Molecular hydrogen in IllustrisTNG galaxies: carefully comparing signatures of environment with local CO and SFR data

Massachusetts Institute of Technology

Molecular hydrogen in IllustrisTNG galaxies: carefully comparing signatures of environment with local CO & SFR data

Adam R. H. Stevens,[1,2][*] Claudia del P. Lagos,[1,2] Luca Cortese,[1,2] Barbara Catinella,[1,2]
Benedikt Diemer,[3] Dylan Nelson,[4,5] Annalisa Pillepich,[6] Lars Hernquist,[7] Federico Marinacci[8]
and Mark Vogelsberger[9]

[1] *International Centre for Radio Astronomy Research, The University of Western Australia, Crawley, WA 6009, Australia*
[2] *Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)*
[3] *Department of Astronomy, University of Maryland, College Park, MD 20742, USA*
[4] *Max-Planck-Institut für Astrophysik, 85741 Garching, Bayern, Germany*
[5] *Institut für theoretische Astrophysik, Zentrum für Astronomie, Universität Heidelberg, 69120 Heidelberg, Baden-Württemburg, Germany*
[6] *Max-Planck-Institut für Astronomie, 69117 Heidelberg, Baden-Württemburg, Germany*
[7] *Institute for Theory and Computation, Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA*
[8] *Department of Physics & Astronomy, University of Bologna, 40129 Bologna, Italy*
[9] *Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

21 December 2020

ABSTRACT

We examine how the post-processed content of molecular hydrogen (H_2) in galaxies from the TNG100 cosmological, hydrodynamic simulation changes with environment at $z = 0$, assessing central/satellite status and host halo mass. We make close comparisons with the carbon monoxide (CO) emission survey xCOLD GASS where possible, having mock-observed TNG100 galaxies to match the survey's specifications. For a representative sample of host haloes across $10^{11} \lesssim M_{200c}/M_\odot < 10^{14.6}$, TNG100 predicts that satellites with $m_* \geq 10^9 \, M_\odot$ should have a median deficit in their H_2 fractions of ~ 0.6 dex relative to centrals of the same stellar mass. Once observational and group-finding uncertainties are accounted for, the signature of this deficit decreases to ~ 0.2 dex. Remarkably, we calculate a deficit in xCOLD GASS satellites' H_2 content relative to centrals of 0.2–0.3 dex, in line with our prediction. We further show that TNG100 and SDSS data exhibit continuous declines in the average star formation rates of galaxies at fixed stellar mass in denser environments, in quantitative agreement with each other. By tracking satellites from their moment of infall in TNG100, we directly show that atomic hydrogen (H I) is depleted at fractionally higher rates than H_2 on average. Supporting this picture, we find that the H_2/H I mass ratios of satellites are elevated relative to centrals in xCOLD GASS. We provide additional predictions for the effect of environment on H_2 — both absolute and relative to H I — that can be tested with spectral stacking in future CO surveys.

Key words: galaxies: evolution — galaxies: haloes — galaxies: interactions — galaxies: ISM — galaxies: structure — ISM: molecules

1 MOTIVATION AND BACKGROUND

Complementing pioneering works that date back nearly half a century (see reviews by Haynes, Giovanelli & Chincarini 1984; Blanton & Moustakas 2009), there has been a flurry of studies in the last decade in particular that highlight how galaxies' cold gas can be affected by their environment, be it in terms of their integrated content with statistical samples from observations (Chung et al. 2009; Cortese et al. 2011; Catinella et al. 2013; Boselli et al. 2014; Jaffé et al. 2016; Stark et al. 2016; Brown et al. 2017; Janowiecki et al. 2017) or simulations/models (Tecce et al. 2010; Marasco et al. 2016; Stevens & Brown 2017; Cora et al. 2018; Stevens et al. 2018,

2019a), or in resolved detail with small numbers (Moretti et al. 2018; Boselli et al. 2019; Ramatsoku et al. 2019; Jáchym et al. 2019; Yun et al. 2019). The vast majority of these works have focussed on the atomic hydrogen (H I) in galaxies — the bulk of a galaxy's cold gas — with a generalized consensus that local galaxy density anti-correlates with galaxy H I content. While some attention has also been given to molecular gas, studies of its relation with galaxy environment are less advanced. This is understandable, given that molecular emission is more challenging to measure at low redshift. While H I is readily studied through its ubiquitous radio emission at a wavelength of ~ 21 cm, H_2 (molecular hydrogen) lacks a permanent dipole, and is notoriously difficult to detect from direct electromagnetic emission or absorption. This means relying on emission of other molecules thought to trace H_2:

in particular, carbon monoxide (CO — see the review by Bolatto, Wolfire & Leroy 2013).

To understand how galaxy environment affects any galaxy property, it is vital to first distinguish between centrals and satellites. Central galaxies reside in the global minimum of a halo's potential well (i.e. at the halo's *centre*) and are the most massive galaxy in that halo. In practice, they are often identified observationally as the galaxy with the greatest stellar mass or luminosity in a linked group (various algorithms exist for doing the linking — e.g. Yang et al. 2007; Robotham et al. 2011; Tempel et al. 2016; Tinker 2020). Satellite galaxies make up the remaining, subordinate population of galaxies in haloes. They were once centrals themselves, but have since fallen into haloes greater than their own, per the hierarchical assembly of structure in the Universe under the ΛCDM cosmological model (c.f. White & Rees 1978). Centrals can have any number of associated satellites (including zero). A central and its environment evolve in tandem, while satellites suffer environmental effects that centrals are scarcely affected by. These effects include, but are not limited to, ram-pressure stripping, starvation/strangulation of fresh gas, and tidal stripping (see Gunn & Gott 1972; Larson, Tinsley & Caldwell 1980; Moore et al. 1999, respectively).

Environmental processes that strip the interstellar media of galaxies are expected to impact the atomic-gas content of those galaxies more than their molecular gas. This is because the outskirts of galaxies are more susceptible to tides and ram pressure, where the atomic-to-molecular ratio is almost always much higher (Leroy et al. 2008; Bigiel & Blitz 2012; Stevens & Brown 2017). Nevertheless, once stripping becomes strong enough, H_2 in satellites is expected to be lost as well.

While the number of works targeting variation in molecular-gas content with environment pale in comparison to those of atomic gas, there were efforts to study the effect of cluster environments on galaxies' molecular gas at low redshift as far back as ∼30 years ago (e.g. Kenney & Young 1989; Casoli et al. 1991). Early conclusions about whether these galaxies are H_2-poor or not was later brought into question, in part due to selection effects and uncertainty surrounding the CO-to-H_2 conversion factor (see Boselli et al. 2014, and references therein). Debate has since ensued as to whether dense environments typically produce H_2-poor galaxies. At $z = 0$, based on the fact that the most 'H I deficient' galaxies in the *Herschel* Reference Survey also exhibit some deficiency in H_2, Boselli et al. (2014) argue that cluster satellites are at least partially stripped of H_2. Evidence for direct, resolved stripping of CO has also been observed in a handful of galaxies (Lee & Chung 2018; Moretti et al. 2018; Cramer et al. 2020; Jáchym et al. 2019), while additional cases of molecular disturbances have been found in Virgo (Lee et al. 2017a). Broadly speaking, at intermediate and higher redshifts, reports are even more varied, with some suggestions that (proto)clusters favour H_2-poor galaxies (Jablonka et al. 2013; Castignani et al. 2018; Coogan et al. 2018), others that they favour H_2-rich galaxies (Noble et al. 2017; Hayashi et al. 2018), others that they favour H_2-normal galaxies (Dannerbauer et al. 2017; Lee et al. 2017b; Rudnick et al. 2017; Darvish et al. 2018; Wu et al. 2018), or a mass-dependent/phase-space-dependent combination thereof (Hayashi et al. 2017; Wang et al. 2018; Tadaki et al. 2019).

Part of the challenge of drawing empirical conclusions about the interplay between galaxy environment and H_2 con-

tent comes from data in the literature being heterogeneous and/or limited in sample size. To make robust theoretical predictions for an environment–H_2 relation is also non-trivial, as it ideally involves cosmological, hydrodynamic simulations of significant volume (box lengths of $\gtrsim 100$ Mpc), sufficient resolution (baryonic mass elements of $\lesssim 10^6\,M_\odot$), and an accounting of the multi-phase nature of galactic gas that differentiates between its ionized, atomic, and molecular components. In recent years especially, there has been significant development in the gas-phase decomposition of hydrodynamic simulations, leading to a string of model predictions for the atomic- and molecular-gas content of galaxies (Davé et al. 2013, 2020; Tomassetti et al. 2014, 2015; Lagos et al. 2015; Rafieferantsoa et al. 2015, 2019; Rahmati et al. 2015; Bahé et al. 2016; Marasco et al. 2016; Crain et al. 2017; Marinacci et al. 2017; Diemer et al. 2018, 2019; Rodríguez Montero et al. 2019; Stevens et al. 2019a,b; Schäbe et al. 2020). All these simulations offer laboratories of a sort that allow us to systematically study the causal impact between a wide range of galaxy properties. To that point, this paper is focussed on the relationship between the environment and H_2 content of galaxies, making use of the TNG100 simulation of the IllustrisTNG[1] suite. Where possible, we compare and contrast our simulation results with data from observational molecular gas studies, most notably the xCOLD GASS program (Saintonge et al. 2017), and otherwise make predictions that future surveys may be able to test.

This paper serves as a sequel to Stevens et al. (2019a, hereafter S19). Where S19 targeted the effect of environment on the H I properties of galaxies, here we focus equivalently on H_2, using the same simulation (TNG100) and an overlapping observational galaxy sample (xCOLD GASS). We describe both the simulation and data in Section 2, including our method for 'mock observing' the simulated galaxies consistently with the survey specifications. In Section 3, we compare the overall H_2 properties of galaxies from TNG100 and xCOLD GASS at $z = 0$ as a function of stellar mass, serving as a reference point for our main results, which are subsequently presented in Section 4. There, we slice the data by environment in several ways, not only presenting changes to H_2 with these slices, but also how changes to H I and star formation scale with them. We discuss several caveats in Section 5, and offer final remarks and conclusions in Section 6. Ancillary information regarding our methods and results is provided in Appendices A & B.

2 DATA AND METHODS

2.1 The TNG100 simulation

The focus of this paper is on the results of the TNG100 simulation (Pillepich et al. 2018b; Nelson et al. 2018, 2019a; Marinacci et al. 2018; Naiman et al. 2018; Springel et al. 2018), part of the IllustrisTNG magnetohydrodynamic, cosmological simulation suite. The simulation was run with the AREPO code (Springel 2010), and follows the standard ΛCDM cosmological model, with parameters based on the Planck Collaboration XIII (2016) results: $\Omega_m = 0.3089$, $\Omega_\Lambda = 0.6911$, $\Omega_b = 0.0486$, $h = 0.6774$, $\sigma_8 = 0.8159$, $n_s = 0.9667$. The TNG

---

[1] Illustris: The Next Generation — http://www.tng-project.org/

model includes a range of sub-grid prescriptions that are designed to accommodate relevant astrophysical processes in the formation and evolution of galaxies: for example, gas cooling, star formation, massive black-hole growth, and feedback from both stars and active galactic nuclei (Weinberger et al. 2017; Pillepich et al. 2018a). These are based on the earlier Illustris simulation's sub-grid models (see Vogelsberger et al. 2013, 2014a,b; Genel et al. 2014; Torrey et al. 2014). TNG100 employs a periodic box of length $75\,h^{-1} \simeq 110\,\mathrm{cMpc}$, containing 1820^3 dark-matter particles of mass $7.5 \times 10^6\,\mathrm{M_\odot}$, and 1820^3 initial gas cells of typical mass $\sim 1.4 \times 10^6\,\mathrm{M_\odot}$. The minimum gravitational softening scale for gas — achieved only in galaxy centres — is $0.19\,\mathrm{kpc}$.

Gravitationally bound structures in the simulations were found with the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009). These 'subhaloes' exist within haloes that are built with a friends-of-friends (FoF) particle group finder. The most massive subhalo in a FoF group is dubbed the 'central', and the remainder are 'satellites'. We follow the nomenclature of S19 when describing the integrated properties of galaxies in the simulation. That is, 'inherent' galaxy properties are defined as the subset of particles/cells in a subhalo that also lie within the spherical aperture given by the 'baryonic mass profile' criterion ('BaryMP' for short) of Stevens et al. (2014). 'Mock' properties depend on the dataset being compared to. In this paper, these relate to the xCOLD GASS survey and Sloan Digital Sky Survey (SDSS — York et al. 2000). We describe the mocking process for this in Section 2.4.

All gas cells in the simulation are post-processed in order to calculate their mass fractions in the form of atomic and molecular hydrogen. The method is described in S19, which builds from Lagos et al. (2015) and Diemer et al. (2018). In this paper, we follow the same three prescriptions used by S19, namely those based on the works of Gnedin & Kravtsov (2011, hereafter GK11), Krumholz (2013, hereafter K13), and Gnedin & Draine (2014, hereafter GD14). All three prescriptions rely principally on the local UV field strength (which we model for each galaxy in three dimensions — see Diemer et al. 2018 for details), local metallicity (a proxy for dust), and local surface density (translated from volumetric gas densities through the Jeans approximation) to calculate the H I/H₂ mass ratio of each gas cell.

For TNG100 results at $z = 0$, the same galaxy sample analysed in S19 is analysed in this paper. That is, only subhaloes with $m_* \geq 10^9\,\mathrm{M_\odot}$ and dark-matter fractions above 5% are included.

2.2 The xCOLD GASS survey

The xCOLD GASS[2] program comprises two observational surveys of the CO(1–0) line emission in galaxies with the IRAM[3] 30-m telescope. The surveys were designed to measure (or place upper limits on) the H₂ content of a representative sample of galaxies over $>2\,\mathrm{dex}$ in stellar mass. The original COLD GASS survey (Saintonge et al. 2011) targeted 350 galaxies with $m_* \geq 10^{10}\,\mathrm{M_\odot}$ in the redshift range

$0.025 \leq z \leq 0.05$. These galaxies were a subset of those observed in 21-cm H I emission in the GASS survey (Catinella et al. 2010). Each galaxy was observed sufficiently long to either detect the CO(1–0) line or place an upper limit of $\sim 0.015\,m_*$ on $m_{\mathrm{H_2}}$. The second survey — COLD GASS-low (Saintonge et al. 2017) — observed galaxies with $10^9 < m_*/\mathrm{M_\odot} < 10^{10}$ at $0.01 \leq z \leq 0.02$, with detections in CO for $m_{\mathrm{H_2}} \gtrsim 0.025\,m_*$. Most galaxies in COLD GASS-low, but not all, also have H I data from GASS-low (Catinella et al. 2018). In total, xCOLD GASS comprises 532 galaxies, 477 of which are also in xGASS (GASS + GASS-low). As per Catinella et al. (2018), we refer to the subsample of galaxies in both surveys as 'xGASS-CO'. All H₂ masses from xCOLD GASS shown in this work assume the CO-to-H₂ conversion factor of Accurso et al. (2017), which depends on both the metallicity and star formation rate (SFR) of a galaxy.

All xCOLD GASS data were sourced from the publicly available catalogue. We note that tags for galaxies being centrals and satellites are not included in this dataset, but they are for xGASS (see Section 2.3). Where we present centrals and satellites for xCOLD GASS, we therefore use the xGASS-CO subset. We also note that H₂ masses given in the public dataset and presented in Saintonge et al. (2017) are actually $1.36\,m_{\mathrm{H_2}} \simeq m_{\mathrm{H_2}} + m_{\mathrm{He}}$ (see section 1 of Accurso et al. 2017). In this work, $m_{\mathrm{H_2}}$ means the mass of molecular hydrogen *only*. We therefore subtract the helium contribution from all xCOLD GASS measurements.

Measured fluxes must undergo 'beam corrections' (often also referred to as 'aperture corrections' in the literature) to estimate the intrinsic flux of an object. The beam corrections assume that (i) the face-on, two-dimensional H₂ half-mass radius, $r_{\mathrm{H_2}}^{\mathrm{half}}$, and half-SFR radius, $r_{\mathrm{SFR}}^{\mathrm{half}}$, are equivalent for each galaxy (where the latter is measured independently); (ii) the H₂ surface density of every galaxy falls exponentially with radius; and (iii) the IRAM beam shape is a Gaussian. We can analytically write the beam correction factor as

$$f_{\mathrm{corr}}^{\mathrm{beam}} = \left[2\pi\,r_d^2\,\cos(i)\right]^{-1} \times$$
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{x^2 + y^2}{2\sigma_{\mathrm{beam}}^2} - \frac{\sqrt{x^2 + y^2\,\sec^2(i)}}{r_d}\right) \mathrm{d}x\,\mathrm{d}y\,, \tag{1a}$$

$$r_d = 0.596\,r_{\mathrm{SFR}}^{\mathrm{half}}\,, \tag{1b}$$

$$\sigma_{\mathrm{beam}} = \mathrm{FWHM}\,/\sqrt{8\ln(2)}\,, \tag{1c}$$

where i is the galaxy's inclination, and 'FWHM' refers to the full width at half maximum of the beam — often simply referred to as the 'beam size' — where the beam size of IRAM is 22 arcsec at the frequency of the CO(1–0) observations. Further details regarding the beam correction method can be found in section 2.2 of Saintonge et al. (2012). Both the beam-corrected and uncorrected H₂ masses are used in this work. For a given galaxy, we assume the conversion factor from CO luminosity to $m_{\mathrm{H_2}}$ is the same for the corrected and uncorrected quantities.

