, and tracking the evolution of the cosmic star formation budget during the first 3 Gyr (e.g. Reddy & Steidel 2009; Bouwens et al. 2012, 2015; Coe et al. 2013; Ellis et al. 2013; Finkelstein et al. 2013, 2015; Schenker et al. 2013; Oesch et al. 2013).

Given that star formation depends on the availability of cold gas, a natural explanation for this observed variation of SFR density with time is that it tracks changes in the amount of cold gas in galaxies that can partici-

pate in the star formation process. An alternative explanation would be that the gas contents of galaxies remain roughly unchanged over time, but that the star formation efficiency varies strongly. Observations of molecular gas in galaxies from the local Universe to $z \sim 3$ support the first picture; the cold gas contents of galaxies was significantly higher at the peak of cosmic star formation history (Daddi et al. 2010; Tacconi et al. 2010, 2013; Genzel et al. 2015; Schinnerer et al. 2016; Scoville et al. 2016, 2017; Tacconi et al. 2018; Freundlich et al. 2019; Aravena et al. 2019) and traces the redshift evolution of the star formation rate, modulo a weak evolution of star formation efficiency (Saintonge et al. 2013). These results are also reproduced by simple analytical “equilibrium” models where the SFR is regulated by the mass of the gas reservoir, which is depleted by outflows and replenished by inflowing gas (White & Frenk 1991; Bouché et al. 2010; Davé et al. 2011, 2012; Dekel et al. 2013; Lilly et al. 2013).

The implication is that understanding the cosmic star formation history relies on understanding how the cold gas reservoirs of galaxies evolve, both in terms of their total mass, and in the balance between the different phases (cold, warm and hot). In this study, we focus our attention on the cold atomic and molecular gas phases, given their direct link with star formation. In particular, we use the state-of-the-art molecular gas survey, xCOLD GASS (Saintonge et al.

* E-mail: a.saintonge@ucl.ac.uk

2017), to accurately calculate the mass function and cosmic abundance of molecular hydrogen.

In the local Universe, the interstellar medium of galaxies is dominated by the cold atomic phase; the atomic-to-molecular mass ratio ($M_{\text{HI}}/M_{\text{H}_2}$) is on average a factor of $\sim 3 - 4$ in massive galaxies and increases to ~ 10 in lower mass galaxies ($M_* = 10^9 M_\odot$) (Catinella et al. 2018). Despite this dominance of the atomic phase, it is nonetheless crucial to accurately determine Ω_{H_2} : (1) Star formation is triggered in cold, dense molecular clouds and therefore Ω_{H_2} tells us about the abundance of star-forming gas, unlike Ω_{HI} which traces large extended reservoirs of atomic gas that are one step removed from the star formation process. For example, galaxies with HI excesses are found to have very low star formation rates (Geréb et al. 2018). (2) There are indications that the $\Omega_{\text{H}_2}/\Omega_{\text{HI}}$ ratio may be changing with redshift, with Ω_{H_2} rising quickly with redshift (Decarli et al. 2016, 2019; Riechers et al. 2019) whilst Ω_{HI} rises more gradually (Zafar et al. 2013; Rhee et al. 2018). This picture is also supported by simulations and semi-analytic models (SAMs) (Obreschkow & Rawlings 2009b; Power et al. 2010; Lagos et al. 2011; Popping et al. 2015), making the calibration of the local value critical to anchor other studies.

The cosmic abundance of cold atomic gas, Ω_{HI} , and the HI mass function (HiMF) have been measured locally with high accuracy through large blind surveys (e.g. Rosenberg & Schneider 2002; Zwaan et al. 2005), sometimes exploring the specific environments of groups and clusters (Kilborn et al. 2009; Kovač et al. 2009). The state-of-the-art HI mass function has been produced by the ALFALFA survey; with a final catalog of ~ 31500 HI-detected galaxies with $z < 0.06$ (Haynes et al. 2018), it has the combination of depth and surface area that allows for the accurate determination of both the low-mass and high-mass ends of the HiMF (Martin et al. 2010; Jones et al. 2018).

A large flux-limited blind survey such as ALFALFA is ideal to build a robust and representative gas mass function. The unavailability of a dataset with all these characteristics probing the molecular gas contents of galaxies explains why there have been comparatively fewer measurements of the H_2 mass function (H2MF) and Ω_{H_2} . The molecular gas mass of galaxies is most commonly measured from the luminosity of emission lines of the CO molecule, assuming a specific function for the conversion to total molecular gas mass (the so-called CO-to- H_2 conversion function, α_{CO}). Consequently, the H_2 mass function is most often derived from targeted surveys for CO.

For example, Keres et al. (2003) (hereafter K+03) used the FCRAO (Five College Radio Astronomy Observatory) Extragalactic CO survey (300 observations including 236 CO detections) (Young et al. 1995) to construct a CO luminosity function (COLF) and calculate the value of Ω_{H_2} often used as the reference for the local Universe. From the full FCRAO sample, K+03 selected galaxies with $S_{60} > 5.24$ Jy for inclusion in their COLF. This relatively high flux cut-off reduced the sample down to 200 galaxies, with a bias towards infrared-bright and/or nearby galaxies. The consequence of this is an over-representation of starbursting and merging galaxies in the sample, while in fact these rare objects contribute $< 15\%$ of the star formation budget of the Universe (Rodighiero et al. 2011; Sargent et al. 2012).

To circumvent such selection biases in the determination

Table 1. Details of the high and low mass xCOLD GASS samples. In total 532 galaxies were observed for the xCOLD GASS survey.

Catalogue	Low Mass	High Mass
Stellar Mass	$9.0 < M_* < 10.0$	$10.0 < M_* < 11.5$
Redshift	$0.01 < z < 0.02$	$0.025 < z < 0.05$
Total galaxies	166	366
CO Detections	117	216
CO Non-Detections	49	150

of the H2MF, a solution is to use a blind, flux-limited sample. With facilities such as NOEMA, ALMA and the JVLA, relatively small (but deep) fields can be blindly searched for CO emission, making this a viable option only for high redshift studies (e.g. Walter et al. 2014, 2016; Pavesi et al. 2018; Decarli et al. 2019). Measuring the H2MF through a flux-limited survey is also possible if one chooses to infer the molecular gas mass of galaxies through their far infrared dust emission rather than CO emission lines (e.g. Berta et al. 2013; Vallini et al. 2016).

In this paper, we make use of a volume-limited CO survey, xCOLD GASS (Saintonge et al. 2017), to accurately determine the mass function and cosmic abundance of molecular gas at $z \sim 0$. The idea of using a local volume-limited survey to determine the H2MF has previously been explored by Andreani et al. (2018) for the case of the Herschel Reference Survey. The method we adopt here takes advantage on the strengths of the xCOLD GASS sample selection and observing strategy. First, the sample is stellar-mass selected, and therefore representative of the entire galaxy population with $M_* > 10^9 M_\odot$. With over 500 galaxies, it also has the required statistics to robustly determine the H2MF above the survey's completeness limit. With H_2 masses derived from CO(1-0), it provides an independent measure from the dust-based studies, where inferred gas masses are sensitive to the dust emissivity, the dust-to-gas abundance ratio and contamination along the line of sight.

