The local PNG bias of neutral Hydrogen, H_I

Alexandre Barreira

Excellence Cluster ORIGINS,
Boltzmannstraße 2, 85748 Garching, Germany

Ludwig-Maximilians-Universität,
Schellingstraße 4, 80799 München, Germany

E-mail: alex.barreira@origins-cluster.de

Abstract. We use separate universe simulations with the IllustrisTNG galaxy formation model to predict the local PNG bias parameters b_ϕ and $b_{\phi\delta}$ of atomic neutral hydrogen, H_I. These parameters and their relation to the linear density bias parameter b_1 play a key role in observational constraints of the local PNG parameter f_{NL} using the H_I power spectrum and bispectrum. Our results show that the popular calculation based on the universality of the halo mass function overpredicts the $b_\phi(b_1)$ and $b_{\phi\delta}(b_1)$ relations measured in the simulations. In particular, our results show that at $z \lesssim 1$ the H_I power spectrum is more sensitive to f_{NL} compared to previously thought (b_ϕ is more negative), but is less sensitive at other epochs (b_ϕ is less positive). We discuss how this can be explained by the competition of physical effects such as that large-scale gravitational potentials with local PNG (i) accelerate the conversion of hydrogen to heavy elements by star formation, (ii) enhance the effects of baryonic feedback that eject the gas to regions more exposed to ionizing radiation, and (iii) promote the formation of denser structures that shield the H_I more efficiently. Our numerical results can be used to revise existing forecast studies on f_{NL} using 21 cm line-intensity mapping data. Despite this first step towards predictions for the local PNG bias parameters of H_I, we emphasize that more work is needed to assess their sensitivity on the assumed galaxy formation physics and H_I modeling strategy.

Keywords: cosmological simulations, hydrodynamical simulations, physics of the early universe, power spectrum

ArXiv ePrint: 2112.03253

Contents

1 Introduction

The level of non-Gaussianity of the distribution of the energy density fluctuations generated during inflation holds key information about the particle content and physics of the early Universe [1, 2]. The most popular and well-studied type of this *primordial non-Gaussianity* (PNG) is called *local-type* PNG, where the primordial gravitational potential ϕ is written as [3]

$$\phi = \phi_G + f_{NL} \left[\phi_G^2 - \left\langle \phi_G^2 \right\rangle \right], \tag{1.1}$$

with ϕ_G being a Gaussian random field and f_{NL} a constant that quantifies the leading-order departure from Gaussianity ($\langle \cdots \rangle$ indicates ensemble averaging). Different models of inflation make different predictions for f_{NL}, and so constraining its value observationally lets us distinguish between them. The current tightest bound comes from the analysis of the cosmic microwave background (CMB) data from the Planck satellite, which constrained $f_{NL} = -0.9 \pm 5.1 \ (1\sigma)$ [4]. Using large-scale structure data, the tightest constraint to date is $f_{NL} = -12 \pm 21 \ (1\sigma)$ [5], and was obtained using quasars from the eBOSS survey (see also ref. [6]). This bound is still looser than the CMB one, but the volumes spanned by next-generation large-scale structure surveys should let us reach precisions of order $\sigma_{f_{NL}} \sim 1$, which would mark, even without a detection, an important landmark in our ability to distinguish between competing models of inflation [7, 8].

Some of these future constraints on f_{NL} will come from studies of the spatial distribution of (atomic) neutral hydrogen (H_I) in the Universe [9–19], which will be mapped by 21 cm line-intensity mapping experiments such as BINGO [20], CHIME [21], HIRAX [22] and SKA [23]. Compared to traditional galaxy surveying where individual galaxies are identified on the sky and their redshifts are estimated spectroscopically or photometrically, the line-intensity mapping approach targets the integrated emission from all sources (resolved or not) in a given region of the Universe, with the radial information inferred through measurements in different frequency bands [24]. In the case of H_I, the relevant line is the 21cm radiation emitted by its spin-flip transition. After the epoch of reionization $z \lesssim 6$, most of the hydrogen

in the Universe is ionized, but a significant amount of H_I is still present in galaxies inside gas clouds that are sufficiently dense to shield the H_I from the ionizing radiation. By skipping the need to resolve individual galaxies, line-intensity mapping surveys can scan very rapidly large areas of the sky with good redshift coverage, which is key to probe the large distance scales where f_{NL} contributes the most; when completed, some of these surveys will have mapped the H_I out to redshifts $z \sim 3$ over about half of the sky. These experiments are also subject to distinct observational systematics compared to traditional galaxy surveys (e.g. the large foreground emission that needs to be subtracted), which will strongly establish the robustness of large-scale structure constraints on f_{NL} if the different surveying techniques give consistent results [25–27].

The key observational signatures of local PNG on the H_I distribution arise through a set of *bias parameter* terms that are generated if $f_{NL} \neq 0$. In general, bias parameters are a set of redshift-dependent numbers that describe how the local abundance of a given tracer of the large-scale structure (dark matter halos, galaxies, H_I, etc.) is modulated by different types of long-wavelength ($\gtrsim$ 50–100 Mpc) perturbations (see ref. [28] for a comprehensive review). For the case of H_I, and specifying already to the bias parameters we focus on in this paper, we can write [29–32]

$$\rho_{H_I}(\boldsymbol{x}, z) \supset \bar{\rho}_{H_I}(z)\Big[1 + b_1(z)\delta_m(\boldsymbol{x}, z) + b_\phi(z)f_{NL}\phi(\boldsymbol{q}) + b_{\phi\delta}(z)f_{NL}\phi(\boldsymbol{q})\delta_m(\boldsymbol{x}, z)\Big], \quad (1.2)$$

where $\rho_{H_I}(\boldsymbol{x}, z)$ is the local H_I energy density, $\bar{\rho}_{H_I}(z)$ is its cosmic average, $\delta_m(\boldsymbol{x}, z)$ is a large-scale total matter density fluctuation and $\phi(\boldsymbol{q})$ is a large-scale primordial gravitational potential fluctuation ($\boldsymbol{q}$ is the initial Lagrangian coordinate associated with the final Eulerian coordinate $\boldsymbol{x}$). This equation makes apparent the physical meaning of the bias parameters b_1, b_ϕ and $b_{\phi\delta}$ as the *response* of the local H_I energy density to the presence of the perturbation each multiplies, i.e.,

$$b_1 = \frac{1}{\rho_{H_I}}\frac{\partial \rho_{H_I}}{\partial \delta_m} \quad ; \quad b_\phi = \frac{1}{\rho_{H_I}}\frac{\partial \rho_{H_I}}{\partial(f_{NL}\phi)} \quad ; \quad b_{\phi\delta} = \frac{1}{\rho_{H_I}}\frac{\partial^2 \rho_{H_I}}{\partial(f_{NL}\phi)\partial\delta_m}. \quad (1.3)$$

For example, b_1 quantifies how much more ($b_1 > 0$) or less ($b_1 < 0$) H_I exists inside large-scale mass overdensities. These three parameters are the most important ones for observational constraints on f_{NL} using the H_I power spectrum (2-point function) and bispectrum (3-point function). Concretely, for the case of the H_I power spectrum, and analogously to the case of the galaxy power spectrum [29], the f_{NL} signature is $\propto b_1 b_\phi f_{NL}/k^2$, where k is the wavenumber.[1] Similarly, the leading-order contributions to the bispectrum include a series of terms proportional to $f_{NL}b_1^3$, $f_{NL}b_\phi b_1^2$ and $f_{NL}b_{\phi\delta}b_1^2$, whose amplitude is what can be used to constrain f_{NL} (see e.g. figure 1 of ref. [33] for a visualization of these contributions).

Since the observational signatures of f_{NL} are effectively degenerate with the bias parameters, a solid understanding of the latter is critical to obtain the best possible constraints on local PNG from large-scale structure data (see refs. [33–35] for recent discussions). However, as the bias parameters effectively describe how the long-wavelength environment impacts the H_I distribution, they depend on the complicated interplay between gravity, reionization, star formation and stellar/black hole feedback, and are thus extremely challenging to predict theoretically. While b_1 can be fitted for directly from the data since it enters also through terms that do not multiply f_{NL}, the same is not true for the local PNG parameters b_ϕ and

---

[1]This is often referred to as the *scale-dependent* bias effect, which can be a misleading name since what is scale-dependent is not any bias parameter, but the relation between potential and mass perturbations $\phi \sim \delta_m/k^2$.

$b_{\phi\delta}$, which require assumptions to be made on in order to constrain $f_{\rm NL}$ competitively. The standard way around this problem is based on the hope that, while the parameters b_ϕ and $b_{\phi\delta}$ themselves are expected to be extremely sensitive to the details of structure formation, their relation to b_1 may be more robust and easier to predict theoretically. Thus, if b_ϕ and $b_{\phi\delta}$ can be fixed in terms of b_1, then determining the latter from the data breaks the degeneracy with $f_{\rm NL}$, which can then be constrained.

For example, the majority of the constraint studies on $f_{\rm NL}$ in the literature (including those based on H$_{\rm I}$ data) assume relations based on the universality of the halo mass function, from which it follows e.g. that $b_\phi \propto (b_1 - 1)$. Despite its widespread adoption however, there is no reason to expect these relations to hold exactly for real tracers of the large-scale structure like the H$_{\rm I}$, which motivates simulation-based works to refine these simple theoretical priors. Indeed, even for the case of dark matter halos in gravity-only simulations, we now know that $b_\phi(b_1)$ [36, 37] and $b_{\phi\delta}(b_1)$ [35] are not perfectly described by the corresponding universality expressions. Further, using hydrodynamical simulations of galaxy formation, refs. [35, 38] showed recently that the $b_\phi(b_1)$ and $b_{\phi\delta}(b_1)$ relations of galaxies selected by a variety of criteria (including stellar-mass, black-hole mass accretion rate and color) are also not adequately described by the universality assumption, which needs to be revisited.