Because xCOLD GASS is not complete in terms of stellar mass (even though it is representative), whenever we calculate statistics such as running percentiles or means, we assign weights to each galaxy to compensate for the relative galaxy

---

[2] eXtended Carbon monOxide Legacy Database for the *Galaxy evolution explorer* Arecibo 'Sloan digital sky survey' Survey
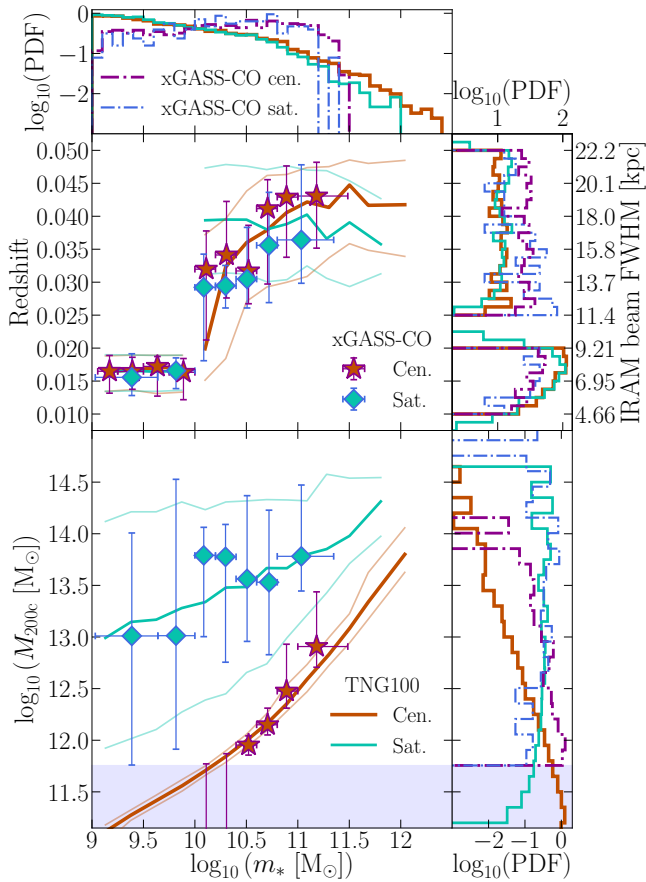[3] Institute for Radio Astronomy in the Millimetre range

Figure 1: Distributions of stellar masses, parent halo masses, and redshifts of galaxies used in this work, split into centrals and satellites. Points and thick lines in the main panels are respective medians for xGASS-CO and the 'mock observed' TNG100 galaxies, per the procedure summarized in Section 2.4 (excluding Section 2.4.3 in this case). Thin lines and vertical error bars give the corresponding 16th and 84th percentiles. Horizontal error bars for xGASS-CO points cover the full bin width. For each marked redshift, the equivalent physical scale on which H_2 masses are measured is given on the right-hand side (FWHM = full width at half-maximum). Smaller panels give the one-dimensional projected probability distribution functions (PDFs) along each axis. The shaded region in the bottom panels is where halo masses are not explicitly computed for xGASS-CO galaxies. For this reason, points for xGASS-CO centrals drop off at low stellar masses; there do exist galaxies at those masses, but they cannot meaningfully be shown on these axes.

count in adjacent 0.1-dex slices of stellar mass. The expected frequency is based on the ratio of galaxy counts in TNG100 for those same bins. This method follows S19, which is very similar to that used in Catinella et al. (2018).

We note that xCOLD GASS was *not* designed for environmental galaxy studies. In truth, neither the sample size nor detection threshold are ideal for this use. Nevertheless, there is a sufficiently representative range of environments covered by the two surveys (see the bottom panels of Fig. 1), and they presently remain unrivalled in their combined number counts and depth of CO surveys at $z \simeq 0$. While we stress that we push these data to their usage limit, xCOLD GASS persists as the most relevant dataset in the literature for a point of comparison for this paper.

2.3 SDSS data

All stellar masses and SFRs from observational data used in this paper (i.e. for xCOLD GASS and the additional SDSS sample we describe below) originate from the MPA–JHU[4] catalogue from SDSS Data Release 7 (Abazajian et al. 2009). These properties have been adjusted for a Chabrier (2003) stellar initial mass function (IMF) and Planck Collaboration XIII (2016) cosmology as necessary (to be consistent with the TNG model). Stellar masses are based on fits to each galaxy's spectral energy distribution, as derived from SDSS's broadband photometry (Salim et al. 2007; also see Kauffmann et al. 2003). SFRs rely on fibre spectroscopy, but have had a correction applied for the small fibre aperture (Salim et al. 2007). SFRs for actively star-forming galaxies use $H\alpha$ emission as their primary indicator, which canonically probe historical time-scales of order 10^7 yr. For passive galaxies, estimated SFRs are related to the 4000-Å break (Brinchmann et al. 2004), which probes an average time-scale of order 10^9 yr. Central/satellite assignments and host halo masses are based on the Yang et al. (2007) group finder. For xGASS, this includes tweaks outlined in Janowiecki et al. (2017). Otherwise, we use the 'modelB' catalogue.

In Section 4.4 and Appendix A1, we use a subsample of galaxies from SDSS that is much larger than xCOLD GASS to look at environmental effects on galaxies' specific star-formation rates (sSFRs). This sample is volume-limited over the redshift range $0.02 \leq z \leq 0.05$, completely covers stellar masses from 10^9 to $10^{11.5}\,M_\odot$, and only includes galaxies with SFR_FLAG_tot = 0. This is taken from the sample used in Brown et al. (2017), which was also used in S19.

2.4 Mock observations

2.4.1 H_2 properties

A limiting factor in single-dish surveys like xCOLD GASS is that there is a finite beam size that can often be smaller than the galaxies of interest. And as raised in Section 2.2, beam corrections are model-dependent; that is, one must assume how CO (and H_2) is distributed in galaxies to estimate the fraction of flux that the beam is insensitive to. In principle, when comparing to simulations, this assumption can be circumvented by instead applying the same beam form to the simulated galaxies, using their known gas distribution, and comparing with *directly* observed fluxes of equivalent galaxies. We refer to this strategy as 'mock observing'.

Our method for mock-observing galaxies in TNG100 is based on S19; for H_2 masses, the process is very similar to that used for calculating mock H I masses in S19. First, we need a means of translating between angular scales and physical scales. To this end, each FoF group in the $z = 0$ snapshot of TNG100 is given a random redshift that is drawn from the probability distribution function of central galaxies in the xGASS survey at the appropriate stellar mass. This redshift is directly turned into a cosmological distance. Distances to satellite galaxies in each FoF group are taken as that of their central plus their inherent z-coordinate displacement from their central in the simulation. Where the central has $m_* \geq 10^{10}\,M_\odot$ but has associated satellites with $m_* <$

---
[4] Max Planck institute for Astrophysics–Johns Hopkins University

$10^{10}\,\mathrm{M}_\odot$, the host halo is assigned a second redshift, used exclusively for those satellites. This second redshift is drawn from the distribution of *all satellites with* $m_* < 10^{10}\,\mathrm{M}_\odot$ in the xGASS survey. This ensures that the hard mass transition from galaxies in the COLD GASS and COLD GASS-low surveys is mimicked in the mock. Note the subtle difference here from S19; up to three redshifts were assigned to TNG100 haloes in S19 to account for the overlapping mass range in GASS and GASS-low. The mass overlap in COLD GASS and COLD GASS-low totals less than a handful of galaxies, and is therefore negligible for our considerations.

For our mock IRAM beam, we maintain the assumption that H_2 is captured with a 100% response at the beam's centre, and that the response falls off with projected radius as a Gaussian with a standard deviation of 9.36 arcsec (FWHM = 22 arcsec). For each galaxy in TNG100, we use its assigned distance to convert this to physical units. All gas cells in the FoF group are then assigned a weight based on this response and their xy displacement from the galaxy centre (taken as the baryonic centre of potential) to measure the integrated H_2 mass of the galaxy of interest (this is done for each galaxy in the FoF group individually). All gas cells that are outside a threshold line-of-sight (always the same as the z-direction) relative velocity to the galaxy centre are then eliminated (equivalent to reassigning their mass weight to zero), as these would not contribute to the theoretical CO line. This threshold is set by a Tully–Fisher relation that considers inclination (equation 6 of S19; based on Tully & Fisher 1977; Reyes et al. 2011).

For completeness, we show the two-dimensional redshift–stellar mass plane for the observational data and our TNG100 mock in Fig. 1 (the 'mock' properties incorporate Section 2.4.2). Consideration of the redshift distribution of observational sources is important, not because we expect significant redshift evolution over the probed range, but because the beam implies significantly different physical 'apertures' across the redshift range. The upper panels of Fig. 1 are directly comparable to those of fig. 2 in S19. The main differences here are that (i) we are restricted to xGASS-CO rather than the full xGASS sample, (ii) mock redshifts for TNG100 satellites with $m_*/\mathrm{M}_\odot \in (10^{10}, 10^{10.23})$ have changed, and (iii) we note the relatively much smaller IRAM beam size on the right-hand side of the y-axis. We caution that about half the galaxies with $m_* < 10^{10}\,\mathrm{M}_\odot$ have a mock beam of FWHM < 8 kpc applied to them (with an absolute minimum FWHM of 4.66 kpc). The potential concern is that we are approaching the scale of the simulation's force resolution; for force calculations involving stellar and dark-matter particles, the softening length is 0.74 kpc. This effectively means the distribution of mass within the inner 2 kpc of a TNG100 galaxy is unreliable (Pillepich et al. 2018b).

To help visualize our mock process, we present images of five TNG100 galaxies in Fig. 2. These galaxies were selected at random within a set of criteria that showcase the significance of the mock process by design. Each frame covers the full spatial extent of baryons that contribute to the inherent properties of the galaxies (Section 2.1). The H_2 structure of each galaxy (in its inherent orientation) is shown in the second row of images. The code to create the gas images has

been placed in a publicly available function.[5] Overlaid are a dashed ellipse and a solid circle; the former is the inclined two-dimensional H_2 half-mass radius, the latter the physical beam size applied in the mock. We re-image the H_2 after applying the mock beam-response weights to each gas cell in the bottom panels. Comparison of the pairs of H_2 images highlights the gas that is 'missed' by the mock beam. We emphasize that these images are purely meant to help visualize the mock process, but images are in no way actually involved in the mock process. In effect, the beam 'sees' a single pixel with an intensity equal to the integral of the images in the bottom panels. Extra panels are provided in Fig. 2 for context, including images of stellar content, and plots for their stellar mass, H_2 fraction, and mock-to-inherent H_2 mass ratio. Fig. 2 is discussed further in Section 3.

The mock-observed H_2 properties of TNG100 galaxies are only compared to the uncorrected H_2 masses from xCOLD GASS. We discuss the relative merits of this approach versus trusting the beam corrections and comparing to the inherent TNG100 properties at length in Section 3.

2.4.2 Other properties

When presenting 'mock' stellar masses for TNG100 galaxies, we approximately account for expected (random) observational uncertainties. In S19, we calculated this by adding a random number drawn from a Gaussian distribution of standard deviation 0.08 dex to each inherent logarithmic stellar mass. In retrospect, this was an underestimate of the typical observational m_* uncertainty. Even with an advanced SED-fitting code, Robotham et al. (2020) highlight that 0.1 dex is a "very optimistic lower limit" on the true uncertainty of an SED-derived stellar mass. For this work, we have revised the induced uncertainty on mock stellar masses to 0.2 dex (consistent with e.g. Donnari et al. 2019).[6]

We note that these new mock stellar masses were calculated after the mock redshifts had already been drawn for each galaxy. This is part of the reason why the redshift distribution at fixed stellar mass for centrals does not align perfectly with xGASS-CO in Fig. 1 (but it is still close). While this departure from S19 means there is greater horizontal smoothing in figures where m_* is on the x-axis, we have tested that it is not significant enough to affect any of the discussion or conclusions of this paper. We have also tested that using a fixed-aperture stellar mass (such as the commonly adopted 30-kpc radius – see e.g. section 5.1.1 of Schaye et al. 2015 for its relevance) before applying this uncertainty makes little difference to our results.[7]

---

[5] See the `build_gas_image_array()` function at https://github.com/arhstevens/Dirty-AstroPy/blob/master/galprops/galplot.py.

[6] Because our sample has a hard cut of $m_* \geq 10^9\,\mathrm{M}_\odot$ for *subhaloes*, we *reflect* induced uncertainties at $10^9\,\mathrm{M}_\odot$ to avoid (i) throwing galaxies away and (ii) artificially over-diluting low-mass bins of actual low-mass galaxies. For example, if a galaxy with inherent $m_* = 10^{9.1}\,\mathrm{M}_\odot$ draws a value of -0.15 dex for its uncertainty, its mock stellar mass is set to $10^{9.05}\,\mathrm{M}_\odot$ rather than $10^{8.95}\,\mathrm{M}_\odot$.

[7] The typical inherent aperture radius exceeds 30 kpc for TNG100 galaxies with $m_* \gtrsim 10^{10.5}\,\mathrm{M}_\odot$, meaning there would be a decrease in the (mock) stellar masses of these galaxies with the application of a 30-kpc aperture, but this turns out to be irrelevant for this
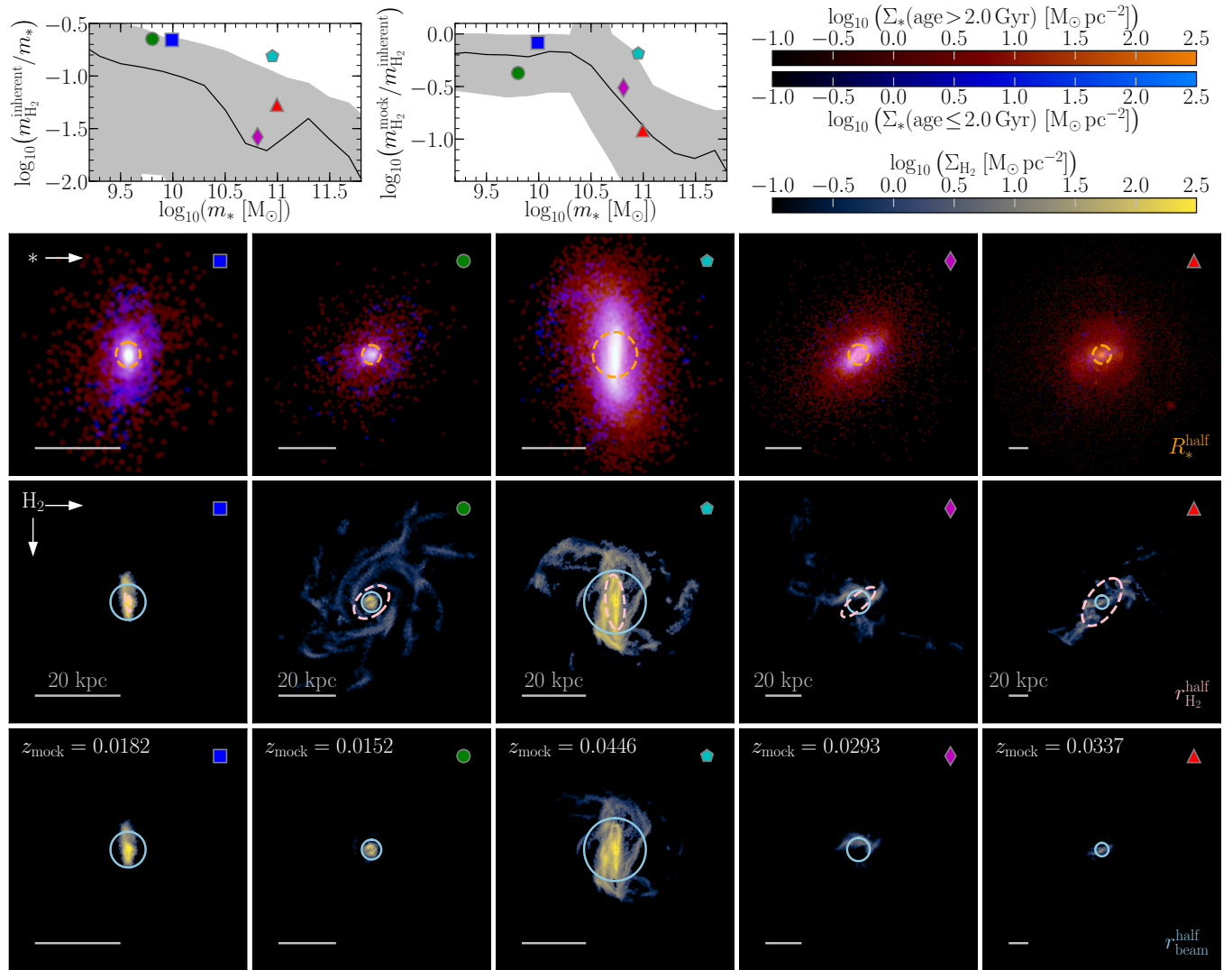
Figure 2: Images of five example central galaxies from TNG100. Two were selected to have stellar masses just under $10^{10}\,\mathrm{M_\odot}$, the other three just under $10^{11}\,\mathrm{M_\odot}$. All were selected to have an H_2 fraction above the median for their stellar mass, as indicated in the top-left panel. The final selection was to ensure a range of differences between the mock and inherent H_2 masses, as seen in the top-centre panel. In these two panels, the black line is the median for all TNG100 galaxies, while the shaded region covers the 16th–84th percentile range. The top-left panel is directly comparable to the middle panel of Fig. 3. Each coloured symbol represents a specific galaxy, which can be connected to each column of images. The first row of images shows the galaxies' stellar content. Two colours are used here: red to indicate stellar particles more than $2\,\mathrm{Gyr}$ old, and blue for younger stars. Dashed, orange circles have radii that match the three-dimensional stellar half-mass radius of each galaxy. The middle row of images shows all the H_2 associated with the galaxy. Dashed, pink ellipses are the face-on, two-dimensional H_2 half-mass radii of the galaxies, each inclined and oriented based on the galaxy's inherent inclination. The diameter of each solid, sky-blue circle equals the full width at half-maximum for the mock IRAM beam. The bottom row of images apply the beam response to the H_2. These are therefore not 'mock images', but rather a visualization of the mass that contributes to the mock H_2 masses. For visual simplicity, we approximate both stellar particles and gas cells as uniform spheres. The radius assumed for stellar particles equals the softening scale. For gas, each cell's recorded volume (mass ÷ density) is conserved. Each frame has dimensions $2R_{\mathrm{BaryMP}} \times 2R_{\mathrm{BaryMP}}$ (see Stevens et al. 2014 and Section 2.1 of this paper). The GD14 $H\,\textsc{i}/H_2$ prescription is used throughout this figure. From left to right, the subhalo IDs of these galaxies as recorded in the TNG100 database are 482614, 465842, 424255, 408935, and 308060.