Throughout this work we assume a cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.3$, $\Omega_k = 0$ and $\Omega_\Lambda = 0.7$.

2 DATA AND METHODS

2.1 Sample and Measurements

xCOLD GASS is a CO(1-0) survey undertaken with the IRAM-30m telescope, targeting 532 mass-selected SDSS galaxies with $M_* > 10^9 M_\odot$. Details about the survey, observations and data products are in Saintonge et al. (2011) and Saintonge et al. (2017), with basic properties of the sample summarised in Table 1.

The survey meets key requirements for building a mass/luminosity function. First, the xCOLD GASS sample is large and homogeneous, as all the measurements were carried out with the same instrument and a precise observing strategy; all galaxies were observed until the CO(1-0) line was detected, or until a sensitivity to a gas fraction M_{H_2}/M_* of $\sim 2\%$ was reached (Saintonge et al. 2011). This means that even for non-detections, we are able to place stringent upper limits on the molecular gas mass which can be included in our derivation of the H2MF.

Second, the sample is representative of the $z \sim 0$ galaxy

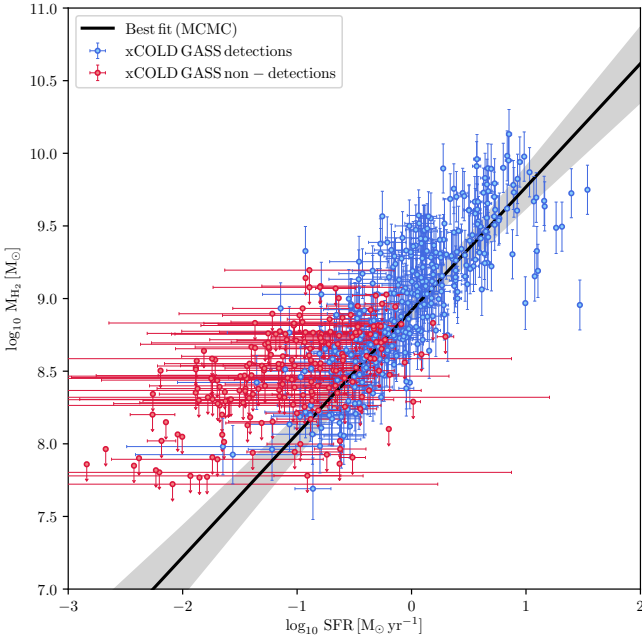


Figure 1. Linear trend between SFR and M_{H_2} for xCOLD GASS, with detections as blue circles and non-detections as red circles. The black line is the line of best fit, and the shading represents the uncertainty around this fit.

population with stellar masses larger than $10^9 M_\odot$. The 532 galaxies were selected randomly out of the larger parent sample of all SDSS galaxies in the mass and redshift range of the survey (see Tab. 1). Therefore unlike in the case of K+03, our sample is not biased towards particularly gas-rich and star-forming galaxies. While the xCOLD GASS observations are still targeted, unlike in an ideal scenario of a very deep, flux-limited blind survey, the sample selection and observing strategy allow us to probe a large range of M_{H_2}/M_* values and correct for incompleteness issues. In addition, when selecting the sample, we only required it to have a flat distribution in stellar mass, allowing us to study with similar statistics the galaxy population over more than two orders of magnitude in M_* . This has the advantage of allowing us to constrain the high mass end of the H2MF with better statistics, and we can account for the resulting selection bias using the stellar mass function of the SDSS parent sample. Further details on the survey design are can be found in (Saintonge et al. 2011).

For each of the 532 xCOLD GASS galaxies we calculate a total molecular gas mass, M_{H_2} . For the 333 galaxies where the CO(1-0) line was detected with $S/N > 3$, we calculate M_{H_2} via the CO line luminosity, $L'_{\text{CO}(1-0)}$ and the CO-to- H_2 conversion function of Accurso et al. (2017),

$$\log \alpha_{\text{CO}} (\pm 0.165 \text{ dex}) = 14.752 - 1.623[12 + \log(\text{O}/\text{H})] + 0.062 \log \Delta(\text{MS}), \quad (1)$$

where $12 + \log(\text{O}/\text{H})$ is the metallicity calculated using the prescription outlined in Pettini & Pagel (2004), and $\Delta(\text{MS})$ is the offset from the main sequence. The error on the predicted value of α_{CO} is ± 0.165 dex. The non-

detections are treated in three different ways. They are assigned (1) $M_{\text{H}_2} = 0$, which is the most conservative choice, (2) the M_{H_2} value corresponding to the 5σ upper limit, the most optimistic option, and (3) the M_{H_2} value expected for the galaxy given its SFR. The third option is motivated by the very well known close relation between molecular gas and star formation, which we parameterise here using the xCOLD GASS sample as shown in Fig. 1. We fit the relation between $\log_{10} M_{\text{H}_2}$ and $\log_{10} \text{SFR}$ as a first order polynomial, with Gaussian intrinsic scatter around this line with standard deviation Λ in the $\log_{10} M_{\text{H}_2}$ direction. This is accomplished using the Markov Chain Monte Carlo (MCMC) affine-invariant ensemble sampler *emcee* (Goodman & Weare 2010; Foreman-Mackey et al. 2013). Unless otherwise stated, throughout this work *emcee* is run using 100 walkers and 1000 steps. The chain for all the parameters is inspected to check for convergence, a burn-in of 500 steps is discarded and the final 100 steps are used. Once the chain has converged, the best-fit relation is:

$$\log_{10} M_{\text{H}_2} = (0.85 \pm 0.03) \log_{10} \text{SFR} + (8.92 \pm 0.02), \quad (2)$$

with an intrinsic scatter of $\Lambda = 0.26 \pm 0.02$. In the case of upper limits, the probability distribution function was integrated from $-\infty$ to the value of the upper limit (see the Appendix of Sawicki (2012) who use the same method). We also use this relation combined with individual metallicity-dependent conversion factors (α_{CO}) to calculate the expected $L'_{\text{CO}(1-0)}$ for each galaxy for case (3). Using the detections, we have verified that this relation holds for galaxies with low M_{H_2} , SFR and gas fractions. We note that for 87% of non-detections the estimated H_2 mass calculated using equation 2 results in a M_{H_2}/M_* gas fraction of $< 2\%$, in agreement with the observing strategy outlined in Saintonge et al. (2011).