This motivates our main goal in this paper, which is to use hydrodynamical simulations to study the bias parameters b_ϕ and $b_{\phi\delta}$ of the H$_{\rm I}$ distribution. The methodology we adopt in this paper is similar to that of refs. [35, 38], who used separate universe simulations of galaxy formation with the IllustrisTNG model to study the same bias parameters, but for the halo and galaxy distribution. One of our main results is that the $b_\phi(b_1)$ and $b_{\phi\delta}(b_1)$ relations of the H$_{\rm I}$ are generically overpredicted by the results that are typically encountered in the literature assuming the above-mentioned universality expressions. We show this can be explained physically by the impact that long-wavelength $f_{\rm NL}\phi$ perturbations have on the H$_{\rm I}$ content of dark matter halos, which can be understood as a new *bias* parameter (or response function) in the halo model. We will show results for just a single galaxy formation model and H$_{\rm I}$ modeling recipe, and as a result, given the uncertainties we still have on these areas, and despite the intuition this will let us build already, the analysis in this paper should be regarded as the first step towards robust theoretical priors on the $b_\phi(b_1)$ and $b_{\phi\delta}(b_1)$ relations of the H$_{\rm I}$. We stress that these relations play a key role in observational constraint studies of $f_{\rm NL}$ using H$_{\rm I}$ data, which strongly motivates more works like this. The rest of this paper is organized as follows. In section 2, we describe the galaxy formation simulations, the H$_{\rm I}$ modeling strategy, and the numerical methods that we use to estimate the H$_{\rm I}$ bias parameters b_1, b_ϕ and $b_{\phi\delta}$. We show and discuss our numerical results in section 3, and summarize and conclude in section 4.

2 Methodology

In this section we describe the N-body simulations, the modeling of the H$_{\rm I}$ distribution and the methods we use to estimate the H$_{\rm I}$ bias parameters b_1, b_ϕ and $b_{\phi\delta}$ defined in eqs. (1.2) and (1.3).

2.1 Numerical simulation specifications

We obtain our numerical results using separate universe simulations run with the moving-mesh code AREPO [39] and IllustrisTNG as the galaxy formation model [40, 41]. This model is characterized by a set of *sub-grid* prescriptions for physics including gas cooling, star formation, metal enrichment, stellar winds, supernovae feedback, and black hole feedback,

that broadly reproduces a series of key observations including the low-redshift galaxy stellar mass function, the cosmic star formation history, the sizes and colors of galaxies and the gas fractions in galaxies and galaxy groups [42–48]. The fiducial cosmological parameters are: mean total matter density today $\Omega_{m0} = 0.3089$, mean baryon density today $\Omega_{b0} = 0.0486$, mean dark energy density today $\Omega_{\Lambda0} = 0.6911$, dimensionless Hubble rate $h = 0.6774$, primordial scalar power spectrum amplitude $\mathcal{A}_s = 2.068 \times 10^{-9}$ (at $k_{\rm pivot} = 0.05/{\rm Mpc}$, corresponding to $\sigma_8(z = 0) = 0.816$), and primordial scalar spectral index $n_s = 0.967$.

As we describe below, we measure the H$_{\rm I}$ bias parameters b_ϕ and $b_{\phi\delta}$, respectively, as the response of the H$_{\rm I}$ abundance and its linear bias parameter b_1 to long-wavelength $f_{\rm NL}\phi$ perturbations. Under the separate universe approach, this is equivalent to the response to changes in the cosmological parameter $\mathcal{A}_s$, and so we consider also two additional cosmologies called High$\mathcal{A}_s$ and Low$\mathcal{A}_s$, which share the same parameters as the fiducial except that $\mathcal{A}_s \to \mathcal{A}_s[1 + \delta_{\mathcal{A}_s}]$, where $\delta_{\mathcal{A}_s} = +0.05$ for High$\mathcal{A}_s$ and $\delta_{\mathcal{A}_s} = -0.05$ for Low$\mathcal{A}_s$. When we vary $\mathcal{A}_s$, we keep the sub-grid parameters of the IllustrisTNG model fixed also to their fiducial values in order to interpret our results on the impact of $f_{\rm NL}\phi$ perturbations at fixed galaxy formation physics.

We will show results obtained at two numerical resolutions. One is called TNG300-2 and is characterized by a cubic box with size $L_{\rm box} = 205\,{\rm Mpc}/h$ and $N_{\rm p} = 2 \times 1250^3$ initial dark matter and gas resolution elements. The other is higher-resolution, it is called TNG100-1.5, and it is characterized by $L_{\rm box} = 75\,{\rm Mpc}/h$ and $N_{\rm p} = 2 \times 1250^3$. These separate universe simulations have been presented for the first time in ref. [38] (to which we refer the reader for more details), and used in previous works already to study galaxy bias [35] and halo occupation distribution responses [49]. In our results below we analyse the simulation outputs at redshifts $z = 0, 0.5, 1, 2, 3$. We have also used the publicly available IllustrisTNG data for the resolutions dubbed TNG300-1 ($L_{\rm box} = 205\,{\rm Mpc}/h$, $N_{\rm p} = 2 \times 2500^3$) and TNG100-1 ($L_{\rm box} = 75\,{\rm Mpc}/h$, $N_{\rm p} = 2 \times 1820^3$) to carry out some numerical convergence and consistency checks of our modeling at the fiducial cosmology.

When we analyse the H$_{\rm I}$ content in halos, we consider gravitationally-bound objects found using a friends-of-friends algorithm with linking length $b = 0.2$ times the mean dark matter interparticle distance. We quote as halo mass $M_{\rm h}$ and total H$_{\rm I}$ mass $M_{\rm H_{\rm I}}$ the summed contribution from all halo member particles and cells, and consider objects that comprise at least $N_{\rm cell} \geq 50$ gas cells.

2.2 Modeling of H$_{\rm I}$

Our strategy to model the H$_{\rm I}$ distribution in our simulations follows largely that of ref. [50]. In IllustrisTNG, the hydrogen gas mass fraction of each Voronoi cell starts off at its primordial value of $f_{\rm H} = 0.76$, and it decreases with time as stars form and enrich the interstellar medium with heavier metals. The first part of the modeling of H$_{\rm I}$ involves determining the fraction $f_{\rm H_N}$ of all hydrogen that is neutral. For non-star-forming gas (which for the star formation model in IllustrisTNG [51] is all gas with local hydrogen number density $n_{\rm H} < 0.106/{\rm cm}^3$), we use the split into neutral and ionized fractions that is calculated self-consistently during the simulation taking into account the redshift-dependent UV radiation background, attenuation due to self-shielding in higher-density gas regions, and ionizing radiation by local active galactic nuclei (AGN) [52].

The situation is more complicated for star-forming gas because $f_{\rm H_N}$ is not computed self-consistently: the simulation output values are based on a mass-weighted temperature of the cold and hot gas phases, which is not a well-defined physical temperature. The authors

of ref. [50] go about this complication by setting the temperature of all star-forming gas cells to $T = 10^4$K, and repeating the default calculation f_{H_N} at post-processing with this temperature value. Here, based on the discussion in ref. [53], we follow an even simpler approach and consider *all hydrogen of star forming cells to be neutral.* This is based on the observation that the hot gas in star-forming regions is expected to be entirely ionized, the cold component is mostly neutral, and that the star formation model of ref. [51] predicts this cold fraction to be within 0.9 and 1 (cf. the top right panel of figure 1 in ref. [53] for an illustration of how good an approximation $f_{H_N} = 1$ is for star-forming gas).

Given the neutral hydrogren fraction, the second part of the modeling involves determining the fraction of H_N that is in atomic H_I and molecular H_2 form. As in ref. [50], we use the Krumholz-McKee-Tumlinson (KMT) model [54–56] to determine the fraction f_{H_2} of all neutral Hydrogen that is molecular:

$$f_{H_2} = \begin{cases} 1 - 0.75s/(1 + 0.25s) & \text{if } s < 2 \\ 0 & \text{if } s \geq 2 \text{ or gas is non-star-forming,} \end{cases} \tag{2.1}$$

with

$$s = \frac{\ln\left(1 + 0.6\chi + 0.01\chi^2\right)}{0.6\tau_c}, \tag{2.2}$$

$$\chi = 0.756\left(1 + 3.1Z^{0.365}\right), \tag{2.3}$$

$$\tau_c = \frac{3}{4}\frac{\Sigma_{\text{gas}}\sigma_d}{\mu_H}, \tag{2.4}$$

and where Z is the gas metalicity in solar units, $\sigma_d = Z \times 10^{-21}$ cm^2 is an estimate of the cross-section of dust, μ_H is the mean hydrogen nucleus mass and Σ_{gas} is the gas surface density. The latter can be calculated as $\Sigma_{\text{gas}} = \lambda_J\rho_{\text{gas}}$, where λ_J is the gas Jeans length (used here as a proxy for the size of gravitationally-bound gas clouds) and ρ_{gas} is the gas volume density, or using a more sophiscated approach based on actually projecting the three-dimensional gas distribution in the simulation [53, 57]. Here, we follow the approximate approach of ref. [50] and evaluate the gas surface density as $\Sigma_{\text{gas}} = R\rho_{\text{gas}}$, with $R = (3V_{\text{cell}}/(4\pi))^{1/3}$ and V_{cell} the volume of the Voronoi gas cell in AREPO. Finally, the atomic neutral hydrogen fraction we are interested in is given by $f_{H_I} = 1 - f_{H_2}$, and we compute the H_I energy density as $\rho_{H_I} = f_{H_I}f_{H_N}\rho_H$, where ρ_H is the total hydrogen density in each gas cell.