For mock halo masses, we mimic an abundance-matching procedure. That is, we sum the mock stellar masses of galaxies within each host halo to calculate a total stellar mass for that halo, then reassign halo virial masses such that each halo has the same rank when ordered by either mass. When presenting satellites and centrals for TNG100 in Fig. 1, we treat the galaxy with the highest mock stellar mass in each halo as the central. For all other plots, the description in Section 2.4.3 is followed.

paper. An extensive comparison of the relevant apertures (and others) is explored in Stevens et al. (2014).

Mock SFRs are designed to average over the historical time-scale that observational SFRs would. First, we calculate the 20-Myr historically averaged SFR for each galaxy, based

on the birth mass and birth time of each star particle. We calculate a mock sSFR by taking the ratio of this historical SFR to the mock stellar mass, including the induced uncertainty on the latter. We then add a random Gaussian error of $0.013\,\mathrm{dex}$ on each sSFR, which is typical of the sSFR uncertainty given in the SDSS database for star-forming galaxies in our subsample described in Section 2.3. For returned sSFRs less than $10^{-11}\,\mathrm{yr}^{-1}$, the SFR is remeasured on a 1-Gyr time-scale, and a random Gaussian error of $0.025\,\mathrm{dex}$ is then applied to the new sSFR (typical of those given for non-starforming systems in the SDSS subsample). Our induced errors here are another departure from the mock procedure in S19, which better serve the comparisons in this paper. Even after doing this, 17.4% of TNG100 galaxies still have SFR $= 0$. Were these observed in a survey like SDSS, their SFRs would unlikely be recorded as zero, though. In the spirit of finding the 'best match', we fit a power law to the passive population of SDSS, measure its scatter (assumed to be Gaussian), and then assign random SFRs drawn from this sequence to the TNG galaxies that still had an SFR lower than the non-zero minimum in the SDSS sample described in Section 2.3. More information is provided in Appendix A1.

No attempt is made to mock the SDSS fibre aperture for SFRs here. We have instead opted to take the SDSS aperture corrections at face value. The reason for the difference in philosophy versus our H_2 mock measurements lies in the physical scales involved. The SDSS fibre is 3 arcsec in diameter. At $z = 0.05$ — the highest redshift in the interval we compare against — this translates to $\sim 3\,\mathrm{kpc}$, i.e. a radius of $\sim 1.5\,\mathrm{kpc}$. This *maximal* radius is only twice the stellar gravitational softening scale used in TNG100, and is smaller than the smallest beam size used for H_2 mock measurements (cf. Fig. 1). While fibre-like quantities have been measured and adopted in previous TNG100 analyses (e.g. Nelson et al. 2018; Donnari et al. 2019), our choice to avoid them is twofold: (i) those previous works used a larger radius than the maximal $1.5\,\mathrm{kpc}$ that is relevant here, and (ii) arduous resolution tests would be required to trust results at this relatively small scale. This means there are potentially strong systematic uncertainties in the SDSS data that we have not explicitly account for/modelled.

Finally, mock H I masses are calculated very similarly to H_2. The only difference is the beam size, which for H I is 3.5 arcmin, designed to be consistent with Arecibo and therefore xGASS. The mock redshifts and method for applying the mock beam are otherwise identical to Section 2.4.1.

2.4.3 Cross-contamination of centrals and satellites

As is often the case with comparisons between simulations and observations, even when mock observations for the former are done, typically not all the effects or assumptions that go into the quantities derived from observations are mimicked in the simulation properties. For example, the central/satellite assignment in the xGASS database (which originates from the group catalogue of Yang et al. 2007) is not pure; that is, many galaxies tagged as satellites will actually be centrals, and vice versa. Stevens & Brown (2017) have shown that the cross-contamination of satellites and centrals in the Yang et al. (2007) catalogue can potentially fully explain why the mean (stacked) H I fractions of satellites and centrals ($\forall m_* > 10^9\,\mathrm{M}_\odot$) in the ALFALFA survey (Giovanelli

$\log_{10}\left(M_{200c}^{\mathrm{parent}}/\mathrm{M}_\odot\right)$	<12	$[12, 13)$	$[13, 14)$	>14
Central purity	0.95	0.90	0.80	0.70
Satellite purity	0.50	0.65	0.70	0.75

Table 1: Purities assumed when artificially cross-contaminating centrals and satellites for mock TNG results. These vary with galaxy type (central or satellite: rows) and parent halo mass ($M_{200c}^{\mathrm{parent}}$: columns). Cf. Campbell et al. (2015).

et al. 2005) are so similar, despite a greater difference being predicted by cosmological models of galaxy formation (specifically DARK SAGE in that case).[8] Relatedly, Davies et al. (2019) have also shown that group finders with less cross-contamination lead to stronger differences between the quenched fractions of centrals and satellites.

When presenting 'mock' results for centrals and satellites (with the exception of Fig. 1), we intentionally crosscontaminate the two populations in TNG100. Bravo et al. (2020) have shown that assuming a fractional contamination between centrals and satellites closely mimics the procedure of running a group catalogue in a simulated galaxy lightcone and assessing contamination there. In a similar vein, we use weights when calculating statistical quantities (means, medians, and other percentiles) for TNG100 satellites or centrals, which are assigned to ensure we get back a desired effective purity. We fold in a dependence on host halo mass to these purities, loosely drawn from fig. 6 of Campbell et al. (2015) for the Yang et al. (2007) group finder. These are summarized in Table 1. Additionally, we apply these weights for *each* stellar-mass bin. For example, when calculating the running statistics for centrals, we first isolate all galaxies with stellar masses in the bin being considered, then break those centrals into halo masses, assign the appropriate weight to each central ($w_{\mathrm{cen},i}$: the number from the relevant column in the 'centrals' row of Table 1), and finally calculate the weights for satellites as

$$w_{\mathrm{sat}} = \frac{1}{N_{\mathrm{sat}}} \sum_{i=1}^{4} (1 - w_{\mathrm{cen},i}) N_{\mathrm{cen},i} . \tag{2}$$

Here, $N_{\mathrm{cen},i}$ refers to the number of centrals in that stellar-*and* halo-mass bin, while N_{sat} refers to the *total* number of satellites in the stellar-mass bin (irrespective of their parent halo mass). This effectively means that centrals of a given halo mass are contaminated with satellites of all parent halo masses. This same process is repeated for every stellar-mass bin. All steps are then done again to calculate the running statistics of satellites (swapping 'cen' and 'sat' in Equation 2).[9]

This contamination process is another feature that was not included in the mock results of S19. Because there are several

---

[8] S19 showed that this problem can be also be solved by mockobserving simulated galaxies (from TNG100) to account for the sky area considered in the ALFALFA all-sky data cube for each galaxy when the stack was made. But, evidently, the equivalent xCOLD GASS mock observations do not offer a solution for H_2 here.

[9] As a point of clarification for the bottom-left panel of Fig. 11, satellites in a given halo mass bin are contaminated with centrals of all halo masses (of the same stellar mass), but there is no contamination from *other* satellites in haloes of different masses. This is in line with the contamination procedure used in Stevens & Brown (2017).

new features to the mock process in this paper, we update the key xGASS-related results from S19 in Appendix A2 to be consistent with the methods described here.

2.5 Conventions

Throughout this paper, we adhere to the following conventions.

• Whenever we refer to an 'H$_2$ fraction', we exclusively mean the ratio of H$_2$ mass to stellar mass of a galaxy (and likewise for 'H I fractions').

• All cosmology-dependent quantities — both simulated and observational — assume (or have been adjusted to assume) $\Omega_m = 0.3089$, $\Omega_\Lambda = 0.6911$, and $h = 0.6774$ (Planck Collaboration XIII 2016).

• Likewise, all IMF-dependent quantities assume a Chabrier (2003) IMF. When converting from a Kroupa (2001) IMF, we assume $m_{*,\mathrm{Chabrier}} = \frac{0.61}{0.66} m_{*,\mathrm{Kroupa}}$ and $\mathrm{SFR}_{\mathrm{Chabrier}} = \frac{0.63}{0.67} \mathrm{SFR}_{\mathrm{Kroupa}}$ (see Madau & Dickinson 2014).

• Zeroes are never discarded when calculating statistical quantities such as means, medians, and other percentiles.

• Logarithms are always taken *after* the statistical quantity of interest is computed.

• Upper-case R refers to three-dimensional/spherical radii. Lower-case r refers to two-dimensional/cylindrical radii.

• Upper-case M (or $\mathcal{M}$) refers to the mass of haloes. Lower-case m refers to the mass of galaxies.

• When binning data, we use bins of nominal (but not strict) width 0.2 dex. The absolute least number of galaxies we use in a bin is 15. No galaxies are discarded when binning. When necessary, bin widths automatically adapt to ensure both latter criteria are fulfilled.

• There are no 'helium contributions' to H I or H$_2$ masses.

• Whenever we refer to 'halo mass', we always mean that of the 'host' or 'parent' halo. We use these terms interchangeably.

3 A CENSUS OF MOLECULAR GAS IN GALAXIES AT REDSHIFT ZERO

Before addressing effects of environment, let us devote this section to setting the scene for what the xCOLD GASS survey and TNG100 simulation tell us about the H$_2$ properties of galaxies at $z = 0$ on the whole (and in a comparative sense).

3.1 H$_2$ relative to stellar content

In the top two panels of Fig. 3, we show the variation in the H$_2$ fraction of galaxies from both TNG100 and xCOLD GASS as a function of stellar mass, including the relation's scatter (cf. fig. 3 of Diemer et al. 2019). The top panel makes no correction to the CO observations for the finite beam, and instead applies the same beam to the simulations to derive comparable 'mock observed' H$_2$ masses (as outlined in Section 2.4). Forward-modelled uncertainties on stellar masses are also included. By contrast, the second panel *does* employ the Saintonge et al. (2012) beam correction to the observations, while the inherent (gravitationally bound) properties of TNG galaxies are plotted. In both cases, non-detections in xCOLD GASS assume their upper-limit measurements for

the percentiles shown with hexagons and error bars, and are instead set to zero for the medians shown by the plusses. Loosely, both datasets in both panels suggest that higher-mass galaxies have lower H$_2$ fractions, with significant scatter at all masses. It is important that we show both these panels, as it highlights the difficulty and care required to compare simulations with observations; the level of agreement is notably different between the two. At the low-mass end, the uncorrected and mock quantities seem better matched. But at the high-mass end, the choice of comparison either leads to the conclusion that TNG100 galaxies are systematically too gas-rich or contrarily that they are *suggestively* too gas-poor[10] (modulo *systematic* uncertainties in xCOLD GASS masses, which are not shown).

Fig. 2 provides a deeper understanding of why there can be such significant differences between the mock and inherent H$_2$ properties of some TNG100 galaxies, even while other galaxies show negligible differences. The first two columns of images show two galaxies with comparable stellar masses and H$_2$ fractions, yet vastly different H$_2$ structure. If the first galaxy were observed face-on, its two-dimensional H$_2$ half-mass radius would be measured as $r_{\mathrm{H_2}}^{\mathrm{half}} = 2.2\,\mathrm{kpc}$, well within the mock beam size. In fact, nearly all the H$_2$ in this galaxy fits within the mock beam. By contrast, the second galaxy has $r_{\mathrm{H_2}}^{\mathrm{half}} = 7.3\,\mathrm{kpc}$, which exceeds the beam size. The outer half of its H$_2$ mass also lies in extended spiral arms, all of which is missed by the beam. The next three galaxies all have higher stellar masses, each closer to $10^{11}\,\mathrm{M}_\odot$. Again, these highlight a variety of situations where anywhere from 35% to 88% of the galaxies' H$_2$ is missed by the mock beam. The most extreme of these three examples is actually representative of the *average* galaxy at this stellar mass, as discerned from the top-middle panel of Fig. 2.

All of this motivates the question: is one form of comparison (mock/uncorrected or inherent/corrected) more meaningful than the other? An argument can be made that, because the beam correction for xCOLD GASS requires a model for how H$_2$ in galaxies should be distributed (i.e. that it is distributed exponentially, with a half-mass radius equal to the radius enclosing half the galaxy's star formation rate), the mock/uncorrected comparison is the most direct. Unfortunately, this implicitly relies on H$_2$ being distributed in TNG galaxies the same way it is in reality; otherwise the applied beam would exclude a different fraction of the true H$_2$ mass of the galaxy. As it happens, Diemer et al. (2019, see their fig. 7) have shown that H$_2$ in TNG galaxies can extend to higher radii than found in observations (cf. Bolatto et al. 2017). While the difference in extent is only of order tens of percent on average, there is a tail of galaxies that stretches out to factors of a few. We have since learned that higher-mass galaxies lie in this tail at greater frequency. Fig. 2 tells a consistent story, where several examples of galaxies with H$_2$ half-mass radii notably exceeding their stellar half-mass radii are shown. Based on this then, if the goal is to compare the integrated H$_2$ masses of galaxies, one can make a counter-argument that the inherent/corrected comparison might be more meaningful.

The truth is that neither form of comparison is perfect. But

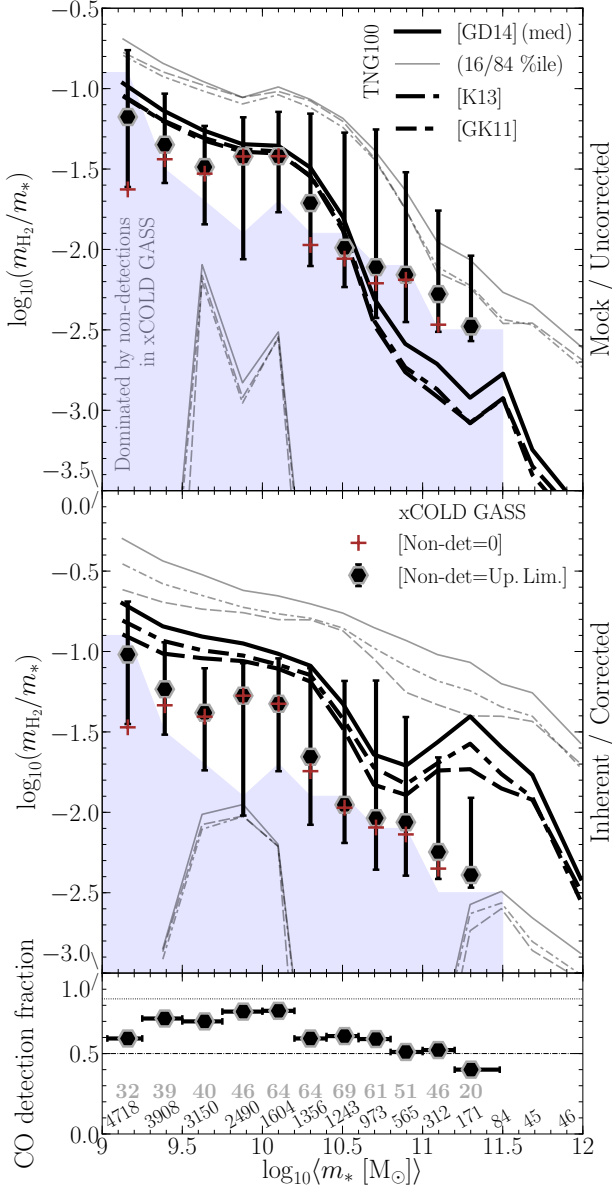[10] The detection fraction of xCOLD GASS is too low to claim this definitively.

Figure 3: Molecular-hydrogen fraction of all galaxies (i.e. centrals and satellites together) as a function of stellar mass. **The top panel** compares mock-observed properties from TNG100 with the data from xCOLD GASS that have *not* been beam-corrected. **The second panel** compares the inherent properties of TNG100 galaxies with the beam-corrected xCOLD GASS values. In each case, the running 16th, 50th, and 84th percentiles of both the simulation and survey data are given by lines and hexagons+error bars, respectively. Hexagons represent medians when non-detections are assumed to take their upper-limit value, whereas the plus symbols represent medians when non-detections are set to have $m_{H_2} = 0$. To mitigate redundancy, we exclude a second set of error bars for the plusses; these would extend from the top of the black error bars down to $-\infty$. The different line styles for TNG100 correspond to different H I/H₂ prescriptions. The shaded region in these panels covers where the observations are dominated by non-detections. **The bottom panel** shows the detection fraction of xCOLD GASS galaxies, with horizontal error bars covering the full bin width. Printed numbers are the galaxy counts in each bin for xCOLD GASS (grey, bold) and TNG100 (black, angled). For reference, dotted and dot-dashed horizontal lines are respectively given at detection fractions of 0.84 and 0.5.
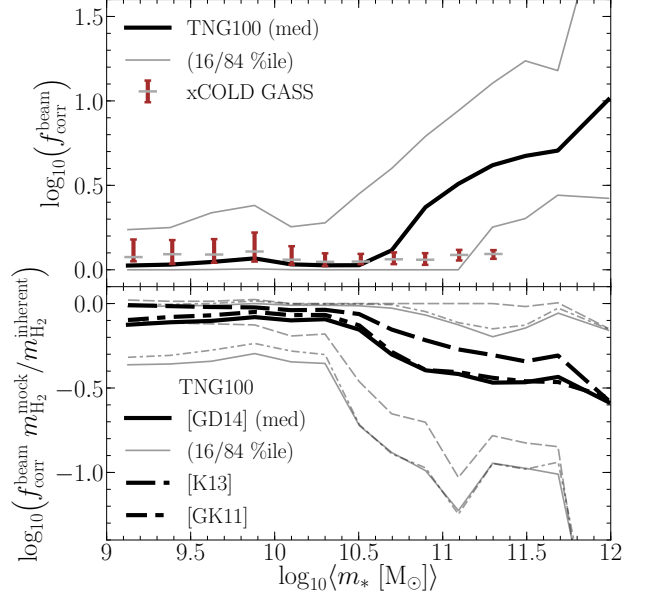

Figure 4: Top panel: Running distributions of beam correction factors for xCOLD GASS and equivalents for TNG100, per Equation (1), as a function of stellar mass (inherent for TNG100). Horizontal dashes for xCOLD GASS are medians, while error bars cover the 16th–84th percentile range. f_{corr}^{beam} has no dependence on m_{H_2}, and thus all detections and non-detections for xCOLD GASS are treated equivalently here, plus there is no H I/H₂ prescription dependence for TNG. **Bottom panel:** Ratio of TNG100 effective beam correction factors to the actual difference factor between the inherent and mock H₂ masses. Running medians, 16th and 84th percentiles are shown for each H I/H₂ prescription. If Equation (1) were consistently accurate, all lines in this panel should be close to 0. Instead, f_{corr}^{beam} is seen to underestimate the 'true' correction factor for a significant fraction of galaxies. The f_{corr}^{beam} quantities for TNG100 galaxies are not used anywhere else in this paper. Galaxies with SFR = 0 *are* included in this plot; they are interpreted to have $f_{corr}^{beam} = m_{H_2}^{inherent}/m_{H_2}^{mock} = 1$.

there is a test we can perform to help inform when one comparison might be more meaningful than the other. Namely, we can calculate an 'effective beam correction factor' for each TNG100 galaxy using Equation (1), and compare the distribution of these factors with (i) those of xCOLD GASS, and (ii) the ratio of inherent to mock H₂ masses of TNG100. Importantly here, Equation (1) is *independent* of any actual H₂ properties of the galaxies. We define r_{SFR}^{half} for TNG galaxies using the instantaneous SFRs of their gas cells. If Equation (1) is an accurate method for beam correction, then f_{corr}^{beam} should be almost the same as $m_{H_2}^{inherent}/m_{H_2}^{mock}$ for each TNG galaxy. To check this, we plot the running distributions of both f_{corr}^{beam} and the ratio of it to $m_{H_2}^{inherent}/m_{H_2}^{mock}$ in Fig. 4 (both quantities are interpreted to be equal to 1 for galaxies with no star formation activity).

