

# Atomic and molecular gas in IllustrisTNG galaxies at low redshift

Benedikt Diemer<sup>1</sup>,<sup>1</sup>★† Adam R. H. Stevens<sup>2,3</sup>, Claudia del P. Lagos<sup>2,3,4</sup>,  
A. R. Calette<sup>5</sup>, Sandro Tacchella<sup>6</sup>,<sup>1</sup> Lars Hernquist,<sup>1</sup> Federico Marinacci<sup>7</sup>,<sup>1,6</sup>  
Dylan Nelson,<sup>7</sup> Annalisa Pillepich,<sup>8</sup> Vicente Rodriguez-Gomez<sup>9</sup>,  
Francisco Villaescusa-Navarro<sup>10</sup> and Mark Vogelsberger<sup>11</sup>

<sup>1</sup>*Institute for Theory and Computation, Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA 02138, USA*

<sup>2</sup>*International Centre for Radio Astronomy Research, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*

<sup>3</sup>*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)*

<sup>4</sup>*Cosmic Dawn Center (DAWN), Niels Bohr Institute, University of Copenhagen, Nørregade, Copenhagen DK-0000-0003-3631-7176, Denmark*

<sup>5</sup>*Instituto de Astronomía, Universidad Nacional Autónoma de México, A. P. 70-264, 04510 Ciudad de México, México*

<sup>6</sup>*Department of Physics and Astronomy, University of Bologna, via Gobetti 93/2, I-40129 Bologna, Italy*

<sup>7</sup>*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str 1, D-85741 Garching, Germany*

<sup>8</sup>*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*

<sup>9</sup>*Instituto de Radioastronomía y Astrofísica, Universidad Nacional Autónoma de México, Apdo. Postal 72-3, 58089 Morelia, Mexico*

<sup>10</sup>*Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*

<sup>11</sup>*Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

Accepted 2019 May 8. Received 2019 April 22; in original form 2019 February 25

## ABSTRACT

We have recently developed a post-processing framework to estimate the abundance of atomic and molecular hydrogen (H I and H<sub>2</sub>, respectively) in galaxies in large-volume cosmological simulations. Here we compare the H I and H<sub>2</sub> content of IllustrisTNG galaxies to observations. We mostly restrict this comparison to  $z \approx 0$  and consider six observational metrics: the overall abundance of H I and H<sub>2</sub>, their mass functions, gas fractions as a function of stellar mass, the correlation between H<sub>2</sub> and star formation rate, the spatial distribution of gas, and the correlation between gas content and morphology. We find generally good agreement between simulations and observations, particularly for the gas fractions and the H I mass–size relation. The H<sub>2</sub> mass correlates with star formation rate as expected, revealing an almost constant depletion time that evolves up to  $z = 2$  as observed. However, we also discover a number of tensions with varying degrees of significance, including an overestimate of the total neutral gas abundance at  $z = 0$  by about a factor of 2 and a possible excess of satellites with no or very little neutral gas. These conclusions are robust to the modelling of the H I/H<sub>2</sub> transition. In terms of their neutral gas properties, the IllustrisTNG simulations represent an enormous improvement over the original Illustris run. All data used in this paper are publicly available as part of the IllustrisTNG data release.

**Key words:** ISM: molecules – galaxies: abundances – galaxies: star formation – galaxies: structure.

## 1 INTRODUCTION

Over the last decade, large-volume cosmological simulations have matured into one of the key tools in the study of galaxy formation (Schaye et al. 2010, 2015; Dubois et al. 2014; Vogelsberger et al. 2014a,b; Davé, Thompson & Hopkins 2016; Davé et al. 2019, see also the review of Somerville & Davé 2015). However, such simulations achieve their large volumes at the cost of relatively

low resolution, meaning that many physical processes need to be implemented through subgrid prescriptions, which are inevitably based on a significant number of free parameters (e.g. Schaye et al. 2015; Pillepich et al. 2018a). These parameters must be calibrated to reproduce the properties of galaxies in the real Universe, such as the stellar mass function at low redshift. Galaxy properties that are not used in the tuning are then the simulations’ predictions, and need to be carefully compared to observations to check whether the physical models create realistic galaxy populations.

In this work, we focus on the IllustrisTNG simulation suite (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018a; Pillepich et al. 2018b; Springel et al. 2018), though we also

\* E-mail: benedikt.diemer@cfa.harvard.edu

† NHFP Einstein fellow.

briefly compare to the original Illustris-1 simulation (Genel et al. 2014; Vogelsberger et al. 2014a,b). Many comparisons between the Illustris or IllustrisTNG simulations and observations have already been undertaken, including topics as wide-ranging as galaxy colours (Sales et al. 2015; Nelson et al. 2018a), sizes (Genel et al. 2018), the mass–metallicity relation (Torrey et al. 2018, 2019), morphology (Snyder et al. 2015; Rodriguez-Gomez et al. 2019; Tacchella et al. 2019), star formation (Sparre et al. 2015; Diemer et al. 2017; Donnari et al. 2019), merger rates (Rodriguez-Gomez et al. 2015), black holes (Weinberger et al. 2018; Habouzit et al. 2019), and the abundance of shell galaxies (Pop et al. 2018).

However, one category of observations is noticeably underrepresented in this list: the abundance, phase, and structure of gas in galaxies. For the original Illustris model, Vogelsberger et al. (2014b) showed that the neutral fraction was broadly compatible with the trends seen in the Arecibo Legacy Fast ALFA (ALFALFA) survey. For the IllustrisTNG model, the total gas fraction in large groups and clusters was used as a constraint (Pillepich et al. 2018a, see also Vogelsberger et al. 2018 and Barnes et al. 2018). In addition, there have been studies of the larger scale gas distribution around galaxies, namely of absorption lines due to the circum-galactic medium (Suresh et al. 2017; Nelson et al. 2018b), the abundance of damped Lyman-alpha systems (DLAs) and the Lyman-alpha forest (Bird et al. 2014; Gurvich, Burkhart & Bird 2017), hot gaseous haloes (Bogdán et al. 2015), and other gas-related phenomena such as jellyfish galaxies (Yun et al. 2019).

Nevertheless, we are still missing a complete census of the bulk of the observable gas in IllustrisTNG galaxies, namely, of atomic and molecular hydrogen (denoted H I and H<sub>2</sub> hereafter). The molecular (or simply, cold) component, especially, is of interest because it is thought to provide the fuel for star formation. While the IllustrisTNG simulations reproduce numerous properties of the observed stellar population, star formation is governed by an equilibrium interstellar medium (ISM) model that does not directly predict the abundance of H<sub>2</sub> (Springel & Hernquist 2003). Thus, a detailed comparison of the gas and stellar properties of galaxies can tell us whether the relationship between gas and star formation is physically correct, or whether it emerges due to the tuning of the ISM model. There are, however, significant obstacles to H I and H<sub>2</sub> comparisons on both the observational and theoretical sides.

Observationally, H I in galaxies is detected either by the 21-cm emission from its spin-flip transition or by absorption in quasar or radio galaxy spectra (Prochaska, Herbert-Fort & Wolfe 2005; Zwaan et al. 2005; Allison et al. 2019). The former technique has provided tight constraints on the local H I abundance and its distribution in thousands of nearby galaxies (Martin et al. 2010; Haynes et al. 2018). At higher redshift, quasar sightlines provide rough constraints on the overall abundance of H I, but we are lacking detailed information about H I in galaxies (e.g. Zafar et al. 2013). H<sub>2</sub> is, arguably, even more difficult to observe because the molecule lacks a permanent dipole moment (Draine 2011). Instead of observing H<sub>2</sub> directly, we rely on tracers such as CO, which are subject to uncertain conversion factors (Bolatto, Wolfire & Leroy 2013). Nevertheless, molecular observations are now being pushed to high redshifts and high resolution by instruments such as ALMA, offering a chance to study gas in high-redshift galaxies in great detail (e.g. Walter et al. 2016).

From a theoretical perspective, H I and H<sub>2</sub> comparisons are challenging because the IllustrisTNG simulations predict the fraction of all gas that is in neutral hydrogen, but not whether this gas is atomic or molecular. The H I/H<sub>2</sub> transition is governed by the balance between H<sub>2</sub> formation on dust grains and destruction by

Lyman–Werner band ultraviolet (UV) photons, leading to complex dependences on the density structure, star formation rate (SFR), and metallicity of the ISM (e.g. Spitzer & Zweibel 1974; Black & Dalgarno 1976; Elmegreen 1989, 1993; Sternberg & Dalgarno 1989; Krumholz, McKee & Tumlinson 2009; Draine 2011). These dependences have been captured in a number of models, including observational correlations, calibrations from simulations, and analytic models (e.g. Blitz & Rosolowsky 2006; Gnedin & Kravtsov 2011; Krumholz 2013). Such formulae can then be used to roughly constrain the molecular fraction in large-volume cosmological simulations in post-processing (e.g. Duffy et al. 2012). Lagos et al. (2015) applied this technique to the EAGLE simulation and showed that it reproduces the observed overall H<sub>2</sub> abundance, mass function, gas fraction, and depletion time. Based on similar modelling, Bahé et al. (2016) investigated the abundance, size, and structure of H I discs in EAGLE and found them to be broadly compatible with observations, though their conclusions depended somewhat on the chosen H I/H<sub>2</sub> model (for related results, see Marasco et al. 2016; Lagos et al. 2016; Crain et al. 2017; Marinacci et al. 2017).

In Diemer et al. (2018, hereafter *Paper I*), we undertook a systematic study of models that predict the H I and H<sub>2</sub> fractions and applied them to the IllustrisTNG simulations. We extended the cell-by-cell modelling of Lagos et al. (2015) by introducing a projection-based method and found the predictions of most models to be broadly compatible, as long as care is taken when computing the intermediate quantities they rely on. Stevens et al. (2019) used very similar modelling to investigate the dependence of H I on environment, carefully mock-observing the simulation to take observational systematics into account. Popping et al. (2019) use a version of the H I/H<sub>2</sub> modelling that, among other differences, assumes a simpler approach to the UV field, and compare the high-redshift abundance of H<sub>2</sub> to ALMA observations (see also Popping, Somerville & Trager 2014).

In this work, we compare the results of *Paper I* to observations of galaxies at low redshift. We restrict ourselves to six observational metrics: the overall abundance of H I and H<sub>2</sub>, their mass functions, gas fractions as a function of stellar mass, the correlation between H<sub>2</sub> and SFR, the spatial distribution of H I and H<sub>2</sub> within galaxies, and correlations with morphology. Whenever possible, we carefully consider the observational systematics of the respective observational samples, attempting to faithfully mock-observe our simulated galaxies. All H I/H<sub>2</sub> data used in this paper are publicly available as part of the IllustrisTNG data release (Nelson et al. 2019a, [tng-project.org/data](https://tng-project.org/data)).

The paper is structured as follows. We give an overview of all observational data used in Section 2. We briefly review the simulation data in Section 3 (referring the reader to *Paper I* for details) and discuss our techniques for matching simulated galaxies to observations. We present the results of our comparisons in Section 4, including a brief review of the original Illustris-1 simulation. We further discuss particular successes and tensions in Section 5 before summarizing our conclusions in Section 6. We describe our experiments with morphology in Appendix A and provide detailed tests of our methodology in Appendix B.

We follow the notation of *Paper I*, denoting the masses of all gas as  $M_{\text{gas}}$ , of all hydrogen as  $M_{\text{H}}$ , of atomic hydrogen as  $M_{\text{H I}}$ , of molecular hydrogen as  $M_{\text{H}_2}$ , and of neutral hydrogen as  $M_{\text{H I}+\text{H}_2}$ . The same subscripts are used for volumetric mass densities (e.g.  $\rho_{\text{H}}$ ), surface densities (e.g.  $\Sigma_{\text{H}}$ ), number densities (e.g.  $n_{\text{H}}$ ), and column densities (e.g.  $N_{\text{H}}$ ). The neutral fraction,  $f_{\text{H I}+\text{H}_2}$ , refers to the fraction of all gas (including helium and metals) that is in neutral hydrogen.

## 2 OBSERVATIONAL DATA

In this section, we give an overview of the observational data used in our comparison. Each subsection discusses one of our six observational metrics.

### 2.1 Overall abundance of H I and H<sub>2</sub>

As a first-order check on the abundance of neutral gas in IllustrisTNG, we use observations of the total H I abundance as a function of redshift (though, for the remainder of the paper, we will mostly be concerned with  $z \approx 0$ ). Our data compilation is largely based on that of Rhee et al. (2018), with a few sources added and removed. In particular, we use data from the blind ALFALFA 21-cm survey and similar observations (Zwaan et al. 2005; Martin et al. 2010; Freudling et al. 2011; Hoppmann et al. 2015; Obuljen et al. 2018), from 21-cm observations of low-redshift galaxy samples (Lah et al. 2007; Braun 2012; Delhaize et al. 2013; Rhee et al. 2013, 2016, 2018), and from quasar absorption line studies that constrain the number of DLAs and Lyman-limit systems (LLSs) at high redshift (Prochaska et al. 2005; Rao, Turnshek & Nestor 2006; Noterdaeme et al. 2009, 2012; Zafar et al. 2013; Neeleman et al. 2016; Bird, Garnett & Ho 2017; Rao et al. 2017).

The H<sub>2</sub> abundance of the Universe at low redshift can currently not be determined in blind surveys because the emission of the CO tracer is simply too weak. Similar caveats affect all H<sub>2</sub>-related comparisons in this paper: H<sub>2</sub> is never observed directly, and the relationship between H<sub>2</sub> and CO is probably much more complicated than suggested by constant conversion factors (Wolfire, Hollenbach & McKee 2010; Glover & Mac Low 2011; Leroy et al. 2011; Narayanan et al. 2011, 2012; Feldmann, Gnedin & Kravtsov 2012; Bolatto et al. 2013). Thus, all H<sub>2</sub> abundances are to be interpreted as estimates that are likely accurate to factors of a few.

Subject to this caveat, we compare our simulations to low-redshift constraints based on targeted CO observations of galaxies (Keres, Yun & Young 2003; Obreschkow & Rawlings 2009; Saintonge et al. 2017). While the first blind surveys are beginning to probe the abundance of H<sub>2</sub> at high redshift (Decarli et al. 2016; Walter et al. 2016; Pavesi et al. 2018; Riechers et al. 2019), we do not use their results in this study because the small volumes probed lead to significant sample variance and other complications (see Popping et al. 2019, for a detailed comparison). In recent years, the H<sub>2</sub> abundance is increasingly being inferred from the dust continuum emission, e.g. as measured by ALMA, leading to different systematics (Scoville et al. 2017; Tacconi et al. 2018). However, such measurements are currently possible only at high redshift, and are thus not the focus of this paper.

### 2.2 Mass functions

The H I mass function has been measured to exquisite accuracy by ALFALFA (Giovanelli et al. 2005; Martin et al. 2010; Haynes et al. 2011; Huang et al. 2012b; Haynes et al. 2018; Jones et al. 2018). The final  $\alpha$ .100 catalogue contains about 30 000 sources below  $z = 0.06$ . Our primary constraint for the H<sub>2</sub> mass function is the work of Boselli et al. (2014b), specifically their estimate based on a fixed CO-to-H<sub>2</sub> conversion factor (which we convert to  $\alpha_{\text{CO}} = 2.0 \text{ M}_{\odot}/(\text{K km s}^{-1} \text{ pc}^2)$  as in Popping et al. 2019). The underlying data are taken from the *Herschel* Reference Survey, a sample of 323 galaxies observed with the SPIRE instrument (Boselli et al. 2010). The galaxies in this survey are local, with distances

between 15 and 25 Mpc. For a comparison, we also show the H<sub>2</sub> mass function of Obreschkow & Rawlings (2009), which is based on Keres et al. (2003). The underlying data set is the FCRAO survey (Young et al. 1995), which measured the CO emission for 300 local galaxies. However, due to the range of distances and observing strategies, this survey is somewhat difficult to mock-observe.

### 2.3 Gas fractions

By gas fractions, we mean the ratio of the mass in H I or H<sub>2</sub> divided by stellar mass. These ratios are commonly constrained observationally and help us determine whether IllustrisTNG produces an appropriate amount of gas in galaxies of a certain stellar mass. The scatter in the gas fractions is known to be sizeable, meaning that both the median fraction and the distribution around the median are of interest. However, comparing simulations to observed gas fractions is non-trivial because of numerous survey-specific aspects such as the galaxy selection, aperture (or beam), and other assumptions used in the processing of the observations. Stevens et al. (2019) recently demonstrated that the total H I fractions in IllustrisTNG can differ significantly from mock observations that take the observational systematics into account.

To avoid having to mock-observe many heterogeneous surveys, we rely on the recent compilation of Calette et al. (2018, hereafter C18). They quantify the H I and H<sub>2</sub> fractions as a function of stellar mass, including fitting functions for the mass-dependent distribution of gas fractions. C18 amalgamate H I and H<sub>2</sub> masses from a large range of radio observations that have been combined with stellar luminosities or masses from optical or infrared observations. In particular, their H I data include the Updated Nearby Galaxy Catalogue (Karachentsev, Makarov & Kaisina 2013), GASS (Catinella et al. 2013, 2018), the *Herschel* Reference Survey (Boselli et al. 2010; Boselli, Cortese & Boquien 2014a; Boselli et al. 2014b,c), ATLAS<sup>3D</sup> (Serra et al. 2012), the Nearby Field Galaxy Survey (Jansen et al. 2000b,a; Kannappan et al. 2013), THINGS (Leroy et al. 2008), ALFALFA (Huang et al. 2012a), UNAM-KIAS (Hernández-Toledo et al. 2010), AMIGA (Lisenfeld et al. 2011), and a number of other compilations (Geha et al. 2006; Stark et al. 2013; Bradford, Geha & Blanton 2015). Their H<sub>2</sub> data are taken from the *Herschel* Reference Survey, COLD GASS (Saintonge et al. 2011a), ATLAS<sup>3D</sup>, HERACLES (Leroy et al. 2008), ALLSMOG (Bothwell et al. 2014), EGN0G (Bauermeister et al. 2013), and Stark et al. (2013). These data sources are discussed in detail in appendices A and B of C18.

Using the C18 compilation offers a number of advantages. First, C18 homogenize the survey data to a Chabrier (2003) stellar initial mass function (IMF), the same IMF used in IllustrisTNG. Secondly, numerous survey biases and upper limits are taken into account, the latter via a Kaplan–Meier estimator. For example, surveys can easily be biased towards gas-rich galaxies (see fig. 1 in C18 for an example of a fitting function that overestimates the full, debiased data set). Thirdly, all CO observations are converted to H<sub>2</sub> masses based on the same mass-dependent  $\alpha_{\text{CO}}$  factor.