2.2 Building the Mass Function

Since xCOLD GASS is not a flux-limited survey, nor a purely volume-limited sample, we must use an appropriate method to build any luminosity or mass function. First, we assign to each xCOLD GASS galaxy a weight to correct for the flat stellar mass distribution of the sample (see Sec.2.2 and Fig. 3 in Saintonge et al. 2017). Taking these weights into account, and treating non-detections in one of the three ways outlined in Section 2.1, the measured values of M_{H_2} or M_{HI} are binned and counted to produce a mass ‘‘histogram’’ that is representative of the entire galaxy population with $M_* > 10^9 M_\odot$.

To then turn this histogram into a mass function (with the correct units of $\text{Mpc}^{-3} \text{dex}^{-1}$), we compute the effective volume of the survey using the stellar mass function of Baldry et al. (2012). The stellar mass function is integrated over the mass range of xCOLD GASS (i.e. $M_* > 10^9 M_\odot$), giving a total inverse volume for galaxies in this range. Finally, to get the total effective volume probed by xCOLD GASS, the total number of galaxies in the survey (532) is divided by this integral. Normalising the mass ‘‘histogram’’ by this effective volume produces a well calibrated mass function.

The uncertainty on each point of the luminosity or mass function is calculated using bootstrapping with replacement.

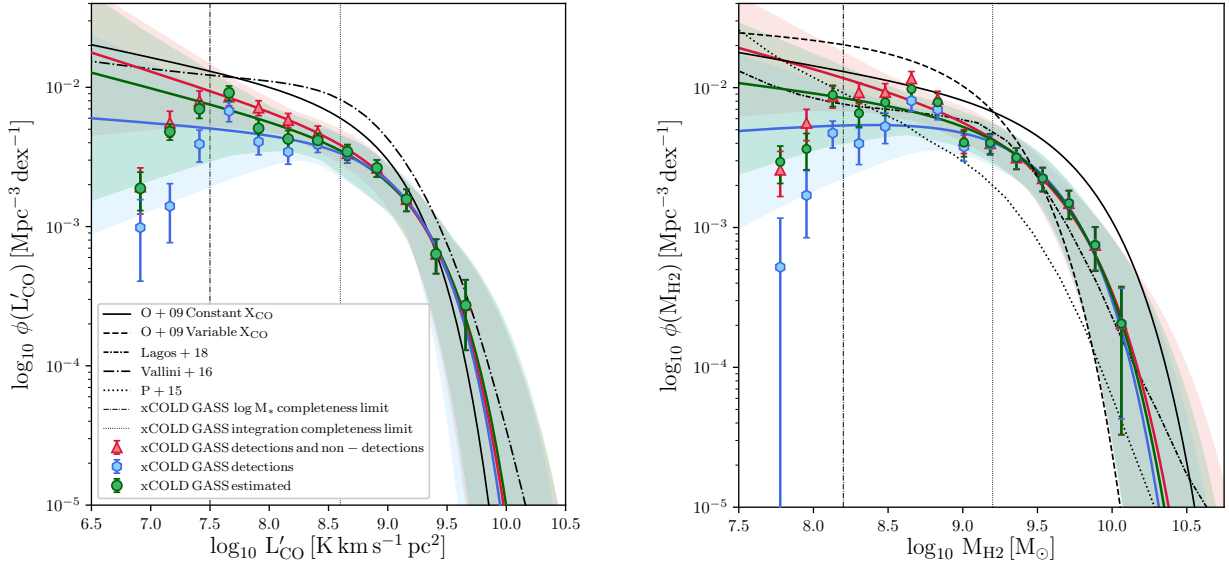


Figure 2. The xCOLD GASS CO luminosity function (COLF, *left*) and H₂ mass function (H2MF, *right*). The CO(1-0) non-detections are treated in three different ways: they are set to zero (blue hexagons), they are set to the 5σ upper limit (red triangles), or an estimated value of $L'_{\text{CO}(1-0)}$ and M_{H_2} based on the empirical relation shown in Fig. 1 (green circles). The latter is considered our best and default version. The H₂ mass for each galaxy was calculated from the xCOLD GASS L'_{CO} measurements using the conversion factor (α_{CO}) in Equation 1 from Accurso et al. (2017). The solid coloured lines and shaded regions show the best fit Schechter function and associated 1σ error respectively, for each case. For comparison the O+09 H₂ Schechter functions using both constant and variable conversion factors are shown (solid black and dashed black lines respectively). A COLF empirically derived from the infrared LF is shown (black dash-dotted line) (Vallini et al. 2016), as well as H₂ mass functions predicted from a semi-empirical model P+15 (black dotted line) and from a semi-analytic model (Lagos et al. 2018a) (black densely dash-dotted line). The stellar mass and integration completeness limits are shown as vertical dot-dashed and dotted lines, respectively.

For a given sample of N galaxies, we resample N times, allowing for repetitions in the new sample. We also resample each selected galaxy using its measurement error to include the effect of observational errors. Each time we resample the original sample with replacement we also calculate the number density of galaxies in each mass bin and record these values. This procedure is repeated 1000 times and we use the 1σ standard deviation of the resamples in each mass bin as the error. Once the mass function is built and errors are determined, the best-fitting Schechter function (Schechter 1976) parameters are inferred using *emcee*. The median of the posterior from the converged *emcee* chain is taken as the best fit for each parameter, and the 1σ error is calculated from the 16th and 84th percentiles. We use the median rather than the mean as the mean can be outside of the 1σ confidence interval if there is large skewness (Hogg & Foreman-Mackey 2018). It should be noted that the median of the posterior probability density function for each parameter, if taken together as a ‘best-fit’ result (M_{H_2} , ϕ^* , α), does not necessarily result in a best-fit. This is because the median for each parameter is a marginalisation of the posterior probability density function in one dimension. In cases where there is curvature in the parameter space, as is true with Schechter function fits, taking the medians for each parameter collectively can result in a fit that falls in a low-probability region of the posterior probability density function (see Hogg & Foreman-Mackey (2018) for a more detailed discussion).

For each luminosity/mass function we build, we deter-

mine two completeness limits, which are the consequence of the xCOLD GASS sample selection and observing strategy. First, there is a ‘stellar mass completeness’ limit, because xCOLD GASS only targeted galaxies with $M_* > 10^9 M_\odot$. Given the positive correlation between M_* and M_{H_2} (and L'_{CO}), this completeness limit is set at the highest value of M_{H_2} found in xCOLD GASS for galaxies near the stellar mass cutoff of $> 10^9 M_\odot$. Galaxies with lower stellar masses, which are not present in the xCOLD GASS sample, contribute insignificantly to the luminosity/mass function above that limit. Secondly, the ‘integration completeness’ limit marks the value of M_{H_2} (and L'_{CO}) above which all the galaxies are guaranteed to have a 5σ detection of the CO(1-0) line, based on the observing strategy of the survey. This is set to $M_{\text{H}_2} = 0.02M_*$ for $M_* = 10^{11} M_\odot$, the mass of the most massive galaxies in the xCOLD GASS sample.

