Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **456,** 3553–3570 (2016) doi:10.1093/mnras/stv2904

Neutral hydrogen in galaxy clusters: impact of AGN feedback and implications for intensity mapping

Francisco Villaescusa-Navarro,[1,2]★ Susana Planelles,[3,4] Stefano Borgani,[1,2,4] Matteo Viel,[1,2] Elena Rasia,[1,5] Giuseppe Murante,[1] Klaus Dolag,[6,7] Lisa K. Steinborn,[6] Veronica Biffi,[1,4] Alexander M. Beck[6] and Cinthia Ragone-Figueroa[8]

[1]*INAF, Osservatorio Astronomico di Trieste, via Tiepolo 11, I-34131 Trieste, Italy*
[2]*INFN – National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy*
[3]*Departament d'Astronomia i Astrofisica, Universitat de Valencia, c/ Dr. Moliner 50, E-46100 Burjassot, Valencia, Spain*
[4]*Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy*
[5]*Department of Physics, University of Michigan, 450 Church St., Ann Arbor, MI 48109, USA*
[6]*Universitäts-Sternwarte München, Scheinerstr. 1, D-81679 München, Germany*
[7]*Max-Plank-Institut für Astrophysik, Karl-Schwarzschild Strasse 1, D-85740 Garching, Germany*
[8]*Instituto de Astronomia Teorica y Experimental (IATE), Consejo Nacional de Investigaciones Cientificas y Tecnicas de la Republica Argentina (CONICET), Observatorio Astronomico, Universidad Nacional de Cordoba, Laprida 854, X5000BGR Cordoba, Argentina*

Accepted 2015 December 8. Received 2015 December 8; in original form 2015 October 19

ABSTRACT

By means of zoom-in hydrodynamic simulations, we quantify the amount of neutral hydrogen (H I) hosted by groups and clusters of galaxies. Our simulations, which are based on an improved formulation of smoothed particle hydrodynamics, include radiative cooling, star formation, metal enrichment and supernova feedback, and can be split into two different groups, depending on whether feedback from active galactic nuclei (AGN) is turned on or off. Simulations are analysed to account for H I self-shielding and the presence of molecular hydrogen. We find that the mass in neutral hydrogen of dark matter haloes monotonically increases with the halo mass and can be well described by a power law of the form $M_{\rm H\,I}(M, z) \propto M^{3/4}$. Our results point out that AGN feedback reduces both the total halo mass and its H I mass, although it is more efficient in removing H I. We conclude that AGN feedback reduces the neutral hydrogen mass of a given halo by ∼50 per cent, with a weak dependence on halo mass and redshift. The spatial distribution of neutral hydrogen within haloes is also affected by AGN feedback, whose effect is to decrease the fraction of H I that resides in the halo inner regions. By extrapolating our results to haloes not resolved in our simulations, we derive astrophysical implications from the measurements of $\Omega_{\rm H\,I}(z)$: haloes with circular velocities larger than ∼25 km s^{-1} are needed to host H I in order to reproduce observations. We find that only the model with AGN feedback is capable of reproducing the value of $\Omega_{\rm H\,I} b_{\rm H\,I}$ derived from available 21 cm intensity mapping observations.

Key words: methods: numerical – galaxies: clusters: general – cosmology: miscellaneous.

1 INTRODUCTION

The formation and evolution of galaxies is an extremely complicated process that we do not fully understand yet. In the standard picture, gas falls within the gravitational potential wells of dark matter (DM) haloes, where it cools down and eventually forms stars. Galaxies grow mainly by accreting gas from the intergalactic medium (IGM). That gas, which is mainly ionized, becomes neutral once its density is high enough to be self-shielded against the exterior radiation. Neutral hydrogen (H I) is one of the major gas components in galaxies, and it has been shown that strong correlations show up among the H I content and stellar mass (Catinella et al. 2010; Cortese et al. 2011; Huang et al. 2012). This is surprising because stars form from the collapse and fragmentation of molecular hydrogen (H_2) clouds, which are created, under proper conditions, from H I and, therefore, the presence of neutral hydrogen does not imply that star formation is taking place.

★E-mail: villaescusa@oats.inaf.it

From observations we know that galaxies can be classified into two different groups. One group is dominated by blue star-forming galaxies, which are rich in cold gas (late-type) whereas the other group (early-type) contains red passive galaxies which host, on average, low gas fraction. The physical mechanisms responsible for this bimodality are not yet fully understood. Among the different mechanisms that can quench star formation there are mergers (Toomre & Toomre 1972), gas stripping by ram pressure (Gunn & Gott 1972; Yoon & Rosenberg 2014) and active galactic nucleus (AGN) feedback (Ho, Darling & Greene 2008; Hughes & Cortese 2009; Leslie et al. 2016). The relative importance of the different processes is also not well known.

In order to improve our understanding of the physical processes responsible for galaxy formation and evolution, it is thus very important to perform observations which not only tell us about the star formation rate and the stellar mass, but also which inform us on the mass in neutral and molecular hydrogen, being those the main constituents of the interstellar medium. One of the purposes of this paper is to investigate the impact of AGN feedback on the H I content in groups and clusters of galaxies.

Besides using H I to improve our understanding of galaxy formation and evolution, neutral hydrogen plays a key role in cosmology, since it can be used as a tracer of the large-scale structure of the Universe. Therefore, it is very important to properly model the spatial distribution of neutral hydrogen, in the post-reionization epoch, as it will be sampled, either by H I-selected galaxies (Santos et al. 2015a; Yahya et al. 2015) or via intensity mapping (Battye, Davies & Weller 2004; Chang et al. 2010; Masui et al. 2013; Switzer et al. 2013; Santos et al. 2015b), by radio telescopes as the Giant Metrewave Radio Telescope (GMRT),[1] the Ooty Radio Telescope (ORT),[2] the Low-Frequency Array (LOFAR),[3] the Murchison Widefield Array (MWA),[4] the Canadian Hydrogen Intensity Mapping Experiment (CHIME),[5] the Five hundred metre Aperture Spherical Telescope (FAST),[6] ASKAP (The Australian Square Kilometre Array Pathfinder),[7] MeerKAT (The South African Square Kilometer Array Pathfinder)[8] and the future SKA (The Square Kilometre Array).[9]

Intensity mapping is a new technique to sample the large-scale structure of the Universe which consists in performing a low angular resolution survey where the 21 cm emission from individual galaxies is not resolved. By using this technique, the radio telescopes will just measure the integrated emission from neutral hydrogen from many unresolved galaxies. The underlying idea is that on large scales, fluctuations in the integrated 21 cm signal will follow that of the underlying matter (Bharadwaj & Sethi 2001; Bharadwaj, Nath & Sethi 2001; Battye et al. 2004; McQuinn et al. 2006; Chang et al. 2008; Loeb & Wyithe 2008; Bull et al. 2015; Santos et al. 2015b; Villaescusa-Navarro, Bull & Viel 2015a). Thus, intensity mapping represents a different way to sample the large-scale structure of the Universe and it is expected to revolutionize cosmology given the very large volumes it can sample and the spectroscopic nature of the measurements (Alonso et al. 2015; Bull et al. 2015; Santos et al. 2015b; Villaescusa-Navarro et al. 2015a).

The function $M_{\rm H\,I}(M, z)$, which represents the average neutral hydrogen mass hosted by a DM halo of mass M at redshift z, plays a key role in 21 cm cosmology, since the shape and amplitude of the 21 cm power spectrum, on large, linear scales, are completely determined by that function. We now briefly explain the reason of this.

On large scales, the amplitude of the 21 cm power spectrum can be written as (see Villaescusa-Navarro et al. 2015a for a detailed discussion)

$$P_{21\,\rm cm}(k, z) = \overline{\delta T_{\rm b}}^2(z) b_{\rm H\,I}^2(z) \left(1 + \frac{2}{3}\beta(z) + \frac{1}{5}\beta^2(z)\right) \times P_{\rm m}(k, z), \quad (1)$$

where $\beta(z) = f(z)/b_{\rm H\,I}(z)$ is the redshift-space distortion parameter and the third factor on the right-hand side of the above equation arises from the Kaiser formula (Kaiser 1987). $b_{\rm H\,I}(z)$ represents the bias of the neutral hydrogen, $P_{\rm m}(k, z)$ denotes the linear matter power spectrum and the mean brightness temperature, $\overline{\delta T_{\rm b}}(z)$, is given by

$$\overline{\delta T_{\rm b}}(z) = 189 \left(\frac{H_0(1+z)^2}{H(z)}\right) \Omega_{\rm H\,I}(z) h \ {\rm mK}, \quad (2)$$

where $H(z)$ and H_0 are the value of the Hubble parameter at redshifts z and 0, respectively. h represents the value of H_0 in units of $100\,{\rm km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_{\rm H\,I}(z)$ is the ratio of the comoving neutral hydrogen density at redshift z to the Universe critical density at $z = 0$, $\rho_{\rm c}^0$.

Therefore, for a given cosmological model, the amplitude of the 21 cm signal depends on the amount of neutral hydrogen, via $\Omega_{\rm H\,I}(z)$, but also on the way the H I is distributed among different haloes, i.e. on the neutral hydrogen bias, $b_{\rm H\,I}(z)$. Contrary to what observers usually measure (the H I mass within galaxies), the important quantity for cosmology is the function $M_{\rm H\,I}(M, z)$, which gives the average H I mass hosted by a halo of mass M at redshift z. Given that function, it is straightforward to compute both $\Omega_{\rm H\,I}(z)$ and $b_{\rm H\,I}(z)$ as

$$\Omega_{\rm H\,I}(z) = \frac{1}{\rho_{\rm c}^0} \int_0^\infty n(M, z) M_{\rm H\,I}(M, z)\, {\rm d}M, \quad (3)$$

$$b_{\rm H\,I}(z) = \frac{\int_0^\infty n(M, z) b(M, z) M_{\rm H\,I}(M, z)\, {\rm d}M}{\int_0^\infty n(M, z) M_{\rm H\,I}(M, z)\, {\rm d}M}, \quad (4)$$

where $n(M, z)$ and $b(M, z)$ are the halo mass function and halo bias, respectively. In other words, for a given cosmological model, the function $M_{\rm H\,I}(M, z)$ completely determines the shape and amplitude of the 21 cm power spectrum on large scales.

The H I mass residing in galaxy clusters plays a fundamental role in determining the value of the H I bias (see Yoon & Rosenberg 2014, for an observational study on the cold atomic gas content of galaxies in groups and clusters). The reason is that clusters of galaxies are very biased objects, and therefore, if a significant amount of neutral hydrogen is found within them, they can substantially enhance the value of $b_{\rm H\,I}(z)$. This can be rephrased saying that the high-mass end of the function $M_{\rm H\,I}(M, z)$ is very important since the H I bias strongly depends on it.

[1] http://gmrt.ncra.tifr.res.in/
[2] http://rac.ncra.tifr.res.in/
[3] http://www.lofar.org/
[4] http://www.mwatelescope.org/
[5] http://chime.phas.ubc.ca/
[6] http://fast.bao.ac.cn/en/
[7] http://www.atnf.csiro.au/projects/askap/index.html
[8] http://www.ska.ac.za/meerkat/
[9] https://www.skatelescope.org/

At $z \simeq 0$, the bias of H I-selected galaxies[10] have been measured from the H I Parkes All-Sky Survey (Barnes et al. 2001) obtaining a value of $b_{\rm H\,I,gal} = 0.7 \pm 0.1$, in good agreement with the one measured from the Arecibo Legacy Fast ALFA Survey (Giovanelli et al. 2005; Martin et al. 2012, see however Marín et al. 2010 for analytical estimates of the H I bias at $z = 0$ and its redshift evolution). At $\langle z \rangle = 0.8$, 21 cm intensity mapping observations with the Green Bank Telescope have been used to measure the product $b_{\rm H\,I}(z) \times \Omega_{\rm H\,I}(z)$, finding a value of $6.2^{+2.3}_{-1.5} \times 10^{-4}$ (Chang et al. 2010; Masui et al. 2013; Switzer et al. 2013). Finally, at $\langle z \rangle = 2.3$ the bias of damped Lyman α absorbers (DLAs) has been recently estimated in Font-Ribera et al. (2012) to be $b_{\rm DLAs} = (2.17 \pm 0.20)\beta_{\rm F}^{0.22}$, where $\beta_{\rm F}$ is the Lyα forest redshift distortion parameter whose value is of order 1.

Thus, it is very important to model as best as possible the function $M_{\rm H\,I}(M, z)$ since the signal-to-noise ratio of the 21 cm signal depends on it. On the other hand, the function $M_{\rm H\,I}(M, z)$ contains important astrophysical information, giving us information on how H I is distributed among the different haloes and it can be constrained by combining observations from the Lyα forest, 21 cm intensity mapping and so on.

Theoretical models like the one presented in Bagla, Khandai & Datta (2010) propose a phenomenological function for $M_{\rm H\,I}(M, z)$ making the hypothesis that galaxy clusters do not host a significant amount of neutral hydrogen. This hypothesis relies on the fact that observations point out that galaxies in clusters are H I deficient (Solanes et al. 2001; Gavazzi et al. 2005, 2006; Taylor et al. 2012, 2013; Catinella et al. 2013; Dénes, Kilborn & Koribalski 2014). In Villaescusa-Navarro et al. (2014), it was shown that this simple model is capable of reproducing extremely well the abundance of DLAs at redshifts $z = [2.4–4]$, while in Padmanabhan, Choudhury & Refregier (2015) authors claim that the model successfully reproduces the bias of the H I-selected galaxies at $z \simeq 0$. In this paper, we check the validity of the assumption behind the Bagla et al. (2010) model.