The C18 compilation does, however, introduce one complication: they split the galaxy sample into early- and late-type galaxies (ETGs and LTGs, respectively). This distinction is sensible because the gas content of galaxies strongly correlates with their morphological type (Kannappan et al. 2013; Boselli et al. 2014c). Thus, quantifying the median gas fraction of all galaxies introduces additional scatter compared to the relations split by morphology. In Section 3.5 and Appendix A, we experiment with different morphological selections, but find no criterion that reliably creates late-type and



early-type samples with the properties expected from observations. To avoid the resulting difficulties in interpretation, we reconstruct the distribution of gas fractions in the overall galaxy sample. C18 provide fitting functions for the distribution of gas fractions of early and late types as a function of stellar mass (their equations 3 and 7), as well as the ETG fraction as a function of stellar mass. We add the distributions of the LTG and ETG samples accordingly. For the ETG sample, C18 give a mass-dependent cut-off gas fraction below which the distribution cannot be reliably inferred because the majority of observations are upper limits. We separately constrain the properties of those galaxies with gas fractions above the limit as well as the fraction of galaxies below the limit. We use a slightly updated version of the C18 best-fitting parameters.

C18 find that, once they split into LTGs and ETGs and homogenize the data, the different samples are broadly consistent, except for certain data sets where strong biases are expected. Those are collected in a ‘bronze’ sample, which we do not use in our analysis. C18 show that their gas fractions are compatible with the observed mass functions and that the H I fractions measured in individual LTGs are compatible with the stacked ALFALFA analysis of Brown et al. (2015), indicating that the galaxy sample is not missing any major contributions to the overall H I abundance. Given that some data sources were selected from optical samples such as SDSS, there may be a bias against satellites close to their central. For example, galaxies in the SDSS-I and II spectroscopic surveys could not be closer than 55 arcsec due to fibre collisions (e.g. Zehavi et al. 2002; Guo, Zehavi & Zheng 2012). We will return to this difficulty when discussing the gas fractions of satellites in Sections 4.3 and 5.2. C18 assume  $h = 0.7$  while IllustrisTNG uses  $h = 0.6774$ ; we rescale masses wherever appropriate.

## 2.4 Correlations with SFR

For our comparison of H<sub>2</sub> mass and SFRs, we use the xCOLD GASS survey (Saintonge et al. 2011a, 2017). This data set is composed of two surveys, COLD GASS and xCOLD GASS-low, which combine to cover all stellar masses above  $10^9 M_\odot$ . The CO observations were taken with the IRAM telescope, their stellar properties are taken from SDSS and other surveys. It is well established that the depletion time,  $t_{\text{depl}} \equiv M_{\text{H}_2}/\text{SFR}$ , is roughly constant but depends on redshift (Bigiel et al. 2008, 2011; Leroy et al. 2008, 2013; Daddi et al. 2010; Genzel et al. 2010, 2015; Tacconi et al. 2010, 2013, 2018; Saintonge et al. 2011b,a, 2017). We compare IllustrisTNG to the parametrization of Tacconi et al. (2018),  $t_{\text{depl}} \approx 1 \text{ Gyr} \times (1 + z)^{-0.57}$ .

## 2.5 Spatial distribution of gas

One of the most well-constrained indicators of the spatial distribution of gas in galaxies is the H I radius,  $R_{\text{H I}}$ , commonly defined as the radius where the surface density falls below  $\Sigma_{\text{H I}} < 1 M_\odot \text{ pc}^{-2}$ . This radius forms a tight power-law relationship with  $M_{\text{H I}}$ , at least in the local galaxy samples where it has been observed in resolved 21-cm observations (Broeils & van Woerden 1994; Broeils & Rhee 1997; Verheijen & Sancisi 2001; Wang et al. 2013, 2014; Martinsson et al. 2016; Ponomareva, Verheijen & Bosma 2016; Wang et al. 2016). We use the data compilation of Wang et al. (2016), who combine  $R_{\text{H I}}-M_{\text{H I}}$  data from about 500 galaxies measured by 15 different observational projects. For a comparison, we also show the similar relation of Lelli, McGaugh & Schombert (2016). The measured H I sizes correspond to a face-on orientation.

The resolved H I observations used here do not represent a volume-limited sample but are typically selected to include mostly H I-rich targets. However, all morphological types are present in the surveys and the gas fractions roughly match the averages observed in volume-limited samples, indicating that there are no strong biases in the selection (Lelli et al. 2016). Moreover, the  $R_{\text{H I}}-M_{\text{H I}}$  is extremely robust in that it does not vary as a function of stellar brightness or stellar diameter (Lelli et al. 2016; Wang et al. 2016).

Going beyond  $R_{\text{H I}}$ , we also investigate the radial dependence of the H I surface density. For this purpose, we use the H I profiles from the Bluedisk survey (Wang et al. 2013, 2014). They initially selected 25 galaxies as a control sample and 25 galaxies as an H I-rich sample, but the surface density profiles (with radii scaled to  $R_{\text{H I}}$ ) turn out to be almost identical. The Bluedisk galaxies have  $10^{10} M_\odot < M_* < 10^{11} M_\odot$  and lie on the star-forming main sequence, with SFRs between 0.5 and  $10 M_\odot \text{ yr}^{-1}$  (Cormier et al. 2016).

Finally, we wish to quantify the spatial distribution of H<sub>2</sub>, but there are a number of issues. First, the observational sample sizes are limited by the difficulty of obtaining spatially resolved CO observations. Secondly, the radial profiles of H<sub>2</sub> appear to be diverse (e.g. Young & Scoville 1982; Regan et al. 2001; Leroy et al. 2008, 2009). Thirdly, uncertain (and perhaps radially varying) CO-to-H<sub>2</sub> conversion factors complicate the interpretation of observations (Bolatto et al. 2013, and references therein). Given these difficulties, we delegate a detailed comparison with CO profiles and maps to future work. Instead, we restrict ourselves to a comparison of the observed half-mass radii of CO and stars from the EDGE-CALIFA survey (Bolatto et al. 2017). This data set adds spatially resolved CARMA CO observations to a subset of the optical integral field unit spectroscopy from CALIFA (Sánchez et al. 2012; Walcher et al. 2014). We note that there are numerous other ways to quantify the CO spatial scale, but some of them use relatively high surface densities that are difficult to reproduce in our simulations. For example, Davis et al. (2013) measured  $R_{\text{CO}}$  as the radius where the surface density of H<sub>2</sub> reaches  $15 M_\odot \text{ pc}^{-2}$ , corresponding to their  $3\sigma$  detection limit. This surface density is never reached by two-thirds of TNG100 galaxies with  $M_* > 10^{10} M_\odot$ . As a result, Davis et al. (2013) find very small values for  $R_{\text{CO}}$  that are close to the force resolution scale used in IllustrisTNG.

## 3 SIMULATION DATA

Our modelling of the H I/H<sub>2</sub> transition and the relevant details of the IllustrisTNG simulations were discussed at length in Paper I. Here we briefly review the aspects that are most pertinent to this work and refer the reader to Paper I for details.

### 3.1 The IllustrisTNG simulations

The IllustrisTNG suite of cosmological magneto-hydrodynamical simulations was run using the moving-mesh code AREPO (Springel 2010). In this paper, we use the highest resolution versions of the 100 and 300 Mpc box sizes, hereafter referred to as TNG100 and TNG300 (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018a; Pillepich et al. 2018b; Springel et al. 2018). As the TNG100 simulation has significantly higher spatial and mass resolution than TNG300, we interpret differences in their galaxy properties to be resolution effects, i.e. we trust the TNG100 results over TNG300 (see also Paper I). We use a lower resolution version of TNG100, TNG100-2, to quantify the convergence with resolution.

IllustrisTNG adopts the Planck Collaboration XIII (2016) cosmology:  $\Omega_m = 0.3089$ ,  $\Omega_b = 0.0486$ ,  $h = 0.6774$ , and  $\sigma_8 =$

0.8159; the same cosmology is assumed throughout the paper. The simulations follow a large range of physical processes, including prescriptions for gas cooling, star formation, stellar winds, metal enrichment, supernovae, black hole growth, and active galactic nuclei (AGNs; Weinberger et al. 2017; Pillepich et al. 2018a). This so-called TNG model builds on the original Illustris model (Vogelsberger et al. 2013, 2014a,b; Genel et al. 2014; Torrey et al. 2014; Sijacki et al. 2015). We briefly compare our results to the original simulation in Section 4.7. Dark matter haloes and their galaxies were identified using the SUBFIND algorithm (Springel et al. 2001).

### 3.2 Modelling the H I/H<sub>2</sub> fractions

In Paper I, we considered the H I/H<sub>2</sub> models of Leroy et al. (2008), Gnedin & Kravtsov (2011), Krumholz (2013), Gnedin & Draine (2014), and Sternberg et al. (2014) (hereafter L08, GK11, K13, GD14, and S14, respectively). These models represent three separate categories: (i) observational correlations with the mid-plane pressure of galaxies (L08), (ii) fitting functions trained on high-resolution simulations with radiative transfer and chemical networks (GK11, GD14), and (iii) analytical models of simplified, individual clouds that were modified to represent an average over a larger ISM region (K13, S14). Models in the latter two categories demand the Lyman–Werner band UV field strength as an input parameter, and we thus modelled the UV field by propagating light from young stars in an optically thin fashion.

Conventionally, H I/H<sub>2</sub> models have been computed cell by cell, i.e. by assigning a molecular fraction to each gas element. The main issue with this method is that all models listed above rely on surface densities rather than volume densities. For a 3D gas cell, such surface densities are commonly estimated by multiplying with the Jeans length, roughly representing the size of a self-gravitating system in equilibrium (e.g. Schaye 2001). In Paper I, we showed that this approximation does not, in general, reproduce the actual surface density of our galaxies. We proposed an alternative way to compute the models in projection, where each galaxy is rotated into a face-on position, projected on to a 2D map, and the model equations are solved in two dimensions.

In this work, we remain agnostic as to which method and which H I/H<sub>2</sub> model are most accurate and view their spread as a systematic uncertainty. We consider both the cell-by-cell and projected versions, with the exception of the cell-by-cell L08 model, which was shown to be unphysical (Paper I). Thus, the H I and H<sub>2</sub> masses of each galaxy are estimated by a total of nine different methods. In the forthcoming figures, we shade the region between the models but do not show the individual models to avoid crowding. The differences between the model predictions are shown in the corresponding figures in Paper I.

### 3.3 Galaxy selection

In Paper I, we computed the H I/H<sub>2</sub> fractions of all TNG100 galaxies with either  $M_* > 2 \times 10^8 M_\odot$  or  $M_{\text{gas}} > 2 \times 10^8 M_\odot$ . We use the same mass limits in this work, but increase them to  $M_* > 5 \times 10^{10} M_\odot$  and  $M_{\text{gas}} > 5 \times 10^9 M_\odot$  for TNG300. These selections result in stellar-mass-selected samples of 43 213 and 22 109 galaxies for TNG100 and TNG300, respectively, and 140 923 and 449 176 galaxies for the gas-mass-selected samples. We caution that our mass cut-offs in TNG100 are chosen somewhat aggressively, as they correspond to about 200 stellar-population particles or gas cells.

We do not cut the galaxy sample any further at this point, though we will later impose additional cuts to mimic the galaxy selection of

various observational data sets. We include both central and satellite galaxies because they are not generally separated in observations. As a result, our sample contains galaxies from different environments such as field galaxies and cluster members. It is now well established that the H I fractions tend to be significantly lower in the latter, a trend partially captured in the morphology–density relation where ETGs tend to both live in denser environments and contain less gas (Haynes & Giovanelli 1984; Cortese et al. 2011; Catinella et al. 2013; Boselli et al. 2014c; Stevens et al. 2019). Similarly, different gas fractions have been observed in the satellite population by splitting observed samples using group catalogues (Brown et al. 2017). However, C18 do not split their galaxy sample by any isolation criterion, and the mass functions also include both satellites and centrals. Moreover, the IllustrisTNG simulations have been shown to recover the various trends with environment (Stevens et al. 2019).

Finally, we do not cut out objects with unusual ratios of their gas, stellar, and dark matter content. For example, IllustrisTNG contains some objects with barely any dark halo, which can be traced to non-cosmological formation mechanisms (Nelson et al. 2019a). We find that such objects contribute to a population of gas-free galaxies that we discuss further in Sections 4.3 and 5.2.

### 3.4 Mock-observing masses, sizes, and SFRs

The most natural definition for the masses of stars or gas in simulated galaxies is to include all particles or gas cells that are gravitationally bound to the respective halo or subhalo. When comparing to observations, however, this definition rarely applies because the observed data are usually based on a finite beam size or aperture, which can either reduce the measured gas mass due to partial coverage or increase it due to confusion with other objects. Such effects can lead to significant differences in the inferred gas masses (e.g. Stevens et al. 2019).

We compute all aperture masses from linearly interpolated, cumulative mass profiles of the respective species. The profiles contain only matter bound to the respective subhalo and are binned in 50 linearly spaced radii at fixed fractions of the outer radius of our 2D maps (equation 4 in Paper I). If the desired aperture is larger than the largest radial bin, we use the cumulative mass in that final bin. For any quantity that is defined cell-by-cell or particle-by-particle (e.g. stellar mass or gas masses), we could use a spherically averaged (3D) or a projected (2D) profile. For disc-like, anisotropic systems, this choice could make a difference in the computed masses. We choose to compute all masses and sizes from projected profiles to be consistent with those quantities that are computed in projection in the first place (Section 3.2). In such cases, we cannot use 3D profiles because the information about the 3D positions has been lost. We have tested the difference between using 2D and 3D profiles and find it to be insignificant, with no clearly discernible trend in the inferred masses.

We compute sizes in a similar way, namely, we linearly interpolate the projected mass profile to find a density threshold (e.g.  $\Sigma_{\text{H I}} = 1 M_\odot \text{pc}^{-2}$ ) or a fraction of the integrated mass (e.g. the half-mass radius). This method is compatible with the observational samples to which we compare, where sizes correspond to a face-on orientation (Section 2.5).

#### 3.4.1 Stellar masses and sizes

Optical surveys used to measure stellar masses tend to underestimate the true extent of the stellar distribution (e.g. Bernardi et al.

2013; Kravtsov, Vikhlinin & Meshcheryakov 2018). To mimic such effects, we define  $M_*$  as the stellar mass within a fixed aperture of 30 kpc. Schaye et al. (2015) demonstrated that this definition tracks the Petrosian radii used in observations (see also Pillepich et al. 2018b), for an investigation of various definitions in IllustrisTNG). A fixed 30-kpc aperture also matches the BaryMP technique of Stevens et al. (2014) to about 20 per cent accuracy (with larger deviations at  $M_* \gtrsim 10^{11} M_\odot$ ).

In extremely rare cases (0.003 per cent of galaxies in TNG100, 0.02 per cent in TNG300), the radius of the innermost bin of the stellar mass profile is larger than 30 kpc. In those cases, we linearly interpolate the logarithmic value of the first bin from zero. We find that the details of this procedure have no appreciable impact on our results. We have also tested that increasing the number of bins by more than twofold has less than a 1 per cent effect on the computed stellar masses.

Besides the bias related to the size of a galaxy, stellar masses suffer from significant systematic uncertainties such as the assumed IMF and the modelling of the stellar light profiles. For example, while C18 homogenize their data to the same IMF, modelling uncertainties largely remain (see their section 2.1.3). Thus, we add a lognormal scatter of 0.2 dex to the stellar masses from IllustrisTNG (Marchesini et al. 2009; Mancini et al. 2011; Bower, Benson & Crain 2012; Mitchell et al. 2013). When comparing gas fractions at fixed stellar mass, this has two main effects. First, features in the relations are smoothed out, which can hide trends such as rapid changes due to the feedback implementation (see the discussion in Stevens et al. 2019 and Nelson et al. 2019a). Secondly, as the gas fractions typically fall with stellar mass, Eddington bias can shift the gas fractions upwards (Eddington 1913).

### 3.4.2 SFRs

We compare our  $H_2$  masses and SFRs to the xCOLD GASS survey, which combines measurements of H I and CO properties of galaxies with optical surveys (Saintonge et al. 2017). Their SFRs are derived using a combination of UV and infrared (IR) luminosities (Janowiecki et al. 2017). While UV tracers such as the  $H\alpha$  line are indicative of star formation on short time-scales, IR tracers are sensitive to time-scales of a few hundred Myr; their combination typically traces variations over roughly 50–150 Myr (e.g. Caplar & Tacchella 2019).

We use the instantaneous SFRs given by the simulation, but caution that there may be a slight systematic offset (Donnari et al. 2019). To put a simplistic upper limit on this offset, we have compared the instantaneous SFRs of IllustrisTNG galaxies, defined as the sum of the SFRs of a galaxy’s gas cells, to an average over the stars formed within about 200 Myr. Given the time-scales discussed above, this represents a conservative estimate of any differences. For the relevant stellar mass and redshifts,  $z = 0$  and 2, we find that the median instantaneous SFRs of TNG100 galaxies differ from the averaged ones by about 3 per cent, a negligible offset. There is, however, a scatter of 0.15 dex at  $z = 0$  and 0.6 dex at  $z = 2$  to which we return when interpreting our results in Section 4.4.

### 3.4.3 H I masses and sizes

The majority of the H I data described in Section 2 are based on Arecibo observations. The beamwidth of the Arecibo telescope at 21 cm is about 3.4 arcmin (e.g. Jones et al. 2018), corresponding to a physical aperture that depends on redshift. For example, at the median redshift of the GASS survey,  $z = 0.037$ , 3.4 arcmin

corresponds to roughly 140 kpc, leading Bahé et al. (2016) to choose a fixed aperture radius of 70 kpc. The median redshift of the ALFALFA sample is slightly lower, which would demand a slightly smaller aperture. Moreover, only H I within a certain range of relative velocity would be registered in the 21-cm band. However, Bahé et al. (2016) found that a spherical aperture of 70 kpc corresponds very well to a more complex procedure that takes the velocity of the gas into account (see their appendix A). We thus follow their suggestion and use a circular 70 kpc aperture in face-on projection for all our H I masses. As Arecibo is a single-dish telescope, we apply a Gaussian beam instead of a radial cut-off, i.e. we weight the projected H I density profile by a Gaussian with  $\sigma = 70$  kpc. The differences are negligible at all stellar masses except at  $M_* > 10^{11} M_\odot$  where using the Gaussian slightly increases the H I mass. We further discuss the impact of the chosen aperture in Appendix B.

One observational complication that is not taken into account is blending: galaxies sometimes lie behind each other and close enough in velocity space to be confused into a single detection, increasing the measured H I fraction and decreasing the number of objects. However, Jones et al. (2015) investigated this effect for ALFALFA and found that it would change the ALFALFA mass function by less than  $3\sigma$ , a small shift given the extremely tight error bars. In Appendix B, we compare our H I masses to those of Stevens et al. (2019), who take blending into account and find that the differences do not appreciably change our results. For our  $H_2$  masses, blending effects are negligible due to the small spatial extent of  $H_2$  and due to the optical galaxy selection (Sections 2.3 and 3.4.4).