It should be noted that the KMT model is just one of several different approaches to model the H_I and H_2 phases of neutral hydrogen in hydrodynamical simulations. More advanced analytical approaches can account in particular also for the contribution of young stars to the ionizing radiation field in the interstellar medium, and other approaches based on empirical observational correlations and calibration against high-resolution numerical simulations also exist; see refs. [53, 57] for a comparison of such different models using also IllustrisTNG. The simple KMT model can nonetheless be regarded as representative enough of the phenomenology of all these different approaches, and sufficient to our purpose here to begin to build intuition about the local PNG bias parameters of the H_I. In fact, figure 13 of ref. [53] shows that the dependence of f_{H_2} on the local hydrogen density predicted by the KMT model is in between the predictions from many other models, which ref. [57] subsequently showed perform roughly in the same way when compared to the data. This justifies the adoption of the KMT model in this paper, but we note that future analyses of these H_I bias parameters should eventually determine the impact of different H_I modeling techniques.
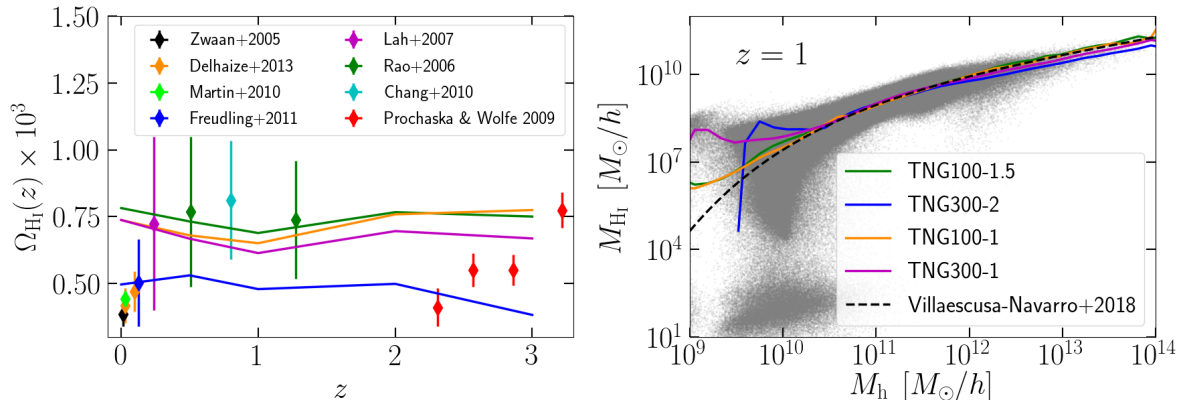
Figure 1. Time evolution of the H$_I$ cosmic energy density Ω_{H_I} (left) and mean H$_I$-to-total halo mass relation $M_{H_I}(M_h)$ at $z = 1$ (right). The result is shown for the fiducial cosmology and for the TNG300-2 (blue), TNG100-1.5 (green), TNG300-1 (purple) and TNG100-1 (orange) resolutions. On the left, the symbols with error bars show a few representative observational estimates compiled in table 2 of ref. [58], and converted to our fiducial cosmology: black from ref. [59], orange from ref. [58], light green from ref. [60], blue from ref. [61], purple from ref. [62], green from ref. [63], cyan from ref. [64] and red from ref. [65] (see ref. [57] for a discussion of the comparison between the simulation and these observational results). On the right, the grey dots mark the actual H$_I$ and total mass of the halos in the TNG300-1 simulation, and the black dashed line shows the three-parameter fit obtained in ref. [50] using their modeling of H$_I$ in TNG100-1.

The left panel of figure 1 shows the time evolution of the total H$_I$ energy density $\Omega_{H_I}(z) = \rho_{H_I}(z)/\rho_{c0}$, where $\rho_{c0} = 3H_0^2/(8\pi G)$ is the critical energy density today. The result is shown for the fiducial cosmology of our TNG300-2 and TNG100-1.5 resolutions (blue and green), as well as for the original TNG300-1 and TNG100-1 simulations (purple and orange). For the latter two cases, our results agree very well with those obtained using the same simulation data in refs. [50, 57]. Recall, in particular, that the modeling in ref. [50] differs from ours in that they recompute f_{H_N} in star-forming gas at post-processing, whereas we simply set it to unity: the good agreement between the two works thus illustrates the weak impact of this approximation. We have also explicitly checked that the H$_I$ content of halos in our simulations agrees very well with that shown in ref. [50]. As an example, we show in the right panel of figure 1, the mean H$_I$-to-total halo mass relation $M_{H_I}(M_h)$ at $z = 1$. The black dashed curve shows the fit of ref. [50] to their relation in the TNG100-1 simulation, which agrees very well with our result for the same resolution (orange line) for $M_h \gtrsim 10^{10}\, M_\odot/h$; the difference at lower halo masses is expected as the simple fit does not describe the details of the low-mass end of the distribution also in ref. [50].

Figure 1 shows also that numerical resolution has a strong impact on the H$_I$ distribution. Generically, Ω_{H_I} decreases with decreasing resolution, as best seen by the markedly lower H$_I$ abundance in TNG300-2 compared to all other resolutions. Note however that TNG100-1.5 is a resolution intermediate to TNG100-1 and TNG300-1, but it has more H$_I$ at $z < 1$. The impact of numerical resolution can be understood at least partly by inspecting the behavior of the $M_{H_I}(M_h)$ relation. For example, the right panel of figure 1 shows that this relation in TNG300-2 cuts off sharply for $M_h \lesssim 5 \times 10^9\, M_\odot/h$. This is because at this resolution, these objects are not as well resolved, and supernovae driven winds, tidal stripping and heating by UV background radiation become very efficient at removing gas, which gets more easily ionized once in the intergalactic medium (see ref. [50] for an indepth discussion of this

low-mass cutoff). Additionally, high-density gas clouds that can shield the H_I from ionizing radiation are also harder to resolve overall in TNG300-2. At the higher mass end, starting from $M_h \gtrsim 10^{12} M_\odot/h$, the halos in TNG300-2 also display a lower amount of H_I, which is partly due to these larger halos comprising also a larger number of poorly resolved subhalos, which for the reasons just discussed are not as H_I rich. The results from the other numerical resolutions can be broadly explained along similar lines. We shall take these considerations around numerical resolution into account when we interpret our numerical findings below.

2.3 Bias parameter measurements

We evaluate the H_I bias parameters b_1, b_ϕ and $b_{\phi\delta}$ in a way that is in all analogous to how ref. [35] evaluated the same parameters for halos and galaxies. Concerning b_1, we estimate it as

$$b_1(z) = \lim_{k \to 0} \frac{P_{mH_I}(k,z)}{P_{mm}(k,z)}, \tag{2.5}$$

where P_{mH_I} and P_{mm} are the H_I-matter cross-power spectrum and matter power spectrum, respectively. The lowest wavenumbers probed by our simulations are $k \gtrsim 0.03h/\text{Mpc}$ for TNG300-2 and $k \gtrsim 0.08h/\text{Mpc}$ for TNG100-1.5, where the scale-dependence of the P_{mH_I}/P_{mm} ratio can still be non-negligible. To make our estimates of b_1 more robust then, we fit this ratio on scales $k < 0.15h/\text{Mpc}$ for TNG300-2 and $k < 0.3h/\text{Mpc}$ for TNG100-1.5 using $b_1 + Ak^2$, where A is a nuisance parameter that absorbs the leading-order impact from scale-dependent corrections.[2] We show as error bars on b_1 the error estimated from the least-squares fitting method.

For the case of b_ϕ, we estimate it using its definition as the response of the H_I energy density to changes in the amplitude of the primordial gravitational potential ϕ in cosmologies with local PNG (cf. eq. (1.3)). Concretely, if ϕ_L is the amplitude of some long-wavelength primordial potential perturbation, then it can be shown using the separate universe ansatz [29, 66] that local structure formation inside this perturbation is equivalent to global structure formation in a cosmology with the primordial scalar power spectrum amplitude $\mathcal{A}_s$ rescaled as $\mathcal{A}_s \to \mathcal{A}_s [1 + \delta_{\mathcal{A}_s}]$, where $\delta_{\mathcal{A}_s} = 4f_{\text{NL}}\phi_L$, i.e.,

$$b_\phi(z) = \frac{\partial \ln \rho_{H_I}(z)}{\partial(f_{\text{NL}}\phi_L)} \equiv 4\frac{\partial \ln \rho_{H_I}(z)}{\partial \delta_{\mathcal{A}_s}}. \tag{2.6}$$

The bias parameter b_ϕ can then be estimated by finite diferencing using our set of fiducial, High$\mathcal{A}_s$ and Low$\mathcal{A}_s$ cosmologies as

$$b_\phi(z) = \frac{b_\phi^{\text{High}\mathcal{A}_s} + b_\phi^{\text{Low}\mathcal{A}_s}}{2}, \tag{2.7}$$

with

$$b_\phi^{\text{High}\mathcal{A}_s} = \frac{4}{+|\delta_{\mathcal{A}_s}|} \left[\frac{\rho_{H_I}^{\text{High}\mathcal{A}_s}(z)}{\rho_{H_I}^{\text{Fiducial}}(z)} - 1 \right], \tag{2.8}$$

$$b_\phi^{\text{Low}\mathcal{A}_s} = \frac{4}{-|\delta_{\mathcal{A}_s}|} \left[\frac{\rho_{H_I}^{\text{Low}\mathcal{A}_s}(z)}{\rho_{H_I}^{\text{Fiducia}}(z)} - 1 \right], \tag{2.9}$$

---

[2]We also calculated b_1 using the original TNG300-1 simulation (not shown), which is closer in resolution to TNG100-1.5, but has a bigger box size. We found the corresponding values of b_1 to agree relatively well, and so that any errors on our b_1 estimates deriving from the small size of TNG100-1.5 do not have a significant impact on our final conclusions.

Redshift	$z = 0$	$z = 0.5$	$z = 1$	$z = 2$	$z = 3$
b_1 (TNG100-1.5)	0.73 ± 0.07	1.01 ± 0.04	1.42 ± 0.02	2.04 ± 0.02	2.5 ± 0.04
b_1 (TNG300-2)	0.62 ± 0.06	0.86 ± 0.04	1.18 ± 0.03	1.80 ± 0.02	2.58 ± 0.06
b_ϕ (TNG100-1.5)	-1.70 ± 0.05	-0.44 ± 0.39	0.47 ± 0.12	2.39 ± 0.12	3.82 ± 0.06
b_ϕ (TNG300-2)	-1.76 ± 0.16	-1.40 ± 0.11	-0.26 ± 0.32	2.08 ± 0.14	1.98 ± 0.22
$b_{\phi\delta}$ (TNG100-1.5)	-3.86 ± 0.68	-2.27 ± 0.47	-1.70 ± 1.71	0.27 ± 1.44	5.37 ± 1.66
$b_{\phi\delta}$ (TNG300-2)	-2.55 ± 0.65	-3.23 ± 0.36	-2.93 ± 0.52	1.06 ± 0.97	1.09 ± 14.5

Table 1. Values of the bias parameters b_1, b_ϕ and $b_{\phi\delta}$ of the H_I distribution measured in this work.

and where the superscripts indicate in which cosmology the H_I energy density is measured (recall $|\delta_{\mathcal{A}_s}| = 0.05$). With only one realization of the initial conditions we cannot quote errors on our measurements in a statistical ensemble sense, and bootstrapping or resampling methods are also not adequate approaches given our relatively small box sizes. We note however that $b_\phi{}^{\text{High}\mathcal{A}_s}$ and $b_\phi{}^{\text{Low}\mathcal{A}_s}$ should be the same up to numerical noise, and so we will use their difference as a rough estimate of the error in our measurements. Note also that since the three cosmologies share the same phases of the initial conditions, sample variance errors on b_ϕ cancel already to a large degree.