For $m_* \lesssim 10^{10.3}\,M_\odot$, the bottom panel of Fig. 4 highlights that Equation (1) does a good job of predicting $m_{H_2}^{inherent}/m_{H_2}^{mock}$ for the median TNG galaxy. However, the 16th percentiles show that f_{corr}^{beam} can underestimate this ratio by a factor of ~ 2 at significant frequency. This notably depends on the H I/H₂ prescription; in the case of K13, H₂ does not extend as far out for each galaxy, and thus f_{corr}^{beam} remains accurate within $\sim 30\%$ (still as an underestimate) for most

galaxies.[11] For the same mass range, the top panel of Fig. 4 shows that the running distributions of $f_{\rm corr}^{\rm beam}$ for TNG100 and xCOLD GASS are fairly similar. In principle, it does not matter if these are quantitatively identical or not; differences between distributions here are the result of differences in the distributions of $r_{\rm SFR}^{\rm half}$.[12] the point of the correction being that it accounts for this. This comparison simply tells us that the SFR sizes — and therefore the gas sizes — of TNG100 galaxies are physically reasonable for the majority with $m_* \lesssim 10^{10.3}\,{\rm M_\odot}$. Given this, and the unsurprising fact that $f_{\rm corr}^{\rm beam}$ is not identical to $m_{\rm H_2}^{\rm inherent}/m_{\rm H_2}^{\rm mock}$, forward-modelling the beam is, objectively, a better way to compare simulations with observations. In other words, the mock/uncorrected comparison should be more meaningful than the inherent/corrected comparison on the basis that it convolves a known gas distribution with a known beam to compare to a directly observed quantity, as opposed to *assuming* how gas is distributed in xCOLD GASS galaxies to effectively deconvolve the beam. The fact that we find $f_{\rm corr}^{\rm beam}$ to frequently underestimate the beam correction factor explains why TNG100 and xCOLD GASS are better aligned at low masses in the top panel of Fig. 3 than in the middle panel.

However, the situation is not so clear-cut at higher masses. The top panel of Fig. 4 shows wild differences in $f_{\rm corr}^{\rm beam}$ between TNG100 and xCOLD GASS for $m_* \gtrsim 10^{10.7}\,{\rm M_\odot}$. This is most likely symptomatic of gas being unrealistically distributed in the simulated galaxies. Moreover, as seen in the bottom panel of Fig. 3, $f_{\rm corr}^{\rm beam}$ and $m_{\rm H_2}^{\rm inherent}/m_{\rm H_2}^{\rm mock}$ start diverging at $m_* > 10^{10.3}\,{\rm M_\odot}$. This implies that the radial gas profiles of these galaxies deviate significantly — i.e. more than one would expect — from a symmetric exponential. The major differences in the mock and inherent H_2 masses seen at these stellar masses are therefore somewhat artificial (or, in other words, not predictive of reality). In light of this, we suggest that the inherent/corrected comparison with xCOLD GASS is a safer choice than the mock/uncorrected comparison at high stellar masses, albeit still an imperfect one.

Given the uncertainties raised above, our philosophy is to present both the mock/uncorrected and inherent/corrected comparisons where relevant throughout this paper. Each requires a grain of salt in its interpretation, but both are needed to provide an as-complete-as-possible picture.

As shown explicitly in the bottom panel of Fig. 3, a significant number of the xCOLD GASS galaxies are not detected in CO; in fact, for the highest-mass bin, the majority are non-detections. This is also reflected in the shaded regions in the top two panels; these cover the area where there are more upper limits in the observational data than detections.[13] Because the detection fraction of xCOLD GASS galaxies is < 0.84 in all mass bins, we cannot show the 16th percentile in Fig. 3 for xCOLD GASS H_2 fractions if non-detections are assumed to have no H_2 (those percentiles would all be zero,

i.e. $-\infty$ in log space). In essence, this means the lower half of the scatter in the H_2 fraction–stellar mass relation is observationally unconstrained. We therefore cannot make meaningful statements about how representative the frequency of properly H_2-poor galaxies in TNG is of reality.

Nevertheless, we note that the 16th percentile of inherent H_2 fractions of TNG100 galaxies shows a dip over $10^{10.2} < m_*/{\rm M_\odot} < 10^{11.3}$, reaching a trough outside the axes of Fig. 3. This lines up with the dip in $H\,\textsc{i}$ fractions identified in S19. This feature effectively disappears in the mock properties, in part because the H_2 fractions of all the most massive galaxies drop after the application of the relatively small beam. As discussed in S19, this feature is most likely related to the onset of kinetic feedback from massive black holes / active galactic nuclei (AGN — for details on the implementation and result of said feedback, see Weinberger et al. 2017; Zinger et al. 2020, respectively).

In general, there is a slightly greater sensitivity to the choice of $H\,\textsc{i}/H_2$ prescription for TNG H_2 results than for $H\,\textsc{i}$ (cf. Diemer et al. 2018, 2019; S19). This is perhaps unsurprising, as $H\,\textsc{i}$ tends to dominate over H_2, meaning small fractional changes in $H\,\textsc{i}$ between prescriptions are reflected by larger fractional changes in H_2 (as the total $H\,\textsc{i}+H_2$ is fixed). The source of this difference is most readily understood by looking at the $m_{\rm H_2}/m_{\rm H\,\textsc{i}}$ ratios of the simulated galaxies, which we assess next.

3.2 H_2 relative to $H\,\textsc{i}$ content

In Fig. 5, we show the $H_2/H\,\textsc{i}$ mass ratios of all TNG100 and xGASS-CO galaxies as a function of stellar mass. As documented by Catinella et al. (2018), the $H_2/H\,\textsc{i}$ mass ratios of xGASS-CO galaxies rise with stellar mass on average. But even over 2.5 dex in stellar mass, this rise is less than (or at least comparable to) the relation's scatter. We again show two panels that compare the mock-observed simulation properties with uncorrected observations (top) and inherent simulation properties with corrected observations (bottom). Although, we note that no beam correction has been applied to the $H\,\textsc{i}$ content of xGASS-CO galaxies in the bottom panel of Fig. 5 (it has only been applied for H_2, as given in the public xGASS and xCOLD GASS databases); see S19 for the potential significance of this (much larger) beam. We also need to be particularly careful about the treatment of non-detections; while some galaxies are detected in both $H\,\textsc{i}$ and CO, some are only detected in one, some the other, and some neither. We still present running percentiles where non-detections (in both phases) are set to their upper limits. But it no longer makes sense to plot the equivalent when non-detections are set to zero; to do this for $m_{\rm H\,\textsc{i}}$ would mean dividing by zero (for the same reason, only TNG100 galaxies with $m_{\rm H\,\textsc{i}} > 0$ are included in Fig. 5). Instead, we show a second set of points (exes) with dashed error bars, which exclusively consider galaxies detected in *both* phases. This again provides two imperfect means of comparing to simulation results. Somewhat reassuringly though, there is little difference between the two sets of observational points.

TNG100 mock results are broadly consistent with observations in the left half of the top panel of Fig. 5, but are again hampered by the overextension of H_2 at high stellar masses in the simulation. Here, the $H_2/H\,\textsc{i}$ mass ratios show a systematic drop from $m_* \gtrsim 10^{10.3}\,{\rm M_\odot}$, before rising again at

---

[11] As a point of comparison, the uncertainty taken for beam corrections in Saintonge et al. (2017, section 2.5) is 15%.
[12] Per Equation (1), the only other thing that could matter is differences in the distribution of $\sigma_{\rm beam}$ in physical units, but these are the same by construction.
[13] While the survey was designed with an H_2 fraction detection limit (mentioned in Section 2.2), in practice, some non-detections have higher upper limits on their H_2 fractions than other detections' measurements at the same stellar mass.
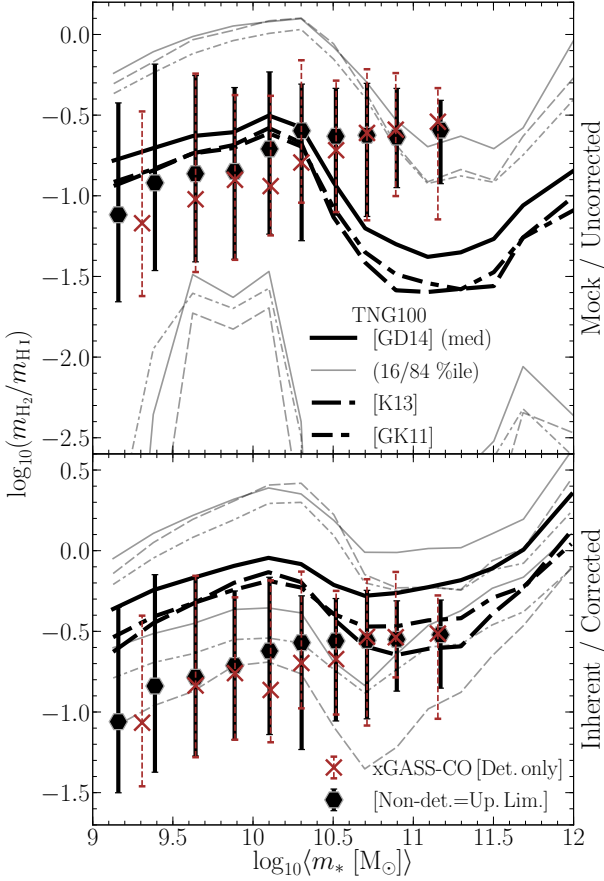
Figure 5: Ratio of molecular to atomic hydrogen in galaxies as a function of their stellar mass, at $z \simeq 0$. Plotting conventions follow Fig. 3 for TNG100 data. For xGASS-CO data, filled symbols (with solid error bars) give the median (plus 16th & 84th percentiles) for the complete overlap in xGASS and xCOLD GASS, where non-detections in either H I or CO assume their upper limits. Exes with dashed error bars *exclusively* use the galaxies detected in *both* H I and CO.

$m_* \gtrsim 10^{11}\,\mathrm{M}_\odot$. While the depth of this feature is tied to our mock procedure, comparison with the bottom panel of Fig. 5 shows that the feature still exists for the inherent galaxy properties, just more weakly. As such, there must be an underlying physical cause for this feature in the simulation. This suggests that gas from the galaxies' centres — i.e. where the majority of H₂ but only a fraction of H I is (see e.g. Leroy et al. 2008; Obreschkow et al. 2009; Stevens et al. 2019b) — is being removed or depleted. Based on differences in mock/inherent H I masses alone, especially when contrasting satellites and centrals, S19 hypothesized that AGN feedback was likely responsible for displacing gas from galaxies' centres in TNG100 at the mass scale in question. Having imaged several galaxies, we can confirm there are clear holes and asymmetries in the gas structure of many galaxies at these masses, as would be expected from ejective feedback at each galaxy's centre (for a related analysis of galactic outflows in the TNG model, see Nelson et al. 2019b). The fourth galaxy in Fig. 2 is a prime example of this. Most of the H₂ on one side of the galaxy has been blown out, with the other side clearly disrupted, as seen by its lack of alignment with the stellar disc. This physical effect is then compounded when mock measurements are

made, as it is then *only* the central region of the galaxy that would contribute meaningfully to the H₂ mass, but that gas has been displaced.

While we cannot claim that TNG100 is a 100% faithful representation of reality when it comes to galaxies' H₂ content — indeed, we have identified some issues — Figs 3 & 5 at least give us confidence that the simulation is 'in the right ball park', so to speak (especially for $10^9 \lesssim m_*/\mathrm{M}_\odot \lesssim 10^{10.2}$). Similar statements could be made about the EAGLE[14] simulations (cf. fig. 5 of Lagos et al. 2015; also see Davé et al. 2020). It will likely be years before next-generation simulations and CO surveys will allow us to make comparative statements that are more grand and robust than this. In the interim, the simulations that we have on hand still provide worthwhile synthetic laboratories to study the effects of environment on H₂, bearing in mind the caveats that carry through from the above.

4 MOLECULAR GAS IN GALAXIES BY ENVIRONMENT

4.1 Central galaxies versus satellites

As outlined in Section 1, one of the main aims of this paper is to investigate how centrals and satellites differ in their H₂ content at fixed stellar mass. In Fig. 6, we show precisely this for xGASS-CO (i.e. the xCOLD GASS galaxies that have central/satellite status determined as part of xGASS, per Section 2.2) and TNG100 at $z = 0$. Similar to Fig. 3, we compare the mock/uncorrected values in the top panels, and the inherent/corrected values in the middle panels. Because we are now subsampling the observational data, we are dealing with weaker statistics. This, combined with there being visually more information in the figure, means we have shown the running medians and 84th percentiles in separate columns. We exclude a column of panels for the 16th percentile here primarily (but not exclusively — see below) due to the low detection fraction in the observations. Importantly, we can trust that the distribution of parent halo masses of satellites in xGASS-CO is very similar to that of TNG100 (Fig. 1), meaning their gas properties at fixed stellar mass should be directly comparable.

Looking at the observational data alone in any of the main four panels of Fig. 6, there is little evidence to support the idea that satellites are *significantly* depleted of their H₂ at a particular mass scale. But some degree of systematic difference does appear to be present, as 5–6 of the seven stellar-mass bins show satellites to have lower median and 84th-percentile H₂ fractions than centrals. While the bootstrapping errors on these percentiles imply this is not statistically significant for any singular bin, the likelihood of randomly observing this result for *all* bins simultaneously is low (we elaborate below). Nevertheless, the results of Fig. 6 contrast with the directly comparable plot for H I in xGASS (fig. 5 of S19), where there is a clearer distinction between the distribution of H I fractions of satellites and centrals for all fixed $m_* \lesssim 10^{10.75}\,\mathrm{M}_\odot$ (see Appendix A2). We have tested that those H I results remain when we select xGASS-CO galaxies

_____

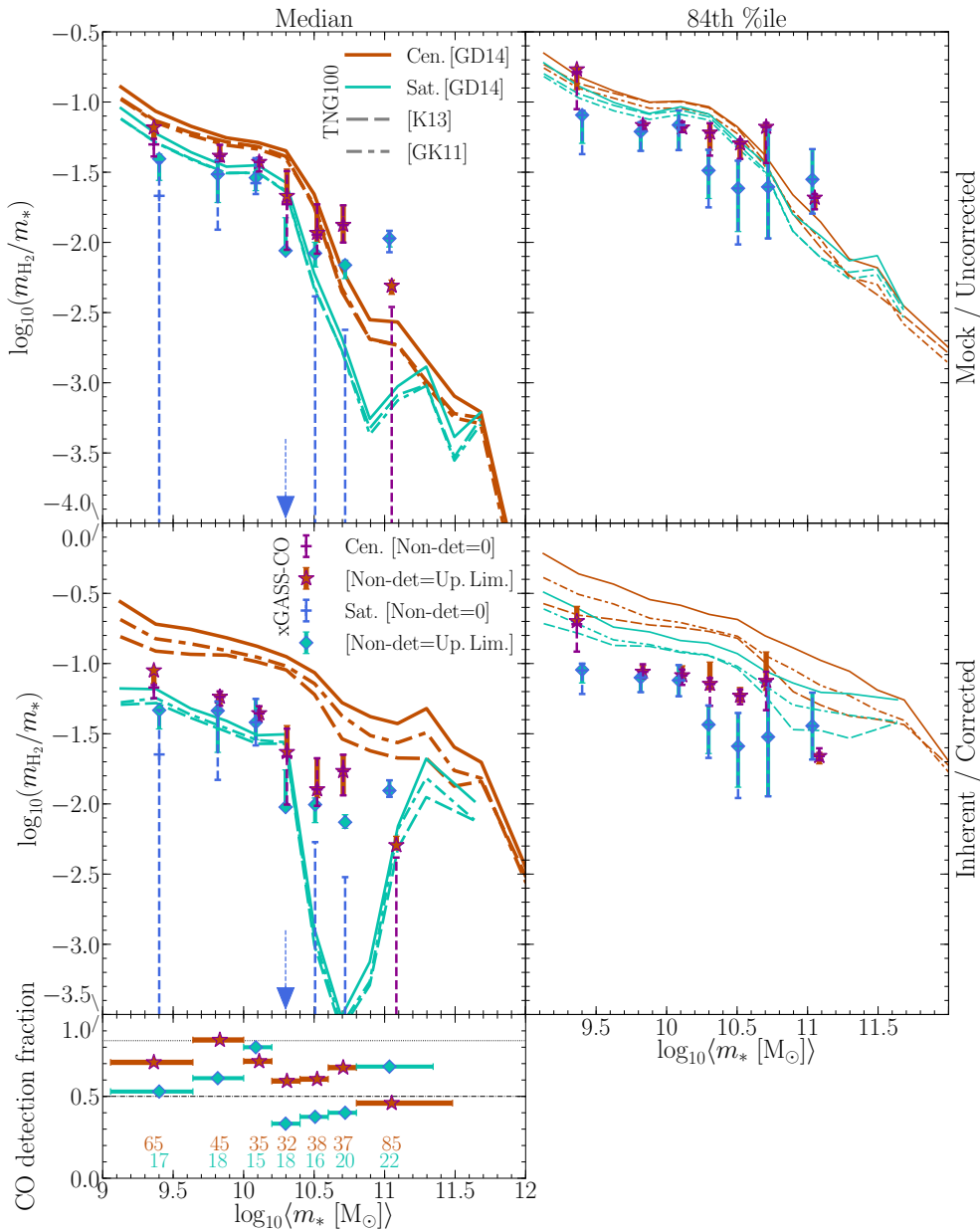[14] Evolution and Assembly of GaLaxies and their Environments

Figure 6: Molecular-hydrogen fraction as a function of stellar mass for central and satellite galaxies presented separately. The layout of this figure differs from Fig. 3. Here, the left-hand panels only show the median trend for TNG100 and xGASS-CO, while the right-hand panels show the running 84th percentile. Error bars for xGASS-CO percentiles are 68% confidence intervals *of those percentiles* from bootstrapping — points accompanying the error bars still represent the raw percentiles from the data. Two sets of error bars exist in each panel: one for when we assume that non-detections take their upper limit value (solid bars with either pentagram or diamond points), and one for when we assume non-detections have zero H_2 (dashed bars with plusses). Downward arrows are used when the entire 68% confidence interval falls below the axes. The same bins are used for both centrals and satellites. The bottom-left panel shows the detection fraction, width, and number of galaxies in each xGASS-CO bin.

exclusively, implying that the *relative* lack of observed difference in H_2 for centrals and satellites cannot be fully explained away by the subsample being small and/or potentially unrepresentative.