Finally, in Section 4.5, we use the H I radius  $R_{H I}$ , defined as the radius where the surface density of H I falls below  $1 M_\odot \text{pc}^{-2}$ . We measure this radius from projected one-dimensional density profiles because the observational measurements are face-on equivalents, i.e. they are corrected for inclination. As discussed in Section 2.5, the data sets are heterogeneous but are generally very well resolved and include the entire H I disc. Thus, we do not impose an aperture when measuring  $M_{H I}$  for use in the  $R_{H I}$ – $M_{H I}$  relation.

### 3.4.4 $H_2$ masses and sizes

For  $H_2$ , the chosen aperture turns out to be much more important than for H I. The data used in C18 are dominated by the xCOLD GASS survey, which consists of 532 measurements of the CO luminosity taken with the IRAM 30-m telescope (Saintonge et al. 2017). The 3-mm channel (which contains the CO 1–0 transition) has a field of view of 22 arcsec. Over the survey’s redshift range,  $0.01 < z < 0.05$ , this corresponds to between 4.6 and 20 kpc. Based on the high-mass end of this redshift range, we use 20 kpc as an approximation for the typical beam of IRAM. The instrument used is a single-pixel bolometer, meaning that the beam is Gaussian in shape. The quoted beam size corresponds to the full width at half-maximum, or  $2.355\sigma$ , meaning that the standard deviation of our Gaussian aperture is  $\sigma \approx 8.5$  kpc. Saintonge et al. (2017) apply a correction factor for inclination, which increases the measured CO luminosity by a median factor of 1.17 (see also Saintonge et al. 2012). Since we mock-observe galaxies in face-on orientation, this correction should roughly match our procedure.

The *Herschel* Reference Survey observations that underlie our  $H_2$  mass function use a rather different instrument, SPIRE (Boselli et al. 2010). The SPIRE field of view is 4 arcmin by 8 arcmin, which we approximate by a single aperture of 6 arcmin. The survey sources lie between 15 and 25 Mpc; we take 20 Mpc as a typical distance. At



this separation, 6 arcmin corresponds to 36 kpc or a radial aperture of 18 kpc. Rather than a single beam such as IRAM, SPIRE has a resolution of 30 arcsec. Thus, we use a hard cut-off radius of 18 kpc rather than a Gaussian aperture, though the difference is insignificant.

Returning to the IRAM-based observations used in the gas fraction measurements, we might worry that the  $H_2$  mass of our sample could be overestimated because a fraction of the galaxies lies at lower redshift and experiences even smaller apertures. However, there are also reasons to fear that our measurement might underestimate the observations because of the spatial extent of  $H_2$  in IllustrisTNG. In Section 4.5.3, we demonstrate that the  $H_2$  in our simulated galaxies tends to be more spread out than suggested by CO observations, both in relation to the stellar component and in absolute units. This disagreement begs the question of whether we wish to mock-observe exactly what a telescope would see given the simulation data, or whether we are interested in quantifying the overall  $H_2$  mass of the simulated galaxies. For the purposes of gas fractions and mass functions, we are interested in the latter. Thus, our results should probably be seen as a lower limit on the  $H_2$  mass one would observe given the simulations. We quantitatively investigate the effect of the  $H_2$  aperture in Appendix B.

### 3.5 Separation into early and late types

As discussed in Section 2.3, we base our comparisons of gas fractions on the observational compilation of C18, who split galaxies into LTGs and ETGs. This distinction is physically meaningful because the  $H I$  and  $H_2$  fractions of the two groups differ significantly, with ETGs containing less gas than LTGs at fixed stellar mass. To assess whether IllustrisTNG galaxies match this trend, we split our sample into LTGs and ETGs based on the concentration of stellar mass,  $C_{80,20} > 4.9$ . This split reproduces the C18 ETG fraction for both centrals and satellites. Here, we define concentration as  $C_{80,20} \equiv 5 \log_{10}(r_{80}/r_{20})$ , where  $r_{80}$  and  $r_{20}$  are the radii enclosing 80 and 20 per cent of the gravitationally bound stellar mass, respectively. This parameter (observationally defined through stellar light rather than mass) is known to strongly correlate with the morphological type of galaxies (Kent 1985; Bershad, Jangren & Conselice 2000; Lotz, Primack & Madau 2004). The concentration in IllustrisTNG appears to match observations reasonably well (Rodriguez-Gomez et al. 2019), especially when a scatter in stellar mass is applied. In Appendix A, we describe our experiments with morphology and investigate the correlation between various morphological indicators and other galaxy properties.

## 4 RESULTS

We are now ready to compare the gas content of IllustrisTNG galaxies to observations. We split this section into six observational metrics, namely the overall abundance of  $H I$  and  $H_2$  (Section 4.1), mass functions (Section 4.2), gas fractions (Section 4.3), correlations with SFR (Section 4.4), the spatial distribution of gas (Section 4.5), and the correlation between neutral gas content and morphology (Section 4.6). Throughout the paper, we adhere to a consistent plotting scheme where the blue lines and shapes correspond to TNG100, orange to TNG300, and grey or black elements indicate observational data. Wherever we show  $H I/H_2$ -related quantities, we indicate the variation between the nine different models as a dark shaded area. We omit the individual model lines to avoid confusion and refer the interested reader to Paper I where the detailed differences between the models are

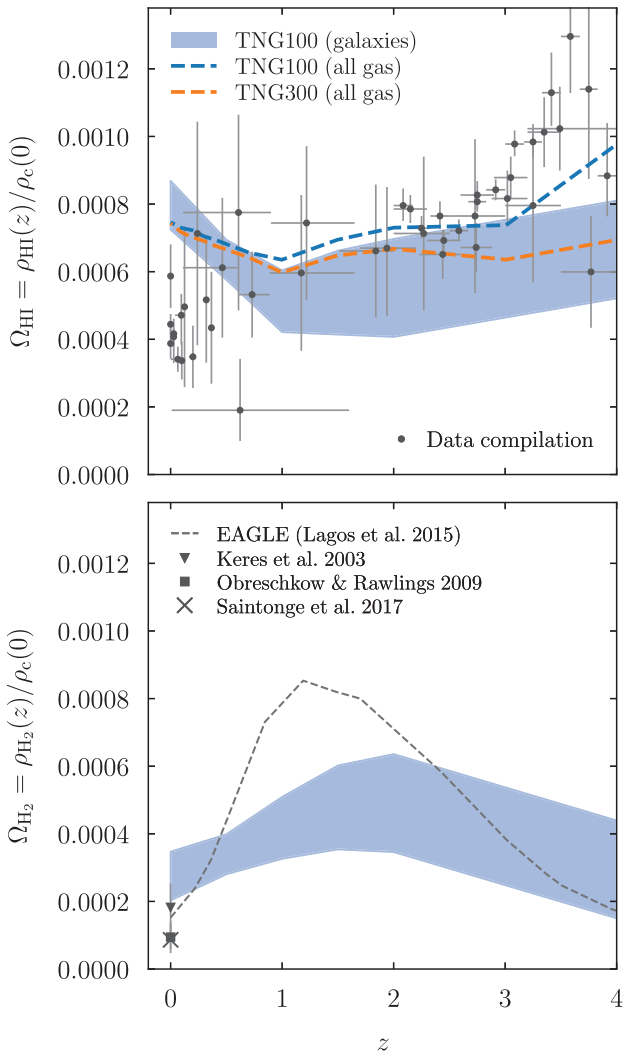
shown. Where applicable, the 68 per cent scatter is shown as a lighter shaded area of the same colour. In this case, it would not make sense to show the maximum scatter of the models; instead, we take the mean of the 16th and 84th percentiles of all the nine models. We focus on median rather than mean quantities to avoid issues with outliers and galaxies that contain no or very little gas. To avoid overcrowding the figures in this section, we do not show results from the original Illustris-1 simulation or from the lower resolution TNG100-2 simulation. However, we provide full sets of figures comparing TNG100 to Illustris-1 and to TNG100-2 at benediktdiemer.com/data. We briefly review the results for Illustris-1 in Section 4.7.

### 4.1 Overall abundance of $H I$ and $H_2$

To set a baseline expectation for the match between the neutral gas content in IllustrisTNG galaxies and the real Universe, we first consider the overall abundance of  $H I$  and  $H_2$  as a function of redshift as shown in Fig. 1. These abundances, denoted  $\Omega_{H I}$  and  $\Omega_{H_2}$ , are defined with respect to the critical density of the Universe at  $z = 0$ . The TNG100 result is to be understood as a lower limit because it does not include the  $H I$  and  $H_2$  in galaxies that were not included in our sample selection (Section 3.3). For this reason, we do not show TNG300 because the much higher stellar and gas mass limits would mean that significant gas reservoirs would be excluded. Furthermore, our calculation does not include any neutral gas not bound to haloes and subhaloes, although such gas should be negligible because virtually all non-bound gas is ionized by the UV background (hereafter UVB). We assess the importance of these effects by comparing to the  $\Omega_{H I}$  calculation of Villaescusa-Navarro et al. (2018). While their method used a simpler prescription for the  $H_2$  fraction, it included all gas cells in the simulation. The small difference to our galaxy sample indicates that the vast majority of  $H I$  resides within or near galaxies, at least below  $z \approx 2$ . We caution that the simulation results are not fully converged, especially at high redshift (fig. 26 in Villaescusa-Navarro et al. 2018). Moreover, the  $H I/H_2$  separation becomes less certain at high  $z$  because all our models were calibrated at low  $z$ , as evident from their increasing dispersion.

As this paper focuses on observational comparisons at  $z = 0$ , we are most interested in low-redshift constraints on  $\Omega_{H I}$  and  $\Omega_{H_2}$ . We assume that the observational data represent the total amount of  $H I$  and  $H_2$  in the Universe, with no severe effects due to explicit or implicit selection functions (e.g. Obreschkow et al. 2013). As shown in the top panel of Fig. 1, observations agree that  $\Omega_{H I}$  falls to about  $4 \times 10^{-4}$  at  $z = 0$ , about half of the  $H I$  we find in TNG100 (regardless of whether galaxies or all gas is used). Thus, we expect the  $H I$  mass function and gas fractions to also overestimate observations. For  $H_2$ , the picture is similar (bottom panel of Fig. 1). Here we expect virtually no molecular gas outside of galaxies (see also Lagos et al. 2015), though some galaxies might fall below our mass selection thresholds. Moreover,  $\Omega_{H_2}$  depends more strongly on the  $H I/H_2$  model than  $\Omega_{H I}$ . Despite these caveats, the  $H_2$  abundance in TNG100 appears to be somewhat higher than observed. However, some  $H I/H_2$  models fall within the observational uncertainty, especially given the additional systematic uncertainties due to the CO-to- $H_2$  conversion factor and due to sample selection.

The redshift evolution of  $\Omega_{H I}$  is characterized by relatively weak changes between  $z = 4$  and 1. Given the large uncertainties on the data, this result does not present a conflict with observations, particularly since we have not taken into account potential selection



**Figure 1.** Overall abundance of H I (top) and H<sub>2</sub> (bottom) as a function of redshift. As in all the following figures, the shaded blue region encloses the curves due to the different H I/H<sub>2</sub> models. The TNG100 result represents a lower limit because it does not include galaxies outside of our selection [ $\max(M_{\text{gas}}, M_*) \geq 2 \times 10^8 M_\odot$ , Section 3.3] or gas that is not associated with any subhalo. However, the dashed lines in the top panel demonstrate that using all gas particles gives similar results (Villaescusa-Navarro et al. 2018). We expect the amount of H<sub>2</sub> outside galaxies to be negligible. See Section 2.1 for details on the H I data compilation.

effects. At  $z < 0.5$ , however,  $\Omega_{\text{H I}}$  is observed to sharply decline, whereas the H I abundance in TNG100 increases. In contrast,  $\Omega_{\text{H}_2}$  displays an evolution of a factor of  $\approx 2$ – $3$ . Its peak at  $z \approx 2$  is close to the peak of the cosmic SFR density (e.g. Madau & Dickinson 2014), consistent with a picture where the differences between  $\Omega_{\text{H I}}$  and the cosmic SFR are owed to a changing H I-to-H<sub>2</sub> conversion efficiency, driven primarily by a combination of gas column density and metallicity (e.g. Obreschkow & Rawlings 2009; Lagos et al. 2014). The bottom panel of Fig. 1 also shows the evolution of  $\Omega_{\text{H}_2}$  in EAGLE according to Lagos et al. (2015). Their H<sub>2</sub> abundance rises to a higher level around  $z = 1$  but falls to a slightly lower level at  $z = 0$ . With rapidly improving CO data at high redshift, this evolution will be a useful constraint on galaxy formation models. In Fig. 1, we have omitted high- $z$  CO surveys because

their small volume and their selection functions complicate the interpretation. We refer the reader to Popping et al. (2019) for such a comparison.

In summary, TNG100 seems to exhibit a reasonable redshift evolution of  $\Omega_{\text{H I}}$  and  $\Omega_{\text{H}_2}$ , but contains about twice as much neutral gas as observed at  $z = 0$ . We further discuss this tension in Section 5.1.

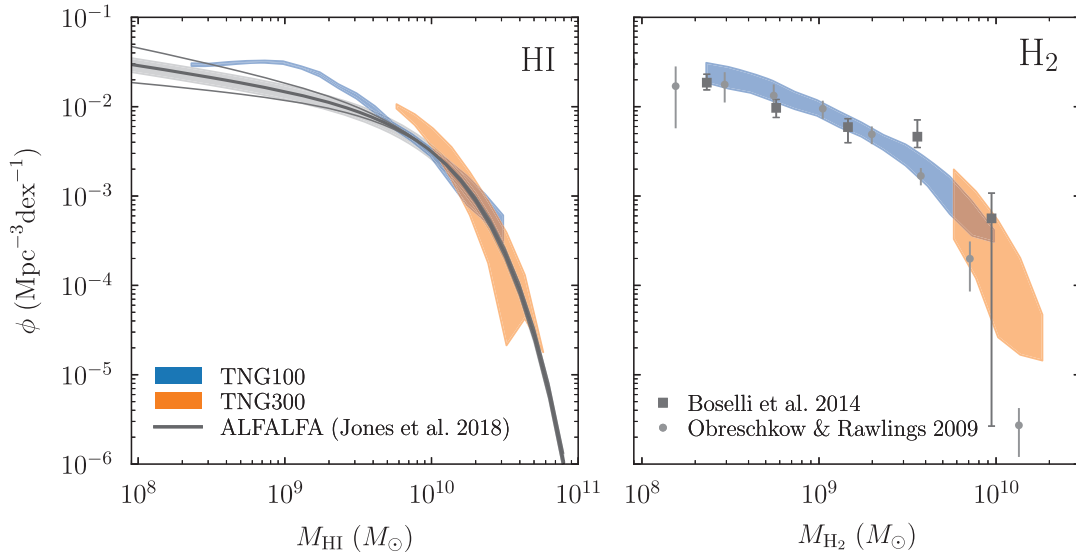
## 4.2 Mass functions

Fig. 2 shows the H I and H<sub>2</sub> mass functions from TNG100 and TNG300 as well as a number of observational constraints. Overall, the simulations match the data fairly well, especially at the high-mass end. The H I mass function is compared to the extremely well-measured mass function from the ALFALFA survey (Jones et al. 2018). The shaded uncertainty region around the ALFALFA mass function represents the systematic error on their normalization due to the uncertainty in the overall flux calibration; the thin lines show the mass function with low-mass slopes that are  $1\sigma$  lower and higher than the fiducial mass function. We cannot combine these uncertainties because Jones et al. (2018) do not provide covariances. There is no discernible tension above  $M_{\text{H I}} \gtrsim 5 \times 10^9 M_\odot$ , and the TNG100 and TNG300 results agree reasonably well. Between masses of about  $3 \times 10^8 M_\odot$  and  $5 \times 10^9 M_\odot$ , however, IllustrisTNG contains about 2–3 times too many objects. This prediction is robust to the choice of H I/H<sub>2</sub> model because the H<sub>2</sub> fraction is low, meaning that all H I/H<sub>2</sub> models basically predict the same amount of H I. Moreover, the choice of aperture has very little impact on H I and on this mass range in particular (Fig. B1). Given the results shown in Fig. 1, the disagreement is hardly surprising: since  $\Omega_{\text{H I}}$  exceeds the  $z = 0$  measurement from ALFALFA (Obuljen et al. 2018), the mass function must also exceed its ALFALFA counterpart at some masses.

Crain et al. (2017) showed that a similar excess in the EAGLE H I mass function is largely due to an unphysical, resolution-dependent population of objects that contain very little dark matter or stars. We have checked that this is not the case in IllustrisTNG, i.e. that the galaxies around  $M_{\text{H I}} \approx 10^9 M_\odot$  do not exhibit any obvious unphysical characteristics. Comparing to a lower resolution version of TNG100, however, we do find a similar excess in the H I mass function that is shifted to higher masses. A comparison with the upcoming TNG50 simulation (Nelson et al. 2019b; Pillepich et al. 2019) will reveal whether residual resolution effects change the TNG100 H I mass function.

While the H I mass function is more tightly constrained observationally, the H<sub>2</sub> mass function provides a more interesting point of comparison for our simulation because H<sub>2</sub> is a smaller fraction of neutral gas, meaning that the form of the H<sub>2</sub> mass function is less directly tied to the overall neutral gas content. The right-hand panel of Fig. 2 compares the H<sub>2</sub> mass functions in TNG100 and TNG300 to the observations of Obreschkow & Rawlings (2009) and Boselli et al. (2014b). The aperture chosen for the simulated galaxies was matched to the Boselli et al. (2014b) observations (Section 3.4.4); the Obreschkow & Rawlings (2009) points are shown to highlight the systematic observational uncertainty on the H<sub>2</sub> mass function. As with the H I mass function, TNG100 and TNG300 are well converged and agree with the observations at the high-mass and low-mass ends. Any apparent disagreements with the observed data are not truly significant, given that our H I/H<sub>2</sub> modelling is uncertain to at least a factor of 2 (Paper I) and that the observations rely on an uncertain CO-to-H<sub>2</sub> conversion factor (see also Popping et al. 2019).