3 RESULTS

3.1 The CO(1-0) Luminosity Function and Molecular Gas Mass Function at $z \sim 0$

The CO luminosity function from xCOLD GASS, computed using the method described above, is presented in Fig. 2 (left panel). The three different treatments of the non-detections are used, and the two completeness limits are represented as vertical lines. Above the integration completeness limit, the three versions of the COLF are in agreement, as expected. Below this limit, they begin to clearly diverge. The effect

of the stellar mass selection can be clearly seen in the incompleteness below the mass completeness limit. The “blue” COLF, which has the non-detections all set to $L'_{\text{CO}(1-0)} = 0$, represents a lower limit on the luminosity function, while the “red” version is a slight overestimation since it puts all the non-detections at their 5σ upper limits. We adopt as our best COLF the “green” version from Fig. 2, where we use empirical predictions of $L'_{\text{CO}(1-0)}$ for the non-detections.

The best-fitting¹ Schechter function above the mass completeness limit is determined with *emcee* and has parameters: $\log_{10} L'_{\text{CO}}/\text{K km s}^{-1} \text{pc}^2 = 9.29^{+0.14}_{-0.12}$, $\alpha = -1.25^{+0.10}_{-0.10}$ and $\phi^*/10^{-3}\text{Mpc}^{-3} = 1.26^{+0.45}_{-0.38}$.

For comparison, Fig. 2 shows the COLF from the FCRAO survey derived from the H2MF presented in Obreschkow & Rawlings (2009a) (hereafter O+09) which uses the same sample as K+03, as well as a COLF derived from the infrared luminosity function (Vallini et al. 2016). These COLFs only appear to agree for the highest luminosities. The xCOLD GASS result suggests a much lower abundance of intermediate luminosity objects ($7.5 < \log_{10} L'_{\text{CO}} < 9$) yet a steeper faint end slope compared to the results from O+09 and Vallini et al. (2016).

Similarly, the H_2 mass function from xCOLD GASS is shown in Fig. 2 (right-hand panel). If we were using a constant value of α_{CO} for all galaxies, then the COLF and H2MF would be identical in shape; in particular they would have the same Schechter slope and normalisation parameters α and ϕ respectively. Indeed, in Figure 2, we can see that the mass function using a constant conversion factor from O+09 has the same shape as the COLF from O+09. However, we must take into account that galaxies with sub-solar metallicities are under-luminous in CO, which we accomplish by using the variable α_{CO} prescription of Accurso et al. (2017) to convert from $L'_{\text{CO}(1-0)}$ to M_{H_2} . We also compare our mass function to a H2MF derived using a semi-analytic model (Lagos et al. 2018a) and one derived using a semi-empirical method from Popping et al. (2015) (hereafter P+15) which combines a sub-halo abundance matching model with a model to indirectly estimate the HI and H_2 masses of each galaxy. The $z = 0$ mass function data for this model is available for download². The P+15 mass function does not agree with our results or the O+09 results around both the knee and the high-mass end. However, this serves to illustrate the range of H_2 mass functions in the literature.

Using this method and again using the “green” mass function where non-detections are estimated, the best-fitting Schechter function above the mass completeness limit is determined with *emcee* and has parameters: $\log_{10} M_{\text{H}_2}/M_{\odot} = 9.59^{+0.11}_{-0.10}$, $\alpha = -1.18^{+0.11}_{-0.11}$ and $\phi^*/10^{-3}\text{Mpc}^{-3} = 2.34^{+0.72}_{-0.61}$.

Comparing our best-fitting Schechter function to the mass function determined using a constant conversion factor from O+09 which uses the FCRAO sample (Young et al. 1995), we see that the high-mass end tails off sooner, despite the fact the two bright-ends of the COLFs were in good agreement. This is because the variable conver-

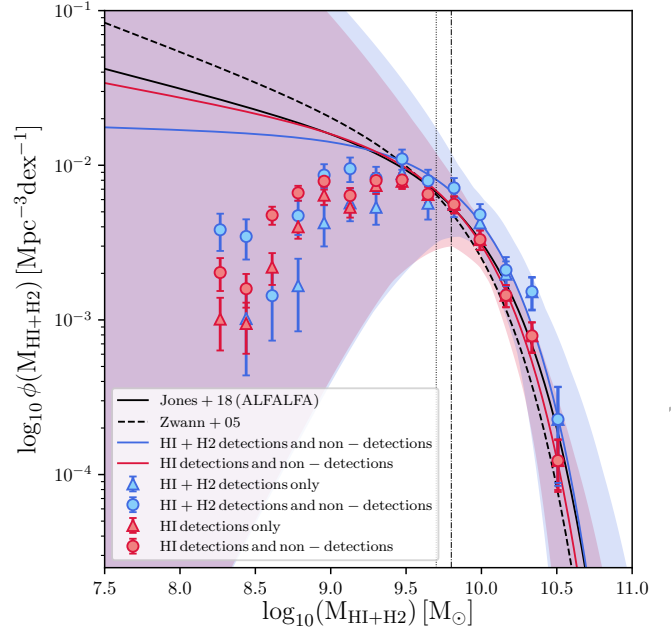


Figure 3. The HI mass function (red) and HI + H_2 mass function (blue) for the overlapping xGASS and xCOLD GASS sample. All detections and non-detections are shown as filled circles, while detections only are shown as triangles. The HI mass function derived by Zwaan et al. (2005) and the ALFALFA HiMF (Jones et al. 2018) are shown for comparison.

sion factor prescription decreases with increasing metallicity, giving ($\alpha_{\text{CO}} < 4.36 M_{\odot}(\text{K km s}^{-1})^{-1}$) for galaxies with $12 + \log(\text{O}/\text{H}) \gtrsim 8.6$ (Accurso et al. 2017), which are typically more massive due to the mass-metallicity relation. Conversely, all galaxies in O+09 were given a much higher conversion factor of ($\alpha_{\text{CO}} = 6.5 M_{\odot} (\text{K km s}^{-1})^{-1}$).

3.2 The Total Cold Gas (HI + H_2) Mass Function

Having both HI and H_2 mass measurements the xCOLD GASS galaxies allows us to determine the mass function of cold gas (HI + H_2) in the local Universe. Additionally, by building an HI mass function from the xCOLD GASS sample and comparing it to the HiMF from the ALFALFA survey (Jones et al. 2018), we can validate our methodology (as presented in Section 2.2).

The HI data are retrieved from the xGASS catalog (Catinella et al. 2018), which contains Arecibo HI observations for 1179 SDSS-selected galaxies in the mass range $9.0 < \log_{10} M_*/M_{\odot} < 11.5$. Of these galaxies, 477 are also in the xCOLD GASS sample. This is the sub-sample of xCOLD GASS that we use in this section to derive the total HI + H_2 mass function.