Finally, we evaluate $b_{\phi\delta}$ as

$$b_{\phi\delta}(z) = \left[\frac{\partial \ln b_1(z)}{\partial(f_{\text{NL}}\phi)} + b_\phi(z)\right] b_1(z) \equiv \left[4\frac{\partial \ln b_1(z)}{\partial \delta_{\mathcal{A}_s}} + b_\phi(z)\right] b_1(z), \qquad (2.10)$$

where the second equality follows again from the separate universe equivalence between long-wavelength $f_{\text{NL}}\phi$ perturbations and changes to $\mathcal{A}_s$. This equation can be derived from the definitions of b_1 and $b_{\phi\delta}$ in eq. (1.3) [35], noting that $\rho_{H_I} b_{\phi\delta} = \partial(\rho_{H_I} b_1)/\partial(f_{\text{NL}}\phi)$. We evaluate the derivative of b_1 in eq. (2.10) via finite differences (analogously to eqs. (2.7)–(2.9)) using the values of b_1 estimated from the fiducial, High$\mathcal{A}_s$ and Low$\mathcal{A}_s$ simulations using eq. (2.5). Note that $b_{\phi\delta}$ could also be estimated with separate universe simulations that account simultaneously for changes to $\mathcal{A}_s$ and the mean cosmic matter density, but at the cost of running additional simulations to the ones we use in this paper.

3 Results

In this section we present and discuss our main H_I bias results. We discuss first the parameter b_ϕ and its relation to b_1, and then do the same for the $b_{\phi\delta}$ parameter. Our numerical values for these three bias parameters are listed in table 1.

3.1 The b_ϕ parameter of H_I

The left panel of figure 2 shows the redshift evolution of the b_ϕ values of H_I from the TNG100-1.5 and TNG300-2 resolutions, as labeled. Both resolutions predict that b_ϕ increases with increasing redshift, and that it is negative at $z \lesssim 1$. The two resolutions are in relatively good agreement in their predictions, although there are visible differences at $z = 0.5$, $z = 1$ and $z = 3$. The right panel of figure 2 shows b_ϕ plotted against the corresponding values of b_1, where we note a similar level of agreement between the two resolutions, except perhaps again at $z = 3$ (rightmost point) where TNG300-2 displays a markedly lower value of b_ϕ.

3.1.1 Halo model interpretation of the results

The halo model is a useful tool to guide the interpretation of our numerical results. Assuming that all of the H_I is inside halos with some mass M_h, then b_ϕ can be written as [49]

$$b_\phi^{H_I}(z) = \frac{\int dM_h n_h(M_h) M_{H_I}(M_h) \left[b_\phi{}^h(M_h) + R_\phi^{H_I}(M_h) \right]}{\int dM_h n(M_h) M_{H_I}(M_h)}, \tag{3.1}$$

where $n_h(M_h)$ is the differential halo mass function, $M_{H_I}(M_h)$ is the mean H_I mass in halos of mass M_h (cf. figure 1), $b_\phi{}^h(M_h)$ is the local PNG bias of the halos and

$$R_\phi^{H_I}(M_h) = \frac{\partial \ln M_{H_I}(M_h)}{\partial(f_{NL}\phi)} \tag{3.2}$$

is the response of the $M_{H_I}(M_h)$ relation to long-wavelength perturbations $f_{NL}\phi$. This response function is an ingredient that is often ignored in the literature, but which must be present in general to account for the modulation of $M_{H_I}(M_h)$ by $f_{NL}\phi$, in the same way that $b_\phi{}^h$ accounts for the same modulation of the halo mass function; see ref. [49] for more details. We have evaluated eq. (3.1) using its ingredients measured from our TNG100-1.5 and TNG300-2 simulations. In doing so, we considered all halos with at least $N_{cell} = 50$ gas cells, which for TNG100-1.5, accounts for over 99% of all of the H_I at $z < 2$ and 96% at $z = 3$, and for TNG300-2, it accounts for over 99% at $z < 1$, 97% at z=2 and 91% at z=3. The result is shown by the dashed lines in the left panel of figure 2, which agree extremely well with the simulation measurements, as expected.[3]

To focus on the impact of the response function $R_\phi^{H_I}(M_h)$, we show as dotted lines on the left panel of figure 2 the outcome of eq. (3.1), but with the contribution from $R_\phi^{H_I}(M_h)$ set to zero; by comparing to the dashed line, which shows the full result, one can thus directly appreciate the importance of this response function. Concretely:

1. At $z \leq 1$, the dashed lines are below the dotted lines, which indicates a negative net effect from $R_\phi^{H_I}$. That is, the halos become generically H_I poorer inside $f_{NL}\phi$ perturbations.

2. At $z = 2$, the dashed and dotted lines are comparable, which signals a smaller contribution from $R_\phi^{H_I}$. That is, the H_I content of halos responds less strongly to $f_{NL}\phi$ perturbations.

3. At $z = 3$, for TNG100-1.5, the net effect from $R_\phi^{H_I}$ is positive, i.e., the $f_{NL}\phi$ perturbations make the halos slightly H_I richer (this is to a smaller extent already visible at $z = 2$). For TNG300-2, however, $R_\phi^{H_I}$ has a very negative net effect (cf. dashed blue line below the dotted one).[4]

---

[3]The poorer performance for TNG300-2 at $z = 3$ has to do with the fact that the halos with $N_{cell} \geq 50$ account only for 91% of all of the H_I in the simulation box, and thus the starting assumption of the halo model breaks down slightly.

[4]This is likely due to the lower resolution of TNG300-2, where in the less well resolved low-mass halos, the enhanced supernovae feedback and tidal stripping effects caused by the $f_{NL}\phi$ perturbation (i.e. an increase in $\mathcal{A}_s$) may be unrealistically too efficient at ejecting gas out to lower density regions more exposed to the ionizing radiation.
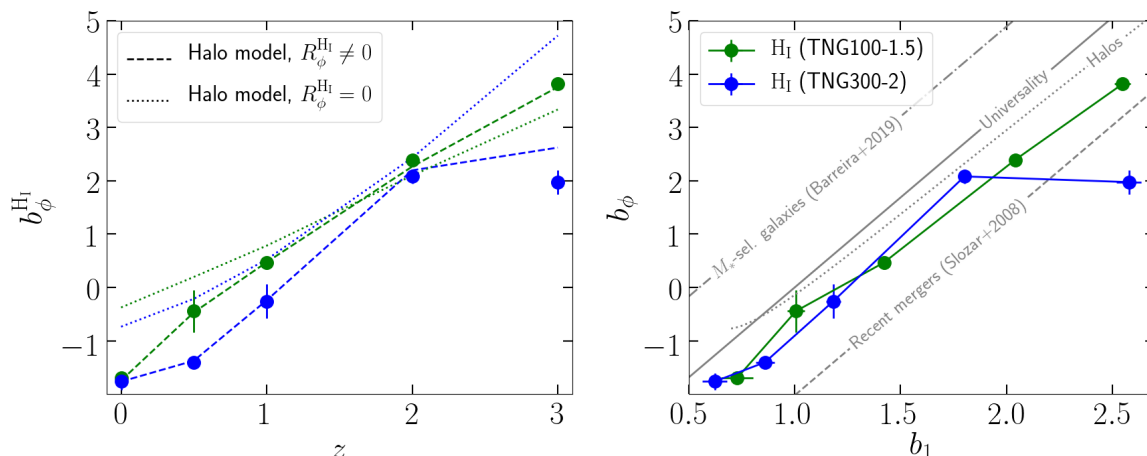
Figure 2. Time evolution of the b_ϕ parameter of H_I (left) and its relation to b_1 (right). The result is shown by the filled circles with error bars for the two resolutions TNG100-1.5 (green) and TNG300-2 (blue). On the left panel, the dashed and dotted lines show the predictions of the halo model, respectively, with and without the response $R_\phi^{H_I}$ of the H_I-to-total halo mass relation taken into account in eq. (3.1). On the right, the solid grey line marks the universality relation, the dotted grey line is the expected relation for halos, the grey dashed line shows the recent-merger relation of ref. [66], and the relation for stellar-mass selected galaxies in IllustrisTNG from ref. [38] is shown by the grey dot-dashed line.

We do not show all of the response function measurements for brevity, and also because they are still relatively noisy and more simulations are needed to aid a more detailed inspection of their mass- and redshift-dependence. Just as a single illustrative case, however, the left panel of figure 3 shows the response function $R_\phi^{H_I}$ measured using our separate universe simulations at $z = 1$. Indeed, and as anticipated in point 1 above, the values of $R_\phi^{H_I}$ are either negative or compatible with zero on the mass range that contributes sizeably to the integral of eq. (3.1).

3.1.2 The connection to the total hydrogen distribution

To understand further these results, we investigated also the response of the total hydrogen content of the halos, i.e., $R_\phi^H = \partial \ln M_H(M_h)/\partial(f_{NL}\phi)$, where $M_H(M_h)$ is the mean mass in hydrogen (not just H_I) in halos of mass M_h. This is shown in the middle panel of figure 3 for $z = 1$, where we observe that it has a shape and size that is similar to the response of the H_I content on the left. That is, the suppression of the H_I content in halos inside $f_{NL}\phi$ perturbations at $z = 1$ follows to some degree the suppression of the total amount of hydrogen H in those halos. Physically, this behavior of R_ϕ^H is expected for at least two reasons, both associated with the enhanced structure formation inside $f_{NL}\phi$ perturbations (or equivalently, given an increase in $\mathcal{A}_s$):