**Figure 2.** Mass functions of atomic (left) and molecular (right) hydrogen at  $z = 0$ . The grey lines and points correspond to various observational measurements (Section 2.2), namely the ALFALFA H I mass function of Jones et al. (2018) and the H<sub>2</sub> mass function of Boselli et al. (2014b). The simulation data were mock-observed to match their respective apertures. The H<sub>2</sub> data of Obreschkow & Rawlings (2009) are shown only for comparison. TNG100 and TNG300 are relatively well converged and generally match the data. At low masses, however, the H I mass function is overestimated by a factor of 2–3. Given the  $1\sigma$  observational uncertainties in the normalization and low-mass slope (shown as a shaded area and as thin grey lines, respectively), the disagreement is statistically significant. Any disagreement with the observed H<sub>2</sub> mass functions is not significant given the systematic uncertainties in our modelling and in the CO-to-H<sub>2</sub> conversion.

### 4.3 Gas fractions

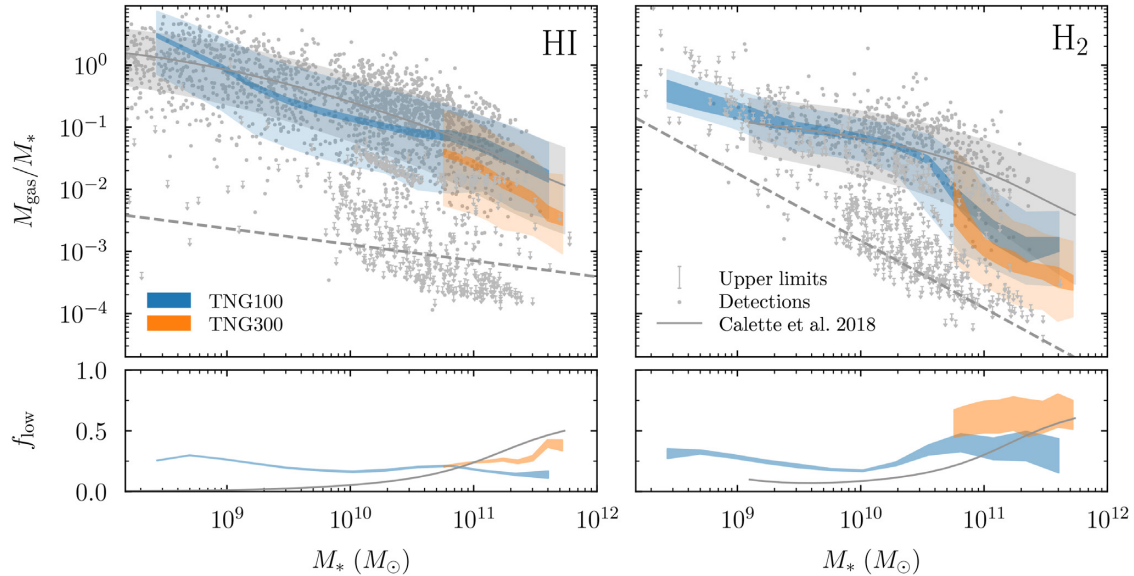
We now turn to the connection between the stellar and gaseous content of galaxies in IllustrisTNG, quantified by the fractions of H I and H<sub>2</sub> mass with respect to stellar mass (hereafter simply referred to as ‘gas fractions’ or  $f_{\text{gas}}$ ). Fig. 3 shows the gas fractions of H I and H<sub>2</sub> as a function of stellar mass. Only stellar mass bins with at least 50 galaxies are shown. When comparing gas fractions, simple means or medians are not appropriate measures of the respective distributions because a number of galaxies will contain zero gas cells, or at least lie below some observational detection threshold (for example, the sensitivity limit of radio telescopes). For this reason, we separately consider two quantities: the fraction of galaxies below some threshold gas fraction, and the distribution of gas fractions above this limit.

In defining the threshold gas fraction, we follow C18, who introduce a mass-dependent lower limit to the distributions for ETGs (dashed lines in Fig. 3). As we are combining the ETG and LTG samples in Fig. 3, we apply the same threshold to both ETGs and LTGs. C18 infer the distribution of gas fractions using Kaplan–Meier estimation, taking into account upper limits, and parametrize their results as separate fitting functions for the distribution of LTG and ETG gas fractions. We reconstruct the distribution of the full sample by summing the LTG and ETG contributions given the ETG fraction at each stellar mass (Fig. A1). For both the simulated and observed distributions, we now count the fraction of galaxies that lie below the threshold,  $f_{\text{low}}$ , and compute the median and 68 percent scatter of those gas fractions that lie above the threshold. The exact value of the threshold is unimportant because it is applied to both simulation and observations. We note that  $f_{\text{low}}$  is not the fraction of galaxies without a gas detection, but rather the fraction of galaxies whose gas fraction falls below the threshold that is imposed a posteriori. Thus, we do not need to worry about

whether a simulated galaxy would or would not have been detected by a given observational survey. If its gas fraction lies below the threshold, it is registered as such by C18. The individual observed data points shown in Fig. 3 highlight the importance of applying an estimator that takes into account upper limits and selection effects.

We first consider the distribution of  $f_{\text{gas}}$  above the threshold. Overall, the H I fraction matches the results of C18 well, including a realistic 68 percent scatter. The differences between the H I/H<sub>2</sub> models are relatively small for H I. While the median dips below the observed median by a factor of about 2 at intermediate stellar masses, it is unclear whether this difference is significant. TNG300 exhibits systematically lower gas abundances than TNG100. This undesirable resolution dependence could be resolved with a rescaling procedure (see e.g. Pillepich et al. 2018b, for stellar masses), but we do not undertake such an exercise here.

The median H<sub>2</sub> fraction is almost perfectly matched at stellar masses below  $2 \times 10^{10} M_{\odot}$ ; at the highest masses, it falls below the observations by a factor of about 4–10 depending on the H I/H<sub>2</sub> model. However, it is not clear how significant this difference is for a number of reasons. First, as discussed in Section 2.1, the observations are systematically uncertain due to the poorly known CO-to-H<sub>2</sub> conversion factor. Secondly, the low density of observations above  $2 \times 10^{11} M_{\odot}$  means that the distribution inferred by C18 is less reliable than at lower masses. Finally, the estimated H<sub>2</sub> fraction becomes more model-dependent at high stellar masses (see also Paper I). Part of this effect is due to the different spatial distribution predicted by the models, which, given a limited observational aperture, can cut out significant fractions of the total H<sub>2</sub> mass and leads to the sharp drop in the H<sub>2</sub> fraction at high masses. In summary, there may be some tension between the observed and simulated H<sub>2</sub> fractions at the high-mass end, but it is hard to quantify this tension exactly.



**Figure 3.** Gas fractions of atomic (left) and molecular hydrogen (right) as a function of stellar mass at  $z = 0$ . The grey lines and shaded areas show the inferred median gas fractions and 68 per cent scatter according to the compilation of C18. These values refer only to galaxies whose gas fractions lie above the dashed lines. The distributions of lower gas fractions could not be reliably determined in C18; their fractional contribution is shown in the bottom panels. As a result, the median gas fractions in the top panels can appear different from those shown in C18, especially at high stellar masses where many galaxies fall below the threshold. The simulation data are analysed in the same way, i.e. by separately counting galaxies that fall below the lower limits. The darker shaded areas show the area covered by all nine H I/H<sub>2</sub> models used and can be interpreted as a systematic uncertainty. The lighter shaded area indicates the mean 68 per cent scatter, i.e. the mean of the scatter contours according to the different models. The detections and upper limits used in the C18 analysis are shown as grey dots and arrows (combining both their LTG and ETG samples). See Section 4.3 for a detailed discussion of these results.

We now turn to the fraction of galaxies with  $f_{\text{gas}}$  below the lower limit. According to C18, this fraction increases with stellar mass, reaching about 50 per cent of both the H I and H<sub>2</sub> distributions at the highest stellar masses (bottom panels of Fig. 3). The IllustrisTNG galaxies do not follow this trend. In the H I distribution, especially, the fraction of low-gas galaxies decreases slightly at high masses. Most notably, the simulations contain a population of at least 25 per cent below the limit at all stellar masses. Virtually all of those, about 85–90 per cent depending on the H I/H<sub>2</sub> model, are satellites. A small fraction of those turn out to be satellites that contain no gas cells at all, presumably because they were stripped away. The high satellite fraction in the low-gas galaxies can perhaps partially explain the disagreement with C18 because some observational samples are likely biased against close satellite pairs due to effects such as fibre collisions (Section 5.2). However, in Section 5.2, we discuss this issue further and conclude that fibre collisions cannot explain the bulk of the excess of galaxies with low gas fractions. Similarly, blending can assign some H I mass to gas-free galaxies, but this effect impacts only a few per cent of our galaxies (Appendix B).

We conclude that the gas fractions in IllustrisTNG broadly match the observations, but that there is an excess of satellites with very low gas fractions. By considering LTGs and ETGs as one combined sample, we have sidestepped the question of how the gas fractions correlate with the morphological type of galaxies; we return to this question in Section 4.6.

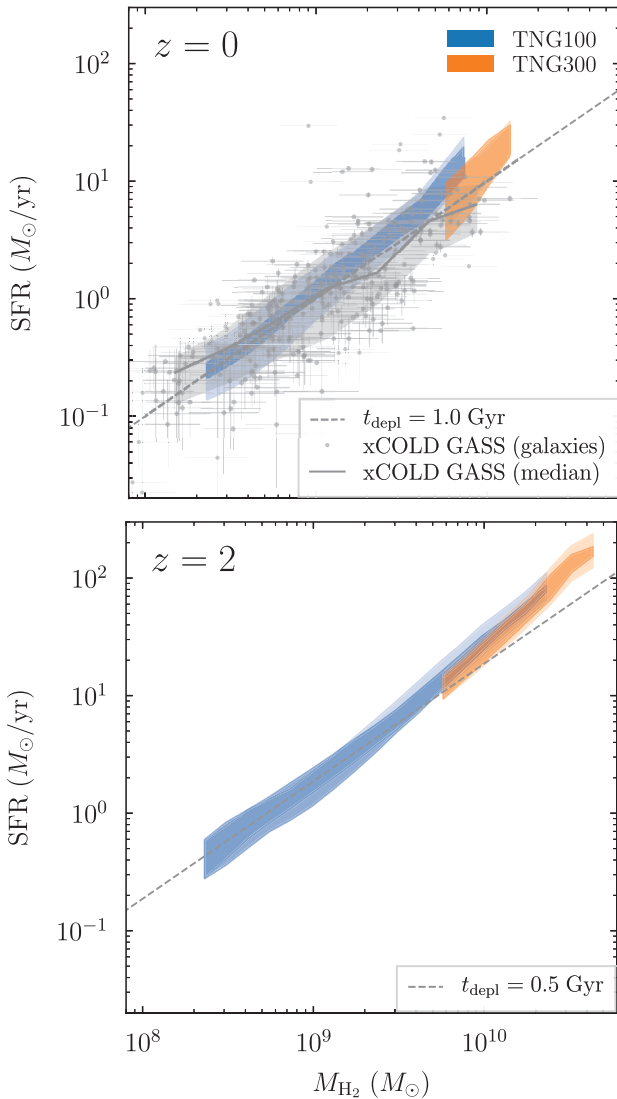
#### 4.4 Correlations with SFR

Given the reasonable gas fractions and the generally realistic star formation activity in IllustrisTNG (Donnari et al. 2019), we expect that the molecular gas reservoirs of galaxies should tightly correlate

with their SFR. This trend is well established observationally (Fig. 4). We note that the  $M_{\text{H}_2}$ –SFR correlation is not quite the same as the Kennicutt–Schmidt relation (Schmidt 1959; Kennicutt 1998) because we have not normalized the mass and SFR by area.

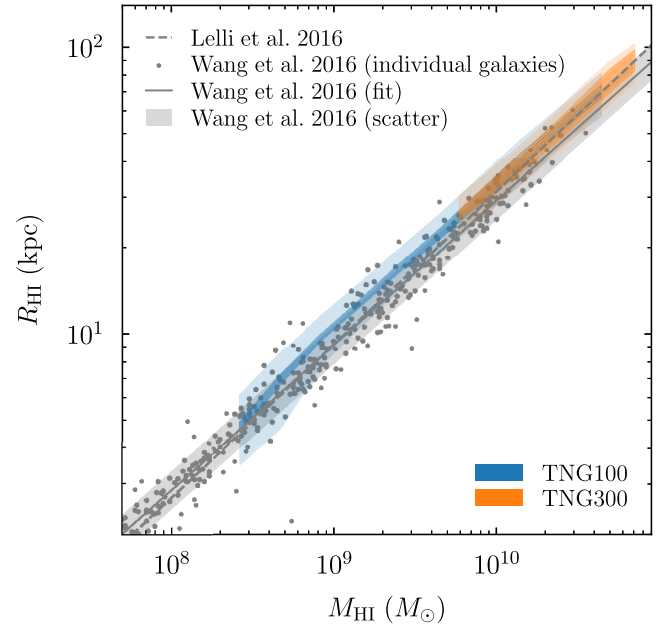
The top panel of Fig. 4 shows the median SFR as a function of  $M_{\text{H}_2}$  for both observations and simulations at  $z = 0$ . The xCOLD GASS galaxy sample is representative in the sense that it was randomly selected from a particular stellar mass range (Saintonge et al. 2012, 2017), and that galaxies were observed until a certain H<sub>2</sub> fraction was detected or a certain upper limit was reached. In particular, galaxies in the COLD GASS sample ( $M_* > 10^{10} M_\odot$ ) were guaranteed to be detected in CO if  $M_{\text{H}_2}/M_* > 1.5$  per cent, while galaxies in the lower mass COLD GASS-low sample ( $10^9 < M_* < 10^{10} M_\odot$ ) were detected if  $M_{\text{H}_2}/M_* > 2.5$  per cent. To match this selection, we have chosen all simulated galaxies in the respective stellar mass ranges that reach the limiting H<sub>2</sub> mass. Given the mixed sources of the SFR measurements, there is no clear detection limit on the SFR (Saintonge et al. 2017). We take the lowest SFR in the sample,  $10^{-2.8} M_\odot \text{ yr}^{-1}$ , as a guideline and neglect all IllustrisTNG galaxies whose SFR falls below this value. In practice, this cut makes no discernible difference.

The median trend at  $z = 0$  is matched well by IllustrisTNG, to a factor of 2 or better at all masses. For TNG300, the situation is less clear because the data become somewhat sparse at the highest masses. The scatter in TNG100 appears to be slightly smaller than observed, but the observed scatter constitutes an upper limit because it does not take into account the sizeable error bars on the individual SFR measurements. Moreover, the observed SFRs correspond to an average over a few hundred Myr and might therefore be subject to added scatter compared to the instantaneous SFRs used for the simulation data (Section 3.4.2).



**Figure 4.** Relation between molecular gas mass and SFR at  $z = 0$  (top) and  $z = 2$  (bottom). In the top panel, the grey points show individual galaxies from xCOLD GASS (Saintonge et al. 2017), and the solid line and shaded region show their median and 68 per cent scatter, respectively. IllustrisTNG matches the median relation well, to better than a factor of 2 at all masses. In TNG300, the scatter is smaller than the variation due to the  $\text{H I}/\text{H}_2$  models. The relation closely approximates a constant depletion time as parametrized by Tacconi et al. (2018, dashed grey lines).

An alternative measure by which to understand the relation between gas mass and star formation is the depletion time,  $t_{\text{depl}} \equiv M_{\text{H}_2}/\text{SFR}$ , i.e. the time over which a gas reservoir would be exhausted given the current SFR. Observationally, the depletion time is thought to be more or less independent of mass but to evolve with redshift. The dashed lines in Fig. 4 correspond to a constant depletion time as parametrized by Tacconi et al. (2018),  $t_{\text{depl}} \approx 1 \text{ Gyr} \times (1+z)^{-0.57}$ . IllustrisTNG reproduces this relation between  $z = 0$  and 2, although the highest mass galaxies have higher SFRs than expected by a factor of about 2. The comparison to Tacconi et al. (2018) should be taken as approximate because we have not matched our galaxy selection to theirs, which includes only galaxies on the star-forming main sequence. However, we have



**Figure 5.** The  $\text{H I}$  mass–size relation at  $z = 0$ , where  $R_{\text{H I}}$  is defined as the radius where  $\Sigma_{\text{H I}}$  falls below  $1 \text{ M}_{\odot} \text{ pc}^{-2}$ . The grey points show the data compilation of Wang et al. (2016), and the solid grey line shows their best linear fit. For a comparison, the dashed line shows the fit from Lelli et al. (2016).

checked that the  $M_{\text{H}_2}$ –SFR connection does not significantly differ between ETGs and LTGs (defined as described in Section 3.5).

Lagos et al. (2015) reported a similar result for the EAGLE simulation, namely a roughly constant depletion time of  $t_{\text{depl}} \approx 10^9$  Gyr at  $z = 0$ . They argued that an overly tight  $M_{\text{H}_2}$ –SFR relation is perhaps to be expected because the star formation in simulations is directly based on the density of gas, and thus the density of cold gas in the ISM models.

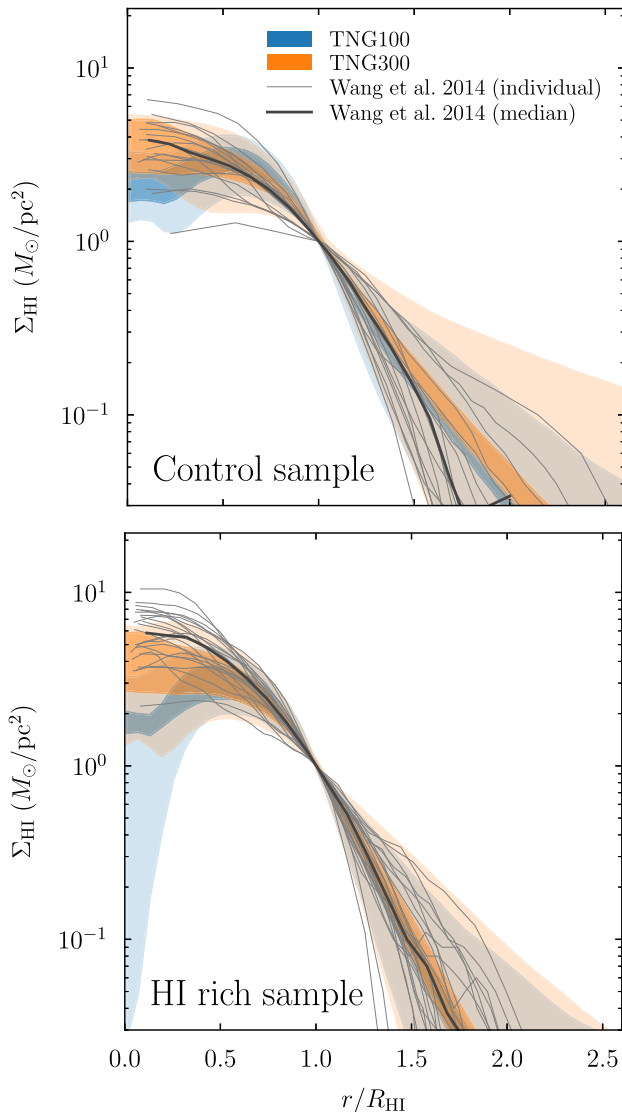
#### 4.5 Spatial distribution

Having investigated how atomic and molecular gas is distributed between galaxies of different stellar masses and SFRs, we now turn to the distribution of gas within galaxies.