In Figure 3 we show both the HiMF and total HI + H_2 mass functions derived from xCOLD GASS. In both cases, we give the mass functions for two different treatments of the non-detections: (1) setting them to zero for the HiMF and excluding galaxies where at least one of the HI or H_2 measurements is a non-detection for the HI + H_2 mass function, or (2) assigning them the gas mass equivalent to the 5σ upper limit. The completeness limits arising from the stel-

¹ The converged chains for this and all other inferred Schechter function parameters can be found at https://github.com/tomjf/Omega_H2_chains.

² http://www.mpa.de/homes/popping/data/gas_sham.data.tar.gz

lar mass cutoff and the depth of the observations are shown as before. We fit both mass functions above the stellar mass completion limit with a Schechter function using *emcee*. The best-fit parameters are summarized in Table 2.

The agreement between the xCOLD GASS and ALFALFA HiMF of Jones et al. (2018) above our completeness limit is excellent, despite our HiMF being built from a far smaller sample that is mass-limited rather than flux-limited. This further confirms that our methodology for building mass functions from a volume-limited sample (Sec. 2.2) is valid, and that our H2MF is accurate in both shape and normalisation. Additionally, it indicates that it is possible to use the ALFALFA HI estimate for Ω_{HI} to calculate the total $\Omega_{\text{HI}+\text{H}_2}$ cold gas abundance in the local Universe, since it is compatible with the xCOLD GASS survey above the integration completion limit.

3.3 The Abundance of Cold Gas at $z \sim 0$

The mass functions shown in Fig. 3 make it very clear that most of the cold gas in the nearby Universe is in atomic rather than molecular form; we quantify this here by calculating the overall abundance of gas in the different phases. First, in Fig. 4, we look at the distribution of ρ_{H_2} , which is the product of $\phi(M_{\text{H}_2})$ and M_{H_2} . The uncertainties and best fit relations from Fig. 2 are translated over. The total value of ρ_{H_2} (and thus Ω_{H_2} , after dividing by the critical density) can be found by integrating the area underneath the curve in Fig. 4. For our best treatment of the non-detections, using the empirical scaling relation in Equation 2 to predict the true H_2 masses of non-detections, we find a value of $\Omega_{\text{H}_2} = (5.34 \pm 0.47) \times 10^{-5} h^{-1}$ (the values for the other two cases are given in Table 2). For comparison, summing the bin values derived from xCOLD GASS, which are affected by mass incompleteness, results in a value for Ω_{H_2} which is approximately 95% of the value quoted above.

In Figure 5, we compare our new value of Ω_{H_2} with a range of results for the literature. Our results are within the (large) uncertainties of the early FCRAO work (O+09), but lower by a factor of up to 2. This difference comes down to an overestimation in the previous study of the number density of galaxies with $M_{\text{H}_2} \lesssim 10^9 M_{\odot}$, and different assumptions for the CO-to- H_2 conversion factor.

Similarly, we calculate the abundance of HI from our mass function in Fig. 3, and find $\Omega_{\text{HI}} = 3.51^{+3.41}_{-1.06} \times 10^{-4}$. For the ALFALFA survey, Jones et al. 2018 find $\Omega_{\text{HI}} = (3.9 \pm 0.1 \pm 0.6) \times 10^{-4}$ when correcting for HI self-absorption, where the first and second quoted errors are the random and systematic errors respectively. When HI self-absorption is not accounted for $\Omega_{\text{HI-uncorr}} = (3.5 \pm 0.1 \pm 0.5) \times 10^{-4}$ (Jones et al. 2018). Our result, albeit with significantly larger uncertainties, is consistent with this value as well as with the HIPASS result (Zwaan et al. 2005). Given the higher precision and accuracy of the ALFALFA result, we adopt the Jones et al. 2018 value of Ω_{HI} and combine with our determination of Ω_{H_2} to arrive at our best estimate of the total cold gas abundance in the local Universe: $\Omega_{\text{HI}+\text{H}_2} = (4.66 \pm 0.70) \times 10^{-4} h_{70}^{-1}$.

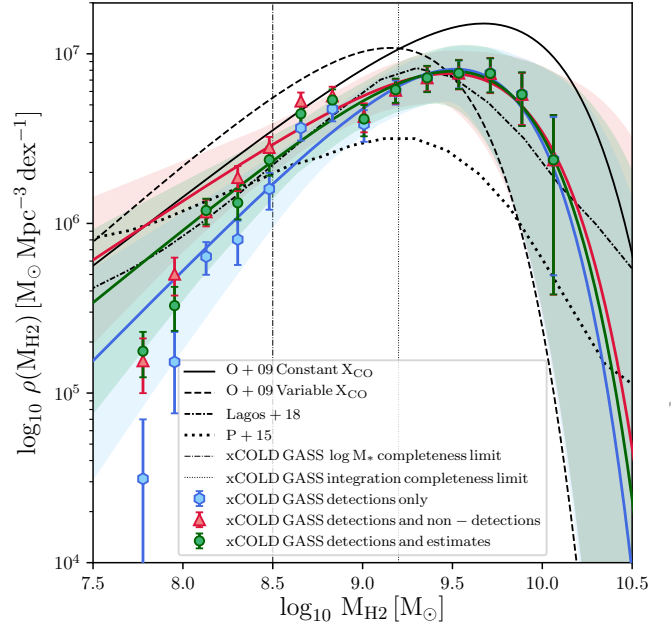


Figure 4. The distribution of ρ_{H_2} with M_{H_2} for the xCOLD GASS sample. Blue hexagons represent the xCOLD GASS sample of CO(1-0) detections only, red triangles represent xCOLD GASS data with both detections and non-detections, and green circles use the xCOLD GASS detections and estimated masses for the non-detections. The solid coloured lines and shaded regions show the best fit Schechter function and associated 1σ error respectively, for each case. The shading around the lines represents the error on the fit. For comparison the O+09 Schechter functions with constant and variable conversion factors are also shown (solid black and dashed black lines respectively) as well as a results derived from a semi-analytical model (Lagos et al. 2018b), black, densely dash-dotted line) and a semi-empirical model (P+15, black dotted line).