The purpose of this paper is to investigate, using state-of-the-art zoom-in hydrodynamic simulations, the amount of neutral hydrogen hosted by groups and clusters of galaxies (i.e. the high-mass end of the $M_{\rm H\,I}(M, z)$ function), its spatial distribution within haloes and its evolution with time. We also test the validity of the Bagla et al. (2010) assumptions by comparing the results of our simulations against the prediction of that model. Moreover, we study the impact of AGN feedback on the neutral hydrogen content of clusters and groups and investigate the implications for 21 cm intensity mapping.

This paper is organized as follows. In Section 2.1, we describe the hydrodynamic simulations used for this work, together with the method used to model the spatial distribution of neutral hydrogen. We investigate in Section 3 the mass in neutral hydrogen hosted by groups and clusters, its spatial distribution within haloes, its redshift evolution and the impact of AGN feedback. In Section 4, we compare our findings against the Bagla et al. (2010) model and study the consequences for 21 cm intensity mapping. We draw the main conclusions and discuss the results of this paper in Section 5. In Appendix A, we discuss the distribution of neutral hydrogen among galaxies belonging to groups and clusters and the possible level of numerical contamination.

[10] Notice that this bias is different from the one defined in equation (4). The bias of H I-selected galaxies is defined as the value of $b^2_{\rm H\,I,gal}(r) = \xi_{\rm H\,I,gal}(r)/\xi_{\rm m}(r)$ on large scales, where $\xi_{\rm H\,I,gal}(r)$ and $\xi_{\rm m}$ represent the two-point correlation function of the H I-selected galaxies and the underlying matter, respectively.

2 SIMULATIONS

In this section, we describe the set of zoom-in hydrodynamic simulations we have used for this work. We then depict the method we employ to model the spatial distribution of neutral hydrogen and the procedure utilized to identify DM haloes and galaxies.

2.1 Hydrodynamic simulations

Here, we provide a short characterization of the set of simulations used in this paper. While we refer to a future work (Planelles et al., in preparation) for a more detailed description, a first analysis of these set of simulations can be found in the recently submitted paper by Rasia et al. (2015). The set of hydrodynamic simulations consists of resimulations of 29 Lagrangian regions centred around the 29 most massive haloes formed in a larger N-body cosmological simulation (see Bonafede et al. 2011, for details on the initial conditions). The simulations, performed with the TreePM–SPH (smoothed particle hydrodynamics) GADGET-3 code (Springel 2005), assume a flat Λ cold dark matter (ΛCDM) cosmology with $\Omega_{\rm m} = 0.24$, $\Omega_{\rm b} = 0.04$, $H_0 = 72$ km s^{-1} Mpc^{-1}, $n_{\rm s} = 0.96$, $\sigma_8 = 0.8$.

The mass resolution for the DM particles and the initial mass of the gas particles are, respectively, $m_{\rm DM} = 8.44 \times 10^8\, h^{-1}\, {\rm M}_\odot$ and $m_{\rm gas} = 1.56 \times 10^8\, h^{-1}\, {\rm M}_\odot$. As for the spatial resolution, gravitational force in the resimulated regions is computed with a Plummer-equivalent softening length of $\epsilon = 3.75\, h^{-1}$ kpc (in physical units at $z < 2$, while fixed in comoving units at $z > 2$).

An improved version of the standard SPH scheme, as discussed in Beck et al. (2016), has been included. This new hydro scheme includes a number of elements (such as an artificial conduction term, a time-dependent artificial viscosity and a Wendland C^4 interpolation kernel) which largely improves the performance of the traditional SPH scheme. Including this new hydro scheme, different sets of simulations, characterized by the inclusion of different sets of baryonic physical processes, have been performed. In this work, our reference simulations, labelled as the AGN simulations, include the effects of the AGN feedback model recently presented in Steinborn et al. (2015). In this model, accretion on to supermassive black holes takes place according to the Bondi formula and is Eddington-limited, but in contrast to the original implementation by Springel, Di Matteo & Hernquist (2005), we compute the accretion rate separately for cold and hot gas using different boost factors (typically, 100 for cold gas and 10 for hot gas accretion). In the particular set of simulations presented in this paper, as in Rasia et al. (2015), we neglect any contribution from hot gas accretion. Furthermore, both mechanical outflows and radiation contribute to the thermal energy. Using variable efficiencies for these two components allows a continuous transition between the quasar and the radio mode. For the radiative feedback, we fix the coupling factor to the surrounding gas to $\epsilon_f = 0.05$. We refer to Steinborn et al. (2015) for a more detailed description of this model.

Besides AGN feedback, our reference model also accounts for the effects of a number of additional processes such as metallicity-dependent radiative cooling, star formation, supernova (SN) feedback and metal enrichment. These processes have been included as described in Planelles et al. (2014). Briefly, radiative cooling and the presence of the UV/X-ray background radiation are included according to, respectively, Wiersma, Schaye & Smith (2009) and Haardt & Madau (2001). The sub-grid model for star formation and

its associated feedback is implemented according to the prescription by Springel & Hernquist (2003). Galactic winds with a velocity of $\sim$350 km s^{-1} characterize the kinetic feedback from SNe. Metal enrichment is also included according to the chemical model by Tornatore et al. (2007). We refer to this simulation set as AGN run.

We emphasize that our AGN reference model has been shown to agree with a number of cluster observations and, therefore, to provide a realistic population of clusters. In particular, in the recent work by Rasia et al. (2015), it has been shown how this set of simulations produce, in a natural way, the coexistence of cool-core and non-cool-core clusters, with entropy and iron abundance profiles in good agreement with observational profiles. A similar good match with observed data is also obtained for other cluster properties such as pressure profiles, gas and baryon mass fractions or X-ray and Sunyaev-Zel'dovich (SZ) scaling relations (see Planelles et al., in preparation and Truong et al., in preparation). Our reference simulation appears to also reproduce the observed scaling between stellar mass of the host galaxy and black hole mass, producing as well a good estimate of brightest galaxy cluster masses.

For completeness, besides our reference AGN simulations, we will also analyse another set of radiative simulations, labelled as CSF, which include the same physical processes as our reference model but for which we have turned off AGN feedback. We use these simulations to investigate the impact of AGN feedback by comparing the results from this set against the one from the AGN simulations.

2.2 Neutral hydrogen distribution

We need to post-process the output of the simulations to account for two critical processes, which are not followed by our simulations, to properly model the distribution of neutral hydrogen: the H I self-shielding and the formation of molecular hydrogen, H_2.

The procedure used to model the spatial distribution of neutral hydrogen is as follows. First of all, the hydrogen mass that is neutral is computed, for all gas particles in the simulation, assuming ionization equilibrium and estimating the photoionization rate using the fitting formula of Rahmati et al. (2013a). Then, we correct the H I mass obtained above by modelling the presence of molecular hydrogen using the model developed in Krumholz, McKee & Tumlinson (2008, 2009) and McKee & Krumholz (2010), the so-called KMT model, assuming that only star-forming particles host H_2.

The H I self-shielding is taken into account by using the fitting formula of Rahmati et al. (2013a), which we use to compute the fraction of hydrogen that is in neutral state (i.e. both neutral hydrogen, H I, and molecular hydrogen, H_2). We now briefly describe the method used here and refer the reader to Rahmati et al. (2013a) for further details. The amount of hydrogen that is neutral is computed, for each gas particle, assuming ionization equilibrium (see for instance appendix A2 of Rahmati et al. 2013a) where the photoionization rate, $\Gamma_{\rm Phot}$, affecting a given gas particle is a function of its hydrogen number density, $n_{\rm H}$,

$$\Gamma_{\rm Phot} = \Gamma_{\rm UVB}\left\{(1-f)\left[1+\left(\frac{n_{\rm H}}{n_0}\right)^{\beta}\right]^{\alpha_1} + f\left[1+\frac{n_{\rm H}}{n_0}\right]^{\alpha_2}\right\}, \tag{5}$$

where $\Gamma_{\rm UVB}$ is the UV background photoionization rate and n_0, α_1, α_2, β and f are free parameters of the fitting formula whose values we take from table A1 of Rahmati et al. (2013a). We interpolate to obtain the value of the above parameters for redshifts not covered by their table A1. We notice that we are neglecting radiation from local sources (Miralda-Escudé 2005; Schaye 2006; Rahmati et al. 2013b) and local X-ray radiation from the hot intracluster medium (ICM) (Kannan et al. 2015) when computing the mass in neutral hydrogen. In Section 5, we discuss the possible consequences of this assumption in our results.

The presence of molecular hydrogen, H_2, is modelled using the KMT analytic model. The molecular hydrogen fraction, $f_{\rm H_2}$, defined as $f_{\rm H_2} = M_{\rm H_2}/M_{\rm NH}$, where $M_{\rm H_2}$ is the mass in molecular hydrogen and $M_{\rm NH} = M_{\rm H\,I} + M_{\rm H_2}$ is the mass in hydrogen which is neutral (computed as explained in the previous paragraph), is estimated using

$$f_{\rm H_2} = \begin{cases} 1 - \frac{0.75s}{1+0.25s} & \text{if } s < 2 \\ 0 & \text{if } s \geqslant 2 \end{cases} \tag{6}$$

where

$$s = \frac{\ln(1 + 0.6\chi + 0.01\chi^2)}{0.6\tau_{\rm c}} \tag{7}$$

with

$$\chi = 0.756(1 + 3.1Z^{0.365}) \tag{8}$$

$$\tau_{\rm c} = \Sigma\sigma_{\rm d}/\mu_{\rm H} \tag{9}$$

where Z is the metallicity of the gas particle, self-consistently described in our simulations, in units of the solar metallicity (Allende Prieto, Lambert & Asplund 2001), $\mu_{\rm H}$ is the mean mass per hydrogen nucleus ($\mu_{\rm H} = 2.3 \times 10^{-24}$ g), $\sigma_{\rm d}$ is the dust cross-section (we take $\sigma_{\rm d} = Z \times 10^{-21}$ cm^2) and Σ is the gas surface density.[11] We assume that the extent and profile of the metallicity are governed by the particle SPH radius and kernel, respectively. Notice that we assume that only star-forming particles (i.e. particles with physical densities higher than $\sim$0.1 H/cm^{-3}) host H_2; thus, the H I mass computed above only needs to be corrected in those particles.

2.3 Identification of DM haloes and galaxies

DM haloes are initially found by applying the Friends-of-Friends algorithm (Davis et al. 1985) with a value of the linking length parameter $b = 0.16$ on top of the simulation snapshots. Next, the centre of a given halo is found by searching the particle with the minimum value of the gravitational potential. Then, the halo radius, R_{200}, and its mass, M, are computed by requiring that the mean density of a sphere centred on the halo centre and with that radius is equal to 200 times the critical density of the Universe at that redshift, $\rho_{\rm c}(z)$, i.e.

$$M = \frac{4\pi}{3}\rho_{\rm c}(z)\triangle_{\rm c}(z)R_{200}^3, \tag{10}$$

where $\triangle_{\rm c}(z) = 200$. The above procedure is also used to find $\{M_{500}, R_{500}\}$ and $\{R_{2500}, M_{2500}\}$, which are defined as above but setting $\triangle_{\rm c}(z) = 500$ and $\triangle_{\rm c}(z) = 2500$, respectively. Finally, galaxies are identified by means of the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009). At $z = 0$, the catalogues of each, the AGN and CSF, simulation contain approximately 450 groups and clusters.

[11] We follow Davé et al. (2013) and compute the surface density of a given gas particle by multiplying its density by its SPH radius.

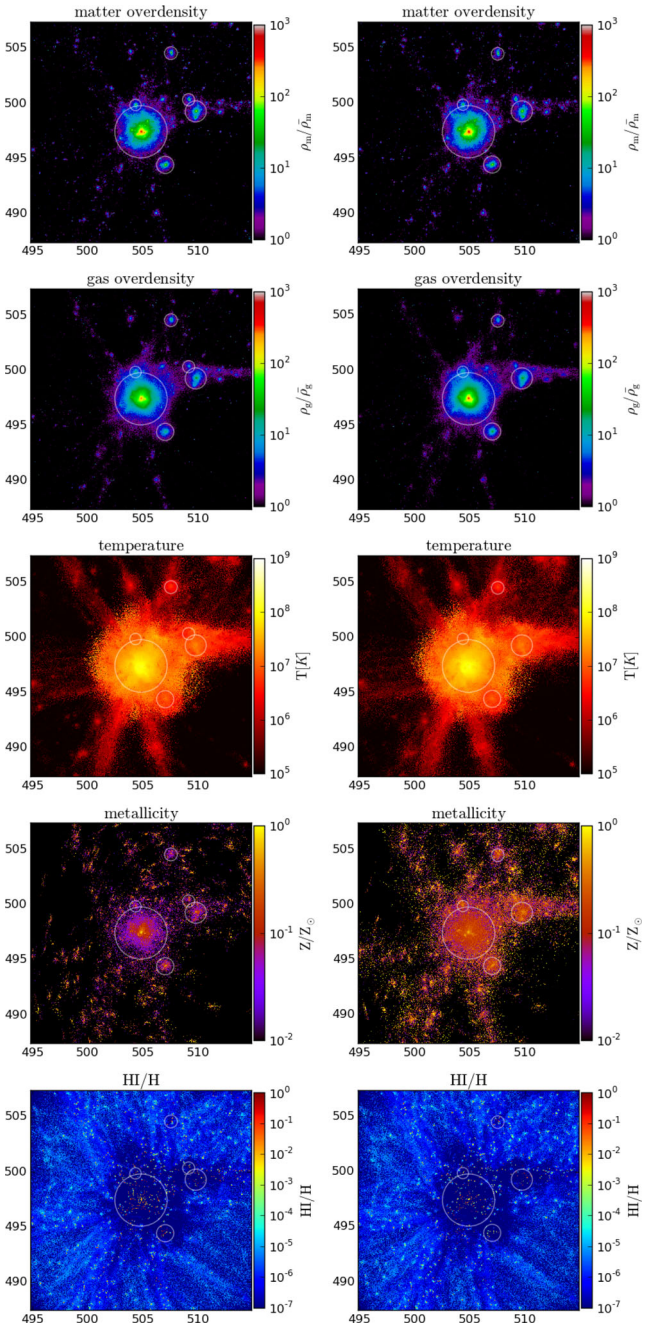

Figure 1. Spatial distribution of matter overdensity (first row), gas overdensity (second row), temperature (third row), metallicity (fourth row) and neutral hydrogen fraction (fifth row) from a slice of 10 h^{-1} Mpc width centred in the most massive halo of one of our simulated regions at $z = 0$. White circles represent the location and radii of the haloes found in that slice. Left- and right-hand panels show the results for the simulations CSF and AGN, respectively. The units of the X- and Y-axes are h^{-1} Mpc.