##### 4.5.1 The $\text{H I}$ mass–size relation

Fig. 5 shows the  $\text{H I}$  mass–size relation as measured in TNG100 and TNG300, and in the observational compilations of Wang et al. (2016) and Lelli et al. (2016). Observationally, the relation is extremely tight, with only 0.06 dex scatter (15 per cent, shaded grey). Its slope is closed to 0.5, the value one would obtain for a disc of constant surface density (0.506 in Wang et al. 2016, 0.535 in Lelli et al. 2016). Both TNG100 and TNG300 match the slope of the relation almost perfectly, but predict very slightly larger  $\text{H I}$  radii than observed, about 14 per cent above the measured relation. The  $\text{H I}/\text{H}_2$  models agree to better than 13 per cent for all bins for TNG100 and to better than 20 per cent for TNG300, a much more well-defined prediction than for gas fractions or mass functions. Interestingly, TNG100 and TNG300 are almost perfectly converged in this metric, unlike in any other quantity we consider. We discuss the physical reasons for the tight relation and the good agreement in Section 5.3.





**Figure 6.** Median radial profiles of the H I surface density. The light shaded areas show the 68 per cent scatter. As the profiles are scaled to  $R_{\text{HI}}$ , they overlap at  $\Sigma_{\text{HI}} = 1 \text{ M}_{\odot} \text{ pc}^{-2}$  by definition. The grey lines show individual galaxies and the median for the Bluedisk sample of Wang et al. (2014). The top panel shows their control sample ( $10^{9.1} < M_{\text{HI}} < 10^{9.8} \text{ M}_{\odot}$ ), and the bottom panel their H I-rich sample ( $M_{\text{HI}} > 10^{9.8} \text{ M}_{\odot}$ ).

#### 4.5.2 H I surface density profiles

Having established that  $R_{\text{HI}}$  matches observations well, we proceed to the radial distribution of H I. Fig. 6 shows the radial H I profiles from the Bluedisk survey (Wang et al. 2013, 2014). Their sample was designed to contain an H I-rich and a control subset (Section 2.5). To mimic their selection, we choose all galaxies with  $10^{10} < M_{*} < 10^{11} \text{ M}_{\odot}$ ,  $0.5 < \text{SFR} < 10 \text{ M}_{\odot} \text{ yr}^{-1}$ , and a stellar half-mass radius of at least 3 kpc (though the latter cut makes virtually no difference). For the control sample, we select galaxies with  $10^{9.1} \text{ M}_{\odot} < M_{\text{HI}} < 10^{9.8} \text{ M}_{\odot}$ , and for the H I-rich sample we apply a minimum of  $M_{\text{HI}} > 10^{9.8} \text{ M}_{\odot}$  (according to the respective H I/H<sub>2</sub> model). Bahé et al. (2016) carefully demonstrated that these relatively simple cuts match the Bluedisk selection closely.

Observationally, the median H I profiles for the two samples turn out to be almost indistinguishable, but the H I-rich IllustrisTNG galaxies do exhibit somewhat different profiles from the control galaxies. The median profiles of the control sample match the data reasonably well, though TNG100 galaxies exhibit a slight deficit at small radii. In both TNG100 and TNG300, the H I profiles in the control sample are perhaps a little too extended at large radii, though the difference is well within the scatter of the observations. In the more massive H I-rich sample, the median outer profiles are matched extremely well but the simulations show clear evidence of central H I holes: up to a factor of 5 deficit of H I at the centre. This trend is more pronounced in TNG100, but both TNG100 and TNG300 show large scatter at the smallest radii. The overall features of the profiles are independent of the H I/H<sub>2</sub> model, which is expected at large radii where H I dominates over H<sub>2</sub>, but not necessarily at small radii where the molecular fraction could be high.

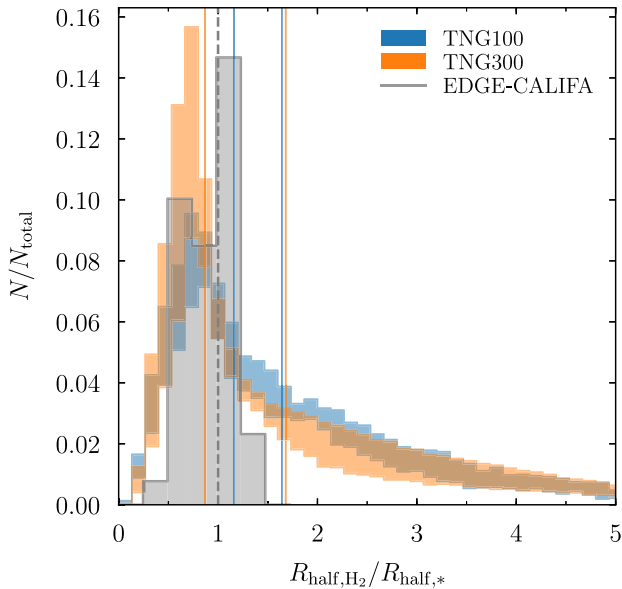
H I holes have also been observed in some galaxies (e.g. Boomsma et al. 2008) but are not prevalent in the Bluedisk sample. We note that the holes in IllustrisTNG galaxies are not likely to be a resolution effect. First, they are more pronounced in the higher resolution TNG100 than in TNG300. Secondly, they are more pronounced in the H I-rich sample, which contains the more massive, and thus more extended, H I discs. According to the median H I mass–size relation,  $R_{\text{HI}} \approx 30 \text{ kpc}$  at the lower mass end of this sample, about 40 times the force resolution length in TNG100. Finally, the holes are not unique to either volumetric or projected models but a robust prediction of both.

The question, then, arises as to what causes the central H I holes. There are two possible processes that can create them. First, gas at the centre can be ionized or ejected, for example due to feedback from AGNs in large galaxies (Weinberger et al. 2018). Secondly, the inner part of the gas disc can have high enough surface densities to almost entirely transition to H<sub>2</sub>. We find evidence for the former mechanism (see also Nelson et al. 2019b). First, the prevalence of H I holes increases with mass, a trend that would be expected for AGN feedback. Secondly, the median neutral fraction,  $f_{\text{HI}+\text{H}_2}$ , drops towards the centre of massive galaxies, indicating that the majority of the gas is ionized in many objects. Thirdly, the prevalence of H I holes depends on the star formation activity: galaxies with a specific SFR (sSFR) above  $10^{-10} \text{ yr}^{-1}$  exhibit barely any H I holes, whereas the vast majority of galaxies below an sSFR of  $10^{-11} \text{ yr}^{-1}$  do exhibit them. Thus, H I holes do not seem to be connected to a dominance of H<sub>2</sub> over H I. The control and H I-rich samples select slightly different SFRs, explaining the different median profiles in Fig. 6.

Bahé et al. (2016) analysed H I profiles in the EAGLE simulation and report that a significant fraction of their galaxies exhibit clear H I holes, with holes becoming more prevalent towards high H I masses (their fig. 5). The holes in EAGLE, however, appear not only at the centre but also throughout the disc, and are traced to their particular feedback implementation (Bahé et al. 2016).

#### 4.5.3 The extent of the H<sub>2</sub> distribution

As discussed in Section 2.5, there are a number of factors that complicate the interpretation of observed molecular profiles. Thus, we restrict our comparison to a simple measure of the relative extent of the molecular and stellar distributions: the ratio of H<sub>2</sub> and stellar half-mass radii, shown in Fig. 7. The grey histogram shows the ratio of CO and stellar half-mass radii from the EDGE-CALIFA survey (Bolatto et al. 2017, see their fig. 12). The CALIFA selection function is based mostly on the optical radii of galaxies (Walcher et al. 2014), but the EDGE selection complicates this picture



**Figure 7.** Distribution of the ratio of  $H_2$  and stellar half-mass radii, a crude test of the extent of  $H_2$  compared to the optical size of the galaxy. The grey histogram shows the distribution of CO-to-stellar half-mass radii from the EDGE-CALIFA survey (Bolatto et al. 2017). Their best-fitting CO-to-stellar scale ratio is unity (dashed line), whereas IllustrisTNG galaxies exhibit somewhat larger median ratios and large scatter (the thin vertical lines bracket the medians according to the different  $H\text{I}/H_2$  models).

further. To mimic the overall selection, we choose all simulated galaxies that fall within the ranges of galaxy properties observed in the final EDGE-CALIFA sample, namely, a projected stellar half-mass radius between 1.3 and 8.8 kpc,  $10^{9.5} M_\odot < M_* < 10^{12} M_\odot$ ,  $M_{H_2} > 10^8 M_\odot$ , and an SFR between 0.01 and  $100 M_\odot \text{ yr}^{-1}$ . We have verified that our results are not sensitive to the exact values of any of these limits.

Bolatto et al. (2017) find a relatively tight relation between the CO and stellar half-mass radii, with a best-fitting ratio of 1.00 (dashed line in Fig. 7). In agreement with these observations, a large fraction of IllustrisTNG galaxies exhibit radius ratios around unity, but there is also a significant tail towards large ratios. These galaxies have an  $H_2$  distribution that is much more extended than the stellar distribution, whereas no such objects were observed in EDGE-CALIFA. As a result, the median ratio is closer to 1.3 in the simulations, though with a large dependence on the  $H\text{I}/H_2$  model. The median ratio tends to increase towards higher  $H_2$  masses. The stellar sizes in IllustrisTNG match observations well (Genel et al. 2018), meaning that the disagreements are likely due to the extent of the  $H_2$  distribution. The  $H_2$  radii appear well converged, with small differences between TNG100 and TNG300, or with the lower resolution version TNG100-2. Splitting the sample into LTGs and ETGs as described in Section 3.5 has only a modest effect on the results.

While half-mass radii have the advantage of being simple to measure, observed CO and stellar profiles are often fitted with an exponential (e.g. Leroy et al. 2008). This procedure tends to give very similar results: when comparing the fitted scale lengths of the CO and stellar profiles, Bolatto et al. (2017) find a best-fitting relation of  $l_{CO} = 0.89 l_*$ , and Leroy et al. (2008) find  $l_{CO} \approx (0.9 \pm 0.2) l_*$ . We thus conclude that our comparison is robust to the exact definition of the scale radii.

There are, however, a number of other caveats. For example, the CO-to- $H_2$  conversion factor could depend on radius, leading to CO

profiles that would systematically differ from  $H_2$  profiles (Bolatto et al. 2013, and references therein). Moreover, the simple CO-to-stellar half-mass ratio shown in Fig. 7 may paint too simplistic a picture because the CO surface density profiles of galaxies have been found to be diverse (e.g. Young & Scoville 1991; Regan et al. 2001; Leroy et al. 2008, 2009). For example, CO emission at large radii has been seen at high redshift (Cicone et al. 2015; Ginolfi et al. 2017; Dannerbauer et al. 2017). While these complications mean that we cannot draw strong conclusions from Fig. 7, there is another piece of evidence that our simulated  $H_2$  distributions are too extended: we showed that IllustrisTNG does not quite match the Kennicutt–Schmidt relation at the high-surface-density end, with low  $\Sigma_{H_2}$  at fixed SFR (fig. 3 in Paper I). This mismatch indicates that star formation in the simulation happens at lower surface densities than observed, which could well be a resolution effect.

In summary, we find tentative evidence that the spatial distribution of  $H_2$  in our modelling is somewhat more extended than in observations. We leave a more detailed comparison for future work.

#### 4.6 Correlation with morphology

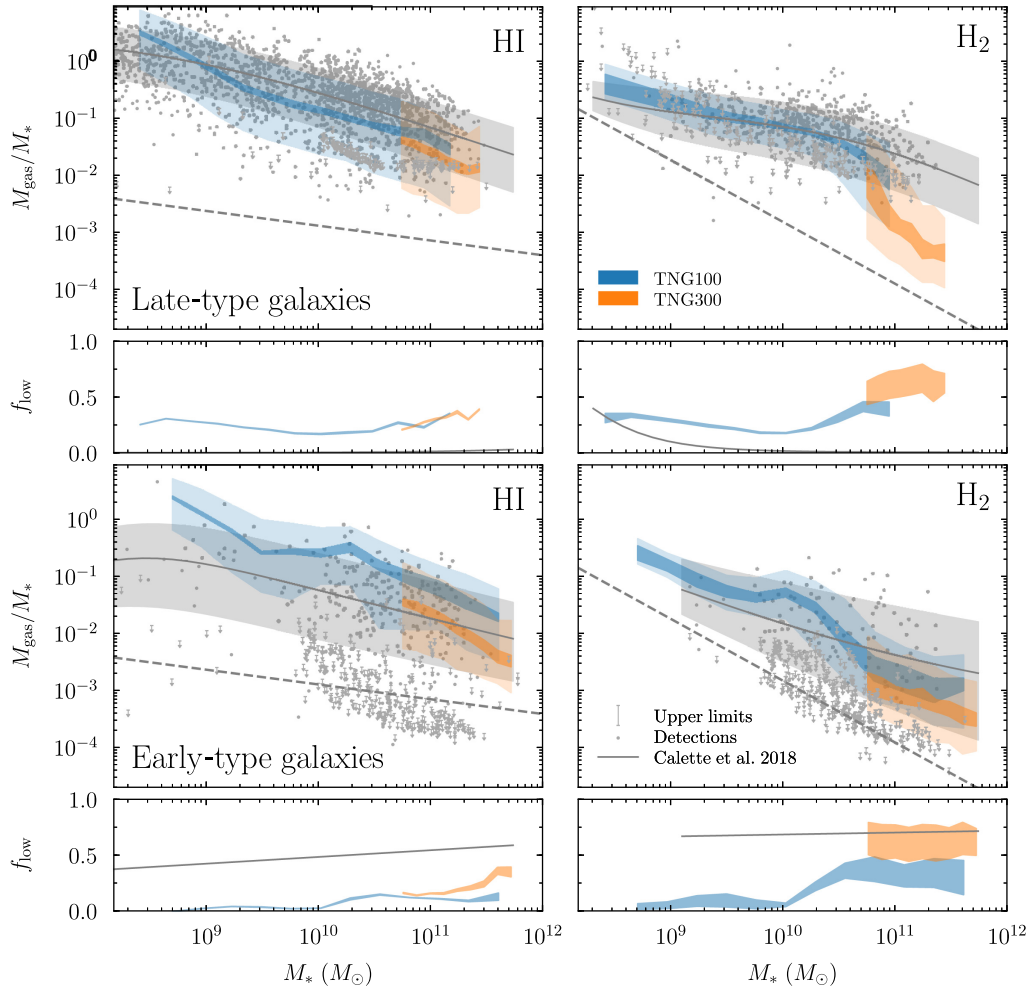
When considering gas fractions as a function of stellar mass in Section 4.3, we combined the C18 results for LTGs and ETGs to recover the distribution of the overall galaxy sample. Considering the types separately, however, tightens the scatter on the relations and allows us to check whether the gas content of IllustrisTNG galaxies correlates with their stellar morphology as expected. To this end, Fig. 8 shows the same comparison of gas fractions as Fig. 3, but split into LTGs (top row) and ETGs (bottom row) according to stellar mass concentration (Section 3.5 and Appendix A).

This comparison reveals a number of issues. First, the fractions of galaxies below the  $f_{\text{gas}}$  threshold trend the wrong way: a significant number of LTGs have no gas when almost no such objects are observed, whereas there are not enough low-gas ETGs. Clearly, a lack of neutral gas does not correlate well with the shape of the stellar distribution as captured by concentration. Secondly, the median  $H\text{I}$  fractions in the ETG sample are too high. While we found that the median gas fractions and scatter of the overall galaxy sample are in good agreement with data, we conclude that these gas fractions are not always assigned to galaxy samples with the expected morphological properties.

However, when splitting the sample by colour, we find much better agreement with the observed gas fractions than in Fig. 8. This conclusion is in line with the work of Rodríguez-Gomez et al. (2019), who found that galaxy morphology and colour in IllustrisTNG do not correlate quite as observed. In Appendix A, we further discuss the connection between gas content, various morphological indicators, and colour. We conclude that the gas content of galaxies correlates much more strongly with colour than with morphology in IllustrisTNG.

#### 4.7 Results for the original Illustris simulation

The IllustrisTNG physics model represents a major improvement over the original Illustris model (Pillepich et al. 2018a). However, it is still informative to evaluate the original simulation with regard to its  $H\text{I}$  and  $H_2$  content. For this purpose, we provide Illustris-1 versions of Figs 1 through 7 at benediktdiener.com/data. In this section, we briefly describe the main conclusions. We omit Fig. 8 from the set and from our discussion because our chosen morphological indicator,  $C_{80,20}$ , takes on such different values



**Figure 8.** Same as Fig. 3, but with the galaxy sample split into LTGs (top row) and ETGs (bottom row) by the concentration of their stellar mass. The C18 constraints are derived directly from their best-fitting distributions. The match to observations is somewhat worse than for the overall sample, highlighting that the gas content of IllustrisTNG galaxies does not correlate with morphology as observed in the real Universe.

in Illustris-1 that it cannot be used for a sensible morphological classification.

Most notably,  $\Omega_{\text{HI}}$  and  $\Omega_{\text{H}_2}$  are massively overestimated in Illustris-1 at  $z = 0$ , by almost an order of magnitude. This neutral gas excess also manifests itself in the HI and H<sub>2</sub> mass functions, which are similarly overestimated by up to an order of magnitude at some masses. It follows that the gas fractions also exceed observations, by a factor of about 5 at the low-mass end. They decrease strongly towards high masses, falling below the C18 relations at  $M_* > 10^{11} M_\odot$ . This trend is also apparent in fig. 3(a) of Vogelsberger et al. (2014b), but the Huang et al. (2012b) ALFALFA gas fractions shown there are somewhat higher than inferred by C18, leading to a seemingly better agreement at the low-mass end. Interestingly, Illustris-1 matches the fraction of galaxies below the threshold better than IllustrisTNG as it contains a much smaller population of very gas-poor objects and follows the expected trends with stellar mass. Given the excess of both HI and H<sub>2</sub>, it is unsurprising that the correlation between  $M_{\text{H}_2}$  and SFR at  $z = 0$  is similarly off, with a large and non-constant depletion time. At  $z = 2$ , however, Illustris-1 does roughly match the constant depletion time found in IllustrisTNG.