3.4 The balance between atomic and molecular gas

The results above show that overall the molecular-to-atomic gas ratio in the local Universe is $19.6\% \pm 3.9\%$. An analysis of the xCOLD GASS data has shown how $M_{\text{H}_2}/M_{\text{HI}}$ is a function of stellar mass, with the most massive galaxies ($10^{11} M_{\odot}$) having on average a $\sim 30\%$ molecular-to-atomic ratio, with the value dropping to $\sim 10\%$ for galaxies with stellar masses of $10^9 M_{\odot}$ (Catinella et al. 2018). To add to this picture, we show in Figure 6 the cumulative distribution of the fractional abundance of both HI and H_2 , normalised by the total cosmic abundance of each gas phase (using our best estimate from xCOLD GASS for Ω_{H_2} and the ALFALFA value for Ω_{HI}). The galaxy population with $M_* > 10^9 M_{\odot}$ accounts for $\sim 89\%$ of the molecular gas in the local Universe, but only $\sim 73\%$ of the atomic gas as traced by HI. The two lines

³ This value for $\Omega_{\text{HI}+\text{H}_2}$ is determined using only data above the mass completeness limit for the combined GASS and xCOLD GASS surveys and thus has large uncertainties. Instead we recommend using $\Omega_{\text{HI}+\text{H}_2} = (4.66 \pm 0.70) \times 10^{-4} h_{70}^{-1}$ which is the sum of our best estimate for Ω_{H_2} and Ω_{HI} from Jones et al. (2018).

Table 2. All Schechter function and density parameter results.

Schechter Function	Knee Units	Knee Value	Normalization $\phi^*/10^{-3}\text{Mpc}^{-3}$	Slope α	Density Parameter
L'_{CO} (DET)	$\log_{10} L'_{\text{CO}}/\text{K km s}^{-1} \text{pc}^2$	$9.18^{+0.12}_{-0.11}$	$1.73^{+0.51}_{-0.45}$	$-1.07^{+0.11}_{-0.11}$	–
L'_{CO} (EST)	$\log_{10} L'_{\text{CO}}/\text{K km s}^{-1} \text{pc}^2$	$9.29^{+0.14}_{-0.12}$	$1.26^{+0.45}_{-0.38}$	$-1.25^{+0.10}_{-0.10}$	–
L'_{CO} (ALL)	$\log_{10} L'_{\text{CO}}/\text{K km s}^{-1} \text{pc}^2$	$9.26^{+0.12}_{-0.11}$	$1.37^{+0.45}_{-0.37}$	$-1.27^{+0.09}_{-0.08}$	–
M_{H_2} (DET)	$\log_{10} M_{\text{H}_2}/M_{\odot}$	$9.49^{+0.10}_{-0.09}$	$3.09^{+0.74}_{-0.67}$	$-0.93^{+0.12}_{-0.12}$	$\Omega_{\text{H}_2} = (4.76 \pm 0.43) \times 10^{-5} h^{-1}$
M_{H_2} (EST)	$\log_{10} M_{\text{H}_2}/M_{\odot}$	$9.59^{+0.11}_{-0.10}$	$2.34^{+0.72}_{-0.61}$	$-1.18^{+0.11}_{-0.11}$	$\Omega_{\text{H}_2} = (5.34 \pm 0.47) \times 10^{-5} h^{-1}$
M_{H_2} (ALL)	$\log_{10} M_{\text{H}_2}/M_{\odot}$	$9.64^{+0.12}_{-0.11}$	$1.99^{+0.69}_{-0.56}$	$-1.30^{+0.10}_{-0.10}$	$\Omega_{\text{H}_2} = (5.82 \pm 0.49) \times 10^{-5} h^{-1}$
M_{HI}	$\log_{10} M_{\text{HI}}/M_{\odot}$	$9.86^{+0.12}_{-0.11}$	$5.53^{+1.50}_{-1.75}$	$-1.18^{+0.56}_{-0.47}$	$\Omega_{\text{HI}} = 3.51^{+3.41}_{-1.06} \times 10^{-4}$
$M_{\text{HI}+\text{H}_2}$	$\log_{10} M_{\text{HI}+\text{H}_2}/M_{\odot}$	$9.89^{+0.14}_{-0.16}$	$6.60^{+1.88}_{-2.37}$	$-1.03^{+0.69}_{-0.54}$	$\Omega_{\text{HI}+\text{H}_2} = 4.11^{+3.30}_{-1.14} \times 10^{-4}$

The labels DET, ALL and EST correspond to mass or luminosity functions built using the three different treatments of non-detections, respectively cases (1), (2) and (3) outlined in Section 2.1.

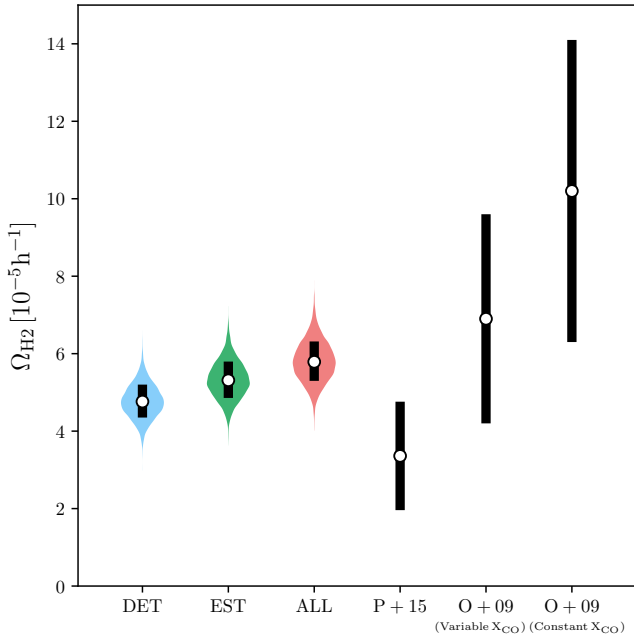


Figure 5. Comparison of Ω_{H_2} calculated in this work and previous estimates from the literature. The coloured violin plots show the probability density of Ω_{H_2} derived from the the best-fitting Schechter functions. The vertical black lines encompass the 16th to the 84th percentiles, whilst the white dot shows the position of the 50th percentile. The estimates using detections, estimated masses and detections and non-detections are shown as blue, green and red respectively. Three estimates from the literature are shown. Two are derived in O+09 using both a constant and variable conversion factor where the third is from P+15 and is derived using semi-analytic models.

showing the cumulative distribution of the fractional abundance of both gas phases show different behaviour, with the HI curve growing constantly whilst the H₂ curve is much steeper than the HI curve for $M_* > 10^{10} M_{\odot}$ before becoming shallower than the HI curve at lower masses. This highlights how the ISM of low mass galaxies is atomic gas-dominated. Indeed, while galaxies with $M_* > 10^{10} M_{\odot}$ account for $\sim 64\%$ of all the molecular gas at $z \sim 0$, we need to push down to $M_* \sim 10^{9.3} M_{\odot}$ to reach the same level of completeness for the abundance of atomic gas.