3 NEUTRAL HYDROGEN IN GROUPS AND CLUSTERS: IMPACT OF AGN FEEDBACK

In this section, we present the main results that we obtain by analysing the hydrodynamic simulations. In order to have a preliminary visual impression of the results, we show in Fig. 1 the spatial distribution of matter, gas, temperature, metallicity and neutral hydrogen fraction around a massive cluster at $z = 0$ in one of our resimulated regions. The left and right columns display the results for the simulations without and with AGN feedback, respectively.

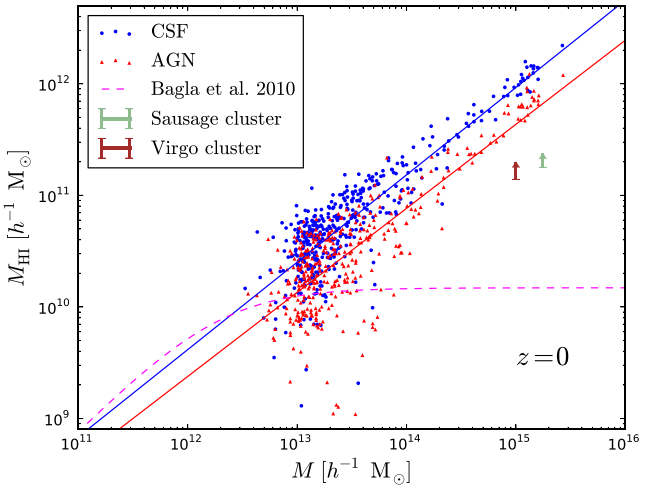

Figure 2. Neutral hydrogen mass within R_{200} as a function of the halo mass, M, for the haloes of the simulations CSF (blue dots) and AGN (red triangles) at $z = 0$. The solid lines show a fit to the results using the function $M_{\rm H\,I}(M, z) = {\rm e}^{\gamma} M^{\alpha}$. The dashed magenta line displays the prediction of the Bagla et al. (2010) model. The green and brown arrows represent a lower limit on the neutral hydrogen mass hosted by the Sausage and Virgo clusters, respectively (see the text for details).

In Section 3.1 we investigate the mass and spatial distribution of H I within groups and clusters of galaxies. We investigate the impact of AGN feedback on the amount and distribution of H I in Section 3.2. In Section 3.3 we study the evolution with time of our results. Finally, we refer the reader to Appendix A for the results on the neutral hydrogen content in galaxies belonging to groups and clusters of galaxies. In that appendix, we also quantify the level of contamination of our results from spurious numerical artefacts in our simulations.

3.1 Neutral hydrogen in groups and clusters at $z = 0$

For each halo of the simulations CSF and AGN, we have computed the total neutral hydrogen mass within R_{200} and in Fig. 2 we show the H I mass versus the total halo mass, M_{200}, at $z = 0$.

We find that, on average, the bigger the halo, the larger the H I mass it hosts. This behaviour takes place independently of whether AGN feedback is switched on or off in the simulations. We find that a simple power law of the form[12] $M_{\rm H\,I}(M, z) = {\rm e}^{\gamma} M^{\alpha}$ can reproduce the mean of our results very well. In Table 1, we show the best-fitting values of the parameters α and γ for the simulations CSF and AGN at different redshifts. The solid lines in Fig. 2 display those fits at $z = 0$. The physical interpretation of the parameters α and γ is straightforward. γ represents an overall normalization: the H I mass of a halo of mass $1\,h^{-1}\,{\rm M}_\odot$ is given by ${\rm e}^{\gamma}\,h^{-1}\,{\rm M}_\odot$. On the other hand, α characterizes the slope of the $M_{\rm H\,I}$–M relation.

The green and brown arrows in Fig. 2 represent a lower limit, although close to the actual value, on the overall H I mass of the Sausage and Virgo clusters estimated from Stroe et al. (2015b) and Gavazzi et al. (2005), respectively.[13] It can be seen that these lower

[12] Notice that in order to give the correct units to all the terms, this relation should be understood as $M_{\rm H\,I}(M, z)/(h^{-1}\,{\rm M}_\odot) = {\rm e}^{\gamma}(M/(h^{-1}\,{\rm M}_\odot))^{\alpha}$.
[13] For the Sausage cluster, we consider that the approximately 100 Hα star-forming galaxies detected (Stroe et al. 2015a) host an average H I mass equal

Table 1. Best-fitting parameters for $M_{\rm H\,I}(M,z) = {\rm e}^{\gamma} M^{\alpha}$ for the two different simulations and for different redshifts. The quoted values of α and γ hold when the masses of both $M_{\rm H\,I}$ and M are in units of $h^{-1}\,{\rm M}_{\odot}$.

Simulation	z	α	γ
CSF	0	0.78 ± 0.02	0.7 ± 0.7
	0.25	0.76 ± 0.02	1.5 ± 0.6
	0.5	0.75 ± 0.02	2.1 ± 0.6
	0.8	0.73 ± 0.02	2.5 ± 0.6
	1	0.71 ± 0.03	3.2 ± 0.8
	1.5	0.69 ± 0.03	4.1 ± 0.8
	2	0.59 ± 0.03	7.3 ± 1.0
AGN	0	0.75 ± 0.02	0.9 ± 0.8
	0.25	0.77 ± 0.02	0.4 ± 0.8
	0.5	0.79 ± 0.02	0.1 ± 0.8
	0.8	0.75 ± 0.02	1.3 ± 0.6
	1	0.73 ± 0.03	2.0 ± 0.9
	1.5	0.73 ± 0.03	2.1 ± 1.0
	2	0.65 ± 0.05	4.7 ± 1.5
	0–1.5	0.75 ± 0.01	1.1 ± 0.4
	0–2	0.75 ± 0.01	1.3 ± 0.4

limits are a factor of ∼3–4 lower than the typical H I masses we find in the haloes of the AGN simulation with similar masses. We remark that the H I masses we measure in simulated galaxy clusters may be partially affected[14] by numerical artefacts that can spuriously increase the predicted mass in neutral hydrogen by a factor of 2 (see Appendix A). Besides, in our calculations we neglect the radiation from local sources and from the hot ICM.

One may wonder whether the H I masses we measure for groups of galaxies are biased since we are selecting those haloes from regions which contain very massive haloes. In order to answer this question, we have selected different regions in which the mass of the most massive halo is different and computed the H I mass within groups. We find that the mass in neutral hydrogen in groups does not depend on the mass of the most massive halo in the region. We thus conclude that our results are not biased by limiting our study to the highly biased regions we simulate in this work.

We have also studied the correlations between the overall gas, stars and H I mass in haloes. For each halo of the CSF and AGN simulations, we have computed the H I, gas and stellar mass within its R_{200} at $z = 0$. In Fig. 3, we show the results. We find that the amount of gas and the stellar mass within haloes increases with the H I mass, which, on the other hand, increases with the mass of the halo. This just reflects the fact that the larger the halo the more CDM, stars, gas and neutral hydrogen it contains.

We have also investigated the spatial distribution of neutral hydrogen within haloes. In each simulation suite, we have selected haloes of approximately the same mass, in order to isolate any dependence of the H I distribution on the halo mass. In particular, we have taken all haloes with masses in three different ranges: $M \in [2\text{–}5] \times 10^{13}\,h^{-1}\,{\rm M}_{\odot}$, $M \in [2\text{–}5] \times 10^{14}\,h^{-1}\,{\rm M}_{\odot}$ and $M > 10^{15}\,h^{-1}\,{\rm M}_{\odot}$. Next, for each halo in a given mass range, we have computed the H I mass within a radius equal to 1/4, 1/2, 3/4 and 5/4 times R_{200}. In Fig. 4, we show the average and dispersion around the mean of the H I mass within those radii normalized to the H I mass within R_{200} as a function of r/R_{200} for the three different mass ranges and for the two different simulations.

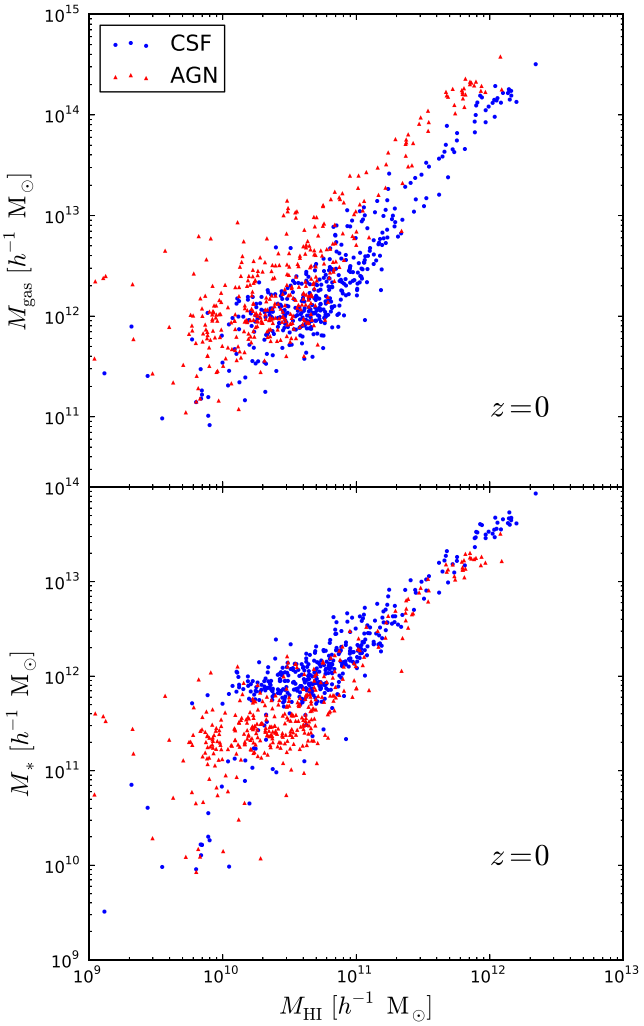

Figure 3. For each halo of the AGN (red) and CSF (blue) simulations at $z = 0$, we have computed the gas, stars and neutral hydrogen mass within R_{200}. The upper panel shows the gas mass versus the H I mass, while the bottom panel displays the stellar mass versus the H I mass.

By construction, the point at $r/R_{200} = 1$ has a value of $M_{\rm H\,I}(r|M,z)/M_{\rm H\,I}(R_{200}|M,z)$ equal[15] to 1 and therefore its dispersion is 0, independently of the simulation setup, redshift and mass range. We find that the results can be well fitted by a law of the form $M_{\rm H\,I}(r|M,z)/M_{\rm H\,I}(R_{200}|M,z) = (r/R_{200})^{\beta}$. In Table 2, we show the best-fitting value of β for the three different mass bins and the two simulations at different redshifts. The dashed lines of Fig. 4 display those fits at $z = 0$ (upper row), $z = 0.5$ (middle row) and $z = 1$ (bottom row). We notice that the fits only use the points with $r < R_{200}$, i.e. the data points with $r = 5R_{200}/4$ are not included in the fit.

Our results point out that for the simulation CSF the value of β, which measures the steepness of the function $M_{\rm H\,I}(r|M,z)$,

to $2.5 \times 10^{9}\ {\rm M}_{\odot}$ (Stroe et al. 2015b, Stroe, private communication). We estimate the H I mass of the Virgo cluster by summing the H I masses of the 296 detected galaxies from table A2 of Gavazzi et al. (2005).

[14] Notice that we estimate the numerical contamination for very massive galaxy clusters to be ∼30 per cent. Thus, it is unlikely that differences among our simulations and observations are only due to spurious H I blobs.

[15] The notation $M_{\rm H\,I}(r|M,z)$ denotes the H I mass within a radius r of a halo of mass M at redshift z.