Despite these significant disagreements, the HI mass–size relation turns out to be robust once again: while slightly shallower

in Illustris, the normalization is matched at  $M_{\text{HI}} \approx 10^9 M_\odot$ . This agreement is somewhat unexpected, given that the stellar sizes of galaxies were about twice too large below  $M_* \lesssim 10^{10.5} M_\odot$  in Illustris-1, an observable that is much improved in TNG100 (Genel et al. 2018; Pillepich et al. 2018a). Thus, the improvement in stellar sizes was not driven by a change in the distribution of neutral gas. This conclusion is confirmed by the HI profiles, which more or less agree between Illustris-1 and IllustrisTNG. Conversely, the median H<sub>2</sub> half-mass radius in Illustris-1 is significantly larger than in IllustrisTNG.

In summary, Illustris-1 galaxies contain a significant excess of neutral gas at  $z = 0$ , leading to offsets in all correlations with other galaxy properties. However, the properties of the excessively large HI discs are internally consistent. Our findings highlight the enormous improvements of the IllustrisTNG galaxy formation model compared to the original model.

## 5 DISCUSSION

In this section, we discuss the physics behind the most striking agreements and tensions with observations that we have discovered. We also consider anecdotal evidence from individual galaxies such as the Milky Way.



### 5.1 Is there an excess of neutral gas in IllustrisTNG?

Perhaps the most significant tension with observations is that IllustrisTNG appears to contain about twice as much neutral gas at  $z = 0$  as expected observationally, despite efficient stellar and AGN feedback. This conclusion is entirely independent of the  $\text{H I}/\text{H}_2$  modelling, which takes the neutral gas abundance as an input from the simulation. In non-star-forming cells (below the density threshold of  $n_{\text{H}} = 0.106 \text{ cm}^{-3}$ ), the neutral fraction is determined numerically by the balance between cooling, the photoionization rate due to the UVB (Faucher-Giguère et al. 2009) and due to nearby AGNs, and self-shielding according to a fitting function (Rahmati et al. 2013a, see Vogelsberger et al. 2013 for details). In cells above the density threshold, the Springel & Hernquist (2003) two-phase ISM model governs the gas physics. There, we assume that all gas in cold clouds is entirely neutral whereas the hot phase is entirely ionized, leading to a neutral fraction of about 90 per cent. If the fraction of star-forming gas was drastically different, we would expect corresponding changes in the stellar properties of IllustrisTNG galaxies, which more or less agree with observations. Given these considerations, and assuming that the observations are accurate, there are a number of possible reasons why the neutral fraction might be overestimated in star-forming cells, in non-star-forming cells, or in both.

In reality, some part of the cold-cloud phase in the Springel & Hernquist (2003) model could be ionized due to radiation from young stars (Rahmati et al. 2013b, see also section 4.3 of Paper I). However, only about 30 per cent of the total neutral gas in TNG100 stems from star-forming cells, meaning that even erasing all neutral gas from them would not reduce the overall neutral abundance by a factor of 2.

In non-star-forming gas, the neutral fraction could be lowered if the gas was heated or experienced stronger photoionization. For example, stronger feedback could remove some gas that is dense enough to self-shield and contain a significant neutral component. Another explanation could be the strength of the UVB: if the model of Faucher-Giguère et al. (2009) underestimated the true background significantly, the neutral fraction would be artificially increased, perhaps without significantly increasing the star formation activity. We have tested this mechanism using a smaller test simulation that uses the Puchwein et al. (2019) UVB instead of Faucher-Giguère et al. (2009). In the test simulation, the neutral gas abundance is reduced by about 10 per cent, indicating that the UVB would need to change somewhat drastically to reduce the neutral gas masses by a factor of 2.

While none of these effects seem likely to account for the entire disagreement in the neutral gas abundance, several of them could conspire. We note that the disagreement in  $\Omega_{\text{H I}}$  is driven largely by an upturn between  $z = 1$  and 0 that is not apparent in observations (Fig. 1). The upturn is partially caused by a decrease in the  $\text{H}_2$  abundance over the same time interval, but the combined neutral abundance does increase in TNG100. An understanding of the reasons for this trend might shed light on the disagreement in  $\Omega_{\text{H I}}$ , an investigation we leave for future work.

### 5.2 What is the origin of gas-free galaxies in IllustrisTNG?

While investigating  $\text{H I}$  and  $\text{H}_2$  fractions in Section 4.3, we discovered that IllustrisTNG contains a significant population of satellites whose gas fractions lie below the detection threshold. Upon closer inspection, some of those galaxies turn out not to contain any gas at all, that is, the halo finder associates no gas cells with them despite

their significant stellar masses. Such gas-free galaxies account for 7 per cent of our TNG100 sample ( $M_* > 2 \times 10^8 M_{\odot}$ ) and are found across the entire range of stellar mass. While gas-free galaxies account only partially for the excess of galaxies below the limit in Fig. 3, they are, at first sight, a strange population that we wish to investigate further.

The vast majority of gas-free galaxies, 94 per cent, are satellites. They can occupy (sub)haloes as massive as  $10^{12} M_{\odot}$ , but live in dark matter haloes that are, on average, about 0.5 dex less massive than those of their gaseous counterparts. About 20 per cent of them contain no dark matter at all, meaning that they are purely stellar clumps. Those objects are mostly tagged as possible halo-finder artefacts (Nelson et al. 2019a, see also Genel et al. 2018 and Pillepich et al. 2018b). Conversely, out of those gas-free galaxies that do have dark haloes, only 3 per cent are tagged. While the average satellite in our sample resides at a median radius of  $0.8R_{200c}$  (where  $R_{200c}$  refers to the halo radius of the host halo), the gas-free satellites live closer to their hosts, at a median distance of  $0.55R_{200c}$ . We find virtually the same distribution of distances for the larger population of satellites that fall into the low-gas category in our comparison of gas fractions (Fig. 3, see also Marasco et al. 2016). From a visual inspection of the stellar distribution around some gas-free satellites, we conclude that some of them are relatively isolated, but that their majority lives in crowded environments, as expected from their radial distribution.

These properties constitute strong evidence that the majority of gas-free galaxies have been stripped of their gas and most of their dark matter by a larger host halo (see e.g. Yun et al. 2019; Stevens et al. 2019, for related investigations in IllustrisTNG). The small fraction of gas-free galaxies that are identified as centrals at  $z = 0$  may well have had a significant encounter in the past. Interestingly, the stellar size distributions of gas-free galaxies are compatible with those of the overall sample. Thus, it appears that the stellar population is relatively immune to the processes that so efficiently strip dark matter and gas from galaxies (cf. Niemiec et al. 2019), presumably because it is much more concentrated than the dark and gaseous components.

In terms of our comparison to observations, the question is whether low-gas satellites would be included in the observational samples used in the C18 compilation. The answer depends somewhat on whether we consider  $\text{H I}$  or  $\text{H}_2$ , and on the stellar mass range. At low masses, where we see a strong excess in the fraction of galaxies with low  $\text{H I}$  mass, the  $\text{H I}$  fractions used in C18 are dominated by the Updated Nearby Galaxy Catalogue of Karachentsev et al. (2013), which explicitly includes objects in crowded environments. For the ALFALFA-based measurements, we expect no explicit bias against satellites as ALFALFA is a blind survey. The only considerable bias against low-gas satellites could stem from blending, where satellites are assigned additional  $\text{H I}$  mass due to confusion with other sources (most likely, their central galaxy). However, we show in Appendix B that blending is a relatively minor effect in our sample.

As there are currently no blind CO surveys at low redshift, many of the  $\text{H}_2$  fraction samples are based on an optical selection using SDSS, meaning that those samples might be biased against close pairs due to fibre collisions (Zehavi et al. 2002). However, the median distance of satellites to their host in our TNG100 sample is 140 kpc, which would correspond to the 55 arcsec fibre collision limit at a redshift of  $z \approx 0.14$ . Since samples such as GASS and COLD GASS have much lower median redshifts, the bulk of their satellite population should not be affected, and there are no further cuts on the proximity to another galaxy (Catinella et al. 2010, 2013;

Saintonge et al. 2011a, 2017). Thus, the impact of fibre collisions remains limited.

In conclusion, IllustrisTNG hosts a population of satellites with very little or no neutral gas, and there is no clear reason why the bulk of such galaxies should not be included in observational samples. Excessive stripping is not likely to be responsible, as Stevens et al. (2019) showed that the reduction in H I mass in TNG100 is compatible with observations (Brown et al. 2017). We conclude that, despite the overall excess in neutral gas discussed in the previous section, too many galaxies in IllustrisTNG are left with small amounts of neutral gas, perhaps due to excessive feedback. Such issues provide motivation to push observations of the H I content of galaxies to lower gas fractions, for example using new telescopes such as FAST (Nan et al. 2011).

### 5.3 What sets the H I mass–size relation?

In Section 4.5.1, we found that IllustrisTNG matches the observed H I mass–size relation with surprising accuracy. The H I/H<sub>2</sub> models agree almost exactly, an indication that the H I radius is generally not set by the H I/H<sub>2</sub> transition (see Wang et al. 2016, for observational arguments to the same effect). Instead, one could imagine that H I sizes were driven by the transition from ionized to neutral gas, and thus by the interplay of the UVB and self-shielding. However, we find that the radial profiles of the neutral fraction fall smoothly, with median neutral fractions between 30 and 60 per cent at  $R_{\text{H I}}$  (depending on stellar mass). There is no sharp break in the neutral gas profiles that would set  $R_{\text{H I}}$ . We have also tested the effect of the UVB explicitly, but the H I mass–size relation does not change in a simulation that uses the Puchwein et al. (2019) UVB prescription.

These considerations hint that the tight H I mass–size relation arises because the H I density profiles take on a universal form. The mass–size slope of roughly 0.5 corresponds to a fixed surface density. Indeed, the surface densities approach more or less constant values at small radii (Fig. 6). Thus, we might speculate that the normalization of the mass–size relation is set by the critical surface density where H I transitions to H<sub>2</sub> (e.g. Bigiel et al. 2008). However, changing this surface density by scaling the UV field in our modelling (see Paper I) does not modify the mass–size relation at all.

We have also investigated the evolution of the H I mass–size relation with redshift. Between  $z = 0$  and 4, the relation approximately maintains its slope, but the normalization (the H I size) decreases with redshift and the scatter increases slightly. In particular, while the TNG100 median relation lies 14 per cent above the Wang et al. (2014) relation at  $z = 0$ , it agrees to better than 1 per cent at  $z = 1$  and lies below the relation by 9 and 11 per cent at  $z = 2$  and 4, respectively (see Obreschkow & Rawlings 2009, for similar simulation results). This evolution constitutes a prediction for future observations of the H I mass–size relation at high redshift.

In summary, the H I mass–size relation appears to be extremely robust in IllustrisTNG (see also Marinacci et al. 2017). The physical reasons for the tight H I mass–size relation will be further investigated in a forthcoming paper (A. Stevens et al., in preparation).

### 5.4 Anecdotal evidence from individual galaxies

The molecular fraction in the Milky Way is observationally known to be about  $M_{\text{H}_2}/M_{\text{H}} \approx 10^9/10^{10} = 0.1$  (e.g. Tumlinson et al. 2002, see also Rachford et al. 2002). When we select Milky Way analogues with an SFR between 0.5 and  $2 M_{\odot} \text{ yr}^{-1}$  and a stellar mass around  $5 \times 10^{10} M_{\odot}$ , the H I/H<sub>2</sub> models predict a median H<sub>2</sub> fraction

between 0.3 and 0.4. A value of 0.1 is outside the 68 per cent scatter of all models, but there are, of course, some galaxies with such relatively low molecular fractions. Moreover, in Paper I we emphasized that our H I/H<sub>2</sub> modelling is systematically uncertain to factors of 2–3.

Zhu et al. (2018) recently highlighted a particular galaxy in TNG100 to be an analogue of the massive, low-surface brightness gas disc Malin 1 (Bothun et al. 1987). Braine, Herpin & Radford (2000) found no CO emission in Malin 1, translating into an upper limit of  $M_{\text{H}_2}/M_{\text{H}} < 0.03$ , given that Malin 1 contains a large H I mass of  $M_{\text{H I}} = 4 \times 10^{10} M_{\odot}$  (Pickering et al. 1997). The Malin analogue in TNG100 has a total neutral hydrogen mass of  $1.4 \times 10^{11} M_{\odot}$ , about three times larger than observationally measured. Like in the real Malin galaxy, this gas is distributed in a massive, low-surface density disc. Our modelling predicts similarly low molecular fractions between 0.03 and 0.1.

## 6 CONCLUSIONS

We have compared the atomic and molecular gas content of galaxies in the Illustris TNG100 and TNG300 simulations to observational data at low redshift. While we largely find good agreement, we uncover a few points of tension. Our main conclusions do not depend on the H I/H<sub>2</sub> model used. They are as follows:

(i) The cosmic abundance of H I at  $z = 0$  is about  $8 \times 10^{-4}$  in units of the critical density, roughly twice the abundance measured by 21-cm surveys. The cosmic abundance of H<sub>2</sub> also appears to be slightly higher than observations at  $z = 0$ , but this disagreement is not conclusive due to selection biases and due to the uncertain CO-to-H<sub>2</sub> conversion factor.

(ii) The H I mass function generally agrees well with observations, but overestimates the observed counts around  $10^9 M_{\odot}$ . The H<sub>2</sub> mass function exhibits no significant disagreements with observations.

(iii) The median H I and H<sub>2</sub> fractions and their scatter are generally well matched by the simulations. While the H<sub>2</sub> fraction is somewhat low at the highest stellar masses, the uncertainties on the corresponding observations are significant. However, IllustrisTNG produces a sizeable population of satellite galaxies (about 25 per cent at some stellar masses) that have little or no neutral gas, significantly more than observationally inferred.

(iv) The relation between H<sub>2</sub> mass and SFR is matched to better than a factor of 2 in IllustrisTNG, with a constant depletion time of 1 Gyr at  $z = 0$ . The trend of decreasing depletion time at higher redshift is also reproduced accurately.

(v) The observed H I mass–size relation is matched to 14 per cent accuracy in IllustrisTNG, and is relatively independent of the H I/H<sub>2</sub> modelling. The radial profiles of H I surface density also agree well, though a fraction of massive galaxies exhibit excessive central H I holes. We find indications that the spatial extent of the H<sub>2</sub> distribution is somewhat larger than observed, though with significant scatter.

(vi) The neutral gas content of simulated galaxies does not show the expected correlations with morphology. As a result, the gas fractions of simulated ETGs are too high, though this statement strongly depends on how morphology is defined.

(vii) The original Illustris-I simulation overestimates the neutral gas content of galaxies at  $z = 0$  by up to an order of magnitude, highlighting that IllustrisTNG represents an enormous improvement in the galaxy formation model.

It is worth noting that our  $\text{H I}/\text{H}_2$  results are almost entirely a prediction of the IllustrisTNG simulations. While the model parameters were calibrated to give reasonable gas fractions in cluster haloes (Pillepich et al. 2018a), it is far from obvious that such tuning would automatically result in the correct amount of neutral gas in galaxies over a large range of masses. Similarly, the star formation and feedback models used in the simulations were calibrated to form realistic amounts of stars in haloes of a given mass, but it is by no means guaranteed that those stars would form from the right amount of neutral gas. Splitting the gas into  $\text{H I}$  and  $\text{H}_2$  introduces additional physical demands on the simulation, for example, because  $\text{H}_2$  forms in high-density regions. Thus, our  $\text{H I}/\text{H}_2$  results represent an impressively accurate set of predictions.

However, we have also discovered some areas of tension with observations that can help to inform future generations of cosmological simulations. For example, the subgrid star formation model of Springel & Hernquist (2003) is currently based on a relatively simplistic picture of the temperature and density structure of the ISM, which determines the abundance of cold molecular clouds. Ideally, the ISM model would directly predict the neutral and molecular fractions, rendering post-processing unnecessary. Such a model would be strongly constrained by some of the data we considered in this paper, for instance, by the  $\text{H}_2$ –SFR relation and the spatial distribution of  $\text{H I}$  and  $\text{H}_2$ . Another area where improved observations could constrain simulation physics is the distribution of  $\text{H I}$  and  $\text{H}_2$  masses at the gas-poor end. As discussed in Section 5.2, the gas fraction is sensitive to a number of physical effects besides star formation and feedback, including the stripping of gas from satellites.

We have left many promising avenues for future work. Most notably, we have focused almost entirely on the local Universe, while there is a rapidly growing body of observations of the gas properties and scaling relations at higher redshift (e.g. Tacconi et al. 2013, 2018; Genzel et al. 2015; Freundlich et al. 2019). One particularly interesting question is how molecular gas reservoirs are related to star formation and quenching, and whether their relation changes over redshift. For example, Suess et al. (2017) recently reported ALMA detections of two galaxies at  $z \approx 0.7$  that host massive reservoirs of molecular gas but exhibit a low SFR. An analysis of similar galaxies in IllustrisTNG could perhaps explain their low star formation efficiency. Finally, we have refrained from a detailed comparison of the spatial distribution of molecular gas, which is increasingly being constrained observationally (e.g. Cormier et al. 2016, 2018; Bolatto et al. 2017).

We emphasize that all simulation data used in this work are publicly available as part of the IllustrisTNG data release (Nelson et al. 2019a). We encourage the community to undertake further analyses and data comparisons.

## ACKNOWLEDGEMENTS

We are grateful to Yannick Bahé, Alberto Bolatto, Martin Bureau, Andreas Burkert, John Forbes, Jonathan Freundlich, Hong Guo, Federico Lelli, Gergő Popping, Greg Snyder, Paul Torrey, and Rainer Weinberger for productive discussions. We thank the referee, Romeel Davé, for his constructive suggestions. This research made extensive use of the python packages NUMPY (Oliphant 2006), SCIPY (Jones et al. 2001), MATPLOTLIB (Hunter 2007), and COLOSSUS (Diemer 2018). Support for Program number HST-HF2-51406.001-A was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated,

under NASA contract NAS5-26555. ARC acknowledges support from CONACyT graduate fellowship. MV acknowledges support through an MIT RSC award, a Kavli Research Investment Fund, NASA ATP grant NNX17AG29G, and NSF grants AST-1814053 and AST-1814259.