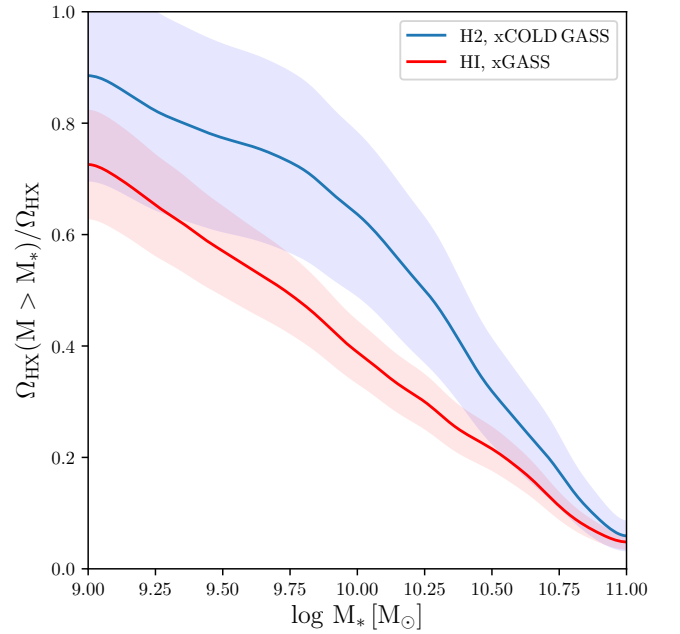


Figure 6. The cumulative fraction of Ω_{H_2} and Ω_{HI} coming from galaxies above a given stellar mass. The blue line shows the result for Ω_{H_2} derived from the xCOLD GASS survey and the red line shows the trend for Ω_{HI} using the xGASS survey. The 1σ error, derived by propagating the bootstrap errors for each mass function at each step in stellar mass, is shown by the shaded regions.

The balance between atomic and molecular gas not only varies with stellar mass but may also vary with redshift. Beyond $z = 0$ there is growing observational evidence from both CO observations (Decarli et al. 2016, 2019; Riechers et al. 2019) and dust-mass tracers (Scoville et al. 2017) that Ω_{H_2} evolves with redshift. Rising to ~ 6.5 times the present day value at $z \approx 1 - 2$ and then either flattening or declining at higher redshifts, mirroring the star-formation history of the Universe. Meanwhile observational evidence (Zafar et al. 2013; Rhee et al. 2018) and simulations (Davé et al. 2017) suggest a more gradual evolution of Ω_{HI} with redshift. Therefore, there is a picture building that at late times $\Omega_{\text{H}_2}/\Omega_{\text{HI}}$ declines, which is also supported by SAMs (Lagos et al. 2011). Molecular gas is depleted faster than it is being re-

plenished by HI reservoirs, which could then explain the sudden downturn in SFR since $z \approx 2$. Our determinations for Ω_{H_2} and $\Omega_{\text{H}_2}/\Omega_{\text{HI}}$ in this paper can help to provide an anchor at $z = 0$ for future studies into the evolution of Ω_{H_2} and $\Omega_{\text{H}_2}/\Omega_{\text{HI}}$.

4 SUMMARY

We present new determinations of the cosmic abundance of molecular gas (Ω_{H_2}) and cold gas ($\Omega_{\text{HI}+\text{H}_2}$) using the xCOLD GASS survey. The homogeneity, depth of the observations and the stellar-mass selected representative sample from the xCOLD GASS survey allow us to determine the H_2 mass function well below the knee, more accurately than before. In addition we account for the fact that galaxies with sub-solar metallicities are under-luminous in CO by applying a metallicity-dependent conversion factor. Non-detections were treated in three different ways, set to zero, estimated based upon their SFR using a scaling relation between SFR and observed gas mass or set at their 5σ upper limits. Best-fitting Schechter functions parameters were then inferred using *emcee*, these parameters are then used to calculate the density parameters for H_2 and cold gas. Using this method our best estimate for the cosmic abundance of H_2 is $\Omega_{\text{H}_2} = (5.34 \pm 0.47) \times 10^{-5} h^{-1}$. For cold gas we find $\Omega_{\text{HI}+\text{H}_2} = (4.66 \pm 0.70) \times 10^{-4} h_{70}^{-1}$. Therefore, we find the ratio of molecular-to-atomic gas as $\Omega_{\text{H}_2}/\Omega_{\text{HI}} = 19.6\% \pm 3.9\%$.

Our results provide more stringent constraints on the functional form of the H_2 mass function at $z = 0$. This is important for cosmological simulations and SAMs which already accurately reproduce the galaxy stellar mass function and its redshift evolution (see [Somerville & Davé 2015](#), for a review). However, to do so they must incorporate heuristic models for feedback, which are not well understood and can vary considerably between simulations. As molecular gas studies push to higher redshifts, increasingly, simulations and SAMs must also predict the gas content of galaxies, which is the fuel of star formation thought to drive galaxy evolution ([Davé et al. 2017](#); [Popping et al. 2019](#)). We can therefore begin to disseminate between the different heuristic models for feedback, which would not be possible by comparing to the observed galaxy stellar mass function alone. Our best-fitting result: $\log M_{\text{H}_2}/M_{\odot} = 9.59^{+0.11}_{-0.10}$, $\alpha = -1.18^{+0.11}_{-0.11}$ and $\phi^*/10^{-3}\text{Mpc}^{-3} = 2.34^{+0.72}_{-0.61}$, for the H_2 mass function and our inferred parameters for the COLF can be used as comparison against predictions made by simulations and SAMs. And this $z = 0$ result can be used alongside Ω_{H_2} determinations derived using H_2 mass functions at higher redshifts from CO observations ([Decarli et al. 2019](#)) or using infrared luminosity as a proxy ([Vallini et al. 2016](#)). Furthermore, we have also provided constraints on the total cold gas mass function.

In addition, our results provide determinations of Ω_{H_2} and $\Omega_{\text{H}_2}/\Omega_{\text{HI}}$ at $z = 0$, giving more precise observational anchors for blind ALMA and NOEMA studies investigating the evolution of H_2 at higher redshifts and how this compares to the evolution of HI and SFR. We have also shown the cumulative fractional abundance of atomic and molecular gas, normalised by their respective total cosmic abundances. The two curves of growth for each gas phase show different behaviour, showing that high-mass galaxies tend to

be dominated by molecular gas whereas low mass galaxies tend to be dominated by HI.

ACKNOWLEDGEMENTS

We thank Barbara Catinella and Luca Cortese for useful discussion, as well as Dusan Keres and Livia Vallini for advice on the comparison with previous work. TF acknowledges support from STFC. AS and AP acknowledge support from the Royal Society. AP received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 818085 GMGalaxies.

DATA AVAILABILITY

The data used in this article are in the public domain: [Saintonge et al. \(2017\)](#) and [Catinella et al. \(2018\)](#). The data for these surveys can be found at <http://www.star.ucl.ac.uk/xCOLDGASS/data.html> and <https://xgass.icrar.org/data.html>.