By contrast, satellites in TNG100 show a clear depletion in their H_2 fractions relative to centrals of the same stellar mass. The inherent results of the simulation find the median satellite to be ~ 0.6 dex (a factor of ~ 4) poorer in H_2 than the median central at fixed $m_* < 10^{10.3}\,M_\odot$. From a theoretical, qualitative stand-point, this is not an unreasonable

result. One can imagine that satellites' molecular gas would continue to be consumed in star formation (and accompanied feedback), but would be unable to efficiently replenish, due to the denial of cosmological gas accretion (similar to the picture outlined by Stevens & Brown 2017). If the depletion of H_2 were to be approximately linear with time, we should see a greater difference between centrals and satellites in log space for low percentiles, and likewise a smaller difference at high percentiles. Indeed, there is only a 0.2–0.3 dex difference (a factor of <2) in the inherent 84th percentile of centrals

and satellites in TNG100 (lower-right panel of Fig. 6). At the same time, for those satellites that experience a pericentric passage close to the halo centre, ram pressure from the relative motion of the intrahalo medium would be so strong that it could entirely remove the interstellar medium (i.e. including H₂). Another reason we opted to exclude panels for the 16th percentile in Fig. 6 is that it equals zero ($-\infty$ in log) for TNG100 satellites $\forall m_* \gtrsim 10^9 \, \mathrm{M}_\odot$.

It is starkly obvious that the separation between centrals and satellites diminishes significantly when we look at the mock-observed simulation outputs. The separation between running medians has been reduced from $\sim$0.6 to $\sim$0.2 dex. From testing (not shown here), we learned that while the application of the mock beam (Section 2.4.1) has an impact on decreasing this separation, the most important factor is the cross-contamination of centrals and satellites (Section 2.4.3). That is, if we apply the contamination to the inherent properties, we find a separation $\sim$0.3 dex. Similarly, had we excluded the cross-contamination from our mock procedure (but kept the other aspects the same), we would have found a separation of $\sim$0.4 dex. Satellite–central cross-contamination is always present in the xGASS-CO data, regardless of whether beam corrections are considered or not. There is no way to correct for this contamination without applying a completely different group-finding algorithm to SDSS; even then, a 'perfect' group finder does not exist to our knowledge, nor might it ever. The best option (arguably, the only option) is to forward-model the contamination, like we do with the mock results of TNG100. As such, the mock/uncorrected comparison is *always* more meaningful than the inherent/corrected comparison when investigating the *relative* influence of environment, despite the concerns about the *absolute* mock H₂ properties at high stellar masses raised in Section 3.1.

A natural question to ask is: do the xGASS-CO data favourably support our prediction that the survey should see a 0.2-dex systematic difference between the H₂ fractions of centrals and satellites? Close inspection of the top panels of Fig. 6 suggest that the data are consistent with this prediction. But the uncertainties in each stellar-mass bin for xGASS-CO make it difficult to read how quantitatively precise this apparent consistency is. By combining the data from all bins, however, we can make a more rigorous comparison. To this end, we calculate the difference in H₂ fraction for each galaxy from the median of *centrals* in each bin, and thus define

$$\Delta \log_{10}(m_{\mathrm{H}_2}/m_*) \equiv \log_{10}(m_{\mathrm{H}_2}/m_*) - \\ \log_{10}\left(\mathrm{Median}\left[\frac{m_{\mathrm{H}_2}}{m_*}(m_*, \mathrm{cen.}) \right] \right) . \quad (3)$$

In Fig. 7, we show the cumulative distributions of this property for both centrals and satellites from each of TNG100 and xGASS-CO. The top panel shows the predicted distributions from TNG100 to lie almost on top of the uncertainty range of xGASS-CO. A Kolmogorov–Smirnov (KS) test of the observational data is enough to rule out the possibility that satellites and centrals follow the same distribution in $\Delta \log_{10}(m_{\mathrm{H}_2}/m_*)$.[15] Not only does this show that the survey

is consistent with the prediction of TNG100 regarding the systematic difference between centrals and satellites, but this also means TNG100 predicts *very* similar scatter in H₂ fractions to what xGASS-CO observes for both satellites and centrals. The bottom panel of Fig. 7 highlights how different the 'true' distributions of H₂ fractions are for TNG100 satellites. We include this panel for completeness, but remind the reader that the top panel is the definitive comparison to be made with observations.

As has already been discussed here and in many other works, H₂ in satellite galaxies is expected to be stripped by ram pressure and/or tides systematically later after infall than H I. Based on this idea, one might expect that the H₂/H I mass ratios of satellites should be higher than centrals of the same stellar mass on average (although, this implicitly neglects other environmental and secular effects that might alter the H I and H₂ reservoirs of satellites differently). We therefore compare the H₂/H I mass ratios of TNG100 centrals and satellites with xGASS-CO data in Fig. 8. At first glance, the median $m_{\mathrm{H}_2}/m_{\mathrm{H\,I}}$ points appear to be higher for satellites than centrals in xGASS-CO, albeit with potential overlap in some bins. To confirm this, we can follow the same procedure as we did for m_{H_2}/m_*, i.e. we can define a quantity $\Delta \log_{10}(m_{\mathrm{H}_2}/m_{\mathrm{H\,I}})$ that is analogous to Equation (3), swapping all instances of 'm_{H_2}/m_*' with '$m_{\mathrm{H}_2}/m_{\mathrm{H\,I}}$'. While we have omitted a figure analogous to Fig. 7 for $\Delta \log_{10}(m_{\mathrm{H}_2}/m_{\mathrm{H\,I}})$ for brevity, we can again rule out the possibility that centrals and satellites have consistent distributions of $\Delta \log_{10}(m_{\mathrm{H}_2}/m_{\mathrm{H\,I}})$ through a KS test. This holds regardless of whether we include non-detections at their upper limit or exclude them entirely.[16] The data therefore support the theory that satellites' H₂/H I mass ratios should be elevated relative to centrals.

Contrary to expectation, the mock-observed H₂/H I mass ratios of TNG100 satellites are *lower* than centrals on average, as seen in the top panel of Fig. 8. This is, however, merely a reflection of the mocking procedure. The applied beam for H I is typically greater in size than the extent of the simulated galaxies' H I, while the opposite is true for H₂. For satellites in particular, this artificially drives $m_{\mathrm{H\,I}}$ up (see S19), and thus drives the $m_{\mathrm{H}_2}/m_{\mathrm{H\,I}}$ ratio down. This is clarified by the bottom panel of Fig. 8, where the inherent $m_{\mathrm{H}_2}/m_{\mathrm{H\,I}}$ ratios of TNG100 galaxies are shown. For fixed H I/H₂ prescription, satellites in fact have systematically higher inherent $m_{\mathrm{H}_2}/m_{\mathrm{H\,I}}$ ratios than centrals of the same mass. This difference is notably greater at $m_* \lesssim 10^{10} \, \mathrm{M}_\odot$ than at higher stellar masses, going from $\sim$0.2 dex to effectively zero eventually.

4.2 Effect of parent halo mass on satellites

Let us now assess the influence of environment on H₂ in a more quantitative sense. Several metrics exist for quantifying galaxy environment. We use host halo mass as our metric, because (i) this is consistent with our previous work with

---

[15] The two-sided, two-sample KS statistic for the two xGASS-CO dot-dashed curves in the top panel of Fig. 7 is 0.31, with a *p*-value of order 10^{-3}.

[16] The two-sided, two-sample KS statistic of $\Delta \log_{10}\left(m_{\mathrm{H}_2}/m_{\mathrm{H\,I}}\right)$ for xGASS-CO centrals versus satellites when we include (exclude) non-detections in CO and H I is 0.31 (0.17), with a corresponding *p*-value of order 10^{-6} (10^{-2}). These values are for the uncorrected H₂ masses, but including beam corrections does not lead to significant difference.
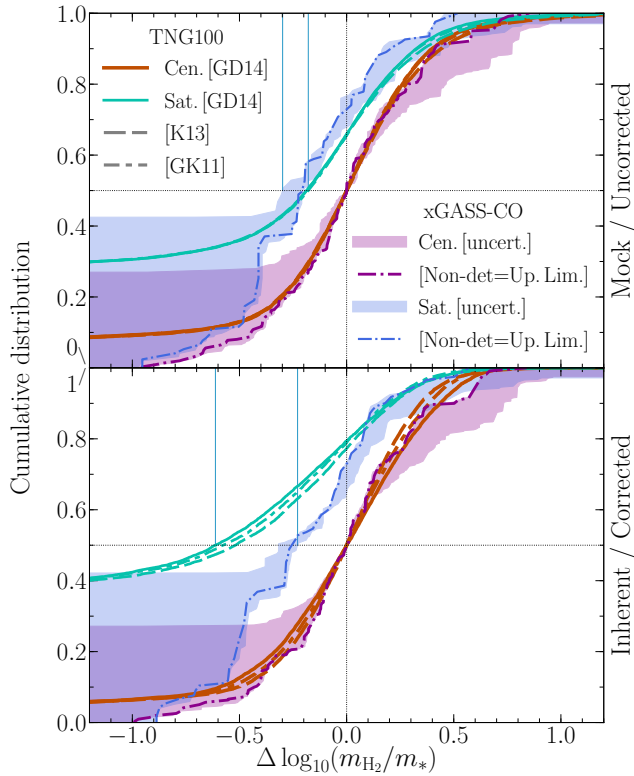
Figure 7: Cumulative distributions of the H_2 fractions of galaxies *relative* to the median H_2 fraction of central galaxies at the same stellar mass (from the same sample; see Equation 3). Distributions for each $H\,\textsc{i}/H_2$ prescription are plotted for TNG100; they are nearly all identical in the top panel. Dot-dashed curves for xGASS-CO assume the upper limits as the H_2 fractions of non-detections. The shaded regions for xGASS-CO show the uncertainty range, accounting for (i) the possibility that non-detections might have no H_2 at all and (ii) the uncertainty on the median H_2 fraction of centrals in each stellar-mass bin (as seen in Fig. 6). Each galaxy in xGASS-CO has a weighted contribution based on the relative frequency of galaxies at its stellar mass in the survey to that in TNG100 (Section 2.2), where weights for centrals and satellites are calculated independently. Lower-mass galaxies therefore dominate these distributions (by virtue of the shape of the stellar mass function). Systematic differences between centrals and satellites are readily identified by the horizontal separation of their respective curves where they intercept the horizontal, dotted line; the vertical, solid lines show the range of these intercepts for satellites from both TNG100 and xGASS-CO. By definition, all curves for centrals intersect at $(0.0, 0.5)$, i.e. where the dotted lines also intersect. Per the top panel, TNG100 predicts that xGASS-CO should find a systematic difference of $\sim0.2\,\mathrm{dex}$ between satellites and centrals. Indeed, xGASS-CO is consistent with this prediction, finding a difference of 0.2–$0.3\,\mathrm{dex}$ within uncertainty. The bottom panel is included for completeness, but — unlike the top panel — it does not provide a fair comparison between TNG100 and xGASS-CO, as no accounting for the cross-contamination of centrals and satellites has been made (see Section 4.1 for details).

TNG100 (Stevens et al. 2019a,b) and (ii) observational trends are evidently stronger with this metric than other popular alternatives, such as those that use the distance to the Nth nearest neighbour (Brown et al. 2017).

In Fig. 9, we break TNG100 satellites into four bins of parent halo mass (in line with S19), plotting the mean H_2 frac-
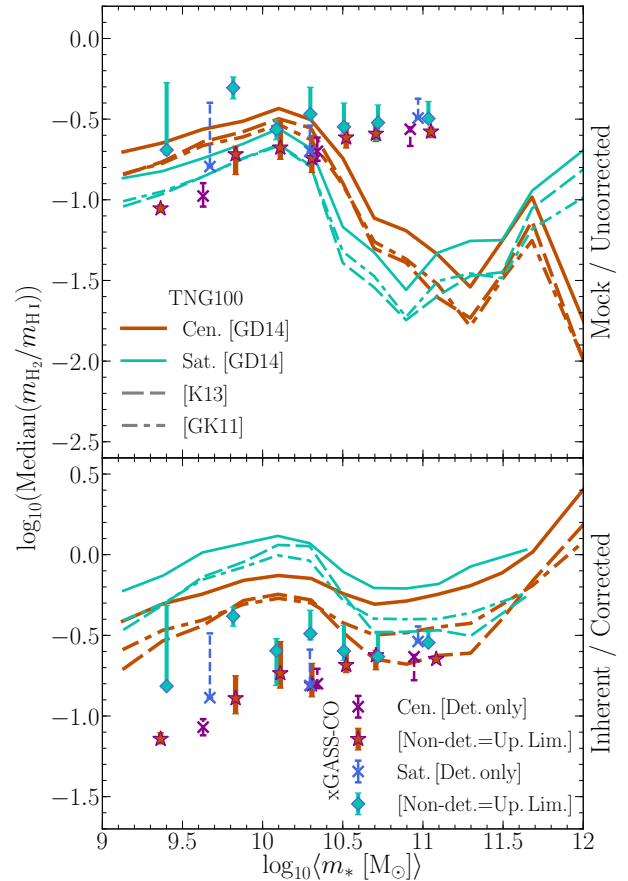
Figure 8: Per Fig. 5, now exclusively showing median trends, with satellites and centrals separated. Error bars on the xGASS-CO data are 68% confidence intervals *on the median* from bootstrapping. The same stellar-mass bins are used for centrals and satellites, ensuring there are at least 15 galaxies of each type in each bin. The low number of detected satellites in both xGASS and xCOLD GASS meant reducing the number of bins from seven with the inclusion of non-detections to three for the 'detection only' data.

tion as a function of stellar mass in the top panel. For a given $H\,\textsc{i}/H_2$ prescription, there is a clear decline in m_{H_2} in higher-mass haloes (denser environments) at fixed stellar mass; satellites hosted in haloes with $M_{200c} \geq 10^{14}\,\mathrm{M}_\odot$ have a factor of ~10 less H_2 than those in haloes with $M_{200c} < 10^{12}\,\mathrm{M}_\odot$. This result is very similar to that found for $H\,\textsc{i}$ in S19 (see their fig. 6), implying that environmental effects in TNG100 have a clear and comparable signature on both atomic and molecular gas.