## REFERENCES

- Allison J. R. et al., 2019, *MNRAS*, 482, 2934  
Bahé Y. M. et al., 2016, *MNRAS*, 456, 1115  
Barnes D. J. et al., 2018, *MNRAS*, 481, 1809  
Bauermeister A. et al., 2013, *ApJ*, 768, 132  
Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, *MNRAS*, 436, 697  
Bershady M. A., Jangren A., Conselice C. J., 2000, *AJ*, 119, 2645  
Bigiel F. et al., 2011, *ApJ*, 730, L13  
Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, *AJ*, 136, 2846  
Bird S., Vogelsberger M., Haehnelt M., Sijacki D., Genel S., Torrey P., Springel V., Hernquist L., 2014, *MNRAS*, 445, 2313  
Bird S., Garnett R., Ho S., 2017, *MNRAS*, 466, 2111  
Black J. H., Dalgarno A., 1976, *ApJ*, 203, 132  
Blitz L., Rosolowsky E., 2006, *ApJ*, 650, 933  
Bogdán Á. et al., 2015, *ApJ*, 804, 72  
Bolatto A. D. et al., 2017, *ApJ*, 846, 159  
Bolatto A. D., Wolfire M., Leroy A. K., 2013, *ARA&A*, 51, 207  
Boomsma R., Oosterloo T. A., Fraternali F., van der Hulst J. M., Sancisi R., 2008, *A&A*, 490, 555  
Boselli A. et al., 2010, *PASP*, 122, 261  
Boselli A., Cortese L., Boquien M., 2014a, *A&A*, 564, A65  
Boselli A., Cortese L., Boquien M., Boissier S., Catinella B., Lagos C., Saintonge A., 2014b, *A&A*, 564, A66  
Boselli A., Cortese L., Boquien M., Boissier S., Catinella B., Gavazzi G., Lagos C., Saintonge A., 2014c, *A&A*, 564, A67  
Bothun G. D., Impey C. D., Malin D. F., Mould J. R., 1987, *AJ*, 94, 23  
Bothwell M. S. et al., 2014, *MNRAS*, 445, 2599  
Bower R. G., Benson A. J., Crain R. A., 2012, *MNRAS*, 422, 2816  
Bradford J. D., Geha M. C., Blanton M. R., 2015, *ApJ*, 809, 146  
Braine J., Herpin F., Radford S. J. E., 2000, *A&A*, 358, 494  
Braun R., 2012, *ApJ*, 749, 87  
Broeils A. H., Rhee M. H., 1997, *A&A*, 324, 877  
Broeils A. H., van Woerden H., 1994, *A&AS*, 107, 129  
Brown T. et al., 2017, *MNRAS*, 466, 1275  
Brown T., Catinella B., Cortese L., Kilborn V., Haynes M. P., Giovanelli R., 2015, *MNRAS*, 452, 2479  
Calette A. R., Avila-Reese V., Rodríguez-Puebla A., Hernández-Toledo H., Papastergis E., 2018, *Rev. Mex. Astron. Astrofis.*, 54, 443  
Caplar N., Tacchella S., 2019, preprint (arXiv:1901.07556)  
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  
Cicone C. et al., 2015, *A&A*, 574, A14  
Cormier D. et al., 2016, *MNRAS*, 463, 1724  
Cormier D. et al., 2018, *MNRAS*, 475, 3909  
Cortese L., Catinella B., Boissier S., Boselli A., Heinis S., 2011, *MNRAS*, 415, 1797  
Crain R. A. et al., 2017, *MNRAS*, 464, 4204  
Daddi E. et al., 2010, *ApJ*, 714, L118  
Dannerbauer H. et al., 2017, *A&A*, 608, A48  
Davé R., Thompson R., Hopkins P. F., 2016, *MNRAS*, 462, 3265  
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *MNRAS*, 482, 486  
Davis T. A. et al., 2013, *MNRAS*, 429, 534  
Decarli R. et al., 2016, *ApJ*, 833, 69  
Delhaize J., Meyer M. J., Staveley-Smith L., Boyle B. J., 2013, *MNRAS*, 433, 1398



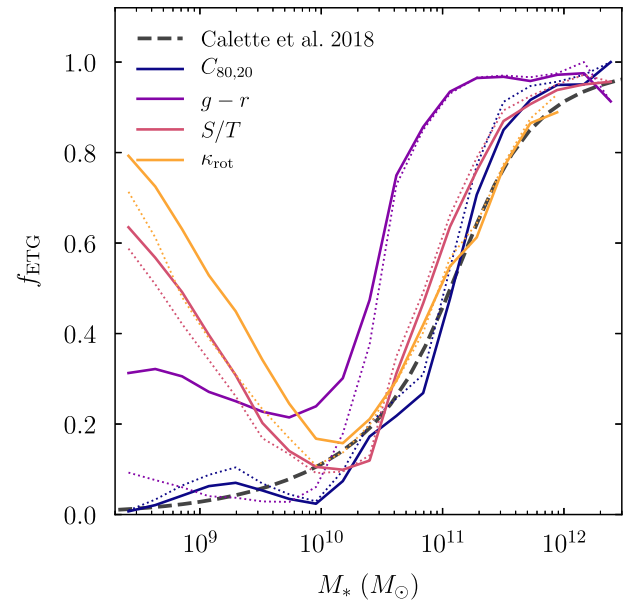
- Diemer B. et al., 2018, *ApJS*, 238, 33
- Diemer B., 2018, *ApJS*, 239, 35
- Diemer B., Sparre M., Abramson L. E., Torrey P., 2017, *ApJ*, 839, 26
- Donnari M. et al., 2019, *MNRAS*, 485, 4817
- Draine B. T., 2011, *Physics of the Interstellar and Intergalactic Medium*. Princeton Univ. Press, Princeton, NJ
- Dubois Y. et al., 2014, *MNRAS*, 444, 1453
- Duffy A. R., Kay S. T., Battye R. A., Booth C. M., Dalla Vecchia C., Schaye J., 2012, *MNRAS*, 420, 2799
- Edgington A. S., 1913, *MNRAS*, 73, 359
- Elmegreen B. G., 1989, *ApJ*, 338, 178
- Elmegreen B. G., 1993, *ApJ*, 411, 170
- Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416
- Feldmann R., Gnedin N. Y., Kravtsov A. V., 2012, *ApJ*, 747, 124
- Freudling W. et al., 2011, *ApJ*, 727, 40
- Freundlich J. et al., 2019, *A&A*, 622, A105
- Geha M., Blanton M. R., Masjedi M., West A. A., 2006, *ApJ*, 653, 240
- Genel S. et al., 2014, *MNRAS*, 445, 175
- Genel S. et al., 2018, *MNRAS*, 474, 3976
- Genel S., Fall S. M., Hernquist L., Vogelsberger M., Snyder G. F., Rodriguez-Gomez V., Sijacki D., Springel V., 2015, *ApJ*, 804, L40
- Genzel R. et al., 2010, *MNRAS*, 407, 2091
- Genzel R. et al., 2015, *ApJ*, 800, 20
- Ginolfi M. et al., 2017, *MNRAS*, 468, 3468
- Giovanelli R. et al., 2005, *AJ*, 130, 2598
- Glover S. C. O., Mac Low M. M., 2011, *MNRAS*, 412, 337
- Gnedin N. Y., Draine B. T., 2014, *ApJ*, 795, 37
- Gnedin N. Y., Kravtsov A. V., 2011, *ApJ*, 728, 88
- Guo H., Zehavi I., Zheng Z., 2012, *ApJ*, 756, 127
- Gurvich A., Burkhart B., Bird S., 2017, *ApJ*, 835, 175
- Habouzit M. et al., 2019, *MNRAS*, 484, 4413
- Haynes M. P. et al., 2011, *AJ*, 142, 170
- Haynes M. P. et al., 2018, *ApJ*, 861, 49
- Haynes M. P., Giovanelli R., 1984, *AJ*, 89, 758
- Hernández-Toledo H. M., Vázquez-Mata J. A., Martínez-Vázquez L. A., Choi Y.-Y., Park C., 2010, *AJ*, 139, 2525
- Hoppmann L., Staveley-Smith L., Freudling W., Zwaan M. A., Minchin R. F., Calabretta M. R., 2015, *MNRAS*, 452, 3726
- Huang S., Haynes M. P., Giovanelli R., Brinchmann J., Stierwalt S., Neff S. G., 2012a, *AJ*, 143, 133
- Huang S., Haynes M. P., Giovanelli R., Brinchmann J., 2012b, *ApJ*, 756, 113
- Huertas-Company M. et al., 2019, preprint ([arXiv:1903.07625](https://arxiv.org/abs/1903.07625))
- Huertas-Company M., Aguerri J. A. L., Bernardi M., Mei S., Sánchez Almeida J., 2011, *A&A*, 525, A157
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Janowiecki S., Catinella B., Cortese L., Saintonge A., Brown T., Wang J., 2017, *MNRAS*, 466, 4795
- Jansen R. A., Franx M., Fabricant D., Caldwell N., 2000a, *ApJS*, 126, 271
- Jansen R. A., Fabricant D., Franx M., Caldwell N., 2000b, *ApJS*, 126, 331
- Jones E. et al., 2001, SciPy: Open Source Scientific Tools for Python. <http://www.scipy.org/>
- Jones M. G., Papastergis E., Haynes M. P., Giovanelli R., 2015, *MNRAS*, 449, 1856
- Jones M. G., Haynes M. P., Giovanelli R., Moorman C., 2018, *MNRAS*, 477, 2
- Kannappan S. J. et al., 2013, *ApJ*, 777, 42
- Karachentsev I. D., Makarov D. I., Kaisina E. I., 2013, *AJ*, 145, 101
- Kennicutt Jr. R. C., 1998, *ApJ*, 498, 541
- Kent S. M., 1985, *ApJS*, 59, 115
- Keres D., Yun M. S., Young J. S., 2003, *ApJ*, 582, 659
- Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, *Astron. Lett.*, 44, 8
- Krumholz M. R., 2013, *MNRAS*, 436, 2747
- Krumholz M. R., McKee C. F., Tumlinson J., 2009, *ApJ*, 699, 850
- Lagos C. d. P. et al., 2015, *MNRAS*, 452, 3815
- Lagos C. d. P. et al., 2016, *MNRAS*, 459, 2632
- Lagos C. D. P., Baugh C. M., Zwaan M. A., Lacey C. G., Gonzalez-Perez V., Power C., Swinbank A. M., van Kampen E., 2014, *MNRAS*, 440, 920
- Lah P. et al., 2007, *MNRAS*, 376, 1357
- Lelli F., McGaugh S. S., Schombert J. M., 2016, *AJ*, 152, 157
- Leroy A. K. et al., 2009, *AJ*, 137, 4670
- Leroy A. K. et al., 2011, *ApJ*, 737, 12
- Leroy A. K. et al., 2013, *AJ*, 146, 19
- Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, *AJ*, 136, 2782
- Lisenfeld U. et al., 2011, *A&A*, 534, A102
- Lotz J. M., Primack J., Madau P., 2004, *AJ*, 128, 163
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Mancini C. et al., 2011, *ApJ*, 743, 86
- Marasco A., Crain R. A., Schaye J., Bahé Y. M., van der Hulst T., Theuns T., Bower R. G., 2016, *MNRAS*, 461, 2630
- Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, *ApJ*, 701, 1765
- Marinacci F. et al., 2018, *MNRAS*, 480, 5113
- Marinacci F., Grand R. J. J., Pakmor R., Springel V., Gómez F. A., Frenk C. S., White S. D. M., 2017, *MNRAS*, 466, 3859
- Martinsson T. P. K., Verheijen M. A. W., Bershadsky M. A., Westfall K. B., Andersen D. R., Swaters R. A., 2016, *A&A*, 585, A99
- Martin A. M., Papastergis E., Giovanelli R., Haynes M. P., Springob C. M., Stierwalt S., 2010, *ApJ*, 723, 1359
- Mitchell P. D., Lacey C. G., Baugh C. M., Cole S., 2013, *MNRAS*, 435, 87
- Naiman J. P. et al., 2018, *MNRAS*, 477, 1206
- Nan R. et al., 2011, *Int. J. Mod. Phys. D*, 20, 989
- Narayanan D., Krumholz M., Ostriker E. C., Hernquist L., 2011, *MNRAS*, 418, 664
- Narayanan D., Krumholz M. R., Ostriker E. C., Hernquist L., 2012, *MNRAS*, 421, 3127
- Neeleman M., Prochaska J. X., Ribado J., Lehner N., Howk J. C., Rafelski M., Kanekar N., 2016, *ApJ*, 818, 113
- Nelson D. et al., 2018a, *MNRAS*, 475, 624
- Nelson D. et al., 2018b, *MNRAS*, 477, 450
- Nelson D. et al., 2019a, *Computational Astrophysics and Cosmology*, 6, 2
- Nelson D. et al., 2019b, preprint ([arXiv:1902.05554](https://arxiv.org/abs/1902.05554))
- Niemiec A., Jullo E., Giocoli C., Limousin M., Jauzac M., 2019, *MNRAS*, 487, 653
- Noterdaeme P. et al., 2012, *A&A*, 547, L1
- Noterdaeme P., Petitjean P., Ledoux C., Srianand R., 2009, *A&A*, 505, 1087
- Obreschkow D., Rawlings S., 2009, *MNRAS*, 394, 1857
- Obreschkow D., Ma X., Meyer M., Power C., Zwaan M., Staveley-Smith L., Drinkwater M. J., 2013, *ApJ*, 766, 137
- Obuljen A., Alonso D., Villacusa-Navarro F., Yoon I., Jones M., 2018, preprint ([arXiv:1805.00934](https://arxiv.org/abs/1805.00934))
- Oliphant T. E., 2006, *Guide to NumPy*. Provo, UT
- Pavesi R. et al., 2018, *ApJ*, 864, 49
- Pickering T. E., Impey C. D., van Gorkom J. H., Bothun G. D., 1997, *AJ*, 114, 1858
- Pillepich A. et al., 2018a, *MNRAS*, 473, 4077
- Pillepich A. et al., 2018b, *MNRAS*, 475, 648
- Pillepich A. et al., 2019, preprint ([arXiv:1902.05553](https://arxiv.org/abs/1902.05553))
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- Ponomareva A. A., Verheijen M. A. W., Bosma A., 2016, *MNRAS*, 463, 4052
- Pop A.-R., Pillepich A., Amorisco N. C., Hernquist L., 2018, *MNRAS*, 480, 1715
- Popping G. et al., 2019, preprint ([arXiv:1903.09158](https://arxiv.org/abs/1903.09158))
- Popping G., Somerville R. S., Trager S. C., 2014, *MNRAS*, 442, 2398
- Prochaska J. X., Herbert-Fort S., Wolfe A. M., 2005, *ApJ*, 635, 123
- Puchwein E., Haardt F., Haehnelt M. G., Madau P., 2019, *MNRAS*, 485, 47
- Rachford B. L. et al., 2002, *ApJ*, 577, 221
- Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013a, *MNRAS*, 430, 2427
- Rahmati A., Schaye J., Pawlik A. H., Raičević M., 2013b, *MNRAS*, 431, 2261

Rao S. M., Turnshek D. A., Nestor D. B., 2006, *ApJ*, 636, 610  
 Rao S. M., Turnshek D. A., Sardane G. M., Monier E. M., 2017, *MNRAS*, 471, 3428  
 Regan M. W., Thornley M. D., Helfer T. T., Sheth K., Wong T., Vogel S. N., Blitz L., Bock D. C. J., 2001, *ApJ*, 561, 218  
 Rhee J., Zwaan M. A., Briggs F. H., Chengalur J. N., Lah P., Oosterloo T., van der Hulst T., 2013, *MNRAS*, 435, 2693  
 Rhee J., Lah P., Chengalur J. N., Briggs F. H., Colless M., 2016, *MNRAS*, 460, 2675  
 Rhee J., Lah P., Briggs F. H., Chengalur J. N., Colless M., Willner S. P., Ashby M. L. N., Le Fèvre O., 2018, *MNRAS*, 473, 1879  
 Riechers D. A. et al., 2019, *ApJ*, 872, 7  
 Rodriguez-Gomez V. et al., 2015, *MNRAS*, 449, 49  
 Rodriguez-Gomez V. et al., 2017, *MNRAS*, 467, 3083  
 Rodriguez-Gomez V. et al., 2019, *MNRAS*, 483, 4140  
 Saintonge A. et al., 2011a, *MNRAS*, 415, 32  
 Saintonge A. et al., 2011b, *MNRAS*, 415, 61  
 Saintonge A. et al., 2012, *ApJ*, 758, 73  
 Saintonge A. et al., 2017, *ApJS*, 233, 22  
 Sales L. V. et al., 2015, *MNRAS*, 447, L6  
 Sánchez S. F. et al., 2012, *A&A*, 538, A8  
 Schaye J. et al., 2010, *MNRAS*, 402, 1536  
 Schaye J. et al., 2015, *MNRAS*, 446, 521  
 Schaye J., 2001, *ApJ*, 559, 507  
 Schmidt M., 1959, *ApJ*, 129, 243  
 Scoville N. et al., 2017, *ApJ*, 837, 150  
 Serra P. et al., 2012, *MNRAS*, 422, 1835  
 Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, *MNRAS*, 452, 575  
 Snyder G. F. et al., 2015, *MNRAS*, 454, 1886  
 Somerville R. S., Davé R., 2015, *ARA&A*, 53, 51  
 Sparre M. et al., 2015, *MNRAS*, 447, 3548  
 Spitzer Lyman J., Zweibel E. G., 1974, *ApJ*, 191, L127  
 Springel V. et al., 2018, *MNRAS*, 475, 676  
 Springel V., 2010, *MNRAS*, 401, 791  
 Springel V., Hernquist L., 2003, *MNRAS*, 339, 289  
 Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726  
 Stark D. V., Kannappan S. J., Wei L. H., Baker A. J., Leroy A. K., Eckert K. D., Vogel S. N., 2013, *ApJ*, 769, 82  
 Sternberg A., Dalgarno A., 1989, *ApJ*, 338, 197  
 Sternberg A., Le Petit F., Roueff E., Le Bourlot J., 2014, *ApJ*, 790, 10  
 Stevens A. R. H. et al., 2019, *MNRAS*, 483, 5334  
 Stevens A. R. H., Martig M., Croton D. J., Feng Y., 2014, *MNRAS*, 445, 239  
 Suess K. A., Bezanson R., Spilker J. S., Kriek M., Greene J. E., Feldmann R., Hunt Q., Narayanan D., 2017, *ApJ*, 846, L14  
 Suresh J., Rubin K. H. R., Kannan R., Werk J. K., Hernquist L., Vogelsberger M., 2017, *MNRAS*, 465, 2966  
 Tacchella S. et al., 2019, preprint ([arXiv:1904.12860](https://arxiv.org/abs/1904.12860))  
 Tacconi L. J. et al., 2010, *Nature*, 463, 781  
 Tacconi L. J. et al., 2013, *ApJ*, 768, 74  
 Tacconi L. J. et al., 2018, *ApJ*, 853, 179  
 Thob A. C. R. et al., 2019, *MNRAS*, 485, 972  
 Torrey P. et al., 2018, *MNRAS*, 477, L16  
 Torrey P. et al., 2019, *MNRAS*, 484, 5587  
 Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, *MNRAS*, 438, 1985  
 Tumlinson J. et al., 2002, *ApJ*, 566, 857  
 Verheijen M. A. W., Sancisi R., 2001, *A&A*, 370, 765  
 Villaescusa-Navarro F. et al., 2018, *ApJ*, 866, 135  
 Vogelsberger M. et al., 2014a, *MNRAS*, 444, 1518  
 Vogelsberger M. et al., 2014b, *Nature*, 509, 177  
 Vogelsberger M. et al., 2018, *MNRAS*, 474, 2073  
 Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, 436, 3031  
 Walcher C. J. et al., 2014, *A&A*, 569, A1  
 Walter F. et al., 2016, *ApJ*, 833, 67