1. First, as found in ref. [38], these perturbations enhance star formation, which accelerates the transformation of the primordial hydrogen into stars and heavier metals. To help visualize this effect, the right panel of figure 3 shows the response $R_\phi^{Metals} = \partial \ln M_{Metals}(M_h)/\partial(f_{NL}\phi)$, where $M_{Metals}(M_h)$ is the mean mass of all metals produced by stars inside halos of mass M_h. Indeed, the positive value of R_ϕ^{Metals}
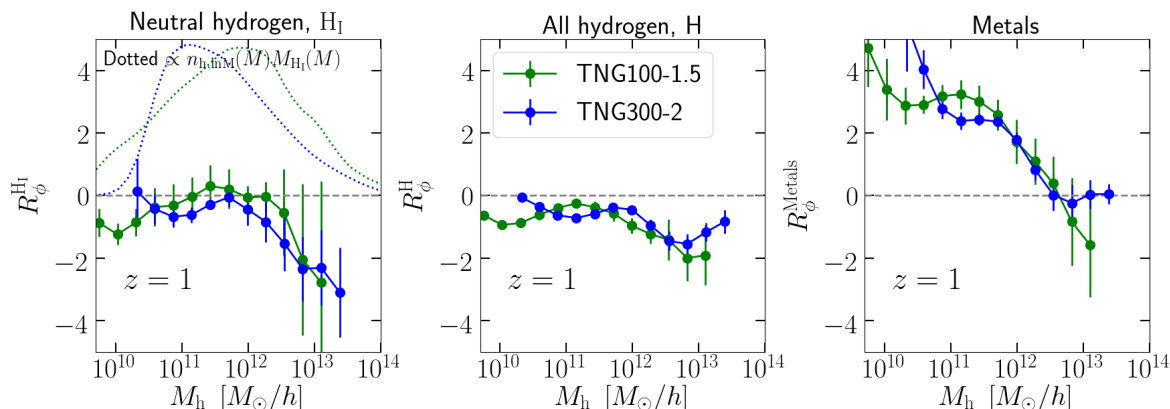
Figure 3. Response functions of the halo gas-to-total mass relation to long-wavelength perturbations $f_{\rm NL}\phi$, $R_\phi^{\rm gas} = \partial{\rm ln}M_{\rm gas}(M_{\rm h})/\partial(f_{\rm NL}\phi)$. The result is shown on the left for atomic neutral hydrogen (gas = H_I), in the middle for all hydrogen (gas = H), and on the right for all metals produced by stars (gas = Metals). All panels are for $z = 1$ and the result is shown for both resolutions TNG100-1.5 and TNG300-2, as labeled. On the left, the dotted lines are $\propto n_{\rm h,lnM}M_{\rm H_I}$, where $n_{\rm h,lnM}$ is the number of halos per logarithmic mass bin; this helps visualize which halo mass scales contribute the most to eq. (3.1).

confirms that $f_{\rm NL}\phi$ perturbations boost the conversion of hydrogen onto heavier metals through star formation.[5]

2. The second reason has to do with the fact that these perturbations also enhance the effects of supernovae (for $M_{\rm h} \lesssim 10^{12}M_\odot/h$) and black hole feedback (for $M_{\rm h} \gtrsim 10^{12}M_\odot/h$) that expel some of the gas out of the halos.

For another perspective into these results, we show in figure 4 the redshift evolution of the response of the total mass, mass in H_I, mass in all hydrogen and mass in metals, found in all halos in the mass bin $M_{\rm h} \in \left[10^{12}, 5 \times 10^{12}\right] M_\odot/h$, as labeled. Focusing first on the TNG300-2 results on the right, the black line shows that the $f_{\rm NL}\phi$ perturbations work to increase the total mass (except at $z = 0$ when it decreases just slightly), and do more so at higher redshift; this is a consequence of there being more halos in the bin, as well as each halo becoming also more massive. The figure shows also that while the mass in H_I can increase for $z \gtrsim 1$ (cf. blue line), this increase is always smaller than that in total mass, which effectively causes a suppression of the H_I-to-total mass relation, i.e., $R_\phi^{H_I} < 0$ (cf. figure 3). At this resolution, the behavior of the H_I response is well traced by that of the total hydrogen shown in cyan color. Note also how the effects of enhanced star formation are visible at $z \geq 1$ by the enhanced mass in metals (cf. orange line), but at lower redshift the stronger effects by black hole feedback inside $f_{\rm NL}\phi$ perturbations eventually work to remove metals from the halos, and can lead to the response of the mass in metals to drop below that of the total mass at $z = 0$.

The results from the TNG100-1.5 resolution on the left of figure 4 agree relatively well with those from TNG300-2, with the main noteworthy difference being that in TNG100-1.5 the behavior of the H_I response is not as well traced by the total hydrogen response. Concretely, at $z \leq 1$ the response of the H_I mass is below that of the total mass, but still

---

[5]This result finds also interesting ramifications to attempts to constrain cosmology and local PNG using intensity mapping observations from the emission lines of some of these heavier elements [67–69].
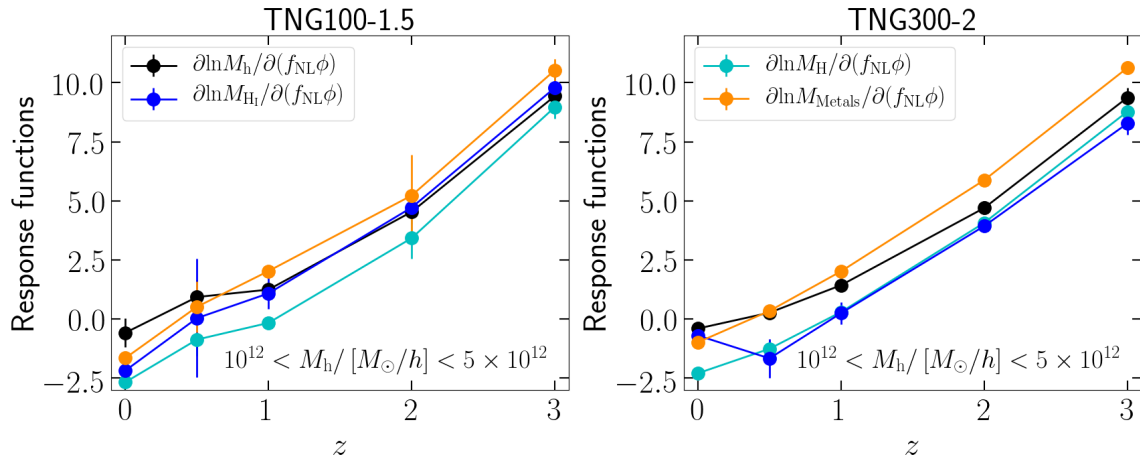
Figure 4. Response of the total mass (black), mass in H_I (blue), mass in all hydrogen (cyan) and mass in metals (orange) found inside all halos in the total mass bin $M_h \in \left[10^{12}, 5 \times 10^{12}\right] M_\odot/h$. The result is shown as a function of redshift, and on the left and right for TNG100-1.5 and TNG300-2, respectively.

visibly above the response of the total hydrogen, i.e., the $f_{NL}\phi$ perturbations still lower the H_I-to-total mass relation of these halos, but lower the H-to-total mass relation even more. Further, at $z \geq 2$ the $f_{NL}\phi$ perturbations can actually lead to an increase in H_I mass that slightly outweighs the increase in total mass. This could be associated with the fact that $f_{NL}\phi$ perturbations promote the formation of denser gas clouds that can shield the hydrogen inside them more efficiently from the ionizing radiation. These dense structures are not as well resolved in TNG300-2, which can explain why the same effect is not visible at this lower resolution.

There are more physical effects that can drive differences between the impact that $f_{NL}\phi$ perturbations have on the H_I and hydrogen mass of halos. For example, related to point 2 above, for the case of the H_I it is actually simply sufficient that the enhanced feedback or tidal stripping effects remove gas from inside high-density regions in the halo (and not completely from inside the halo altogether) for it to be more exposed to the ionizing radiation; this lowers $R_\phi^{H_I}$, but not R_ϕ^H. Further, $f_{NL}\phi$ perturbations also modify the details of the evolution of AGN, and thus their contribution to the local ionizing radiation field in IllustrisTNG (which impacts the H_I directly, but not the total hydrogen). In fact, here one should note that since stars form earlier inside $f_{NL}\phi$ perturbations, reionization begins earlier as well. This effect is not accounted for in our results as we assumed the time-dependent UV background (the main driver of reionization in IllustrisTNG) to be the same in our fiducial, High$\mathcal{A}_s$ and Low$\mathcal{A}_s$ cosmologies. This should not have a critical impact in our results as we focus on redshifts well after reionization is complete, but it is interesting to investigate the impact of this approximation in works more focused around the epoch of reionization $6 \lesssim z \lesssim 10$. We defer to future work a more indepth investigation of the complicated interplay between these and other astrophysical effects, which we note will depend in general on the assumed galaxy formation model and H_I modeling strategy.

3.1.3 Comparison to the universality relation

Finally, we compare the $b_\phi(b_1)$ relation of H_I with other relations in the literature. The solid grey line in the right panel of figure 2 shows the popular universality relation $b_\phi = 2\delta_c (b_1 - 1)$,

where $\delta_c = 1.676$ is the critical density for spherical collapse. The typical approach in the literature (see e.g. refs. [14, 16, 17]) estimates the $b_\phi(b_1)$ relation of the H_I using eq. (3.1) with $R_\phi^{H_I} = 0$ and assuming the universality relation for the halo bias b_h. Since the universality relation is linear, is follows trivially then that $b_\phi(b_1)$ of the H_I is given by the exact same relation, irrespective of the form of $M_{H_I}(M_h)$. This can be checked by assuming $b_\phi^h(M_h) = A + B b_1^h(M_h)$ for the bias of the halos in eq. (3.1) with $R_\phi^{H_I} = 0$, and see that this always results in the same relation $b_\phi^{H_I} = A + B b_1^{H_I}$ for the H_I.

The figure shows that this calculation overpredicts the amplitude of the H_I $b_\phi(b_1)$ relation measured from our IllustrisTNG simulations, which we checked (not shown) is roughly bracketed by $b_\phi = 2\delta_c(b_1 - p)$ with $p \in [1.1, 1.4]$. For $b_1 \lesssim 1.5$ ($z \lesssim 1$), this is largely because of the negative values of $R_\phi^{H_I}$, which push the $b_\phi(b_1)$ relation downwards. For $b_1 \gtrsim 1.5$ ($z \gtrsim 1$) the result is a combination of the impact of $R_\phi^{H_I}$ and the relation of the halos shown by the grey dotted line,[6] which is itself already below the universality relation. Note also that the predictions for the bias parameter b_1 also affect the $b_\phi(b_1)$ relation. For comparison purposes only, the right panel of figure 2 shows also the *recent-merger* relation $b_\phi = 2\delta_c(1-1.6)$ (dashed) derived by ref. [66] and that is usually assumed in constraints using quasars [5, 6], and the relation $b_\phi = 2\delta_c(1-0.55)$ (dot-dashed), which was found in ref. [38] to roughly describe the case for stellar-mass selected galaxies in IllustrisTNG.