Despite the signature of environment appearing similar for both H_2 and $H\,\textsc{i}$ in TNG100, we know from Section 4.1 that the $H_2/H\,\textsc{i}$ mass ratios of satellite galaxies in TNG100 are raised relative to centrals of the same stellar mass. For this to be consistent, we should expect the average $H_2/H\,\textsc{i}$ mass ratios of TNG100 satellites to have a residual dependence on parent halo mass as well. Indeed, this dependence is present, as shown in the middle panel of Fig. 9. The quantity on the y-axis of this panel has a somewhat subtle but deliberate definition. For satellites in each halo- and stellar-mass bin, we measure the mean H_2 and mean $H\,\textsc{i}$ mass, take the ratio of those means, log it, and from this subtract the same log
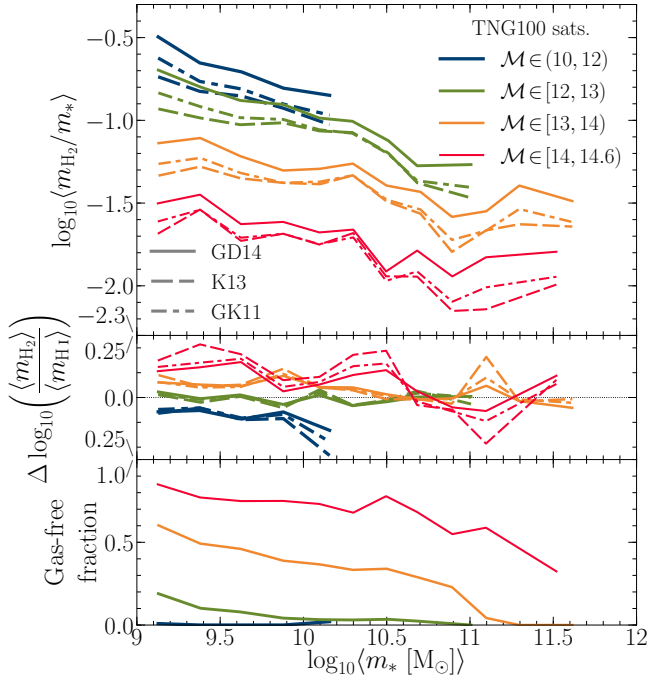
Figure 9: Top panel: Mean inherent H_2 fraction of TNG100 satellites as a function of stellar mass for the same bins of parent halo mass used in S19, where $\mathcal{M} \equiv \log_{10}\left(M_{200c}^{\mathrm{parent}}/M_\odot\right)$. Line colour and thickness relate to halo mass, while line style relates to the HI/H_2 prescription. **Middle panel:** Ratio of the mean inherent H_2 mass to mean inherent HI mass for satellites in each given halo mass bin, *normalised* by the same ratio for *all* satellites at the same stellar mass (see Equation 4). The y-scale has been stretched relative to the top panel for clarity. **Bottom panel:** Fraction of satellites for each halo mass bin that have no gas cells associated with them (and thus have $m_{HI} = m_{H_2} = 0$). The dependence of H_2 fraction on halo mass in the top panel is largely driven by this. HI/H_2 prescriptions are irrelevant for this panel.

ratio for all satellites at the same stellar mass. That is,

$$\Delta \log_{10}\left(\frac{\langle m_{H_2}\rangle}{\langle m_{HI}\rangle}\right) \equiv \log_{10}\left(\frac{\langle m_{H_2}(m_*,\mathrm{sat.},M_{200c})\rangle}{\langle m_{HI}(m_*,\mathrm{sat.},M_{200c})\rangle}\right)$$
$$- \log_{10}\left(\frac{\langle m_{H_2}(m_*,\mathrm{sat.})\rangle}{\langle m_{HI}(m_*,\mathrm{sat.})\rangle}\right). \quad (4)$$

This definition removes systematic differences between the HI/H_2 prescriptions (which can already be seen in the top panel) and allows for all galaxies with gas masses of zero to be naturally included [the latter would *not* be the case were we to plot $\langle m_{H_2}/m_{HI}\rangle$ or $\mathrm{Median}(m_{H_2}/m_{HI})$ instead, for example]. Evidently, parent halo mass only plays a noteworthy role in this quantity for $m_* < 10^{10.5}\,\mathrm{M_\odot}$, and its role is rather noisy for $m_* \gtrsim 10^{9.7}\,\mathrm{M_\odot}$. The former resembles the fact that high-mass satellites are generally less prone to stripping, as they can often be comparable in size to their central. Even for $m_* < 10^{9.5}\,\mathrm{M_\odot}$, $\Delta \log_{10}(\langle m_{H_2}\rangle/\langle m_{HI}\rangle)$ of satellites in the densest environments differs from those in the lowest-mass haloes by only $\sim 0.2\,\mathrm{dex}$.

What really drives the separation (or lack thereof) in the lines of different halo mass in the top two panels of Fig. 9 is the fraction of galaxies that lack gas entirely. This fraction is plotted in the bottom panel. High-density environments in

TNG100 are eventually able to fully remove the gas reservoirs of satellites. An increase in the 'gas-free' fraction has a clear signature in mean H_2 mass, while making it difficult for there to be significant differences between HI and H_2. We should be clear that stripping is not actually binary in the simulation. Rather, once the externally forced removal of gas reduces a galaxy to small number of cells, resolution effects take over, dooming the galaxy to lose its remaining few cells. Up until this point, gas should be stripped smoothly.

Previous studies of galaxies' cold-gas content across environments have typically included much larger samples of observed galaxies than we have presented with xGASS-CO (e.g. Stark et al. 2016; Brown et al. 2017; Stevens & Brown 2017); the important difference is they only had HI data on those galaxies. With only 109 xGASS-CO satellites (including non-detections in CO), it is difficult to split the sample into bins of parent halo mass without losing statistical significance. We therefore leave the results of Fig. 9 as predictions for future observations to test.

4.3 Post-infall changes to satellites' gas

A primary utility of cosmological simulations is that galaxies can be tracked through time: something that is observationally impossible for real galaxies. With TNG100, we can directly see how the HI and H_2 reservoirs of galaxies change after infall (i.e. once they have become satellites). To this end, we select a sample of TNG100 satellite galaxies at $z=0$ with $m_* < 10^{10.2}\,\mathrm{M_\odot}$ (avoiding AGN complications) that reside within $1.5\,R_{200c}$ of their parent halo, have been a satellite for $< 3\,\mathrm{Gyr}$, and have an HI mass that is between 0.4 and $2.0\,\mathrm{dex}$ below the median for galaxies with the same stellar half-mass radius. This last criterion means there is a nominal expectation that some stripping should have taken place (as the galaxies are 'HI deficient'—e.g. Boselli et al. 2014), which we can test, while also ensuring they still have some gas left. The requirement to have only been a satellite for $3\,\mathrm{Gyr}$ minimizes any potential changes to the galaxies' gas properties associated with morphological transformation (see the related discussions of Cortese et al. 2019; Joshi et al. 2020). We track those galaxies back through the baryonic SUBLINK merger trees (Rodriguez-Gomez et al. 2015) for TNG100, calculating their HI and H_2 content at each snapshot.[17]

In the top panel of Fig. 10, we show the ratio of the sample satellites' HI mass at infall to that at redshift zero, versus the same ratio for H_2. Compared in the bottom panel are three example evolutionary tracks of galaxies from their time of infall. We define the 'time of infall' here as that of the first snapshot at which the galaxy was both a satellite according to SUBFIND *and* within $1.5\,R_{200c}$ of its host halo. We chose $1.5\,R_{200c}$ rather than $1.0\,R_{200c}$ because environmental effects are not confined to R_{200c}; in fact, we could have selected an even larger radius, as environmental effects can extend out

---

[17] We note that not all snapshots have the fields saved that we require as input for our gas-phase decomposition methods. We have therefore approximated the values for these missing properties, as described in Appendix B, which we apply to *all* snapshots for Fig. 10 (exclusively) to ensure we can fairly calculate *changes* to the satellites' gas content over time. The uncertainty introduced by these approximations is entirely negligible.
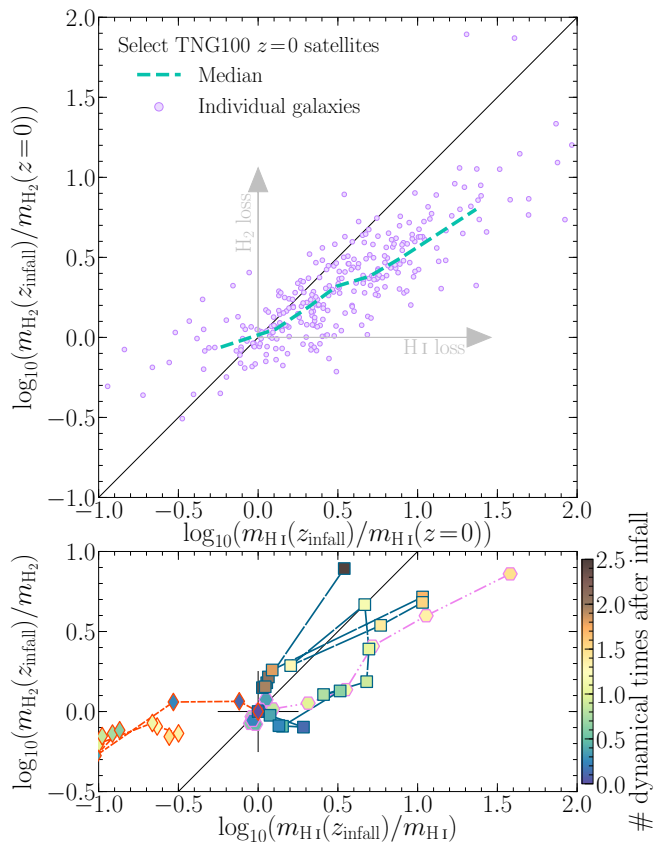
Figure 10: Change in H I mass versus change in H$_2$ mass of TNG100 satellites since the first time they crossed $1.5\,R_{200\mathrm{c}}$ of their parent halo (deemed the time of 'infall'). H I (H$_2$) losses after infall result in rightward (upward) movement in this space. Satellites were selected at $z=0$ to have $m_* < 10^{10.2}\,\mathrm{M}_\odot$, have an 'H I deficiency' of $0.4\,\mathrm{dex}$ (for their stellar half-mass radius), be within $1.5\,R_{200\mathrm{c}}$ of their parent halo, and have been a SUBFIND satellite for no more than $3\,\mathrm{Gyr}$. **Top panel:** Points for individual galaxies and the running median are shown for the sample as it is at $z=0$. **Bottom panel:** Connected points of varied shape show tracks for three example galaxies after crossing $1.5\,R_{200\mathrm{c}}$, coloured by time after infall. The 'dynamical time' for each satellite is calculated as $[10\,H(z_{\mathrm{infall}})]^{-1}$. By definition, these tracks all start at $(0,0)$ (the centre of the crosshair).

to $5\,R_{200\mathrm{c}}$ in clusters (see Bahé et al. 2013; Behroozi et al. 2014; Ayromlou et al. 2019). Galaxies that have lost H I since infall will move to the right, while those that have lost H$_2$ will move upward.

Fig. 10 makes it clear that there is inhomogeneous evolution in satellites' H I and H$_2$ content after infall. A relatively simple example of post-infall evolution is seen by the galaxy represented by hexagons in the bottom panel of Fig. 10; this galaxy maintains its gas for several snapshots before gradually losing H I and H$_2$ at an increasing rate, moving towards the top-right of the panel. Visual inspection of this galaxy (not shown here) clarifies that it is experiencing ram-pressure stripping. The other two galaxies in the panel do not evolve so simply, reminding us that there are many physical processes that take place after infall, not just environmental effects. The galaxy represented by the diamonds was in the process of experiencing a minor merger during infall. This merger brought gas into the galaxy, causing both $m_{\mathrm{H\,I}}$ and m_{H_2} to

increase, shifting the galaxy to the bottom-left of the panel. After this, its gas started to be stripped, causing a reversal in its parameter-space trajectory. The galaxy represented by the squares began to be stripped after infall, but twice collided with intervening gas in the intrahalo medium (which likely would have otherwise been accreted by the central galaxy of the halo). In the first instance, the galaxy was unable to hold the gas for longer than a single snapshot interval. In the second instance, where the galaxy was closer to pericentre, it successfully accreted the gas it collided with, rejuvenating the galaxy for the better part of a dynamical time, before being stripped once again.

Despite the diversity of post-infall changes to satellites' gas, an underlying trend presents itself. This is readily identified by the running median in the top panel of Fig. 10, which is shallower than the 1:1 line, highlighting that losses in H I are generally more pronounced than in H$_2$. In S19, we argued that H$_2$ in TNG100 satellites might be stripped too efficiently, based on their SFRs being too strongly coupled to their H I content. Here, we revise that interpretation somewhat; the preference for stripping H I appears to be present, but it may well be weaker than in the real Universe.

It is important to emphasize that H$_2$ need not be lost from satellites specifically by direct *stripping*. The three examples we described above were selected to be diverse rather than representative. For other galaxies, it could feasibly be that H I is lost from ram pressure, but H$_2$ is consumed in star formation and simply unable to replenish (e.g. Stevens & Brown 2017). However, the facts that (i) stripping manifests as the complete removal of satellites' gas in the densest environments in many cases (bottom panel of Fig. 9), (ii) depletion of H I in TNG100 satellites has a parallel effect on SFR (S19), and (iii) hydrodynamical forces in the simulation have no explicit knowledge of whether gas is predominantly atomic or molecular (as the phase decomposition is calculated in post-processing) imply that stripping affects all gas phases in the simulation. Regardless of the physical mechanism(s) at play, Fig. 10 definitively shows us that H I is more readily depleted than H$_2$ (on average) once galaxies become satellites in TNG100.

4.4 Star formation as a proxy for molecular gas

A wealth of empirical evidence points to a tight connection existing between the amount of molecular gas in a galaxy and its star formation rate, applying not only to integrated properties (captured by a common H$_2$ depletion time — e.g. Saintonge et al. 2017; Tacconi et al. 2018), but also to localised surface densities (Bigiel et al. 2008; Leroy et al. 2008). After all, both molecule formation and star formation should occur in relatively dense, cold gas clouds. Observational measurements of galaxies' star formation rates are orders of magnitude more numerous than those of galaxies' H$_2$ content in the local Universe. Comparing the star formation activity of TNG galaxies against those numerous data provides a less direct, but more statistically meaningful test of the simulated galaxies' molecular-gas content and its relation to galaxy environment.

Because the H I/H$_2$ modelling was done in post-processing, the star formation law applied in TNG and the H$_2$ fractions we derive for the simulated galaxies do not have a precise one-to-one connection. Unsurprisingly however, one does still find
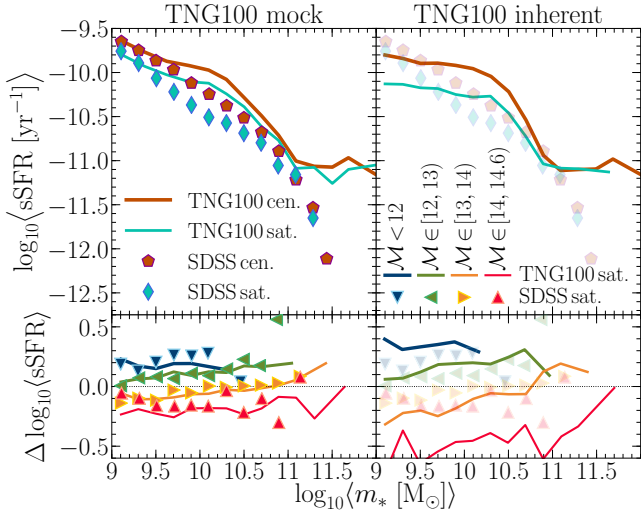
Figure 11: Influence of environment on the stellar mass–specific star formation rate relation for TNG100 and SDSS galaxies. The top panels show the running mean sSFR for centrals and satellites separately. The left column measures TNG100 properties in a more appropriate manner for directly comparing to SDSS (as described in Sections 2.4.2 and 2.4.3). The right-hand column shows the inherent results for TNG100. Faded SDSS median points are shown for reference, but are not intended as a fair comparison in the right-hand panels. The bottom panels break satellites into bins of parent halo mass, where the *y*-axis now shows the *difference* in the mean sSFR of satellites in that halo mass bin from the total satellite population of the respective dataset (see Equation 5). Note that the scale of the *y*-axis in this panel has been stretched for clarity.

that TNG galaxies' star formation rates and H_2 masses are closely correlated (Diemer et al. 2019), as is the case in other cosmological, hydrodynamic simulations (Lagos et al. 2015; Rodríguez Montero et al. 2019; Davé et al. 2020). Comparison with xCOLD GASS shows that the SFR–m_{H_2} relationship in TNG is consistent with the real, local Universe (fig. 4 of Diemer et al. 2019). We have further found that there is no environmental dependence on this relationship in TNG100 (not shown here), which is important for the below.

To provide context for the results that follow, we present and discuss the m_*–SFR plane of TNG100 galaxies in Appendix A1. We also refer the reader to the closely related work by Donnari et al. (2020, 2021), where the quenched fractions of TNG galaxies are investigated in detail, including environmental influences.

In the two large panels of Fig. 11, we show running means separately for satellites' and centrals' sSFRs as a function of stellar mass, for SDSS and both the mock (left panels, as described in Section 2.4.2) and inherent (right-hand panels) properties of TNG100. The left panels remain the definitive comparison with SDSS though. It is clear that satellites have systematically lower sSFRs than centrals of the same stellar mass (or, rather, a greater fraction of satellites are passive), both in the observed and simulated universes. Despite TNG100 generally having a smaller quenched population at most stellar masses (as can be inferred from Fig. A1) — which leads to differences in the shape of the mean m_*–SFR curves between the TNG100 mock predictions and SDSS — both datasets show similar *relative* behaviour between cen-

trals and satellites. In particular, we can quantify the typical central–satellite separation in ⟨sSFR⟩ across the full stellar-mass range as 0.19 and 0.15 dex for SDSS and TNG100, respectively.

The key result of this subsection can be found in the bottom-left panel of Fig. 11, where we split satellites by environment (quantified by their parent halo mass), and show the *difference* between the mean sSFR of those satellites and that of *all* satellites at the same stellar mass:

$$\Delta \log_{10}\langle sSFR \rangle \equiv \log_{10}\left(\frac{\langle sSFR(m_*, sat., M_{200c}) \rangle}{\langle sSFR(m_*, sat.) \rangle} \right). \quad (5)$$

At fixed stellar mass, the mean sSFR of satellites has a visible dependence on parent halo mass both in SDSS and TNG100 (for a directly related assessment of SDSS, see Wetzel, Tinker & Conroy 2012). If we are to trust that sSFR is a fair proxy for H_2 fraction, then Fig. 11 provides evidence — in addition to and independent from Fig. 6 — that environment-driven losses of H_2 in TNG are in quantitative agreement with the real Universe. Comparison of the two bottom panels of Fig. 11 highlights how, between satellites hosted in haloes with $M_{200c} < 10^{12}\,M_\odot$ and those with $M_{200c} \geq 10^{14}\,M_\odot$, a 'true' difference in ⟨sSFR⟩ of almost an order of magnitude can be reduced to a factor of ∼2.5 (∼0.4 dex) through the interpretation of and uncertainties associated with observational data. Comparison of the bottom-right panel of Fig. 11 with the top panel of Fig. 9 also highlights that sSFR and H_2 are indeed very similarly affected by environment in TNG100.

5 ELEPHANTS SWEPT UNDER THE RUG

Before summarizing the findings of this paper, it is worthwhile touching on the underlying limitations at play, the likes of which do not always receive due attention.

5.1 Resolution limitations

Simulation outcomes are always at the mercy of resolution. Of particular relevance here is the effect this can have on gas stripping in satellites, both from tides and ram pressure. For example, a finite resolution limits the accuracy of hydrodynamical-force calculations between the interstellar medium of a satellite and the intrahalo medium it moves through. While idealized simulations with AREPO suggest that the moving-mesh hydrodynamics scheme generally captures effects of ram pressure and induced Kelvin–Helmholtz instabilities more accurately than traditional SPH[18] codes (Springel 2010; Heß & Springel 2012; Sijacki et al. 2012), the median neutral gas cell number (neutral-gas mass divided by the baryonic mass resolution of $1.4 \times 10^6\,M_\odot$) for TNG100 galaxies used in our study is only ∼1200. The gas in many of the TNG100 galaxies is therefore more poorly resolved than the regimes of these tests. The high fraction of gas-free satellites in TNG100 discussed in Section 4.2 could potentially be influenced by this at some level.