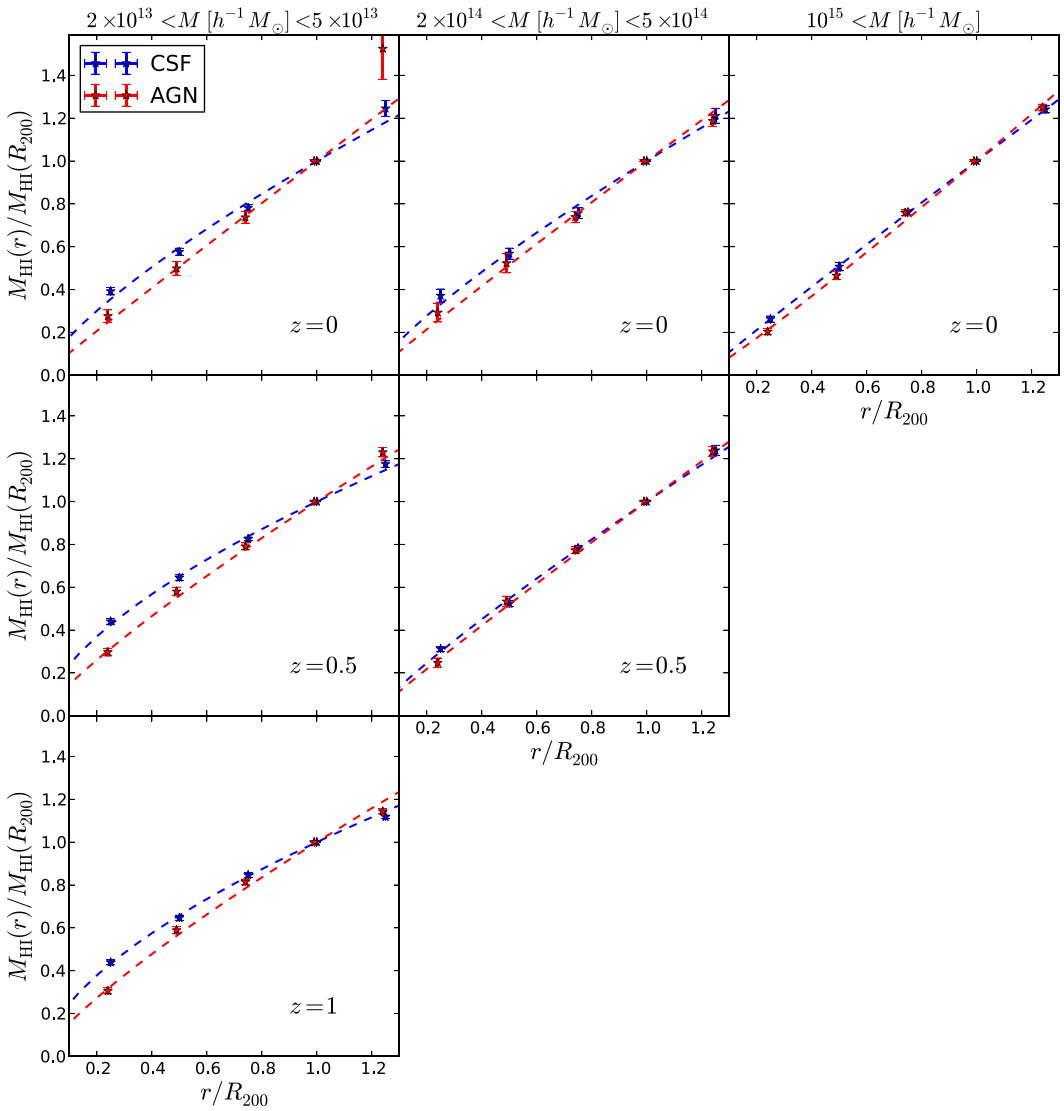

Figure 4. Neutral hydrogen mass within a radius r, normalized by the H I mass within R_{200}, as a function of r/R_{200} for haloes in the mass range $M \in [2\text{–}5] \times 10^{13}\,h^{-1}\,\mathrm{M}_\odot$ (left column), $M \in [2\text{–}5] \times 10^{14}\,h^{-1}\,\mathrm{M}_\odot$ (middle column) and $M > 10^{15}\,h^{-1}\,\mathrm{M}_\odot$ (right column) at $z = 0$ (upper row), $z = 0.5$ (middle row) and $z = 1$ (bottom row) for the haloes of the simulation CSF (blue) and AGN (red). For each halo, we have computed the H I mass within 0.25, 0.50, 0.75, 1.00 and 1.25 times R_{200}, and the points with the error bars display the average and dispersion around the mean of the results. For clearness, we have displaced the results of the AGN simulation by $\Delta(r/R_{200}) = -0.01$. Dashed lines show a fit to the results using the functional form $M_{\rm H\,I}(r)/M_{\rm H\,I}(R_{200}) = (r/R_{200})^{\beta}$. The best-fitting values of β are specified in Table 2.

Table 2. Best-fitting value of β ($M_{\rm H\,I}(r|M,z)/M_{\rm H\,I}(R_{200}|M,z) = (r/R_{200})^{\beta}$) for the two different simulations at redshifts $z = 0$, 0.5 and 1. Entries with '–' indicate that there are not enough haloes in that redshift and mass range to properly derive the value of β.

Simulation	z	Mass range $[h^{-1}\,\mathrm{M}_\odot]$ $[2\text{–}5] \times 10^{13}$	$[2\text{–}5] \times 10^{14}$	$>10^{15}$
CSF	0	0.75 ± 0.05	0.80 ± 0.06	0.966 ± 0.007
	0.5	0.62 ± 0.02	0.87 ± 0.03	–
	1	0.60 ± 0.01	–	–
AGN	0	0.98 ± 0.04	0.96 ± 0.05	1.09 ± 0.05
	0.5	0.83 ± 0.03	0.94 ± 0.04	–
	1	0.81 ± 0.04	–	–

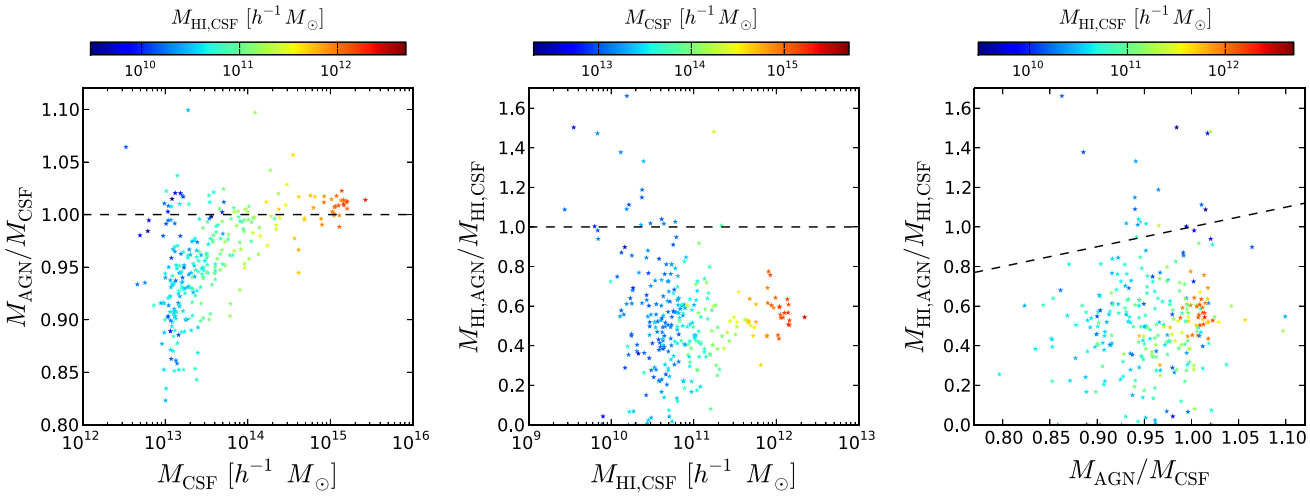

Figure 5. Impact of AGN feedback on the mass and H I content of groups and clusters at $z = 0$. We carry out a halo-by-halo comparison among the simulations CSF and AGN. Left (middle): ratio between the mass (H I mass) of corresponding haloes in the simulations AGN and CSF versus the mass (H I mass) of the CSF halo. The dashed horizontal line represents the case where the masses (H I masses) of the AGN and CSF haloes are the same. Right: ratio between the H I masses of corresponding haloes in the simulations AGN and CSF versus the ratio of their halo masses. The dashed line in that panel represents the curve $M_{\rm H\,I,AGN}/M_{\rm H\,I,CSF} = M_{\rm AGN}/M_{\rm CSF}$. The colour of each point indicates the H I mass of the CSF halo (left- and right-hand panels) and its total mass (middle panel).

increases with the halo mass, while for the simulation AGN the value of β is compatible, within 2σ, with 1 for all mass ranges.

3.2 Impact of AGN feedback

Now we study the impact that AGN feedback induces on the mass and spatial distribution of neutral hydrogen in galaxy groups and clusters. As expected, when AGN feedback is turned on, the mass in neutral hydrogen of a given halo decreases on average (see Fig. 2). This happens because AGN feedback injects energy to the gas, increasing its temperature and therefore avoiding the formation of neutral hydrogen. This can be seen more clearly in the upper panel of Fig. 3 where we show, for each halo in the CSF and AGN simulations, the mass in gas versus the mass in neutral hydrogen. It can be seen that the mass in gas does not change much by switching on or off AGN feedback (see the halo-by-halo comparison below), while haloes in the CSF simulations contain a larger H I mass because the gas in the appropriate conditions to host H I is hotter when AGN feedback is on. By fitting the results to a law of the form $M_{\rm H\,I}(M, z) = {\rm e}^{\gamma} M^{\alpha}$, we find that AGN feedback does not change much the slope of the function $M_{\rm H\,I}(M, z)$ ($\alpha = 0.78$ for CSF and $\alpha = 0.75$ for AGN) but its effect is mainly to shift the overall amplitude of that function ($\gamma = 0.7$ for CSF versus $\gamma = 0.9$ for AGN).

We notice that AGN feedback also affects the halo mass function (Cui, Borgani & Murante 2014; Velliscig et al. 2014; Khandai et al. 2015). In particular, it is expected that haloes in simulations with AGN feedback would be smaller and lighter than those found in simulations with no AGN feedback. Thus, one may wonder whether the H I mass deficit in haloes where AGN is on, with respect to the same haloes where AGN is off, is just due to the fact that those haloes are smaller, i.e. simply because there is less gas in haloes when AGN feedback is switched on. In order to answer this question, we have performed a halo-by-halo comparison between the simulations CSF and AGN. In order to match a halo in the simulation CSF with the corresponding one in the simulation AGN, we require the distance among their centres to be smaller than 15 per cent of the halo radius and, in any case, that the distance is lower than 150 comoving h^{-1} kpc.

In the left-hand panel of Fig. 5, we plot the ratio between the mass of a given halo in the AGN simulation and the mass of the same halo in the simulation CSF as a function of the CSF halo mass. Our results point out that, on average, AGN feedback reduces the total mass of galaxy groups by ∼5–10 per cent, while in clusters its effect is less important: haloes more massive than $\sim 10^{14.5}\,h^{-1}\,{\rm M}_\odot$ have almost the same total mass, independently of whether AGN feedback is switched on or off. The reason of this behaviour is that AGN feedback is more effective at removing baryons in low-mass haloes than in the most massive ones.

The middle panel of Fig. 5 displays the ratio between the neutral hydrogen mass of corresponding haloes of the simulations CSF and AGN as a function of the H I mass of the halo in the CSF simulation. The colour of each point represents the total mass of the CSF halo. We find that AGN feedback can dramatically change the neutral hydrogen content of groups and clusters of galaxies. While in terms of total mass, AGN feedback can decrease the mass content by up to 20 per cent, when focusing on the H I content, AGN feedback can suppress its abundance by more than 95 per cent. We note however that we have found a few haloes where the H I content increases when AGN feedback is on (see the points above the dashed horizontal line in the middle panel of Fig. 5).

Finally, in the right-hand panel of Fig. 5, we show the ratio between the total masses versus the ratio of the H I masses for corresponding haloes in the CSF and AGN simulations. As can be seen in that plot, the majority of points fall below the $y = x$ line, demonstrating that AGN feedback is more efficient at removing H I gas than matter. In other words, the H I deficit we find in haloes where AGN feedback is switched on, in comparison to the same haloes with no AGN feedback, does not arise because those haloes are smaller, but it is mainly due to effects induced by AGN feedback.

Regarding the impact that AGN feedback induces on the spatial distribution of neutral hydrogen within haloes, we can see from Fig. 4 that the steepness of the $M_{\rm H\,I}(r|M, z)/M_{\rm H\,I}(R_{200}|M, z)$ ratio is higher in the simulations where AGN feedback is on. This means

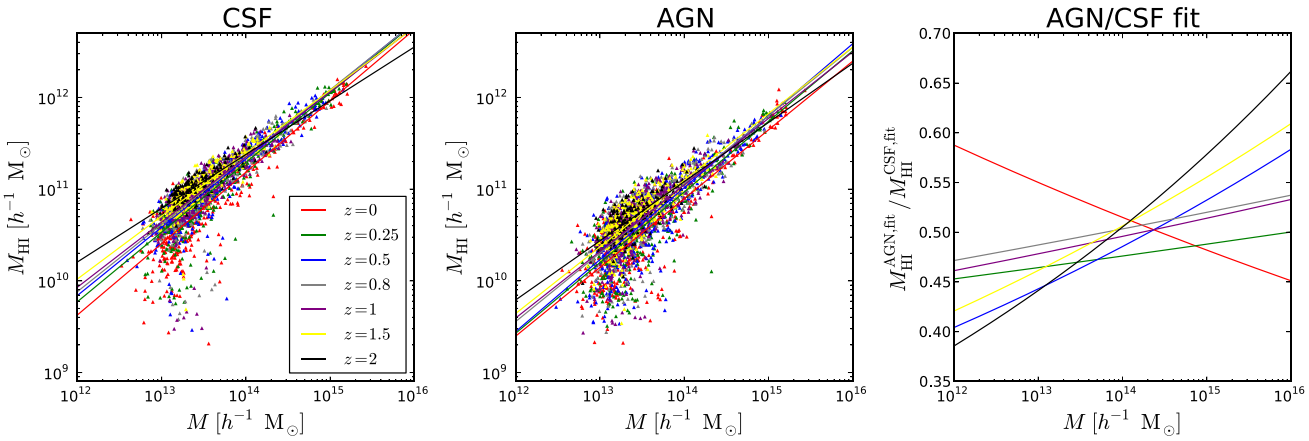

Figure 6. Redshift evolution of the $M_{\rm H\,I}(M, z)$ function for the simulation without and with AGN feedback, left-hand and middle panel, respectively. Each point represents the H I mass inside R_{200} of a halo of mass M, with the colour of the point indicating the redshift (see the legend). For each redshift, we fit the results to a law of the form $M_{\rm H\,I}(M, z) = {\rm e}^{\gamma} M^{\alpha}$. The solid lines represent the best fits at the different redshifts. The right-hand panel shows the ratio between the fit to the $M_{\rm H\,I}(M, z)$ function in the AGN and CSF simulations, as a function of the halo mass, at the different redshifts.

that the fraction of H I residing in the inner regions of the haloes is higher when AGN feedback is not active. This effect increases with decreasing halo mass, pointing out that AGN feedback impacts more strongly on the regions near the halo centre, especially in the group regime.