3.2 The $b_{\phi\delta}$ parameter of H_I

Our results for the $b_{\phi\delta}$ parameter of H_I are displayed in figure 5, which has the same format as figure 2 discussed above for b_ϕ. The first noteworthy point is that our $b_{\phi\delta}$ estimates are noisier compared to b_ϕ, which is expected since $b_{\phi\delta}$ is a second-order bias parameter; this is especially dramatic in our results for TNG300-2 at $z = 3$. Despite the noise, there are still a few trends that one can discern from our results, in particular, that $b_{\phi\delta}$ is generically a growing function of redshift ($b_{\phi\delta} < 0$ at $z \lesssim 2$), and that for $b_1 \lesssim 1.5$, the simulation measurements predict a $b_{\phi\delta}(b_1)$ relation that falls below the corresponding universality expression (grey solid line in figure 5). The latter is in this case given by $b_{\phi\delta}(b_1) = b_\phi - b_1 + 1 + \delta_c[b_2(b_1) - (8/21)(b_1 - 1)]$, where b_2 is the bias parameter associated second-order mass perturbations $\delta_m(\boldsymbol{x}, z)^2$, and which we evaluate using the polynomial fit $b_2(b_1) = 0.412 - 2.143b_1 + 0.929b_1^2 + 0.008b_1^3$ of ref. [72]. This fit was obtained for dark matter halos using separate universe simulations that do not assume universality of the halo mass function, but the impact of this is small for the purpose of our comparisons here (see ref. [72] for a discussion).

Similarly to as discussed in the last section for b_ϕ, here too the halo model is a useful tool to interpret our numerical results. The halo model expression for $b_{\phi\delta}$ is given by [49]

$$b_{\phi\delta}{}^{H_I}(z) = \frac{\int dM_h n_h(M_h) M_{H_I}(M_h) \left[b_{\phi\delta}{}^h(M_h) + b_1^h(M_h) R_\phi^{H_I}(M_h) + b_\phi^h(M_h) R_1^{H_I}(M_h) + R_{\phi\delta}^{H_I}(M_h) \right]}{\int dM_h n(M_h) M_{H_I}(M_h)}, \tag{3.3}$$

where b_1^h, b_ϕ^h and $b_{\phi\delta}^h$ are the bias parameters of the halos, and in addition to $R_\phi^{H_I}$ defined in eq. (3.2), now the result depends also on two extra response functions of the H_I-to-total

---

[6]To evaluate the $b_\phi(b_1)$ relation of the halos, we calculate $b_\phi = 4\partial \ln n^{Tinker}(M_h)/\partial \delta_{A_s}$, where n^{Tinker} is the halo mass function of ref. [70], and obtain b_1 using the fitting formula of ref. [71].
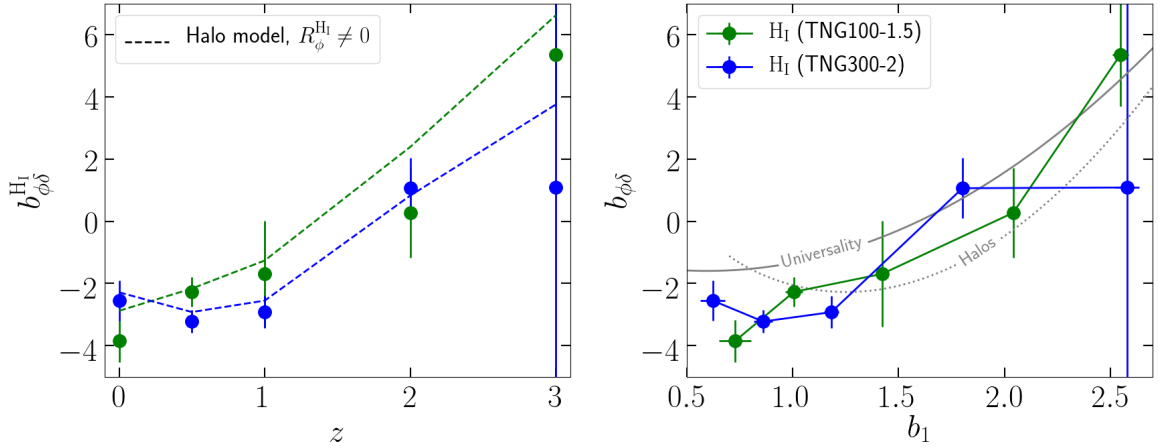
Figure 5. Same as figure 2, but for $b_{\phi\delta}$ instead of b_ϕ. On the left, the dashed lines show the halo model prediction of eq. (3.3) with $R_\phi^{H_I}$ given by the simulation measurements, but with $R_1^{H_I} = R_{\phi\delta}^{H_I} = 0$. On the right, the solid grey line marks the universality relation, and the grey dotted line shows the expected relation for halos obtained with the definition of eq. (2.10) using the fitting formulae from refs. [70, 71].

halo mass relation defined as

$$R_1^{H_I}(M_h) = \frac{\partial \ln M_{H_I}(M_h)}{\partial \delta_m}, \tag{3.4}$$

$$R_{\phi\delta}^{H_I}(M_h) = \frac{1}{M_{H_I}(M_h)} \frac{\partial^2 M_{H_I}(M_h)}{\partial \delta_m \partial(f_{NL}\phi)}. \tag{3.5}$$

The set of separate universe simulations we use in this work does not let us estimate these response functions reliably, which keeps us from being able to evaluate eq. (3.3) entirely. We evaluate it nonetheless with $R_1^{H_I} = R_{\phi\delta}^{H_I} = 0$ and the $R_\phi^{H_I}$ measured from the simulations, and the result is shown by the dashed lines in the left panel of figure 5.[7] The agreement between this halo model prediction and the simulation results is not perfect, but that is not surprising since the halo model result is incomplete; the difference between the dashed lines and the simulation measurements can in fact be used to estimate roughly the size of the combined contribution from $b_\phi^h R_1^{H_I} + R_{\phi\delta}^{H_I}$ that was neglected.

Concerning the comparison to the universality prediction, previous works in the literature [15, 18] have estimated the value of $b_{\phi\delta}$ for the H_I using eq. (3.3) with $R_\phi^{H_I} = R_1^{H_I} = R_{\phi\delta}^{H_I} = 0$ and assuming the universality expression for the bias of the halos $b_{\phi\delta}^h$. Since the second-derivative of the universality relation w.r.t. b_1 is positive, this pushes the corresponding relation of the H_I upwards. This can be seen by assuming $b_\phi^h(M_h) = A + Bb_1^h(M_h) + C\left[b_1^h(M_h)\right]^2$ for the bias of the halos in eq. (3.3), and noting that if $C > 0$, then at fixed numerical value of b_1 one has $b_{\phi\delta}^{H_I} > b_{\phi\delta}^h$. The right panel of figure 5 shows that for $b_1 \lesssim 2$ ($z \lesssim 2$) our simulation results begin to drop below the universality prediction, and so that this way of calculating the $b_{\phi\delta}$ parameter in the literature overestimates (the result becomes less negative) the $b_{\phi\delta}(b_1)$ relation of the H_I in our simulations. We leave a more indepth discussion

---

[7]To perform this calculation we evaluate b_1^h using the fitting formulae of ref. [71] and take $b_{\phi\delta}^h$ to be the dotted line on the right panel of figure 5. All other ingredients in eq. (3.3) are directly measured from the simulations.

of the $b_{\phi\delta}$ parameter of the H_I distribution to future work, which would benefit from larger simulation volumes for improved statistics.

4   Summary and conclusions

Upcoming 21cm line-intensity mapping observations will determine the large-scale distribution of atomic neutral hydrogen H_I, which is a very sensitive probe of the local PNG parameter f_{NL}. These future data have the potential to tighten up the current best constraint on f_{NL} using CMB data, and consequently, to improve on our knowledge of the physics of the early Universe and inflation. The main observational signatures appear in the power spectrum and bispectrum of the H_I distribution and are proportional to terms like $b_1 b_\phi f_{NL}$ and $b_1^2 b_{\phi\delta} f_{NL}$, where b_1, b_ϕ and $b_{\phi\delta}$ are three bias parameters that describe how the H_I energy density depends on large-scale mass overdensities δ_m and gravitational potential perturbations $f_{NL}\phi$ (cf. eqs. (1.2) and (1.3)). Naturally, given the strong degeneracy with f_{NL}, accurate and precise theoretical priors on the expected values of these parameters are of the utmost importance to infer the numerical value of f_{NL} from the data. Our main goal in this paper was to take the first steps towards estimating these bias parameters using hydrodynamical simulations of cosmic structure formation.

Concretely, in this paper we coupled separate universe simulations of the IllustrisTNG galaxy formation model with the analytical KMT model of the split of neutral hydrogen into atomic (H_I) and molecular (H_2) (cf. section 2.2) to predict the $b_\phi(b_1)$ and $b_{\phi\delta}(b_1)$ bias parameter relations of the H_I. We used simulation data obtained at two numerical resolutions called TNG300-2 ($L_{box} = 205\,\mathrm{Mpc}/h$, $N_p = 2 \times 1250^3$) and TNG100-1.5 ($L_{box} = 75\,\mathrm{Mpc}/h$, $N_p = 2 \times 1250^3$), and discussed our numerical findings with the aid of the halo model by inspecting the behavior of the H_I content of dark matter halos. Our main results can summarized as follows:

- The values of b_ϕ and $b_{\phi\delta}$ grow with redshift and are negative at $z \lesssim 1$ and $z \lesssim 2$, respectively (cf. table 1, and figures 2 and 5). Our two numerical resolutions agree also relatively well, despite some visible quantitative differences.

- The popular universality expressions for the $b_\phi(b_1)$ and $b_{\phi\delta}(b_1)$ relations overpredict our simulation measurements (cf. figures 2 and 5), indicating that the most common calculation found in the literature may currently overestimate the values of b_ϕ and $b_{\phi\delta}$.

- The lower values w.r.t. the universality expressions are in part due to the negative response $R_\phi^{H_I}$ of $M_{H_I}(M_h)$ to $f_{NL}\phi$ perturbations at $z \lesssim 1$ (cf. figure 3), i.e., halos become generically H_I poorer inside these perturbations. We discussed how this can be explained by the fact that these perturbations accelerate the conversion of hydrogen to heavier elements by star formation, and enhance the effects of feedback that ejects the gas out to regions where it is more easily ionized.