Furthermore, the potential-well depth of galaxies can be affected by a host of numerical effects, including gravitational

---

[18] Smoothed-Particle Hydrodynamics

softening, the unequal mass partitioning of dark matter and baryons, the imposed density threshold for star formation, and the implementation of feedback (see the recent results of Dutton et al. 2019; Ludlow et al. 2019, 2020). This can artificially alter how susceptible a galaxy is to being stripped. For example, we raised in Section 2.4 that the inner 2 kpc of TNG100 galaxies are effectively 'unresolved'. Yet we showed a galaxy in Fig. 2 with an H_2 half-mass radius of 2.2 kpc, scarcely larger than the 'resolution limit'. Fig. 1 also shows that mock-observed H_2 masses can be sensitive to as little as the inner few kpc for a fraction of low-mass galaxies in TNG100. These examples are by no means representative of all galaxies, but they anecdotally highlight that we are pushing TNG100 close to the limit of its capabilities. In these situations, it also becomes especially difficult to capture environment-driven gas compression, where ram pressure might not only strip gas, but also lead to the conversion of H I into H_2 (e.g. Henderson & Bekki 2016), which empirical evidence shows can happen (e.g. Lee et al. 2017a).

To elaborate on the non-trivial topic of resolution-based force limitations in the context of TNG would be worthy of a paper by itself, if not several. Where me may glean some insight regarding the impact of these limitations in future is through applying a similar analysis in this work (and S19) to the higher-resolution TNG50 simulation (Nelson et al. 2019b; Pillepich et al. 2019).

5.2 Post-processing approximations

Our post-processing models for performing the H I/H_2 breakdown in TNG rely on common assumptions. Namely, (i) they all assume some degree of equilibrium; (ii) they rely on the same principle ingredients: local density, temperature, metallicity, and UV intensity; (iii) they are inherently two-dimensional models that have been translated to three dimensions using the imprecise Jeans approximation; and (iv) they are instantaneous in their application, meaning it is possible for discontinuities in the H_2/H I mass ratio of a given galaxy to occur from snapshot to snapshot. The lattermost of these could artificially amplify the 'random walk' nature of the tracks shown in the bottom panel of Fig. 10. Many of the limitations and assumptions of our H I/H_2 modelling are discussed in greater detail in Diemer et al. (2018).

5.3 Observational uncertainties

It should not be forgotten that observational data also carry potentially important limitations too. Most notably, uncertainties in observed H_2 masses are dominated by systematics associated with the CO-to-H_2 conversion factor. While the conversion factor adopted for xCOLD GASS is physically motivated and empirically calibrated (Accurso et al. 2017), it principally depends on metallicity, and secondarily depends on sSFR relative to the main sequence (which traces the UV field strength). Ideally, the conversion would be done on local scales, but it is instead based on the galaxies' 'integrated' properties, where 'integrated' really means 'inferred from the 3-arcsec SDSS fibre'. The CO-to-H_2 conversion therefore implicitly assumes that the gas-phase metallicity and UV field are uniform in a given galaxy, which is almost certainly false. As discussed in Section 2.4.2, the SDSS fibre size is far smaller

than the beam size used to measure CO, and smaller still relative to the total extent of the xCOLD GASS galaxies. Moreover, the diagnostics used to infer metallicity (oxygen abundance) from the SDSS emission line ratios carry significant systematic uncertainty themselves (see the review by Kewley, Nicholls & Sutherland 2019). Further systematics lie in the assumption that the gas-phase metallicity traces the dust-to-gas ratio (incidentally, this also applies to the H I/H_2 prescriptions used for the simulation).

We note that we neglected to forward-model any H_2 mass uncertainties in our simulation predictions. Had we done this, the scatter in H_2 fractions in the 'mock' results would have increased to reflect the typical statistical error in the xCOLD GASS m_{H_2} measurements. Given that systematics are even more important, it matters less that a predicted line goes through the data, and more that it is close and parallel.

In the spirit of forward-modelling, models that can be overlaid on simulations to infer CO emission provide a means to avoid converting CO flux from observations to H_2 mass entirely. While this requires thoughtful consideration of radiative transfer and chemistry to model in detail (see e.g. Glover et al. 2010), even relatively straightforward methods to forward-model CO emission from cosmological-simulation galaxies can lead to different conclusions than H_2-based comparisons with observations (Davé et al. 2020). Some level of systematic uncertainty will likely still be present, but as with the mock-observation strategy employed in this paper, such a method is at least self-consistent and circumvents assumptions regarding structure and gradients in the observed galaxies.

6 IN A NUTSHELL

We have studied the molecular-gas content of galaxies with $m_* \geq 10^9 \, M_\odot$ at $z = 0$ in the TNG100 simulation, focussing on how it depends on galaxy environment. In addition to taking advantage of earlier work that developed and applied the post-processing framework to derive the H I/H_2 properties of each gas cell in the simulation (Diemer et al. 2018; S19), we have 'mock observed' TNG100 galaxies to allow for the most direct comparison possible with data from the xCOLD GASS survey and SDSS. Our main findings can be summarized as follows.

(i) Were we to take the inherent properties of TNG100 and beam corrections of xCOLD GASS at face value, we would conclude that TNG100 galaxies are systematically too H_2-rich across all $m_* \geq 10^9 \, M_\odot$ (by a factor of $\lesssim 2$ at $m_* \lesssim 10^{10} \, M_\odot$, and greater at higher masses). Most of the tension at low mass is relieved when comparing the mock-observed H_2 masses with the uncorrected xCOLD GASS masses instead, where the former includes the application of a Gaussian beam whose FWHM (consistent with the IRAM 30-m telescope) is smaller than the full extent of H_2 in the galaxies. As outlined in Section 3 and evidenced by Fig. 4, the forward-modelling approach of this mock/uncorrected comparison with observations is more meaningful. At $m_* > 10^{10.5} \, M_\odot$, the relative size of the mock beam is *so* much smaller than the extent of molecular gas that inferred H_2 fractions in TNG100 drop by an order of magnitude. This result is summarized in Fig. 3 with evidence for its reasoning seen in Fig. 2. Sim-

ilar statements can be made about the simulated galaxies' m_{H_2}/m_{H_I} ratios (Fig. 5).

(ii) The onset of AGN feedback in TNG100 galaxies with $m_* \gtrsim 10^{10.5}\,M_\odot$ disrupts a significant fraction of gas discs, sometimes tearing them open from the inside out, displacing much of their neutral gas (e.g. column 4 of Fig. 2). This is more noticeable in satellite galaxies (Fig. 6), which are less able to reaccrete their displaced gas than centrals, and are also more likely to have had merger-induced black-hole accretion in their histories by virtue of their being more clustered on average. This feature was also identified in H I in S19. While the act of mock observation evidently washes this feature out, the finding that the H_2 half-mass radii of high-mass TNG100 galaxies are several times their stellar half-mass radii (when they should be comparable — cf. Bolatto et al. 2017; Diemer et al. 2019; Fig. 4 of this paper) reaffirms that the TNG black-hole feedback model is too efficient at displacing gas as galaxies start to quench.

(iii) In TNG100, the median satellite is ~ 0.6 dex poorer in H_2 than the median central of the same stellar mass, per the simulated galaxies' inherent properties. However, in mock-observing the galaxies, we predicted that this signature would only be measured as ~ 0.2 dex in xCOLD GASS. The most significant aspect of the mocking that causes this smaller signature is the cross-contamination of centrals and satellites in the Yang et al. (2007) group catalogue. In comparing the distribution of xCOLD GASS galaxies' H_2 fractions relative to the median of centrals at the same stellar mass, we showed that xCOLD GASS confirms our prediction (Fig. 7).

(iv) We found that the mean H_2 fraction of TNG100 satellites drops by an order of magnitude when comparing those in host haloes with $M_{200c} < 10^{12}\,M_\odot$ to satellites in hosts with $M_{200c} > 10^{14}\,M_\odot$ (Fig. 9). While we could not compare this in a statistically significant way with xCOLD GASS, we used sSFR as a proxy for H_2 fraction to compare this environmental dependence with a sample of SDSS galaxies. Again through tailored mock-observing, we predicted that SDSS should observe a ~ 0.4 dex difference in the mean sSFR of satellites hosted in $M_{200c} < 10^{12}\,M_\odot$ haloes versus those hosted in $M_{200c} > 10^{14}\,M_\odot$ haloes. This is indeed in quantitative agreement with SDSS (Fig. 11).

(v) The canonical picture of environmental stripping of satellite galaxies' gas is that atomic gas (which is weighted towards the outskirts) is removed before molecular gas (which is deeper in the potential well). By tracking a sample of $z = 0$ satellites back to their point of infall in Fig. 10, we found it is indeed generally true for TNG100 galaxies that H I is lost at a higher fractional rate than H_2 after infall (although, plenty of secular effects continue to operate over this period as well). This is also reflected by $z = 0$ satellites having systematically elevated inherent m_{H_2}/m_{H_I} ratios relative to centrals of the same stellar mass (Fig. 8). There is also evidence that the m_{H_2}/m_{H_I} ratios of satellites at fixed stellar mass continue to increase as one looks at higher parent halo masses in TNG100. But it is difficult to find a strong trend for this, as the fraction of satellites that host zero gas cells in the simulation also increases in denser environments; 75% of satellites with $m_* \geq 10^9\,M_\odot$ hosted in haloes with $M_{200c} \geq 10^{14}\,M_\odot$ are devoid of gas entirely. The cleanest trend we find is at $m_* \lesssim 10^{10}\,M_\odot$, where the ratio of the mean H_2 mass to mean H I mass of satellites is systematically ~ 0.2 dex higher in haloes with $M_{200c} \geq 10^{14}\,M_\odot$ than

those with $M_{200c} < 10^{12}\,M_\odot$ (Fig. 9). This is a prediction that could be tested with stacking of CO and H I spectra targeted at the right galaxies.

To study environmental effects on galaxies' H_2 in greater detail, a large-scale, blind CO survey(s) is ultimately needed. While the near future is bright for H I surveys of this nature, which will lead to a wealth of environment studies (among other topics — e.g. Adams & van Leeuwen 2019; Koribalski et al. 2020; Maddox et al. 2020), equivalents for CO will inevitably lag behind. In the interim, to further advance observational tests of the environmental dependence of galaxies' H_2 from theoretical models and simulations, we might be better off targeting a handful of galaxy clusters, and studying their members in finer detail. Recently completed surveys like VERTICO (Brown et al. in prep.) will prove useful in this regard, offering a three-dimensional view of 51 Virgo cluster galaxies in CO emission.

DATA AVAILABILITY

The IllustrisTNG simulations are publicly available at https://www.tng-project.org/data/, with a full description of those data provided by Nelson et al. (2019a). The H I/H_2 properties in the public database are not identical to those used in this paper, but they are based on the same methodology and are known to agree well (see section 2.3 of S19). Integrated galaxy properties used in this paper are also not directly provided in the public database, but are recoverable through the particle data in principle. Otherwise, data used in this paper can be made available upon request.

The xCOLD GASS survey data are publicly available at http://www.star.ucl.ac.uk/xCOLDGASS/data.html. Similarly, data for xGASS are available at https://xgass.icrar.org/data.html.

SDSS data from the MPA–JHU catalogue are available at https://wwwmpa.mpa-garching.mpg.de/SDSS/DR7/, with improved stellar masses founds at http://home.strw.leidenuniv.nl/~jarle/SDSS/.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Accurso G. et al., 2017, MNRAS, 470, 4750
Adams E. A. K., van Leeuwen J., 2019, Nat. Astron., 3, 188

Ayromlou M., Nelson D., Yates R. M., Kauffmann G., White S. D. M., 2019, MNRAS, 487, 4313

Bahé Y. M., McCarthy I. G., Balogh M. L., Font A. S., 2013, MNRAS, 430, 3017

Bahé Y. M. et al., 2016, MNRAS, 456, 1115

Behroozi P. S., Wechsler R. H., Lu Y., Hahn O., Busha M. T., Klypin A., Primack J. R., 2014, ApJ, 787, 156

Bigiel F., Blitz L., 2012, ApJ, 756, 183

Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846

Blanton M. R., Moustakas J., 2009, ARA&A, 47, 159

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

Bolatto A. D. et al., 2017, ApJ, 846, 159

Boselli A., Cortese L., Boquien M., Boissier S., Catinella B., Gavazzi G., Lagos C., Saintonge A., 2014, A&A, 564, A67

Boselli A. et al., 2019, A&A, 631, A114

Bravo M., Lagos C. d. P., Robotham A. S. G., Bellstedt S., Obreschkow D., 2020, MNRAS, 497, 3026

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

Brown T. et al., 2017, MNRAS, 466, 1275

Campbell D., van den Bosch F. C., Hearin A., Padmanabhan N., Berlind A., Mo H. J., Tinker J., Yang X., 2015, MNRAS, 452, 444

Casoli F., Boissé P., Combes F., Dupraz C., 1991, A&A, 249, 359

Castignani G. et al., 2018, A&A, 617, A103

Catinella B. et al., 2010, MNRAS, 403, 683

Catinella B. et al., 2013, MNRAS, 436, 34

Catinella B. et al., 2018, MNRAS, 476, 875

Chabrier G., 2003, PASP, 115, 763

Chang Y.-Y., van der Wel A., da Cunha E., Rix H.-W., 2015, ApJS, 219, 8

Chung A., van Gorkom J. H., Kenney J. D. P., Crowl H., Vollmer B., 2009, AJ, 138, 1741

Coogan R. T. et al., 2018, MNRAS, 479, 703

Cora S. A. et al., 2018, MNRAS, 479, 2

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

Cortese L. et al., 2019, MNRAS, 485, 2656

Crain R. A. et al., 2017, MNRAS, 464, 4204

Cramer W. J., Kenney J. D. P., Cortes J. R., Cortes P. C., Vlahakis C., Jáchym P., Pompei E., Rubio M., 2020, ApJ, 901, 95

Darvish B., Scoville N. Z., Martin C., Mobasher B., Diaz-Santos T., Shen L., 2018, ApJ, 860, 111

Dannerbauer H. et al., 2017, A&A, 608, A48

Davé R., Katz N., Oppenheimer B. D., Kollmeier J. A., Weinberg D. H., 2013, MNRAS, 434, 2645

Davé R., Crain R. A., Stevens A. R. H., Narayanan D., Saintonge A., Catinella B., Cortese L., 2020, MNRAS, 497, 146

Davies L. J. M. et al., 2019, MNRAS, 483, 5444

Diemer B. et al., 2018, ApJS, 238, 33

Diemer B. et al., 2019, MNRAS, 487, 1529

Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497

Donnari M. et al., 2019, MNRAS, 485, 4817

Donnari M., Pillepich A., Nelson D., Marinacci F., Vogelsberger M., Hernquist L., 2020, preprint (arXiv:2008.00004)

Donnari M. et al., 2021, MNRAS, 500, 4004

Dutton A. A., Macciò A. V., Buck T., Dixon K. L., Blank M., Obreja A., 2019, MNRAS, 486, 655

Genel S. et al., 2014, MNRAS, 445, 175

Giovanelli R. et al., 2005, AJ, 130, 2598

Glover S. C. O., Federrath C., Mac Low M.-M., Klessen R. S., 2010, MNRAS, 404, 2

Gnedin N. Y., Draine B. T., 2014, ApJ, 795, 37 (GD14)

Gnedin N. Y., Kravtsov A. V., 2011, ApJ, 728, 88 (GK11)

Gunn J. E., Gott III J. R., 1972, ApJ, 176, 1

Harris C. R. et al., 2020, Nature, 585, 357

Hayashi M. et al., 2017, ApJ, 841, L21

Hayashi M. et al., 2018, ApJ, 856, 118

Haynes M. P., Giovanelli R., Chincarini G. L., 1984, ARA&A, 22, 445

Henderson B., Bekki K., 2016, ApJ, 822, L33

Heß S., Springel V., 2012, MNRAS, 426, 3112

Hunter J. D., 2007, Comput. Sci. Eng., 9, 90

Jablonka P., Combes F., Rines K., Finn R., Welch T., 2013, A&A, 557, A103

Jáchym P. et al., 2019, ApJ, 883, 145

Jaffé Y. L. et al., 2016, MNRAS, 461, 1202

Janowiecki S., Catinella B., Cortese C., Saintonge A., Brown T., Wang J., 2017, MNRAS, 466, 4795

Joshi G. D., Pillepich A., Nelson D., Marinacci F., Springel V., Rodriguez-Gomez V., Vogelsberger M., Hernquist L., 2020, MNRAS, 496, 2673

Kauffmann G. et al., 2003, MNRAS, 341, 33

Kenney J. D. P., Young J. S., 1989, ApJ, 344, 171

Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, ARA&A, 57, 511

Koribalski B. S. et al., 2020, Ap&SS, 365, 118

Kroupa P., 2001, MNRAS, 322, 231

Krumholz M. R., 2013, MNRAS, 436, 2747 (K13)

Lagos C. d. P. et al., 2015, MNRAS, 452, 3815

Larson R. B., Tinsley B. M., Caldwell C. N., 1980, ApJ, 237, 692

Lee B., Chung A., 2018, ApJ, 866, L10

Lee B. et al., 2017a, MNRAS, 466, 1382

Lee M. M. et al., 2017b, ApJ, 842, 55

Leroy A. et al., 2008, AJ, 137, 2782

Ludlow A. D., Schaye J., Schaller M., Richings J., 2019, MNRAS, 488, L123

Ludlow A. D., Schaye J., Schaller M., Bower R., 2020, MNRAS, 493, 2926

Madau P., Dickinson M., 2014, ARA&A, 52, 415

Maddox N. et al., 2020, preprint (arXiv:2011.09470)