3.3 Redshift evolution

In this subsection, we investigate the time dependence of the functions $M_{\rm H\,I}(M, z)$ and $M_{\rm H\,I}(r|M, z)$.

For each halo of both simulations with and without AGN feedback, at redshifts $z = 0, 0.25, 0.5, 0.8, 1, 1.5$ and 2 we have computed the neutral hydrogen mass hosted by the halo within R_{200}. In Fig. 6, we show the results for the CSF (left) and AGN (middle) simulations. We find that our results can be fitted by the same law we discussed in Section 3.1, $M_{\rm H\,I}(M, z) = {\rm e}^{\gamma} M^{\alpha}$, i.e. a linear relation between the logarithms of $M_{\rm H\,I}$ and M. In Table 1, we show the best-fitting values for α and γ for the two different simulations at the different redshifts. We also show in Fig. 6 with solid coloured lines the fits.

For the simulations with no AGN feedback, we find that the slope of the $M_{\rm H\,I}-M$ relationship, given by α, decreases with redshift, pointing out that the relative differences in the H I content between haloes of different masses are smaller at high redshift than at low redshift. Moreover, the overall normalization of the $M_{\rm H\,I}(M, z)$ function, given by γ, increases its value with redshift. This means that, for haloes with masses lower than $\sim 10^{14}\, h^{-1}\, {\rm M}_\odot$, the neutral hydrogen content of haloes with a fixed mass increases with redshift, being the effect more pronounced in haloes with low masses.

Our results for the simulations with AGN feedback switched on point out that α increases with redshift until $z \sim 0.5$ while at higher redshifts it decreases its value. γ, on the other hand, decreases its value up to $z \sim 0.5$ and increases its value at higher redshifts. We find that a single fit to all results between redshift 0 and redshift 1.5 is a very good description of the data, and we show in Table 1 the best-fitting values.

The right-hand panel of Fig. 6 displays the ratio between the fit to the $M_{\rm H\,I}(M, z)$ function in the simulations with AGN feedback on and off, at different redshifts. We find that for all masses and redshifts, the ratio is always below 1, demonstrating once again the effect of AGN feedback reducing the amount of neutral hydrogen. It is interesting to point out that AGN feedback tends to suppress the H I mass by roughly 50 per cent, with a very weak dependence on redshift and halo mass. Indeed, the ratio $M_{\rm H\,I}^{\rm AGN}/M_{\rm H\,I}^{\rm CSF}$ is very stable around 0.5 while varying the halo mass by four orders of magnitude and from redshift $z = 0$ to 2. We notice that the quantity plotted in that panel is just ${\rm e}^{\gamma_{\rm AGN}} M^{\alpha_{\rm AGN}}/({\rm e}^{\gamma_{\rm CSF}} M^{\alpha_{\rm CSF}})$, where AGN and CSF stand for the value of α and γ for that simulation (see Table 1). The mass dependence of that quantity is therefore given by $M^{\gamma_{\rm AGN}-\gamma_{\rm CSF}}$. At $z > 0$, $\gamma_{\rm AGN} > \gamma_{\rm CSF}$ and thus the ratio increases with mass, while at $z = 0$ the opposite situation takes place, inducing a change of slope as can be seen in the red line in that panel. We note that within 1σ, the value of $\gamma_{\rm AGN} - \gamma_{\rm CSF}$ is also compatible with a positive number at $z = 0$.

We have also investigated the time evolution of the spatial distribution of neutral hydrogen within haloes. In the middle and bottom row of Fig. 4, we show the results of computing the ratio $M_{\rm H\,I}(r|M, z)/M_{\rm H\,I}(R_{200}|M, z)$, as a function of r/R_{200} at redshift $z = 0.5$ (middle) and $z = 1$ (bottom). We notice that there are some situations in which the number of haloes is either 0 or very small (for instance haloes with masses larger than $10^{15}\, h^{-1}\, {\rm M}_\odot$ at $z \geq 1$). In these cases, we do not show the results.

We find that at redshifts higher than $z = 0$, the fraction of the total neutral hydrogen that resides in a given halo is higher in the inner regions in simulations where AGN feedback is switched off. The same conclusions hold at $z = 0$ (see Section 3.1). At $z = 0.5$, we also observe the same trend with halo mass that we find at $z = 0$: the differences between the results in simulations with and without AGN feedback decrease with increasing halo mass.

As with the results at $z = 0$, we find that a very simple law, of the form $M_{\rm H\,I}(r|M, z)/M_{\rm H\,I}(R_{200}|M, z) = (r/R_{200})^{\beta}$, can reproduce very well our measurements. In Table 2, we show the best-fitting values of the parameter β for different halo mass ranges and redshifts. Our results point out that for groups of galaxies ($[2–5] \times 10^{13}\, h^{-1}\, {\rm M}_\odot$) the value of β decreases with redshift, both for simulations with AGN feedback on or off. This means that the fraction of H I that is located in the inner regions of a halo with a fixed mass increases with redshift (at least until $z = 1$). At $z = 0.5$, we find that the value of β increases with the halo mass, independently of the simulation used, a trend that we already found at $z = 0$. Finally, the redshift evolution of β for galaxy clusters is less evident,

with β increasing for the CSF simulations and decreasing for the AGN simulations. Notice however that given the error in the fits, the results at the two different redshifts are compatible at 1σ.

4 COMPARISON WITH THEORETICAL MODELS AND IMPLICATIONS FOR INTENSITY MAPPING

In this section, we compare our results against the predictions of the Bagla et al. (2010) model and study the implications for 21 cm intensity mapping.

As discussed in the introduction, the shape and amplitude of the function $M_{\rm H\,I}(M, z)$ are of primary importance for future surveys that aim at putting constraints on the cosmological parameters using intensity mapping. We now compare our findings with the theoretical model of Bagla et al. (2010), which has been commonly used in the literature to perform forecasts (Camera et al. 2013; Villaescusa-Navarro et al. 2014, 2015a,b; Bull et al. 2015; Carucci et al. 2015) and to create mock 21 cm intensity mapping maps (Seehars et al. 2015).

Bagla et al. (2010) proposed a functional form for the $M_{\rm H\,I}(M, z)$ function as follows:

$$M_{\rm H\,I}(M,z) = \begin{cases} f_3(z)\frac{M}{1+M/M_{\rm max}(z)} & \text{if } \mathrm{M_{min}(z)} \leqslant \mathrm{M} \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

where the values of the free parameters $M_{\rm min}(z)$ and $M_{\rm max}(z)$ are taken in correspondence to the masses of haloes with circular velocities equal to $v_{\rm min} = 30\,{\rm km\,s^{-1}}$ and $v_{\rm max} = 200\,{\rm km\,s^{-1}}$ at redshift z, respectively. $f_3(z)$ is also a free parameter, and its value is chosen to reproduce the H I density parameter $\Omega_{\rm H\,I}(z)$.

With a dashed magenta line, we show in Fig. 2 the prediction of the Bagla et al. (2010) model for the function $M_{\rm H\,I}(M, z)$ at $z = 0$. We find that the Bagla model underestimates the neutral hydrogen content of both clusters and groups of galaxies. The deviations are however more dramatic in galaxy clusters (for haloes of $10^{15}\,h^{-1}\,{\rm M_\odot}$, differences are of two orders of magnitude), with the Bagla model predicting that those haloes should host about the same H I mass while our results indicate that the neutral hydrogen mass increases with the mass of the host halo. Notice that the Bagla model also underpredicts, by almost one order of magnitude, the lower limits on the H I masses of the Sausage and Virgo clusters (see Section 3.1). By extrapolating our results to smaller haloes, we find that the prediction of the Bagla model agrees pretty well with our fitting formula, for haloes with masses lower than $\sim 10^{12}\,h^{-1}\,{\rm M_\odot}$ when AGN feedback is switched off, while when it is switched on the Bagla model overpredicts the neutral hydrogen content of those haloes.

Notice that we are restricting our analysis to the model 3 of Bagla et al. (2010), which physically is the one best motivated among the three different models proposed by those authors. In their model 1, Bagla et al. (2010) proposed that haloes with circular velocities larger than 200 km s^{-1} ($\sim 2 \times 10^{12}\,h^{-1}\,{\rm M_\odot}$ at $z = 0$) do not host any H I at all. This is clearly in tension with our results but also with the lower limits on the H I mass of the Sausage and Virgo clusters. The model 2 of Bagla et al. (2010) lies in between models 1 and 3, and therefore the discrepancies among our results and the observational limits will be larger than those obtained by employing the model 3. We therefore concentrate our analysis in the model 3 of Bagla et al. (2010).

We notice that the fact that the Bagla model underestimates the neutral hydrogen mass in massive haloes has important consequences with what respects the spatial distribution of H I, as we shall see now. As we discussed in the introduction, two of the most important quantities for 21 cm intensity mapping are average neutral hydrogen density in units of the Universe critical density, $\Omega_{\rm H\,I}$, and the bias of that distribution with respect to the one of total matter, $b_{\rm H\,I}$. These quantities can be easily computed once the shape and amplitude of the function $M_{\rm H\,I}(M)$ are known. In Fig. 7, we show with a dashed magenta line the values of $\Omega_{\rm H\,I}$, $b_{\rm H\,I}$ and $b_{\rm H\,I}\Omega_{\rm H\,I}$ that we obtain by using the Bagla model as a function of the minimum mass of haloes that host H I, $M_{\rm min}$ at $z = 0$ (left column) and $z = 0.8$ (right column).

As we have discussed above, the Bagla model has one free parameter, $f_3(z)$, whose value is chosen to reproduce the value of $\Omega_{\rm H\,I}(z)$. In our case, we follow Crighton et al. (2015) and we assume $\Omega_{\rm H\,I}(z) = 4 \times 10^{-4}(1+z)^{0.6}$, in excellent agreement with the observational measurements at $z = 0$ by Zwaan & Prochaska (2006), Braun (2012), Martin et al. (2010) and Delhaize et al. (2013) and at $z \sim 0.8$ by Rao et al. (2006). We show these observational measurements as coloured bands in the upper panels of Fig. 7. Thus, by construction, the value of $\Omega_{\rm H\,I}$ does not depend on $M_{\rm min}$ when using the Bagla model. On the other hand, the value of $b_{\rm H\,I}$ does depend on $M_{\rm min}$, as shown in the middle panels of Fig. 7. We find however that the dependence of $b_{\rm H\,I}$ on $M_{\rm min}$ is very weak, with $b_{\rm H\,I}$ ranging only from $\sim$0.85 to $\sim$0.92 at $z = 0$ and from $\sim$1.03 to $\sim$1.15 at $z = 0.8$ when $M_{\rm min}$ is varied over two orders of magnitude.

The bottom panels of Fig. 7 show the value of $\Omega_{\rm H\,I}b_{\rm H\,I}$ as a function of $M_{\rm min}$. We remind the reader that the relevant quantity for 21 cm intensity mapping is the product $\Omega_{\rm H\,I}b_{\rm H\,I}$ (see equation 1) and not the values of $\Omega_{\rm H\,I}$ and $b_{\rm H\,I}$ separately. Given the fact that $\Omega_{\rm H\,I}(z)$ is fixed in the Bagla model, and that $b_{\rm H\,I}(z)$ barely changes with $M_{\rm min}$, it is not surprising that the quantity $\Omega_{\rm H\,I}b_{\rm H\,I}$ exhibits such a weak dependence with $M_{\rm min}$. At $z \sim 0.8$, 21 cm intensity mapping observations have determined the value of $\Omega_{\rm H\,I}b_{\rm H\,I}$ to be $6.2^{+2.3}_{-1.5} \times 10^{-4}$ (Chang et al. 2010; Masui et al. 2013; Switzer et al. 2013). In the bottom-left panel of Fig. 7, we show with a coloured green band those results. We find that Bagla model reproduces very well those observations, as was already pointed out in Padmanabhan et al. (2015). Unfortunately, no measurements of $\Omega_{\rm H\,I}b_{\rm H\,I}$ are available at other redshifts.

We have also used the fitting function which reproduces the results of our simulations, extrapolating it to halo masses below the resolution limit of our simulations, considering, as in the Bagla model, that only haloes above $M_{\rm min}$ host H I. In other words, we have modelled the function $M_{\rm H\,I}(M, z)$, according to the results of our simulations, as

$$M_{\rm H\,I}(M,z) = \begin{cases} {\rm e}^{\gamma} M^{\alpha} & \text{if } \mathrm{M_{min}(z)} \leqslant \mathrm{M} \\ 0 & \text{otherwise,} \end{cases} \quad (12)$$

where the values of the parameters α and γ are given in Table 1. We take the values of α and γ at $z = 0$ and 0.8 from Table 1. In Fig. 7, we show the value of $\Omega_{\rm H\,I}$, $b_{\rm H\,I}$ and $\Omega_{\rm H\,I}b_{\rm H\,I}$ as a function of $M_{\rm min}$, which we obtain using the above $M_{\rm H\,I}(M, z)$ function for both the CSF and AGN simulations. We also show with coloured bands around the mean value the variation of the results arising from the 1σ uncertainty in the value of α and γ from Table 1.