We note that although the impact of astrophysical uncertainties on f_{NL} constraints with line-intensity mapping data has been investigated in past works [14, 16, 17], this was done only through parametrizations of the gas-to-halo connection, i.e. $M_{H_I}(M_h)$ for the case of H_I. Our finding here that response functions like $R_\phi^{H_I}$ and R_ϕ^{Metals} can be generically nonzero introduces the environmental dependence of the gas-to-halo connection as an additional, new astrophysical ingredient whose impact on f_{NL} constraints needs to be understood as well.

This can be straightforwardly achieved already with the aid of eqs. (3.1) and (3.3), together with parametrizations for response functions like $R_\phi^{\mathrm{H_I}}$ [49].

Taken at face value, is our finding that the universality-based calculations overpredict the $b_\phi(b_1)$ and $b_{\phi\delta}(b_1)$ relations in our simulations good or bad news for f_{NL} constraints using $\mathrm{H_I}$ data? This is hard to answer generically since it depends on the survey setup and redshift range considered. For example, focusing just on the case of $b_\phi(b_1)$ for power spectrum analyses (cf. figure 2), for $b_1 < 1$, which corresponds roughly to $z \lesssim 0.5$–1, the universality relation is negative, and since our b_ϕ measurements are below it, they are larger in absolute value. This means the $\mathrm{H_I}$ distribution is actually more sensitive to the effects of f_{NL}. Conversely, however, starting from $b_1 > 1$, which corresponds roughly to $z \gtrsim 0.5$–1, the universality relation becomes positive, and so by assuming it, one overestimates the values of b_ϕ, and consequently, the true constraining power of the data on f_{NL}. One may argue that most of this constraining power comes from high redshift, where one can access the large distance scales where f_{NL} contributes most sizeably, but these are also the scales that are expected to be more affected by foreground subtraction systematics. The takeaway is that there is strong motivation to revisit the impact of $b_\phi(b_1)$ assumptions in existing forecast studies of f_{NL} using $\mathrm{H_I}$ data. Doing so will also serve the purpose to quantify the accuracy and precision with which future simulation-based works need to determine the local PNG $\mathrm{H_I}$ bias parameters in order for line-intensity mapping surveys to meet a certain target constraining power on f_{NL} (see refs. [33, 35] for discussions).

We emphasize that the modeling of the $\mathrm{H_I}$ in cosmological hydrodynamical simulations is a field that is still under development and therefore with still many associated uncertainties. The results and discussion in this paper provide already very useful intuition and guidance for future works on the subject, but having been obtained at fixed galaxy formation model and $\mathrm{H_I}$ modeling strategy, they should be regarded as just the first step towards more robust theoretical predictions for the b_ϕ and $b_{\phi\delta}$ parameters of the $\mathrm{H_I}$. The importance of these bias parameters in observational searches of local PNG strongly motivates giving continuation to the work started here, including with investigations of the sensitivity of the bias parameters to the assumed galaxy formation physics, and the exploration of more sophisticated $\mathrm{H_I}$ modeling strategies.

Acknowledgments

References

[1] N. Bartolo, E. Komatsu, S. Matarrese and A. Riotto, *Non-Gaussianity from inflation: theory and observations*, *Phys. Rept.* **402** (2004) 103 [`astro-ph/0406398`] [INSPIRE].

[2] X. Chen, *Primordial Non-Gaussianities from inflation models*, *Adv. Astron.* **2010** (2010) 638979 [`arXiv:1002.1416`] [INSPIRE].

[3] E. Komatsu and D.N. Spergel, *Acoustic signatures in the primary microwave background bispectrum*, *Phys. Rev. D* **63** (2001) 063002 [`astro-ph/0005036`] [INSPIRE].

[4] PLANCK collaboration, *Planck 2018 results. IX. Constraints on primordial non-Gaussianity*, *Astron. Astrophys.* **641** (2020) A9 [`arXiv:1905.05697`] [INSPIRE].

[5] E.-M. Mueller et al., *The clustering of galaxies in the completed SDSS-IV extended Baryon oscillation spectroscopic survey: primordial non-Gaussianity in Fourier Space*, `arXiv:2106.13725` [INSPIRE].

[6] E. Castorina et al., *Redshift-weighted constraints on primordial non-Gaussianity from the clustering of the eBOSS DR14 quasars in Fourier space*, *JCAP* **09** (2019) 010 [`arXiv:1904.08859`] [INSPIRE].

[7] M. Alvarez et al., *Testing inflation with large scale structure: connecting hopes with reality*, `arXiv:1412.4671` [INSPIRE].

[8] M. Biagetti, *The hunt for primordial interactions in the large scale structures of the universe*, *Galaxies* **7** (2019) 71 [`arXiv:1906.12244`] [INSPIRE].

[9] S. Camera, M.G. Santos, P.G. Ferreira and L. Ferramacho, *Cosmology on ultra-large scales with HI intensity mapping: limits on primordial non-Gaussianity*, *Phys. Rev. Lett.* **111** (2013) 171302 [`arXiv:1305.6928`] [INSPIRE].

[10] J. Fonseca, S. Camera, M. Santos and R. Maartens, *Hunting down horizon-scale effects with multi-wavelength surveys*, *Astrophys. J. Lett.* **812** (2015) L22 [`arXiv:1507.04605`] [INSPIRE].

[11] D. Alonso, P. Bull, P.G. Ferreira, R. Maartens and M. Santos, *Ultra large-scale cosmology in next-generation experiments with single tracers*, *Astrophys. J.* **814** (2015) 145 [`arXiv:1505.07596`] [INSPIRE].

[12] D. Alonso and P.G. Ferreira, *Constraining ultralarge-scale cosmology with multiple tracers in optical and radio surveys*, *Phys. Rev. D* **92** (2015) 063525 [`arXiv:1507.03550`] [INSPIRE].

[13] J. Fonseca, R. Maartens and M.G. Santos, *Synergies between intensity maps of hydrogen lines*, *Mon. Not. Roy. Astron. Soc.* **479** (2018) 3490 [`arXiv:1803.07077`] [INSPIRE].

[14] J.L. Bernal, P.C. Breysse, H. Gil-Marín and E.D. Kovetz, *User's guide to extracting cosmological information from line-intensity maps*, *Phys. Rev. D* **100** (2019) 123522 [`arXiv:1907.10067`] [INSPIRE].

[15] D. Karagiannis, A. Slosar and M. Liguori, *Forecasts on primordial non-Gaussianity from 21 cm Intensity Mapping experiments*, *JCAP* **11** (2020) 052 [`arXiv:1911.03964`] [INSPIRE].

[16] S. Camera and H. Padmanabhan, *Beyond ΛCDM with H i intensity mapping: robustness of cosmological constraints in the presence of astrophysics*, *Mon. Not. Roy. Astron. Soc.* **496** (2020) 4115 [`arXiv:1910.00022`] [INSPIRE].

[17] R.H. Liu and P.C. Breysse, *Coupling parsec and gigaparsec scales: primordial non-Gaussianity with multitracer intensity mapping*, *Phys. Rev. D* **103** (2021) 063520 [`arXiv:2002.10483`] [INSPIRE].

[18] D. Karagiannis, J. Fonseca, R. Maartens and S. Camera, *Probing primordial non-Gaussianity with the power spectrum and bispectrum of future 21 cm intensity maps*, *Phys. Dark Univ.* **32** (2021) 100821 [`arXiv:2010.07034`] [INSPIRE].

[19] J.-A. Viljoen, J. Fonseca and R. Maartens, *Multi-wavelength spectroscopic probes: prospects for primordial non-Gaussianity and relativistic effects*, *JCAP* **11** (2021) 010 [arXiv:2107.14057] [INSPIRE].

[20] E. Abdalla et al., *The BINGO project I: Baryon acoustic oscillations from integrated neutral gas observations*, arXiv:2107.01633 [INSPIRE].

[21] K. Bandura et al., *Canadian Hydrogen Intensity Mapping Experiment (CHIME) pathfinder*, *Proc. SPIE Int. Soc. Opt. Eng.* **9145** (2014) 22 [arXiv:1406.2288] [INSPIRE].

[22] L.B. Newburgh et al., *HIRAX: a probe of dark energy and radio transients*, *Proc. SPIE Int. Soc. Opt. Eng.* **9906** (2016) 99065X [arXiv:1607.02059] [INSPIRE].

[23] SKA collaboration, *Cosmology with Phase 1 of the Square Kilometre Array: Red Book 2018: technical specifications and performance forecasts*, *Publ. Astron. Soc. Austral.* **37** (2020) e007 [arXiv:1811.02743] [INSPIRE].

[24] E.D. Kovetz et al., *Line-intensity mapping: 2017 status report*, arXiv:1709.09066 [INSPIRE].

[25] S. Cunnington, S. Camera and A. Pourtsidou, *The degeneracy between primordial non-Gaussianity and foregrounds in 21 cm intensity mapping experiments*, *Mon. Not. Roy. Astron. Soc.* **499** (2020) 4054 [arXiv:2007.12126] [INSPIRE].

[26] J. Fonseca and M. Liguori, *Measuring ultralarge scale effects in the presence of 21 cm intensity mapping foregrounds*, *Mon. Not. Roy. Astron. Soc.* **504** (2021) 267 [arXiv:2011.11510] [INSPIRE].

[27] T.L. Makinen et al., *deep21: a deep learning method for 21 cm foreground removal*, *JCAP* **04** (2021) 081 [arXiv:2010.15843] [INSPIRE].

[28] V. Desjacques, D. Jeong and F. Schmidt, *Large-scale galaxy bias*, *Phys. Rept.* **733** (2018) 1 [arXiv:1611.09787] [INSPIRE].

[29] N. Dalal, O. Dore, D. Huterer and A. Shirokov, *The imprints of primordial non-Gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects*, *Phys. Rev. D* **77** (2008) 123514 [arXiv:0710.4560] [INSPIRE].

[30] P. McDonald, *Primordial non-Gaussianity: large-scale structure signature in the perturbative bias model*, *Phys. Rev. D* **78** (2008) 123519 [arXiv:0806.1061] [INSPIRE].

[31] T. Giannantonio and C. Porciani, *Structure formation from non-Gaussian initial conditions: multivariate biasing, statistics, and comparison with N-body simulations*, *Phys. Rev. D* **81** (2010) 063530 [arXiv:0911.0017] [INSPIRE].