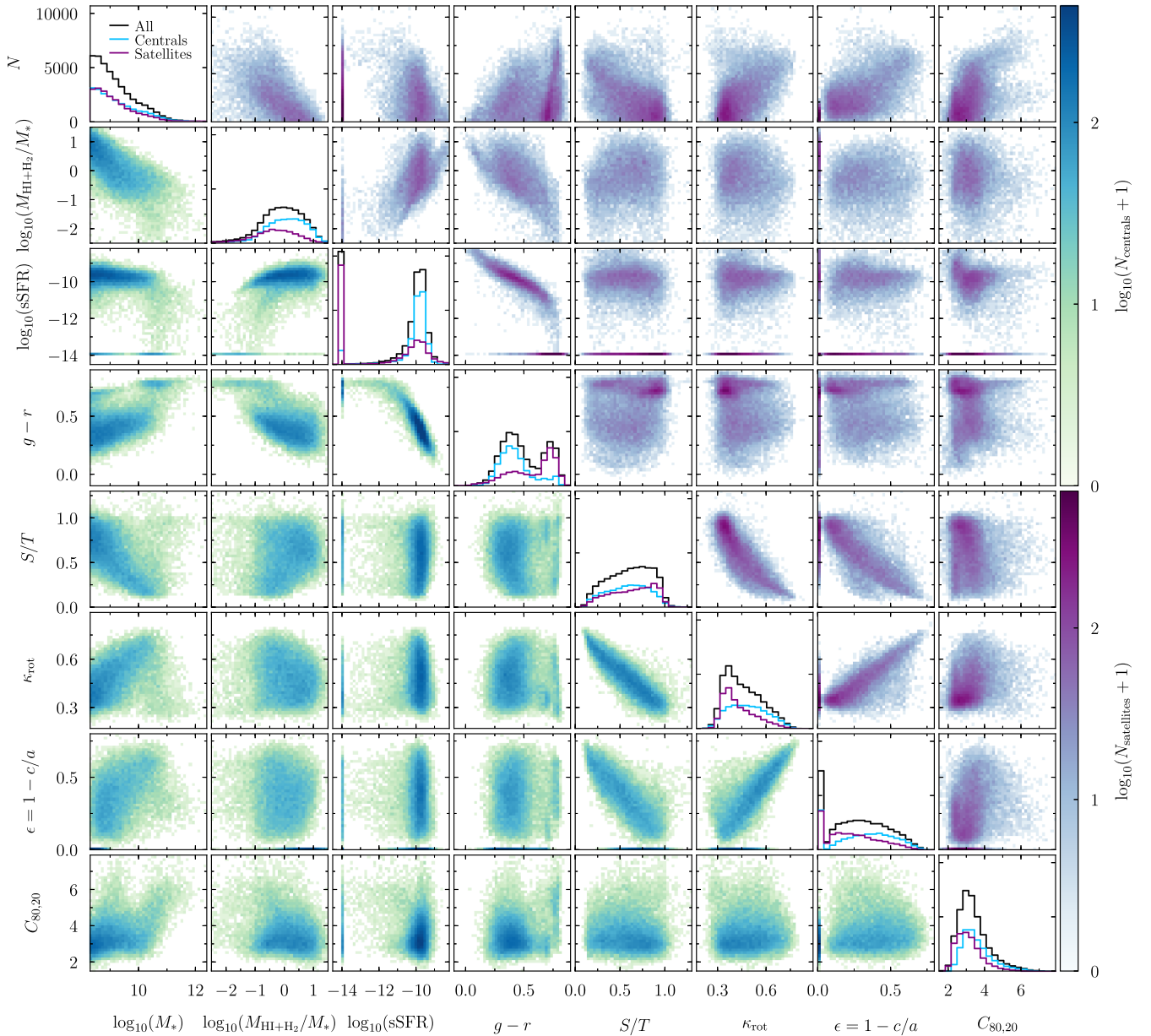
Wang J. et al., 2013, *MNRAS*, 433, 270  
 Wang J. et al., 2014, *MNRAS*, 441, 2159  
 Wang J., Koribalski B. S., Serra P., van der Hulst T., Roychowdhury S., Kamphuis P., Chengalur J. N., 2016, *MNRAS*, 460, 2143  
 Weinberger R. et al., 2017, *MNRAS*, 465, 3291  
 Weinberger R. et al., 2018, *MNRAS*, 479, 4056  
 Wolfire M. G., Hollenbach D., McKee C. F., 2010, *ApJ*, 716, 1191  
 Xu D., Springel V., Sluse D., Schneider P., Sonnenfeld A., Nelson D., Vogelsberger M., Hernquist L., 2017, *MNRAS*, 469, 1824  
 Young J. S. et al., 1995, *ApJS*, 98, 219  
 Young J. S., Scoville N., 1982, *ApJ*, 258, 467  
 Young J. S., Scoville N. Z., 1991, *ARA&A*, 29, 581  
 Yun K. et al., 2019, *MNRAS*, 483, 1042  
 Zafar T., Péroux C., Popping A., Milliard B., Deharveng J. M., Frank S., 2013, *A&A*, 556, A141  
 Zehavi I. et al., 2002, *ApJ*, 571, 172  
 Zhu Q. et al., 2018, *MNRAS*, 480, L18  
 Zwaan M. A., Meyer M. J., Staveley-Smith L., Webster R. L., 2005, *MNRAS*, 359, L30

## APPENDIX A: MORPHOLOGICAL SPLIT

In this appendix, we attempt to categorize IllustrisTNG galaxies according to their morphology in order to investigate the gas content of LTGs and ETGs separately. C18 emphasize that the separation into LTGs and ETGs is not entirely uniform within their sample because different methods were applied to the underlying survey data sets, including visual classification and machine learning, which cannot easily be reproduced in simulations (e.g. Huertas-Company et al. 2011, 2019). Nevertheless, we use the ETG fraction as parametrized by C18 as a guide. Fig. A1 shows how splitting



**Figure A1.** The early-type fraction according to a number of pseudo-morphological classifications. The dashed line shows the ETG fraction as a function of stellar mass as parametrized by C18. The solid coloured lines show the fraction in TNG100 and TNG300 when split on stellar concentration (dark blue),  $g-r$  colour (purple), spheroid-to-total ratio (red), and the fraction of kinetic energy that is in rotation (yellow). The dotted lines show the fractions if only centrals are considered. The split on stellar concentration does best at matching the C18 sample; the similarity of the dotted and solid dark blue lines demonstrates that this split works for both satellites and centrals. We add a scatter of 0.2 dex to all simulated stellar masses to account for the observational uncertainty (Section 3.4.1).



**Figure A2.** Correlations between stellar mass, neutral gas fraction, sSFR,  $g-r$  colour, spheroid-to-total ratio, fractional energy in rotation, flattening parameter, and stellar mass concentration. The green/blue histograms (lower left half) show the density of central galaxies in TNG100 with  $M_* > 2 \times 10^8 M_\odot$  in the respective planes of each pair of quantities. The purple histograms (upper right half) show the same for satellites. The histograms along the diagonal show the one-dimensional distributions of each quantity. In IllustrisTNG, the neutral gas content of galaxies correlates much more strongly with colour than with any morphological indicator.

on a number of different morphology indicators reproduces this fraction. Fig. A2 shows how these indicators correlate with the physical properties of interest, namely, the stellar mass, neutral gas mass, and sSFR.

We find that the most suitable indicator of morphology is the concentration of stellar mass,  $C_{80,20}$ . This concentration can be viewed as a simplified proxy for more complicated profile fitting methods (e.g. Xu et al. 2017). The dark blue lines in Fig. A1 demonstrate that the criterion  $C_{80,20} > 4.9$  reproduces the C18 ETG fraction reasonably well. Concentration correlates very weakly with all of the other indicators we tested (Fig. A2), indicating that it contains unique information. In the real Universe, however, concentration does correlate significantly with colour (Berashady

et al. 2000). We have also tested the relationship between  $C_{80,20}$  and the bulge statistic,  $F(G, M_{20})$  (Lotz et al. 2004), which was derived by Rodriguez-Gomez et al. (2019) using an intricate algorithm well matched to those used in observations.  $C_{80,20}$  and  $F(G, M_{20})$  correlate strongly (Tacchella et al. 2019), an indication that  $C_{80,20}$  is a suitable criterion that can roughly mimic the morphological classifications performed in the C18 sample. When defined in this way, ETGs are dominated by centrals at all stellar masses.

To better understand the connection between gas content and galaxy morphology, we also investigate a number of other pseudo-morphological indicators. First, we consider a cut on the  $g-r$  colour of IllustrisTNG galaxies, using the dust-corrected colours of Nelson et al. (2018a) wherever available. Fig. A2 demonstrates that colour



correlates strongly with neutral gas fraction and sSFR (the latter more or less by construction). By visual inspection of the colour–mass diagram, we find a split of  $g - r = 0.65 + 0.02[\log_{10}(M_*/M_\odot) - 10.28]$  to separate the red and blue galaxy populations. However, this cut leads to a significant excess of ETGs compared to the C18 sample, and an increase in ETGs at low masses due to satellites (purple lines in Fig. A1). Moreover, colour correlates weakly with structural parameters (Fig. A2, see also Rodriguez-Gomez et al. 2019), meaning that it is unlikely to be a good proxy for the morphological classifications in observations. On the other hand, when splitting the sample by colour, we obtain a better fit to the ETG and LTG gas fractions of C18 than when splitting on concentration. This comparison is clearly somewhat unphysical as the colour split selects the wrong number of ETGs and LTGs (Fig. A1), but it highlights that colour correlates more tightly with gas properties than morphology does in IllustrisTNG (Fig. A2).

Another type of criterion could be the kinematics of a galaxy’s stars, e.g. spheroid-to-total ratio, S/T, or the fraction of kinetic energy in rotation,  $\kappa_{\text{rot}}$  (as defined in Rodriguez-Gomez et al. 2017). We calculate the spheroidal component as twice the mass of all stellar-population particles that counterrotate with respect to the total angular momentum axis of all stars (other definitions lead to very similar outcomes). Fig. A1 shows the results of a kinematic classification, where ETGs are defined to have  $\kappa_{\text{rot}} < 0.45$  (yellow) or S/T  $> 0.7$  (red). At masses below  $M_* \approx 10^{10} M_\odot$ , the ETG fraction increases steeply in both cases, in contrast to the C18 sample. Thus, kinematic indicators of morphology cannot, by themselves, be used as a clean indicator of early-type characteristics.

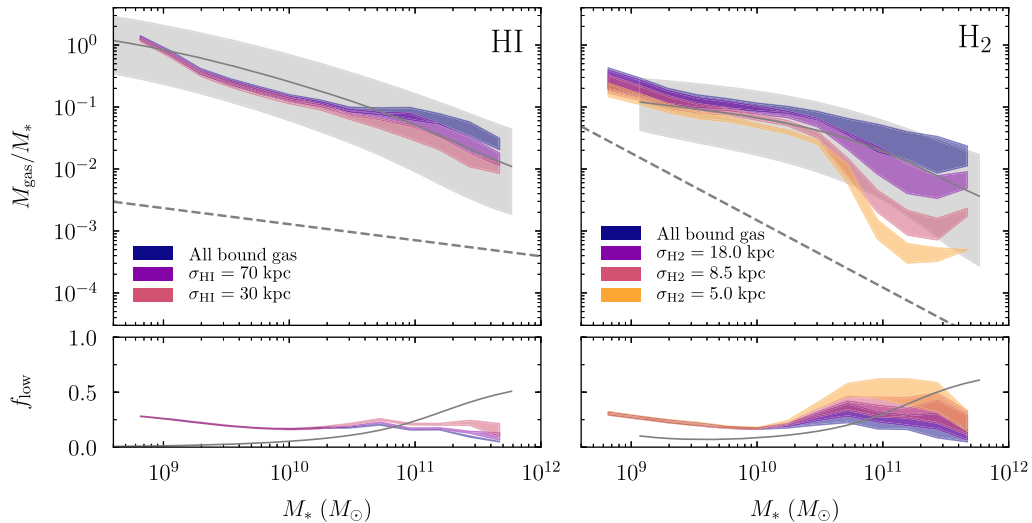
Finally, we consider indicators of the non-isotropic shape of the stellar distribution, such as the flattening parameter,  $\epsilon = 1 - c/a$ , and the triaxiality parameter,  $T = (a^2 - b^2)/(a^2 - c^2)$ . We follow the definitions of Thob et al. (2019), where  $a > b > c$  are the eigenvalues of the inertia tensor of the stellar distribution (computed as in Genel et al. 2015). However, there is no straightforward way to use  $\epsilon$  and  $T$  for our morphological cut because they are non-monotonic

functions of  $M_*$  and depend strongly on whether a galaxy is a central or satellite. As expected,  $\epsilon$  correlates fairly well with S/T and  $\kappa_{\text{rot}}$ . Splitting the sample according to any of these indicators results in LTG and ETG gas fractions that are more or less similar to a concentration split.

In summary,  $C_{80,20}$  appears to provide a reasonable approximation to the ETG fraction of C18 and approximates more sophisticated indicators. However, except for colour, the indicators considered do not correlate well with the neutral gas fraction or sSFR of galaxies (Fig. A2). This lack of correlation provides the explanation for why the C18 gas fractions for LTGs and ETGs are poorly matched compared to the overall sample (Fig. 8). This result is in line with the findings of Rodriguez-Gomez et al. (2019), who showed that the colour–morphology relation does not match Pan-STARRS data, with too many red spirals and blue discs in the simulation. We conclude that the weak correlation between morphology and colour is likely caused by a weak correlation between morphology and gas content.

## APPENDIX B: THE EFFECTS OF APERTURE, BINNING, AND BLENDING

In Section 3.4, we described and justified our choices for the apertures chosen when mock-observing a number of observational surveys. In this appendix, we investigate the quantitative impact of different aperture choices and show that blending does not have a significant impact on our results. Fig. B1 shows mock-observed H I and H<sub>2</sub> fractions using a range of apertures. The differences are particularly large for the H<sub>2</sub> fraction at the high-mass end. As discussed in Section 3.4.4, an even smaller aperture of  $\sigma_{\text{H}_2} = 5$  kpc might be appropriate for the low-redshift galaxies in the xCOLD GASS sample, but such a threshold becomes difficult to interpret because the spatial extent of H<sub>2</sub> gas appears to be broader in IllustrisTNG than in observations (Section 4.5). Moreover, reducing the aperture further has a modest impact below stellar masses of



**Figure B1.** Impact of aperture on the mock-observed gas fractions in TNG100. The panels are the same as in Fig. 3. The shaded regions enclose the different H I/H<sub>2</sub> models; the scatter is not shown to avoid crowding. The dark blue areas show the gas fractions one would infer by including all gas bound to the subhalo as defined by the halo finder. The other colours indicate decreasing apertures, all referring to the standard deviation of a Gaussian beam. At low masses, aperture plays a role only for H<sub>2</sub>, which is less spatially extended than H I. At high masses, our fiducial choices of  $\sigma_{\text{H I}} = 70$  kpc and  $\sigma_{\text{H}_2} = 8.5$  kpc lead to a factor of 2–3 lower H I masses and up to an order of magnitude reduction in the H<sub>2</sub> mass, highlighting the importance of the aperture choice. The difference between the apertures used for H<sub>2</sub> fractions and the H<sub>2</sub> mass function,  $\sigma_{\text{H}_2} = 8.5$  kpc and  $\sigma_{\text{H}_2} = 18$  kpc, is significant.

$2 \times 10^{10} M_{\odot}$ . We conclude that our 8.5-kpc aperture is reasonable but that some systematic uncertainty remains, particularly with respect to the spatial distribution of  $H_2$ . We note that other works have chosen rather different CO apertures. For example, Lagos et al. (2015) chose a fixed aperture of 30 kpc to match their stellar aperture. This choice would increase our mock-observed  $H_2$  masses significantly, bringing them close to the mass of all bound  $H_2$  gas (Fig. B1).

As we derive aperture masses from projected mass profiles, we have also tested for inaccuracies due to the finite number of radial bins. In particular, we have randomly chosen about 1000 galaxies and increased the number of bins in their profiles from 50 to 120. The stellar masses (aperture 30 kpc) are changed by less than 1 per cent for all galaxies. The median H I masses (aperture 70 kpc) shift by less than 0.2 per cent for all H I/ $H_2$  models. In extremely rare cases (less than 2 per cent of galaxies), the profiles exhibit H I holes and thus very steep gradients in their cumulative H I mass profile near the aperture, which can change the mass by about 1 per cent. The  $H_2$  masses are slightly more affected by binning, with the median  $H_2$  mass changing by less than 5 per cent for all H I/ $H_2$  models, and added scatter of about 0.15 dex (for an aperture of 8.5 kpc as used for the gas fraction measurements). When using the 18-kpc aperture with a hard cut-off for mass functions, the median  $H_2$  mass shifts by between 1 and 7 per cent depending on the H I/ $H_2$

model. None of these shifts are relevant to the conclusions of this paper.

Finally, we check the importance of blending. Stevens et al. (2019) carefully considered the redshift distribution of their observed H I samples and modelled the effect of blending, i.e. the confusion of multiple H I sources in a single beam. We have compared our modelling to theirs for a representative sample of galaxies with  $M_* > 10^9 M_{\odot}$ . The following results refer to the volumetric GK11 model, but the results do not strongly depend on the H I/ $H_2$  model. We find that our H I masses (aperture 70 kpc) agree with the Stevens et al. (2019) sample with a median ratio of 0.97 and a 68 per cent interval between 0.72 and 1.06, for all galaxies that host significant H I in both samples. There is, however, a population of 4 per cent of the galaxies that exhibit significant H I content ( $M_{H I} > 10^8 M_{\odot}$ ) in the Stevens et al. (2019) sample but no H I ( $M_{H I} < 10^6 M_{\odot}$ ) in our sample – some of the gas-free objects discussed in Section 5.2, which experience H I contributions from gas in other galaxies. Generally, blending most strongly affects satellites, but our sample is dominated by centrals across the entire mass range. Thus, the statistical shift in our H I masses compared to Stevens et al. (2019) is acceptable.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.