As expected, given the fact that AGN suppresses the H I content of haloes, for a fixed value of $M_{\rm min}$, the value of $\Omega_{\rm H\,I}$ is always lower when using the $M_{\rm H\,I}(M, z)$ function from the AGN simulations than from the CSF simulations. We find that in order to reproduce the value of $\Omega_{\rm H\,I}(z)$ from observations, haloes with circular velocities higher than $\sim 20\,{\rm km\,s^{-1}}$ at $z = 0$ and $\sim 25\,{\rm km\,s^{-1}}$ at $z = 0.8$ are required to host H I when AGN feedback is turned off. On the other hand, when AGN feedback is on, we conclude that haloes with circular velocities higher than $\sim 25\,{\rm km\,s^{-1}}$, at both $z = 0$ and 0.8,

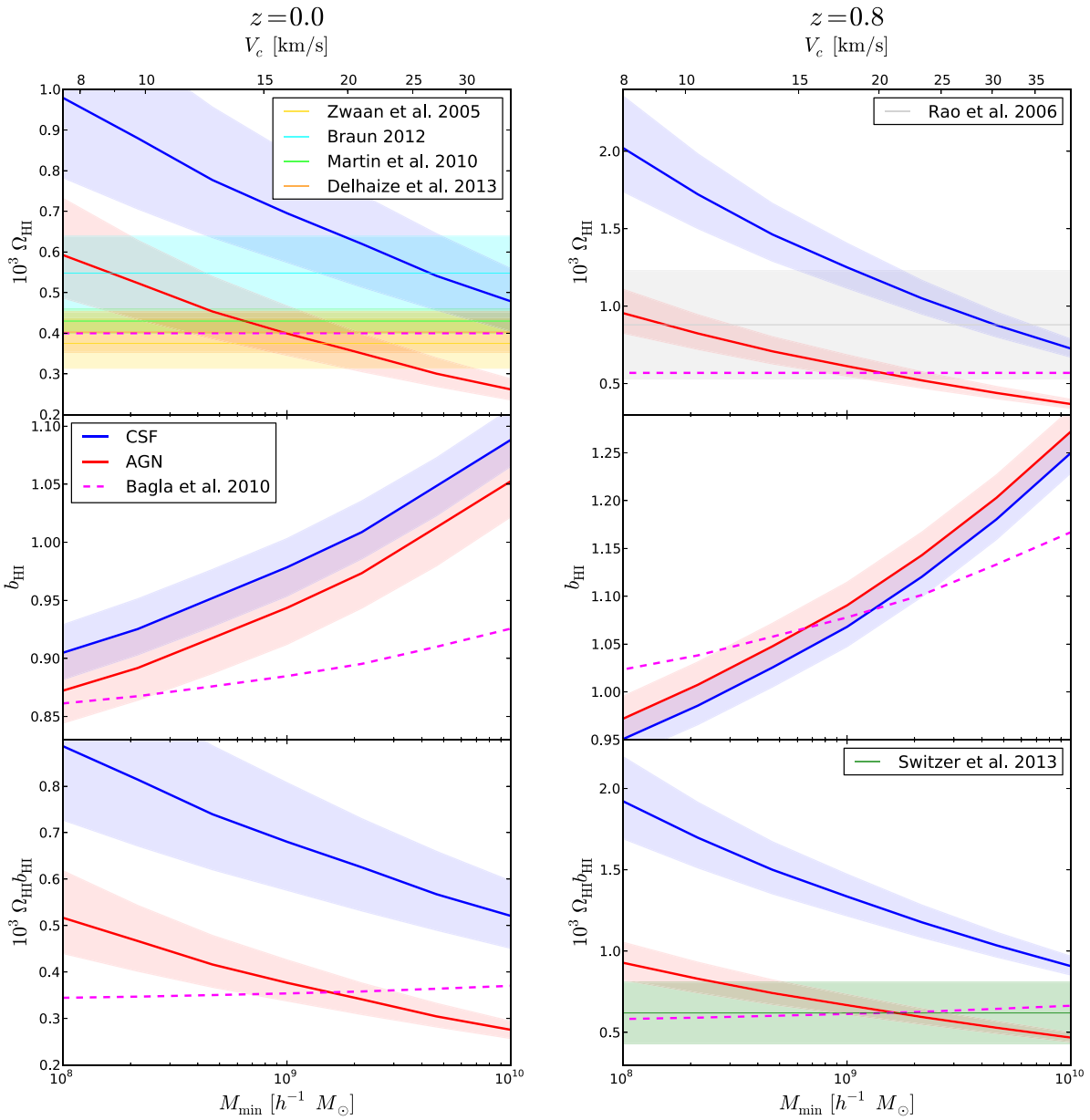

Figure 7. We use the functions $M_{\rm H\,I}(M)$ from the Bagla et al. (2010) model (purple dashed line) and from the fit to the simulations (red and blue solid lines; accompanying coloured areas indicate the variation of the results according to the fitting errors) to compute the value of $\Omega_{\rm H\,I}$ (upper panel), $b_{\rm H\,I}$ (lower panel) and $\Omega_{\rm H\,I}b_{\rm H\,I}$ at $z = 0$ (left column) and $z = 0.8$ (right column). We extrapolate the fits to halo masses lower than the simulation resolution limit assuming that haloes with masses lower than $M_{\rm min}$ do not host any H I. The plots show the results as a function of $M_{\rm min}$. The coloured bands in the upper-left panel display the value of $\Omega_{\rm H\,I}(z \cong 0)$ from observations (Zwaan & Prochaska 2006; Martin et al. 2010; Braun 2012; Delhaize et al. 2013) and in the upper-left panel the value of $\Omega_{\rm H\,I}(z \cong 0.53 \pm 0.38)$ from Rao, Turnshek & Nestor (2006). The green band in the bottom-right panel shows the measurement of $\Omega_{\rm H\,I}b_{\rm H\,I}$ at $z \sim 0.8$ from Switzer et al. (2013).

should host H I. Notice that the standard Bagla model assumes that only haloes with $V_{\rm c} > 30\,{\rm km\,s^{-1}}$ contain H I.

The value of the H I bias, which we computed using equation (4) and making use of Sheth & Tormen (1999) and Sheth, Mo & Tormen (2001) models for the halo mass function and halo bias, respectively, points out that the H I bias is higher in the CSF model at $z = 0$, in comparison with the AGN model, for all values of $M_{\rm min}$, whereas the opposite situation takes place at $z = 0.8$. We notice that the value of the H I bias does not depend on $\Omega_{\rm H\,I}$, but just on the way the H I is distributed among the different haloes. This implies that for the AGN and CSF models, for which we use $M_{\rm H\,I}(M, z) = {\rm e}^{\gamma} M^{\alpha}$, the value of the H I bias only depends on α. Thus, the reason why the H I bias is higher in the CSF model at $z = 0$, in comparison with the AGN model, is simply because α has a higher value in the CSF model at $z = 0$. The same argument holds at $z = 0.8$, where the CSF model has a lower value of α than the AGN model.

By comparing the results of the AGN and CSF models with those of the Bagla model, we find that at $z = 0$, and for the values of $M_{\rm min}$ considered, the H I bias is always higher in the AGN and CSF models. The reason of this behaviour is again the distribution of H I among the different haloes: in the CSF and AGN models, a

significant amount of H I is placed on clusters of galaxies, which are very biased objects and therefore it is not surprising that the value of the H I bias increases in those models. On the other hand, at $z = 0.8$ we find that the value of the H I bias in the AGN and CSF models is higher than the one of the Bagla model only if $M_{\rm min}$ is higher than $\sim 10^9\, h^{-1}\, {\rm M}_\odot$. This is again a consequence of the distribution of H I among the different haloes.

Finally, in the bottom panel of Fig. 7, we show the value of $\Omega_{\rm H\,I} b_{\rm H\,I}$ as a function of $M_{\rm min}$ for the AGN and CSF models. We find that the CSF model predicts a value of $\Omega_{\rm H\,I} b_{\rm H\,I}$ significantly higher than those of the AGN and Bagla models, in disagreement with the observational measurements at $z = 0.8$ by Switzer et al. (2013). On the other hand, the AGN model is capable of reproducing those measurements extremely well for almost all $M_{\rm min}$ masses. It is interesting to notice that for a value of $M_{\rm min} \sim 25\,{\rm km\,s^{-1}}$, for which the AGN model is capable of reproducing the observed value of $\Omega_{\rm H\,I}$ at both $z = 0$ and 0.8, the value of $\Omega_{\rm H\,I} b_{\rm H\,I}$ is very similar among the AGN and Bagla models.

We now discuss the robustness of our results. As we saw in Section 3.1, our results overpredict the H I mass in the most massive haloes by a factor of 3–4. This opens different possibilities. The first one is that our simulations predict correctly the H I mass in groups and small clusters but overpredict the neutral hydrogen mass in the most massive haloes. In that case, the quoted values of $\Omega_{\rm H\,I}$ and $b_{\rm H\,I}$ will barely changed since the amount of H I in those haloes is small.[16] The second possibility is that our simulations overpredict the value of $M_{\rm H\,I}(M, z)$ for all halo masses. In that case, the value of γ will decrease keeping fixed the value of α. Under those circumstances, the value of $b_{\rm H\,I}$ will remain unchanged but the value of both $\Omega_{\rm H\,I}$ and $\Omega_{\rm H\,I} b_{\rm H\,I}$ will decrease by a factor of 3–4, putting the model in difficulties to reproduce the observational results. Finally, it could happen that our simulations overpredict the H I mass in the most massive clusters but underpredict the H I mass in small groups (due to resolution limitations and because numerical contamination is stronger in clusters than in groups). In that situation, the value of α will decrease, affecting both the value of the H I bias, which will decrease, and the value of $\Omega_{\rm H\,I}$, which will increase.

5 SUMMARY AND DISCUSSION

Future radio surveys will sample the large-scale structure of the Universe by detecting the 21 cm emission from cosmic neutral hydrogen. Two different techniques can be employed for cosmological studies: (1) perform an H I galaxy survey, where individual galaxies are detected and (2) use the intensity mapping technique to carry out a low angular resolution survey where the integrated emission from many galaxies is measured. It has been recently pointed out that using an instrument like the future SKA, the constraints on the cosmological parameters will be much tighter if the intensity mapping technique is employed (see Bull et al. 2015).

In order to extract the maximum information from these surveys, exquisite knowledge of the spatial distribution of neutral hydrogen is needed from the theoretical side. A key ingredient for doing cosmology with 21 cm intensity mapping is the function $M_{\rm H\,I}(M, z)$, which represents the average mass in neutral hydrogen that a halo of mass M at redshift z hosts. The reason is that once that function is known, one can compute the bias of the H I, the value of $\Omega_{\rm H\,I}(z)$ and $\delta T_{\rm b}(z)$, and predict the shape and amplitude of the 21 cm power spectrum on large, linear scales.

It is thus very important to model that function as best as possible, in order to properly predict the bias and the amount of H I, which in the end will determine the expected signal-to-noise ratio of a given survey. The function $M_{\rm H\,I}(M, z)$ also carries very important astrophysical information, since neutral hydrogen represents an intermediate state between the highly ionized gas in the IGM and the dense molecular hydrogen that will end up forming stars. The amount of H I in haloes of different mass is largely affected by the astrophysical processes that are undergoing in them, and therefore, we can use the function $M_{\rm H\,I}(M, z)$ to improve our understanding of the physical processes affecting the formation and evolution of galaxies.

In this paper, we have studied the high-mass end of the $M_{\rm H\,I}(M, z)$ function using a set of zoom-in hydrodynamic simulations. Our simulations comprise two different sets, one in which AGN feedback is switched off (CSF) and another one which incorporates AGN feedback (AGN). All simulations incorporate metal-dependent radiative cooling, metal enrichment, SN feedback and star formation. We have post-processed the output of the simulations to account for two critical processes needed to properly model the spatial distribution of H I: the H I self-shielding and the formation of molecular hydrogen. To correct for the former, we use the fitting formula of Rahmati et al. (2013a) while for the latter we employ the KMT model (Krumholz et al. 2008, 2009; McKee & Krumholz 2010).

We find that the larger the mass of a halo, the higher the neutral hydrogen mass it contains, a result that holds at all redshifts studied in this paper ($z \in [0\text{–}2]$) and for haloes in the CSF and AGN simulations. We find that the results of the simulations can be well fitted by a law of the form $M_{\rm H\,I}(M, z) = {\rm e}^{\gamma} M^{\alpha}$. In Table 1, we show the best-fitting values of γ and α at different redshifts and for the two different simulation sets.

We have also investigated the spatial distribution of neutral hydrogen within haloes. In particular, we have studied the function $M_{\rm H\,I}(r|M, z)$, i.e. the function that gives the average H I mass within a radius r for a halo of mass M at redshift z. We find that for both haloes in the CSF and AGN simulations, the function $M_{\rm H\,I}(r|M, z)$ is well described by a law of the form $M_{\rm H\,I}(r|M, z)/M_{\rm H\,I}(R_{200}|M, z) = (r/R_{200})^{\beta}$. Our results point out that the value of β increases with the halo mass, implying that the fraction of H I residing in the inner regions of the halo decreases with the halo mass. We also observe a dependence of β with redshift: for a fixed halo mass, the value of β decreases with redshift.

We have paid special attention to the effect that AGN feedback induces on the amount and spatial distribution of neutral hydrogen by comparing results among the CSF and AGN simulations. Regarding the function $M_{\rm H\,I}(M, z)$, we find that AGN feedback acts by decreasing the total amount of neutral hydrogen hosted by a given halo. Our results point out that the value of the slope of the function $M_{\rm H\,I}(M, z)$ is not significantly changed by AGN feedback, being its main effect to modify its overall normalization. We find that on average, AGN feedback decreases the neutral hydrogen content a halo hosts by $\sim$50 per cent, with a weak dependence on halo mass and redshift.