Marasco A., Crain R. A., Schaye J., Bahé Y., van der Hulst T., Theuns T., Bower R. G., 2016, MNRAS, 461, 2630

Marinacci F., Grand R. J. J., Pakmor R., Springel V., Gómez F. A., Frenk C. S., White S. D. M., 2017, MNRAS, 466, 3859

Marinacci F. et al., 2018, MNRAS, 480, 5113

Moore B., Lake G., Quinn T., Stadel J., 1999, MNRAS, 304, 465

Moretti A. et al., 2018, MNRAS, 480, 2508

Naiman J. P. et al., 2018, MNRAS, 477, 1206

Nelson D. et al., 2018, MNRAS, 475, 624

Nelson D. et al., 2019a, Comput. Astrophys. Cosmology, 6, 2

Nelson D. et al., 2019b, MNRAS, 490, 3234

Noble A. G. et al., 2017, ApJ, 842, L21

Obreschkow D., Croton D., De Lucia G., Khochfar S., Rawlings S., 2009, ApJ, 698, 1467

Pillepich A. et al., 2018a, MNRAS, 473, 4077

Pillepich A. et al., 2018b, MNRAS, 475, 648

Pillepich A. et al., 2019, MNRAS, 490, 3196

Planck Collaboration XIII, 2016, A&A, 594, A13

Rafieferantsoa M., Davé R., Anglés-Alcázar D., Katz N., Kollmeier J. A., Oppenheimer B. D., 2015, MNRAS, 453, 3980

Rafieferantsoa M., Davé R., Naab T., 2019, MNRAS, 486, 5184

Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, MNRAS, 452, 2034

Ramatsoku M. et al., 2019, MNRAS, 487, 4580

Reyes R., Mandelbaum R., Gunn J. E., Pizagno J., Lackner C. N., 2011, MNRAS, 417, 2347

Robotham A. S. G. et al., 2011, MNRAS, 416, 2640

Robotham A. S. G., Bellstedt S., Lagos C. d. P., Thorne J. E., Davies L. J. M., Driver S. P., Bravo M., 2020, MNRAS, 495, 905

Rodriguez-Gomez V. et al., 2015, MNRAS, 449, 49

Rodríguez Montero F., Davé R., Wild V., Anglés-Alcázar D., Narayanan D., 2019, MNRAS, 490, 2139

Rudnick G. et al., 2017, ApJ, 849, 27

Saintonge A. et al., 2011, MNRAS, 145, 32

Saintonge A. et al., 2012, ApJ, 758, 73

Saintonge A. et al., 2017, ApJS, 233, 22

Salim S. et al., 2007, ApJS, 173, 267

Schäbe A., Romano-Díaz E., Porciani C., Ludlow A. D., Tomassetti M., 2020, MNRAS, 497, 5008

Schaye J. et al., 2015, MNRAS, 446, 521

Sijacki D., Vogelsberger M., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 424, 2999

Springel V., 2010, MNRAS, 401, 791

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Springel V. et al., 2018, MNRAS, 475, 676

Stark D. V. et al., 2016, ApJ, 832, 126

Stevens A. R. H., Brown T., 2017, MNRAS, 471, 447

Stevens A. R. H., Martig M., Croton D. J., Feng Y., 2014, MNRAS, 445, 239

Stevens A. R. H., Lagos C. d. P., Obreschkow D., Sinha M., 2018, MNRAS, 481, 5543

Stevens A. R. H. et al., 2019a, MNRAS, 483, 5334 (S19)

Stevens A. R. H., Diemer B., Lagos C. d. P., Nelson D., Obreschkow D., Wang J., Marinacci F., 2019b, MNRAS, 490, 96

Tacconi L. J. et al., 2018, ApJ, 853, 179

Tadaki K. et al., 2019, PASJ, 71, 40

Tecce T. E., Cora S. A., Tissera P. B., Abadi M. G., Lagos C. d. P., 2010, MNRAS, 408, 2008

Tempel E., Kipper R., Tamm A., Gramann M., Einasto M., Sepp T., Tuvikene T., 2016, A&A, 588, A14

Tinker J. L., 2020, preprint (arXiv:2007.12200)

Tomassetti M., Porciani C., Romano-Díaz E., Ludlow A. D., Papadopoulos P. P., 2014, MNRAS, 445, L124

Tomassetti M., Porciani C., Romano-Díaz E., Ludlow A. D., 2015, MNRAS, 446, 3330

Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, MNRAS, 438, 1985

Tully R. B., Fisher J. R., 1977, A&A, 54, 661

Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

Vogelsberger M. et al., 2014a, Nature, 509, 177

Vogelsberger M. et al., 2014b, MNRAS, 444, 1518

Wang T. et al., 2018, ApJ, 867, L29

Weinberger R. et al., 2017, MNRAS, 465, 3291

Weinberger R. et al., 2018, MNRAS, 479, 4056

Wetzel A. R., Tinker J. L., Conroy C., 2012, MNRAS, 424, 232

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wu J. F. et al., 2018, ApJ, 853, 195

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

York D. G. et al., 2000, AJ, 120, 1579

Yun K. et al., 2019, MNRAS, 483, 1042

Zanisi L. et al., 2020, preprint (arXiv:2007.00039)

Zinger E. et al., 2020, MNRAS, 499, 768

APPENDIX A: ANCILLARY RESULTS

A1 The stellar mass–specific star formation rate plane

Similar to the purpose of Section 3 for H₂, to provide context for the effect of environment on SFRs in TNG galaxies, it is informative to briefly review the overall star formation activity of TNG galaxies as compared to galaxy survey data. We thus plot the stellar mass–specific star formation rate

plane for TNG100 alongside SDSS in Fig. A1. While this plane and the related colour–mass plane have been presented for the TNG suite in several papers already (Nelson et al. 2018; Weinberger et al. 2018; Donnari et al. 2019; Pillepich et al. 2019; Zanisi et al. 2020; Zinger et al. 2020), our inclusion of Fig. A1 not only satisfies self-containment, but it is also value-added through its inclusion of SDSS data and a purpose-made 'mock' plane for TNG100, as described in Section 2.4.2. The sample of SDSS galaxies is described in Section 2.3.

The presence of a bimodality in the SDSS m_*–(s)SFR plane is well documented (e.g. Brinchmann et al. 2004; Chang et al. 2015) and is clearly visible in the top panel of Fig. A1. A direct comparison of TNG100 (based on instantaneous SFRs) in the bottom panel suggests the simulation is almost missing a bimodality entirely. This is, however, merely a reflection of the fact that 20.5% of TNG100 galaxies with $m_* \geq 10^9$ M⊙ have an instantaneous SFR = 0. This contrasts to SDSS, where <1.5% of the sample we use do not have a non-zero, measured SFR. We also need to be conscious of the time-scales on which SFRs are measured in SDSS. We aim to match this in the middle panel of Fig. A1 by using mock SFRs (and stellar masses) as described in Section 2.4.2. With this, it is visually much clearer that there is a distinct quenched population of TNG100 galaxies (for a discussion surrounding the *colour* bimodality in TNG — which falls out without artificial assignment as done here — see Nelson et al. 2018). But, loosely speaking, based on the difference in contours between the top and middle panels of Fig. A1, there are apparently still too few quenched galaxies at most stellar masses in TNG100 relative to SDSS. For an elaborate discussion on this topic, see Donnari et al. (2020).

A2 H I and neutral fractions of galaxies

Because several aspects of the mock procedure outlined in Section 2.4 were updated from S19, here we reproduce the key comparisons between TNG100 and xGASS from S19, folding in these new aspects. Fig. A2 is analogous to the left panels of Fig. 6, and shows the running medians for galaxies' neutral-hydrogen and atomic-hydrogen fractions. Because the mock-observed atomic-hydrogen fractions of TNG100 are compared against the full xGASS sample, we retain the mock redshifts used in S19 for this panel (see section 3.1 and the upper panels of fig. 2 of S19). No beam corrections are used for the observations in either panel.

Despite the larger sample size of xGASS, significant uncertainties remain on the satellites medians, as the detection fraction for H I hovers at 50% for satellites of all masses. This is shown in the bottom panel of Fig. A2. Nevertheless, the systematic distinction between the H I content of centrals and satellites is much clearer than was the case for H₂.

APPENDIX B: APPROXIMATE H I/H₂ CALCULATION FOR 'MINI SNAPSHOTS'

Nominally, the neutral fraction of all non-star-forming (NSF) gas cells in TNG comes from the recorded values in the snapshot outputs. However, for the majority of snapshots, only a subset of the particle/cell fields were saved for the sake of
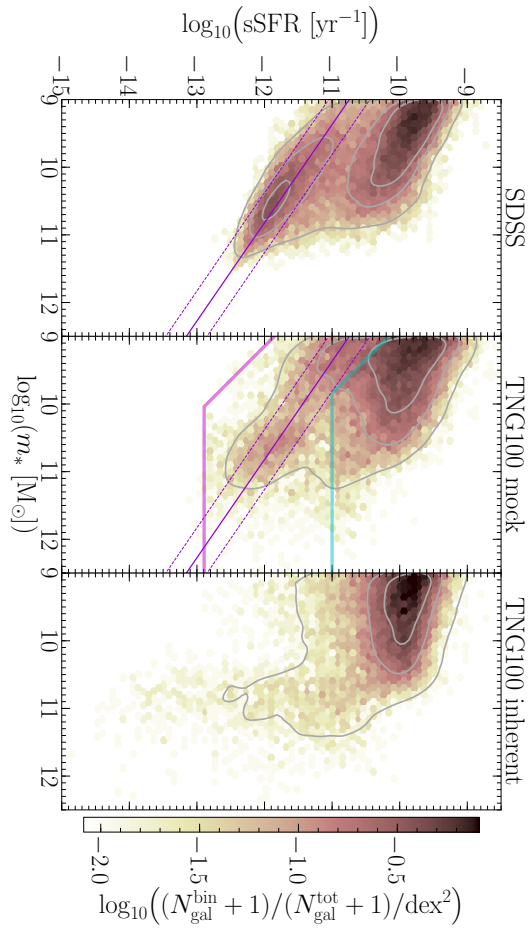
Figure A1: Stellar mass–specific star formation rate plane for SDSS and TNG100 galaxies. Pixel colour indicates the normalised number density of galaxies. Grey contours encompass 38, 68, and 95% of the galaxies contained in the axes (approximately equivalent to 0.5σ, 1σ, and 2σ contours, respectively, based on a two-dimensional Gaussian kernel density estimator applied to the underlying galaxy sample). **The top panel** shows a volume-limited subsample of SDSS galaxies with $m_*/M_\odot \in \left[10^9, 10^{11.5}\right]$ in the redshift interval $[0.02, 0.05]$. The violet, solid line shows a fit to the quiescent population (slope, intercept $= -0.689, -4.547$). The dashed, violet lines show the vertical standard deviation of the population about that fit ($\sigma = 0.305$). **The bottom panel** shows TNG100 galaxies, per their inherent properties, where SFR is summed from the gas cells' instantaneous values. A significant population of galaxies with SFR $= 0$ is not seen in this panel (nor do the contours account for them). **In the middle panel**, TNG100 galaxy properties are remeasured to be more directly comparable with SDSS. First, star-forming galaxies (instantaneous sSFR $\equiv$ SFR$/m_* > 10^{-11}$ yr^{-1}) have their SFR remeasured on a 20-Myr historical time-scale, based on the birth masses of the galaxies' star particles that formed within that look-back period. Any galaxy with either an instantaneous or 20-Myr SFR $< 10^{-11}$ yr^{-1} has its SFR remeasured on a 1-Gyr time-scale. The baryonic mass resolution of TNG100 (taken as 1.4×10^6 M$_\odot$) means any 1-Gyr sSFR must lie rightward of (or close to) the thick, transparent, magenta line. Similarly, any galaxy with a measured 20-Myr SFR must lie rightward of the cyan line. Finally, any remaining galaxies with SFR $= 0$ have a random sSFR pulled from a Gaussian distribution that depends on their stellar mass and matches the fitted SDSS quiescent population; the violet lines are shown again in this panel for reference.
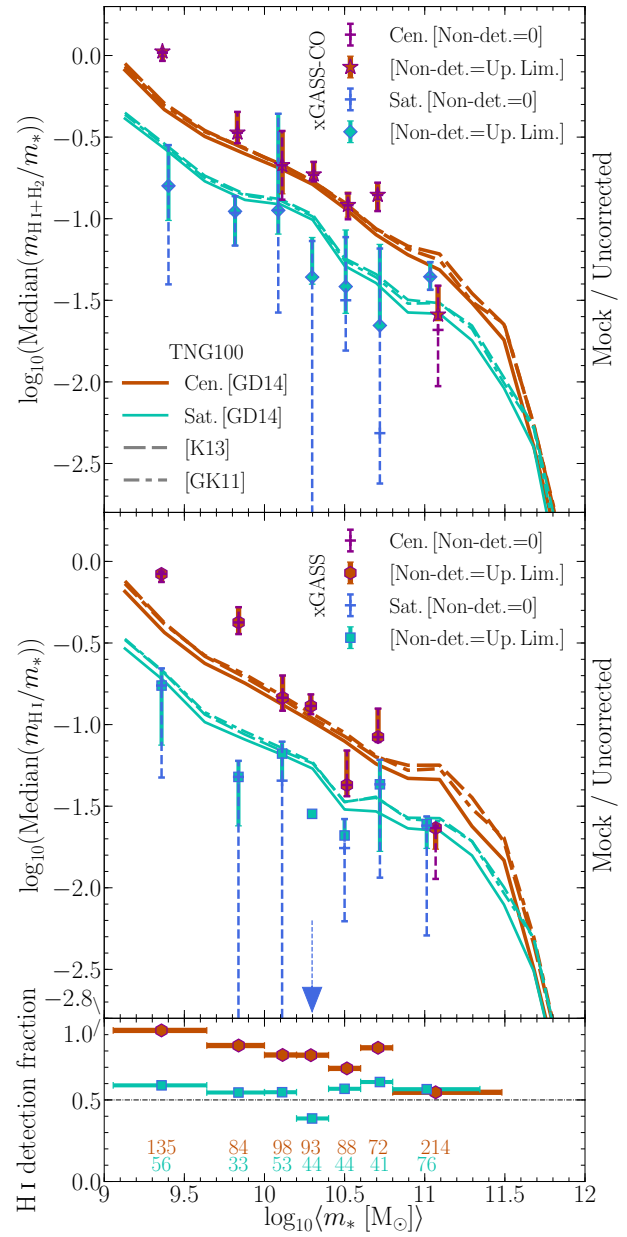
Figure A2: Top panel: Median neutral-hydrogen fraction as a function of stellar mass for TNG100 (at $z = 0$) and xGASS-CO centrals and satellites. **Middle panel:** Median atomic-hydrogen fraction as a function of stellar mass, where now TNG100 is compared against the full xGASS sample. The same bins and plotting style are used as in Fig. 6. Error bars on observational data are uncertainties *on the median* from bootstrapping. All TNG100 results have been updated from S19 to adhere to the mock-observing method in Section 2.4 of this paper (cf. figs 4 & 5 from S19). **Bottom panel:** Detection fraction of the 21-cm line in each xGASS bin.

reducing the simulation's storage requirements. Neutral fraction is one of the fields lost in these so-called 'mini snapshots'. To circumvent this, we find a polynomial fitting function for the neutral-fraction field at $z = 0$ in terms of the 'electron abundance' field, f_e, which *is* saved for all snapshots:

$$f_n^{\mathrm{NSF}} \simeq \begin{cases} \xi_1 + \xi_2 f_e & \forall f_e \leq f_{e0} \\ \xi_3 + \xi_4 f_e + \xi_5 f_e^2 & \forall f_e > f_{e0} \end{cases}, \qquad \text{(B1a)}$$

u	ξ_1	ξ_2	ξ_4	ξ_5	f_{e0}
$<10^{8.7}\,\mathrm{m}^2\,\mathrm{s}^{-2}$	0.995	-0.907	-4.562	1.854	0.986
$\geq 10^{8.7}\,\mathrm{m}^2\,\mathrm{s}^{-2}$	0.998	-0.996	-1.943	0.859	0.984

Table B1: Best-fitting values for Equation (B1), rounded to three decimal places. The first column gives the range of internal energies per unit mass for the cells.

$$\xi_3 = \xi_1 + (\xi_2 - \xi_4)\, f_{e0} - \xi_5\, f_{e0}^2 , \tag{B1b}$$

where f_n is neutral fraction, f_e is formally the ratio of free electrons to hydrogen protons (including free protons + those in hydrogen atoms and molecules, but excluding those in any other chemical species). The best-fitting values for the five parameters are given in Table B1. These parameters have a bimodal dependence on the specific-energy range of the particles. For $u > 10^{8.7}\,\mathrm{m}\,\mathrm{s}^{-1}$, almost all hydrogen (and other species) is ionized, leaving $f_e \gtrsim 1$. Below this energy limit, f_e spans the full possible range.

The abundances of individual chemical species are also not recorded in the mini snapshots. This includes the hydrogen abundance, X. But the total metallicity, Z, is available. We therefore again find an approximate fitting function for $X(Z)$ at $z=0$:

$$X \simeq \begin{cases} 0.753 - 1.26\,Z & \forall Z \geq 0.025 \\ 0.762 - 2.30\,Z + 24.2\,Z^2 & \forall Z < 0.025 \end{cases} . \tag{B2}$$

We ensure an upper limit of $X \leq 0.76$ always, as this was the initial value assumed in the simulation.

Our approximations for f_n and Z both carry an absolute error of less than 0.001 when considering all cells in all haloes used in this work. For gas cells with large neutral masses, Equation (B1) tends to over-predict f_n by 0.005. Even then, the uncertainty introduced by these approximations is definitively negligible. With these approximations, we can otherwise follow the same method for calculating $m_{\mathrm{H\,I}}$ and $m_{\mathrm{H_2}}$ for galaxies per S19. Fig. 10 is the only place in this paper where we have to use this approximation method. PYTHON routines for calculating neutral fractions and metallicity in line with this method are publicly available on Github.[19]

---

[19] See the `neutralFraction_from_electronFraction()` and `HI_H2_masses()` functions at https://github.com/arhstevens/Dirty-AstroPy/blob/master/galprops/galcalc.py.