[32] V. Assassi, D. Baumann and F. Schmidt, *Galaxy bias and primordial non-Gaussianity*, *JCAP* **12** (2015) 043 [arXiv:1510.03723] [INSPIRE].

[33] A. Barreira, *On the impact of galaxy bias uncertainties on primordial non-Gaussianity constraints*, *JCAP* **12** (2020) 031 [arXiv:2009.06622] [INSPIRE].

[34] A. Moradinezhad Dizgah, M. Biagetti, E. Sefusatti, V. Desjacques and J. Noreña, *Primordial Non-Gaussianity from biased tracers: likelihood analysis of real-space power spectrum and bispectrum*, *JCAP* **05** (2021) 015 [arXiv:2010.14523] [INSPIRE].

[35] A. Barreira, *Predictions for local PNG bias in the galaxy power spectrum and bispectrum and the consequences for f_{NL} constraints*, *JCAP* **01** (2022) 033 [arXiv:2107.06887] [INSPIRE].

[36] M. Grossi et al., *Large-scale non-Gaussian mass function and halo bias: tests on N-body simulations*, *Mon. Not. Roy. Astron. Soc.* **398** (2009) 321 [arXiv:0902.2013] [INSPIRE].

[37] M. Biagetti, T. Lazeyras, T. Baldauf, V. Desjacques and F. Schmidt, *Verifying the consistency relation for the scale-dependent bias from local primordial non-Gaussianity*, *Mon. Not. Roy. Astron. Soc.* **468** (2017) 3277 [arXiv:1611.04901] [INSPIRE].

[38] A. Barreira, G. Cabass, F. Schmidt, A. Pillepich and D. Nelson, *Galaxy bias and primordial non-Gaussianity: insights from galaxy formation simulations with IllustrisTNG*, *JCAP* **12** (2020) 013 [arXiv:2006.09368] [ɪɴSPIRE].

[39] V. Springel, *E pur si muove: Galiliean-invariant cosmological hydrodynamical simulations on a moving mesh*, *Mon. Not. Roy. Astron. Soc.* **401** (2010) 791 [arXiv:0901.4107] [ɪɴSPIRE].

[40] R. Weinberger et al., *Simulating galaxy formation with black hole driven thermal and kinetic feedback*, *Mon. Not. Roy. Astron. Soc.* **465** (2017) 3291 [astro-ph/1607.03486].

[41] A. Pillepich et al., *Simulating galaxy formation with the IllustrisTNG Model*, *Mon. Not. Roy. Astron. Soc.* **473** (2018) 4077 [arXiv:1703.02970] [ɪɴSPIRE].

[42] F. Marinacci et al., *First results from the IllustrisTNG simulations: radio haloes and magnetic fields*, *Mon. Not. Roy. Astron. Soc.* **480** (2018) 5113 [arXiv:1707.03396] [ɪɴSPIRE].

[43] A. Pillepich et al., *First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies*, *Mon. Not. Roy. Astron. Soc.* **475** (2018) 648 [arXiv:1707.03406] [ɪɴSPIRE].

[44] J.P. Naiman et al., *First results from the IllustrisTNG simulations: a tale of two elements — chemical evolution of magnesium and europium*, *Mon. Not. Roy. Astron. Soc.* **477** (2018) 1206 [arXiv:1707.03401].

[45] V. Springel et al., *First results from the IllustrisTNG simulations: matter and galaxy clustering*, *Mon. Not. Roy. Astron. Soc.* **475** (2018) 676 [arXiv:1707.03397] [ɪɴSPIRE].

[46] D. Nelson et al., *First results from the IllustrisTNG simulations: the galaxy colour bimodality*, *Mon. Not. Roy. Astron. Soc.* **475** (2018) 624 [arXiv:1707.03395] [ɪɴSPIRE].

[47] D. Nelson et al., *First results from the TNG50 simulation: galactic outflows driven by supernovae and black hole feedback*, *Mon. Not. Roy. Astron. Soc.* **490** (2019) 3234 [arXiv:1902.05554] [ɪɴSPIRE].

[48] A. Pillepich et al., *First results from the TNG50 simulation: the evolution of stellar and gaseous discs across cosmic time*, *Mon. Not. Roy. Astron. Soc.* **490** (2019) 3196 [arXiv:1902.05553] [ɪɴSPIRE].

[49] R. Voivodic and A. Barreira, *Responses of Halo occupation distributions: a new ingredient in the halo model & the impact on galaxy bias*, *JCAP* **05** (2021) 069 [arXiv:2012.04637] [ɪɴSPIRE].

[50] F. Villaescusa-Navarro et al., *Ingredients for 21 cm intensity mapping*, *Astrophys. J.* **866** (2018) 135 [arXiv:1804.09180] [ɪɴSPIRE].

[51] V. Springel and L. Hernquist, *Cosmological SPH simulations: a hybrid multi-phase model for star formation*, *Mon. Not. Roy. Astron. Soc.* **339** (2003) 289 [astro-ph/0206393] [ɪɴSPIRE].

[52] M. Vogelsberger, S. Genel, D. Sijacki, P. Torrey, V. Springel and L. Hernquist, *A model for cosmological simulations of galaxy formation physics*, *Mon. Not. Roy. Astron. Soc.* **436** (2013) 3031 [arXiv:1305.2913] [ɪɴSPIRE].

[53] B. Diemer et al., *Modeling the atomic-to-molecular transition in cosmological simulations of galaxy formation*, *Astrophys. J. Suppl.* **238** (2018) 33 [arXiv:1806.02341] [ɪɴSPIRE].

[54] M.R. Krumholz, C.F. McKee and J. Tumlinson, *The atomic to molecular transition in galaxies. I: an analytic approximation for photodissociation fronts in finite clouds*, *Astrophys. J.* **689** (2008) 865 [arXiv:0805.2947] [ɪɴSPIRE].

[55] M.R. Krumholz, C.F. McKee and J. Tumlinson, *The atomic to molecular transition in galaxies. II: HI and H_2 column densities*, *Astrophys. J.* **693** (2009) 216 [arXiv:0811.0004] [ɪɴSPIRE].

[56] C.F. McKee and M.R. Krumholz, *The atomic-to-molecular transition in galaxies. III. A new method for determining the molecular content of primordial and dusty clouds*, *Astrophys. J.* **709** (2010) 308 [arXiv:0908.0330] [iNSPIRE].

[57] B. Diemer et al., *Atomic and molecular gas in IllustrisTNG galaxies at low redshift*, *Mon. Not. Roy. Astron. Soc.* **487** (2019) 1529 [arXiv:1902.10714] [iNSPIRE].

[58] J. Delhaize, M. Meyer, L. Staveley-Smith and B. Boyle, *Detection of HI in distant galaxies using spectral stacking*, *Mon. Not. Roy. Astron. Soc.* **433** (2013) 1398 [arXiv:1305.1968] [iNSPIRE].

[59] M.A. Zwaan, M.J. Meyer, L. Staveley-Smith and R.L. Webster, *The HIPASS catalogue: Omega(HI) and environmental effects on the HI mass function of galaxies*, *Mon. Not. Roy. Astron. Soc.* **359** (2005) L30 [astro-ph/0502257] [iNSPIRE].

[60] A.M. Martin, E. Papastergis, R. Giovanelli, M.P. Haynes, C.M. Springob and S. Stierwalt, *The Arecibo Legacy Fast ALFA survey: X. The HI mass function and Omega_HI from the 40% ALFALFA survey*, *Astrophys. J.* **723** (2010) 1359 [arXiv:1008.5107] [iNSPIRE].

[61] W. Freudling et al., *Deep 21-cm HI Observations at $z \sim 0.1$: the precursor to the Arecibo ultra deep survey*, *Astrophys. J.* **727** (2011) 40 [arXiv:1011.0877] [iNSPIRE].

[62] P. Lah et al., *The HI content of star-forming galaxies at $z = 0.24$*, *Mon. Not. Roy. Astron. Soc.* **376** (2007) 1357 [astro-ph/0701668] [iNSPIRE].

[63] S.M. Rao, D.A. Turnshek and D.B. Nestor, *Damped lyman alpha systems at $z < 1.65$: the expanded sdss HST sample*, *Astrophys. J.* **636** (2006) 610 [astro-ph/0509469] [iNSPIRE].

[64] T.-C. Chang, U.-L. Pen, K. Bandura and J.B. Peterson, *Hydrogen 21-cm intensity mapping at redshift 0.8*, *Nature* **466** (2010) 463 [arXiv:1007.3709] [iNSPIRE].

[65] J.X. Prochaska and A.M. Wolfe, *On the (Non)Evolution of HI disks over cosmic time*, *Astrophys. J.* **696** (2009) 1543 [arXiv:0811.2003] [iNSPIRE].

[66] A. Slosar, C. Hirata, U. Seljak, S. Ho and N. Padmanabhan, *Constraints on local primordial non-Gaussianity from large scale structure*, *JCAP* **08** (2008) 031 [arXiv:0805.3580] [iNSPIRE].

[67] J. Fonseca, M. Silva, M.G. Santos and A. Cooray, *Cosmology with intensity mapping techniques using atomic and molecular lines*, *Mon. Not. Roy. Astron. Soc.* **464** (2017) 1948 [arXiv:1607.05288] [iNSPIRE].

[68] A. Moradinezhad Dizgah and G.K. Keating, *Line intensity mapping with [CII] and CO(1-0) as probes of primordial non-Gaussianity*, *Astrophys. J.* **872** (2019) 126 [arXiv:1810.02850] [iNSPIRE].

[69] A. Moradinezhad Dizgah, F. Nikakhtar, G.K. Keating and E. Castorina, *Precision tests of CO and [CII] power spectra models against simulated intensity maps*, *JCAP* **02** (2022) 026 [arXiv:2111.03717] [iNSPIRE].

[70] J.L. Tinker et al., *Toward a halo mass function for precision cosmology: The Limits of universality*, *Astrophys. J.* **688** (2008) 709 [arXiv:0803.2706] [iNSPIRE].

[71] J.L. Tinker et al., *The large scale bias of dark matter halos: numerical calibration and model tests*, *Astrophys. J.* **724** (2010) 878 [arXiv:1001.3162] [iNSPIRE].

[72] T. Lazeyras, C. Wagner, T. Baldauf and F. Schmidt, *Precision measurement of the local bias of dark matter halos*, *JCAP* **02** (2016) 018 [arXiv:1511.01096] [iNSPIRE].