[16] We have explicitly checked this by repeating the analysis using the function

$$M_{\rm H\,I}(M, z) = \begin{cases} {\rm e}^{\gamma} M_{\rm max}^{\alpha} & \text{if } {\rm M_{max}} \leqslant {\rm M} \\ {\rm e}^{\gamma} M^{\alpha} & \text{if } {\rm M_{min}(z)} \leqslant {\rm M} \leqslant {\rm M_{max}} \\ 0 & \text{otherwise,} \end{cases} \quad (13)$$

with $M_{\rm max} = 5 \times 10^{14}\, h^{-1}\, {\rm M}_\odot$, and we found that differences are very small. Notice that this function will be able to explain the H I masses of the Virgo and Sausage clusters.

AGN feedback also induces a shift in the value of the parameter β, meaning that the spatial distribution of neutral hydrogen within haloes is also affected by that process. We find that AGN feedback reduces the fraction of H I that is located in the halo inner regions, i.e. AGN feedback tends to increase the value of β.

We have also performed a halo-by-halo comparison among the simulations CSF and AGN. We find that AGN feedback decreases both the H I mass and total halo mass. We then asked ourselves whether the deficit in H I we find in simulations with AGN feedback may just be due to the fact that those haloes are simply smaller and therefore they host less H I. We find that AGN feedback is more effective at reducing the amount of neutral hydrogen than the total mass.

We have investigated the neutral hydrogen content in galaxies belonging to groups and clusters. We find that in groups of galaxies most of the overall H I mass resides within galaxies, while in clusters the fraction of the total H I mass made up of neutral hydrogen within galaxies is small. The reason is that in massive clusters gas can be more efficiently stripped from galaxies, mainly by the interaction with the hot, diffuse, ICM. It is worth reminding that numerical effects may create spurious blobs of H I that may bias our results. We have quantified the importance of these non-physical effects and we find that they are not important for groups, while for clusters they can contribute to the overall H I mass. We conclude that the simulated H I masses within clusters can only be trusted at this stage within a factor of 2.

We have compared our results for the $M_{\rm H\,I}(M, z)$ function against the prediction of the model by Bagla et al. (2010). We conclude that the Bagla et al. (2010) model dramatically underpredicts the amount of neutral hydrogen that resides in clusters of galaxies (by more than two orders of magnitude for haloes with masses of $10^{15}\,h^{-1}\,{\rm M}_{\odot}$), while in groups of galaxies the model and our results agree better, in particular for the simulation with AGN feedback on.

We derive consequences for 21 cm intensity mapping by extrapolating our results for the $M_{\rm H\,I}(M, z)$ function below the mass resolution limit of our simulations. The fact that galaxy clusters host a much higher mass in neutral hydrogen than the one predicted by the Bagla model implies that, at $z = 0$, the bias of the H I will be significantly higher than the one obtained by employing the Bagla model. We stress that the bias of the H I depends only on the slope of the $M_{\rm H\,I}(M, z)$ function and not on its overall normalization. On the other hand, the value of $\Omega_{\rm H\,I}(z)$ depends explicitly on the normalization of the function $M_{\rm H\,I}(M, z)$. We find that in order to reproduce the value of $\Omega_{\rm H\,I}(z)$ from observations, haloes with circular velocities higher than $\sim$20 km s^{-1}, at $z = 0$, and $\sim$25 km s^{-1} at $z = 0.8$ must host H I when AGN feedback is off, while for the simulations where AGN feedback is on we conclude that haloes with circular velocities higher than $\sim$25 km s^{-1} must host neutral hydrogen. We emphasize that the important quantity for 21 cm intensity mapping is $\Omega_{\rm H\,I} b_{\rm H\,I}$, which determines the signal-to-noise ratio of the signal. We find that the CSF model predicts a much higher value of $\Omega_{\rm H\,I} b_{\rm H\,I}$ than the AGN and Bagla models, in tension with observations at $z = 0.8$. On the other hand, the AGN and Bagla models are capable of reproducing those observations very well and at $z = 0$, and assuming that only haloes with circular velocities higher than $\sim$25 km s^{-1} host H I, both models predict a similar value of $\Omega_{\rm H\,I} b_{\rm H\,I}$. We therefore conclude that although the Bagla model underpredicts the mass of H I in galaxy clusters, it is a well-calibrated model for 21 cm intensity mapping.

We emphasize that our AGN simulations overpredict the H I mass of the Virgo and Sausage cluster by a factor of 3–4, and this may bias our conclusions for 21 cm intensity mapping. It could happen that our simulations only overpredict the H I mass of the most massive clusters. In that case, the consequences for intensity mapping will be unchanged, since the H I mass in those haloes is only a small fraction of the overall H I mass. On the other hand, if our simulations overpredict the H I mass of all haloes, very low circular velocity haloes will need to host H I in order to reproduce observations. It may also happen that our simulations overpredict the H I masses of very massive haloes while they underpredict the H I masses of small groups. This situation may arise since numerical contamination is more important in clusters than in groups and due to the limited resolution of our simulations. In that case, the slope of the $M_{\rm H\,I}(M, z)$ function will decrease, impacting the value of the H I bias, which will decrease, and the value of $\Omega_{\rm H\,I}$, which will increase. We plan to use data from observations to study the $M_{\rm H\,I}(M, z)$ function and distinguish among the above different possibilities in a future paper.

Finally, we propose our own model for the $M_{\rm H\,I}(M, z)$ function which comes out from the results of our hydrodynamic simulations with AGN feedback and that by construction will be able to reproduce the observational constraints. We suggest to model the function $M_{\rm H\,I}(M, z)$ as

$$M_{\rm H\,I}(M, z) = \begin{cases} {\rm e}^{\gamma(z)} M^{\alpha(z)} & \text{if } {\rm M_{min}(z)} \leqslant {\rm M} \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

with $\alpha(z) = 3/4$, in agreement with the results of our AGN simulations. $M_{\rm min}(z)$ represents the mass of the smallest halo that is capable of hosting H I at redshift z and $\gamma(z)$ is the value of the overall normalization of the $M_{\rm H\,I}(M, z)$ function. Notice that we are assuming that effects such as radiation from local sources and from the ICM and the limited resolution of our simulations only affect the value of $\gamma(z)$ and not the one of $\alpha(z)$, which may not be the case. In our model, we will further assume that $M_{\rm min}(z)$ does not depend on redshift.

The values of $M_{\rm min}$ and $\gamma(z)$ can be obtained by requiring that the function $M_{\rm H\,I}(M, z)$ reproduce observations of both $\Omega_{\rm H\,I}(z)$ and $b_{\rm H\,I}\Omega_{\rm H\,I}(z = 0.8)$:

$$\Omega_{\rm H\,I}(z) = \frac{{\rm e}^{\gamma(z)}}{\rho_{\rm c}^0} \int_{M_{\rm min}}^{\infty} n(M, z) M^{3/4}\, {\rm d}M \tag{15}$$

$$\Omega_{\rm H\,I} b_{\rm H\,I}(z = 0.8) = \frac{{\rm e}^{\gamma(0.8)}}{\rho_{\rm c}^0} \tag{16}$$

$$\times \int_{M_{\rm min}}^{\infty} n(M, 0.8) b(M, 0.8) M^{3/4}\, {\rm d}M, \tag{17}$$

where $\rho_{\rm c}^0$, $n(M, z)$ and $b(M, z)$ are the actual value of the Universe critical density, the halo mass function and the halo bias, respectively. The value of $\Omega_{\rm H\,I} b_{\rm H\,I}$ at $z = 0.8$ has been measured in Chang et al. (2010), Masui et al. (2013) and Switzer et al. (2013) obtaining a value equal to $6.2^{+2.3}_{-1.5} \times 10^{-4}$. The value of $\Omega_{\rm H\,I}(z)$ has been measured by different authors at different redshifts. A parametrization has been recently proposed by Crighton et al. (2015), $\Omega_{\rm H\,I}(z) = A(1 + z)^{\gamma}$, with $A = (4.00 \pm 0.024) \times 10^{-4}$ and $\gamma = 0.60 \pm 0.05$. Thus, our model for the $M_{\rm H\,I}(M, z)$ function will, by construction, reproduce the observations and at the same time the shape of that function will be in agreement with the results of our hydrodynamic simulations. We finally notice that since the contribution of spurious H I blobs could be more important in clusters than in groups, and since the limited resolution of our simulations may also impact on the H I mass of small groups, the

actual value of $\alpha(z)$ might be lower than 3/4 or alternatively, the $M_{\rm H\,I}(M, z)$ may exhibit a plateau on the high-mass end.

We conclude by pointing out some effects that may impact on our results. First of all, the resolution of our simulations[17] does not allow us to follow the evolution of neutral hydrogen within small galaxies. The stability and robustness of our results should therefore be checked against resolution. We leave this for a future work. Secondly, the value of $\Omega_{\rm b}$ in our simulations is $\sim$20 per cent lower than the one found by *Planck* ($\Omega_{\rm b} = 0.049$; Planck Collaboration XIII 2015). We expect that if there is more gas in the Universe, there will be more neutral hydrogen. Thirdly, we notice that in our analysis we have neglected the radiation from local sources and X-ray emission from the hot ICM. Accounting for these effects would require including a radiative-transfer description of the local radiation field that it is beyond the scope of the present paper. On the other hand, the mismatch between the positions of the H I blobs and the galaxies within clusters in our simulations may significantly decrease the importance of the radiation from local sources. And finally, a more profound and detailed analysis is required to identify the nature (physical/numerical) of the neutral hydrogen clouds outside galaxies we find in our simulations. We plan to investigate these points in a future paper.

ACKNOWLEDGEMENTS

We thank the referee, Alan Duffy, for his useful and constructive report. FVN thanks Andra Stroe, Rhys Taylor and Martha Haynes for useful discussions. FVN and MV are supported by the ERC Starting Grant 'cosmoIGM' and partially supported by INFN IS PD51 'INDARK'. SP, SB and GM are supported by the PRIN-INAF12 grant 'The Universe in a Box: Multi-scale Simulations of Cosmic Structures', the PRINMIUR 01278X4FL grant 'Evolution of Cosmic Baryons' and the INDARK INFN grant. SP acknowledges support by the Spanish Ministerio de Economia y Competitividad (MINECO, grants AYA2013-48226-C3-2-P) and the Generalitat Valenciana (grant GVACOMP2015-227). ER is supported by FP7-PEOPLE-2013-IIF (grant agreement PIIF-GA-2013-627474) and NSF AST-1210973. KD acknowledges the support by the DFG Cluster of Excellence 'Origin and Structure of the Universe'. AMB is supported by the DFG Research Unit 1254 'Magnetisation of Interstellar and Intergalactic Media' and by the DFG Cluster of Excellence 'Origin and Structure of the Universe'. Simulations are carried out using Flux HCP Cluster at the University of Michigan efficiently supported by ARC-TS, and at CINECA (Italy), with CPU time assigned through ISCRA proposals and through an agreement with the University of Trieste. Data post-processing and storage has been done on the CINECA facility PICO, granted us thanks to our expression of interest. We acknowledge partial support from 'Consorzio per la Fisica – Trieste'.

REFERENCES

Allende Prieto C., Lambert D. L., Asplund M., 2001, ApJ, 556, L63
Alonso D., Bull P., Ferreira P. G., Maartens R., Santos M. G., 2015, ApJ, 814, 145
Bagla J. S., Khandai N., Datta K. K., 2010, MNRAS, 407, 567
Barnes D. G. et al., 2001, MNRAS, 322, 486
Battye R. A., Davies R. D., Weller J., 2004, MNRAS, 355, 1339
Beck A. M. et al., 2016, MNRAS, 455, 2110
Bharadwaj S., Sethi S. K., 2001, J. Astrophys. Astron., 22, 293
Bharadwaj S., Nath B. B., Sethi S. K., 2001, J. Astrophys. Astron., 22, 21
Bonafede A., Dolag K., Stasyszyn F., Murante G., Borgani S., 2011, MNRAS, 418, 2234
Braun R., 2012, ApJ, 749, 87
Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, ApJ, 803, 21
Camera S., Santos M. G., Ferreira P. G., Ferramacho L., 2013, Phys. Rev. Lett., 111, 171302
Carucci I. P., Villaescusa-Navarro F., Viel M., Lapi A., 2015, J. Cosmol. Astropart. Phys., 7, 047
Catinella B. et al., 2010, MNRAS, 403, 683
Catinella B. et al., 2013, MNRAS, 436, 34
Chang T. C., Pen U. L., Peterson J. B., McDonald P., 2008, Phys. Rev. Lett., 100, 091303
Chang T. C., Pen U. L., Bandura K., Peterson J. B., 2010, Nature, 466, 463
Cortese L., Catinella B., Boissier S., Boselli A., Heinis S., 2011, MNRAS, 415, 1797
Crighton N. H. M. et al., 2015, MNRAS, 452, 217
Cui W., Borgani S., Murante G., 2014, MNRAS, 441, 1769
Davé R., Katz N., Oppenheimer B. D., Kollmeier J. A., Weinberg D. H., 2013, MNRAS, 434, 2645
Davies J. et al., 2004, MNRAS, 349, 922
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Delhaize J., Meyer M. J., Staveley-Smith L., Boyle B. J., 2013, MNRAS, 433, 1398
Dénes H., Kilborn V. A., Koribalski B. S., 2014, MNRAS, 444, 667
Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497
Duffy A. R., Kay S. T., Battye R. A., Booth C. M., Dalla Vecchia C., Schaye J., 2012, MNRAS, 420, 2799
Font-Ribera A. et al., 2012, J. Cosmol. Astropart. Phys., 11, 059
Gavazzi G., Boselli A., van Driel W., O'Neil K., 2005, A&A, 429, 439
Gavazzi G., O'Neil K., Boselli A., van Driel W., 2006, A&A, 449, 929
Giovanelli R. et al., 2005, AJ, 130, 2598
Gunn J. E., Gott J. R., III, 1972, ApJ, 176, 1
Haardt F., Madau P., 2001, in Neumann D. M., Van J. T. T., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays, Recent results of XMM-Newton and Chandra, XXXVIth Rencontres de Moriond, XXIst Moriond Astrophysics Meeting, Savoie, France
Hess K. M., Wilcots E. M., 2013, AJ, 146, 124
Ho L. C., Darling J., Greene J. E., 2008, ApJ, 681, 128
Huang S., Haynes M. P., Giovanelli R., Brinchmann J., 2012, ApJ, 756, 113
Hughes T. M., Cortese L., 2009, MNRAS, 396, L41
Jáchym P., Combes F., Cortese L., Sun M., Kenney J. D. P., 2014, ApJ, 792, 11
Janowiecki S. et al., 2015, ApJ, 801, 96
Kaiser N., 1987, MNRAS, 227, 1
Kannan R., Vogelsberger M., Stinson G. S., Hennawi J. F., Marinacci F., Springel V., Maccio A. V., 2015, preprint (arXiv:e-prints)
Kent B. R. et al., 2007, ApJ, 665, L15
Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M. S., 2015, MNRAS, 450, 1349
Koopmann R. A. et al., 2008, ApJ, 682, L85
Krumholz M. R., McKee C. F., Tumlinson J., 2008, ApJ, 689, 865
Krumholz M. R., McKee C. F., Tumlinson J., 2009, ApJ, 693, 216
Leslie S. K., Kewley L. J., Sanders D. B., Lee N., 2016, MNRAS, 455, L82
Loeb A., Wyithe J. S. B., 2008, Phys. Rev. Lett., 100, 161301
McKee C. F., Krumholz M. R., 2010, ApJ, 709, 308
McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, ApJ, 653, 815
Marín F. A., Gnedin N. Y., Seo H. J., Vallinotto A., 2010, ApJ, 718, 972
Martin A. M., Papastergis E., Giovanelli R., Haynes M. P., Springob C. M., Stierwalt S., 2010, ApJ, 723, 1359
Martin A. M., Giovanelli R., Haynes M. P., Guzzo L., 2012, ApJ, 750, 38
Masui K. W. et al., 2013, ApJ, 763, L20
Minchin R. et al., 2005, ApJ, 622, L21
Miralda-Escudé J., 2005, ApJ, 620, L91