REFERENCES

- Accurso G., et al., 2017, *MNRAS*, **470**, 4750
 Andreani P., Boselli A., Ciesla L., Vio R., Cortese L., Buat V., Miyamoto Y., 2018, *A&A*, **617**, A33
 Aravena M., et al., 2019, *ApJ*, **882**, 136
 Baldry I. K., et al., 2012, *MNRAS*, **421**, 621
 Berta S., et al., 2013, *A&A*, **555**, L8
 Bouché N., et al., 2010, *ApJ*, **718**, 1001
 Bouwens R. J., et al., 2012, *ApJ*, **754**, 83
 Bouwens R. J., et al., 2015, *ApJ*, **803**, 34
 Catinella B., et al., 2018, *MNRAS*, **476**, 875
 Coe D., et al., 2013, *ApJ*, **762**, 32
 Daddi E., et al., 2010, *ApJ*, **713**, 686
 Davé R., Finlator K., Oppenheimer B. D., 2011, *MNRAS*, **416**, 1354
 Davé R., Finlator K., Oppenheimer B. D., 2012, *MNRAS*, **421**, 98
 Davé R., Rafieferantsoa M. H., Thompson R. J., Hopkins P. F., 2017, *MNRAS*, **467**, 115
 Decarli R., et al., 2016, *ApJ*, **833**, 69
 Decarli R., et al., 2019, *ApJ*, **882**, 138
 Dekel A., Zolotov A., Tweed D., Cacciato M., Ceverino D., Primack J. R., 2013, *MNRAS*, **435**, 999
 Ellis R. S., et al., 2013, *ApJ*, **763**, L7
 Finkelstein S. L., et al., 2013, *Nature*, **502**, 524
 Finkelstein S. L., et al., 2015, *ApJ*, **810**, 71
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, **125**, 306
 Freundlich J., et al., 2019, *A&A*, **622**, A105
 Genzel R., et al., 2015, *ApJ*, **800**, 20
 Geréb K., Janowiecki S., Catinella B., Cortese L., Kilborn V., 2018, *MNRAS*, **476**, 896
 Goodman J., Weare J., 2010, *Communications in Applied Mathematics and Computational Science*, **5**, 65
 Haynes M. P., et al., 2018, *ApJ*, **861**, 49
 Hogg D. W., Foreman-Mackey D., 2018, *ApJS*, **236**, 11
 Jones M. G., Haynes M. P., Giovanelli R., Moorman C., 2018, *MNRAS*, **477**, 2
 Kennicutt Robert C. J., 1998, *ApJ*, **498**, 541
 Keres D., Yun M. S., Young J. S., 2003, *ApJ*, **582**, 659

- Kilborn V. A., Forbes D. A., Barnes D. G., Koribalski B. S., Brough S., Kern K., 2009, *MNRAS*, **400**, 1962
- Kovač K., Oosterloo T. A., van der Hulst J. M., 2009, *MNRAS*, **400**, 743
- Lagos C. D. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011, *MNRAS*, **418**, 1649
- Lagos C. d. P., Tobar R. J., Robotham A. S. G., Obreschkow D., Mitchell P. D., Power C., Elahi P. J., 2018a, *MNRAS*, **481**, 3573
- Lagos C. d. P., Tobar R. J., Robotham A. S. G., Obreschkow D., Mitchell P. D., Power C., Elahi P. J., 2018b, *MNRAS*, **481**, 3573
- Le Floch E., et al., 2005, *ApJ*, **632**, 169
- Lilly S. J., Le Fevre O., Hammer F., Crampton D., 1996, *ApJ*, **460**, L1
- Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, *ApJ*, **772**, 119
- Madau P., Dickinson M., 2014, *ARA&A*, **52**, 415
- Martin A. M., Papastergis E., Giovanelli R., Haynes M. P., Springob C. M., Stierwalt S., 2010, *The Astrophysical Journal*, **723**, 1359
- Obreschkow D., Rawlings S., 2009a, *MNRAS*, **394**, 1857
- Obreschkow D., Rawlings S., 2009b, *ApJ*, **696**, L129
- Oesch P. A., et al., 2013, *ApJ*, **773**, 75
- Pavesi R., et al., 2018, *ApJ*, **864**, 49
- Pettini M., Pagel B. E. J., 2004, *MNRAS*, **348**, L59
- Popping G., Behroozi P. S., Peebles M. S., 2015, *MNRAS*, **449**, 477
- Popping G., et al., 2019, *ApJ*, **882**, 137
- Power C., Baugh C. M., Lacey C. G., 2010, *MNRAS*, **406**, 43
- Reddy N. A., Steidel C. C., 2009, *ApJ*, **692**, 778
- Rhee J., Lah P., Briggs F. H., Chengalur J. N., Colless M., Willner S. P., Ashby M. L. N., Le Fèvre O., 2018, *MNRAS*, **473**, 1879
- Riechers D. A., et al., 2019, *ApJ*, **872**, 7
- Rodighiero G., et al., 2011, *ApJ*, **739**, L40+
- Rosenberg J. L., Schneider S. E., 2002, *ApJ*, **567**, 247
- Saintonge A., et al., 2011, *MNRAS*, **415**, 32
- Saintonge A., et al., 2013, *ApJ*, **778**, 2
- Saintonge A., et al., 2017, *ApJS*, **233**, 22
- Sargent M. T., Béthermin M., Daddi E., Elbaz D., 2012, *ApJ*, **747**, L31
- Sawicki M., 2012, *PASP*, **124**, 1208
- Schechter P., 1976, *ApJ*, **203**, 297
- Schenker M. A., et al., 2013, *ApJ*, **768**, 196
- Schiminovich D., et al., 2005, *ApJ*, **619**, L47
- Schinnerer E., et al., 2016, *ApJ*, **833**, 112
- Scoville N., et al., 2016, *ApJ*, **820**, 83
- Scoville N., et al., 2017, *ApJ*, **837**, 150
- Somerville R. S., Davé R., 2015, *ARA&A*, **53**, 51
- Tacconi L. J., et al., 2010, *Nature*, **463**, 781
- Tacconi L. J., et al., 2013, *ApJ*, **768**, 74
- Tacconi L. J., et al., 2018, *ApJ*, **853**, 179
- Vallini L., Gruppioni C., Pozzi F., Vignali C., Zamorani G., 2016, *MNRAS*, **456**, L40
- Walter F., et al., 2014, *ApJ*, **782**, 79
- Walter F., et al., 2016, *ApJ*, **833**, 67
- White S. D. M., Frenk C. S., 1991, *ApJ*, **379**, 52
- Young J. S., et al., 1995, *ApJS*, **98**, 219
- Zafar T., Péroux C., Popping A., Milliard B., Deharveng J. M., Frank S., 2013, *A&A*, **556**, A141
- Zwaan M. A., Meyer M. J., Staveley-Smith L., Webster R. L., 2005, *MNRAS*, **359**, L30