[17] We notice that according to the resolution tests of Duffy et al. (2012) and their derived criteria for convergence, the profiles, total masses and H I masses of the large majority of our haloes should be converged against resolution.

Padmanabhan H., Choudhury T. R., Refregier A., 2015, preprint (arXiv:e-prints)
Planck Collaboration XIII, 2015, preprint (arXiv:e-prints)
Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, MNRAS, 438, 195
Rahmati A., Pawlik A. H., Raicevic M., Schaye J., 2013a, MNRAS, 430, 2427
Rahmati A., Schaye J., Pawlik A. H., Raičevič M., 2013b, MNRAS, 431, 2261
Rao S. M., Turnshek D. A., Nestor D. B., 2006, ApJ, 636, 610
Rasia E. et al., 2015, ApJ, 813, L17
Santos M., Alonso D., Bull P., Silva M. B., Yahya S., 2015a, Advancing Astrophysics with the Square Kilometre Array (AASKA14), p. 21 (Available at http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=215)
Santos M. G. et al., 2015b, preprint (arXiv:e-prints)
Schaye J., 2006, ApJ, 643, 59
Seehars S., Paranjape A., Witzemann A., Refregier A., Amara A., Akeret J., 2015, preprint (arXiv:e-prints)
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Solanes J. M., Manrique A., García-Gómez C., González-Casado G., Giovanelli R., Haynes M. P., 2001, ApJ, 548, 97
Springel V., 2005, MNRAS, 364, 1105
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Steinborn L. K., Dolag K., Hirschmann M., Prieto M. A., Remus R. S., 2015, MNRAS, 448, 1504
Stroe A. et al., 2015a, MNRAS, 450, 646
Stroe A., Oosterloo T., Röttgering H. J. A., Sobral D., van Weeren R., Dawson W., 2015b, MNRAS, 452, 2731
Sun M., Donahue M., Voit G. M., 2007, ApJ, 671, 190
Switzer E. R. et al., 2013, MNRAS, 434, L46
Taylor R., Davies J. I., Auld R., Minchin R. F., 2012, MNRAS, 423, 787
Taylor R., Davies J. I., Auld R., Minchin R. F., Smith R., 2013, MNRAS, 428, 459
Toomre A., Toomre J., 1972, ApJ, 178, 623
Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, MNRAS, 382, 1050
Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Dalla Vecchia C., 2014, MNRAS, 442, 2641
Villaescusa-Navarro F., Viel M., Datta K. K., Choudhury T. R., 2014, J. Cosmol. Astropart. Phys., 9, 050
Villaescusa-Navarro F., Bull P., Viel M., 2015a, ApJ, 814, 146
Villaescusa-Navarro F., Viel M., Alonso D., Datta K. K., Bull P., Santos M. G., 2015b, J. Cosmol. Astropart. Phys., 3, 034
Wiersma R. P. C., Schaye J., Smith B. D., 2009, MNRAS, 393, 99
Yahya S., Bull P., Santos M. G., Silva M., Maartens R., Okouma P., Bassett B., 2015, MNRAS, 450, 2251
Yoon I., Rosenberg J. L., 2014, Am. Astron. Soc. Meeting #223, #358.10
Zwaan M. A., Prochaska J. X., 2006, ApJ, 643, 675

APPENDIX A: NEUTRAL HYDROGEN IN GALAXIES AND NUMERICAL ARTEFACTS

In this appendix, we examine the neutral hydrogen content in galaxies, its overall contribution to the total H I mass in groups and clusters and the numerical problems our simulations may face.

A very interesting question to address would be: how does the H I mass function (HIMF) change as a function of the halo mass? In other words, how does the distribution of H I mass in galaxies vary with the mass of the halo?

In order to address that question, we have computed, for each subhalo, its stellar and neutral hydrogen mass. In the upper panel of Fig. A1, we show, for each halo in the simulations with AGN feedback, the total number of subhaloes, together with the number

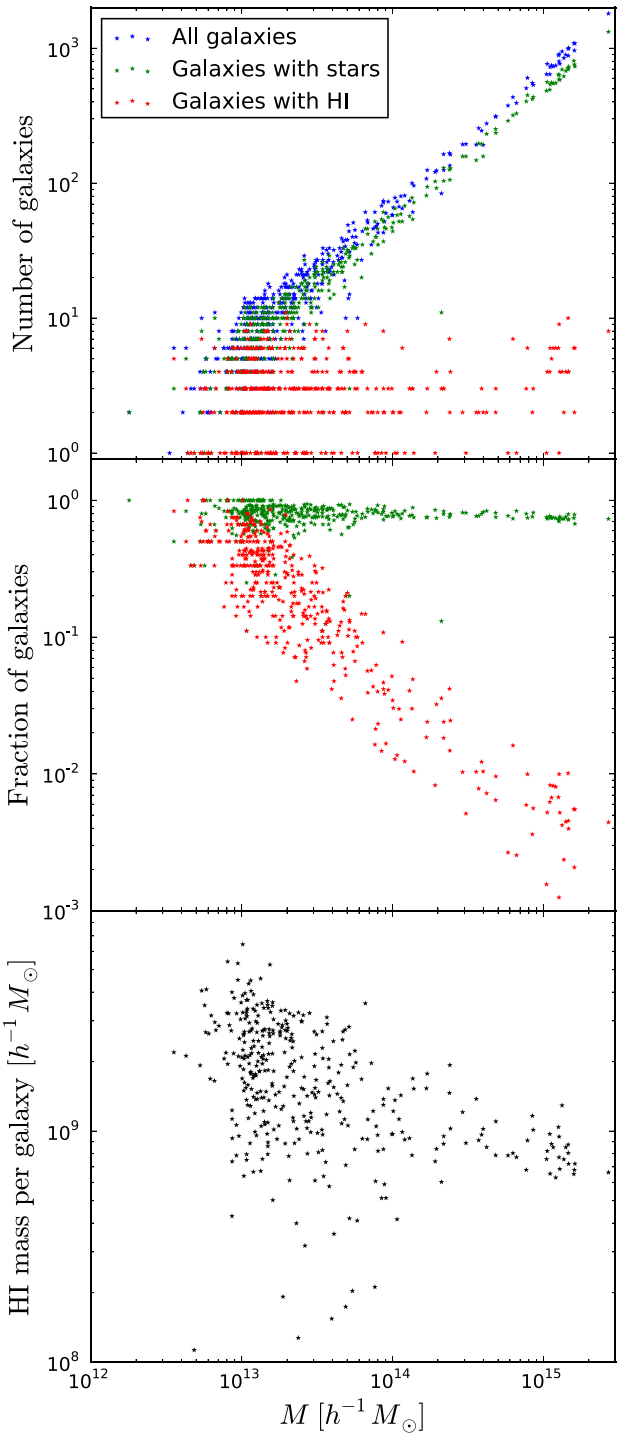

Figure A1. For each halo of the AGN simulations, we compute the total number of subhaloes (blue points), the number of subhaloes with star particles (green points) and the number of subhaloes with neutral hydrogen (red points) at $z = 0$. The upper panel shows the different number of subhaloes as a function of the halo mass, while the middle panel displays the fraction of subhaloes hosting stars and neutral hydrogen. The bottom panel shows the average H I mass per galaxy, computed by dividing the halo H I mass by the total number of galaxies.

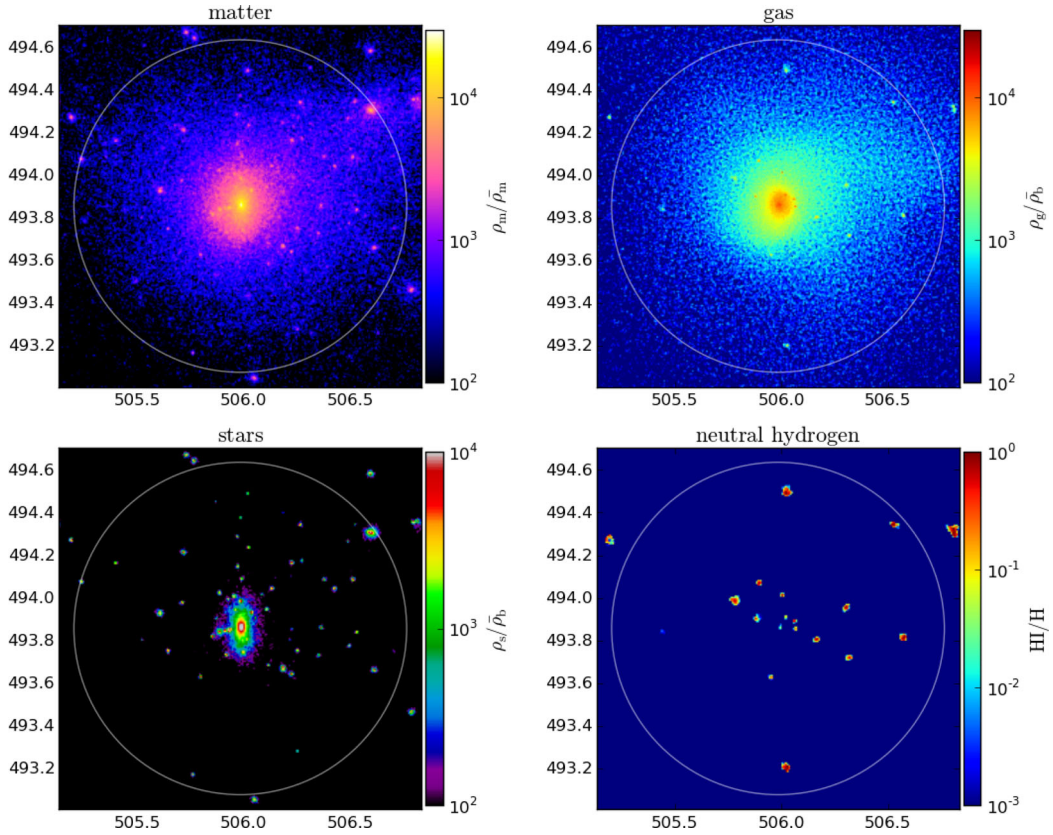

Figure A2. Effect of gas stripping in galaxy clusters. The panels show the spatial distribution of matter (top-left), gas (top-right), stars (bottom-left) and neutral hydrogen (bottom-right) centred on a halo of mass $M = 1.1 \times 10^{14}\,h^{-1}\,\mathrm{M}_\odot$ at $z = 0$. The x- and y-axes show the positions in h^{-1} Mpc; the width of the slices is equal to the diameter of the halo, which is $\sim 1.5\,h^{-1}$ Mpc. It can be noticed that in the majority of the cases, the positions of the H I blobs are slightly displaced with respect to the positions of the galaxies. The white circles represent the halo R_{200}.

of subhaloes hosting stars and the number of subhaloes containing H I.

As expected, we find that the total number of subhaloes increases with the mass of the halo, a trend that is also followed by the number of subhaloes hosting stars. We notice that, in general, the number of galaxies (i.e. subhaloes containing stars) will be smaller than the total number of subhaloes. This happens because SUBFIND identifies subhaloes up to the resolution limit (fixed in subhaloes with at least 32 particles). Thus, very small subhaloes are expected to not contain any star particle.

On the other hand, we find that the number of subhaloes hosting neutral hydrogen does not increase with the halo mass, and that its number is never above $\sim$10 (see also Hess & Wilcots 2013, for a similar analysis with observations). This is mainly due to gas stripping by the ICM.[18] This effect, originated by the interaction of the galaxy gas with the hot, diffuse, ICM, is expected to be more efficient at removing gas of the galaxies residing in the most massive haloes. In the middle panel of Fig. A1, we show the fraction of subhaloes hosting H I as a function of the halo mass. Indeed, we find that this fraction quickly decreases with the mass of the halo, pointing out that gas stripping is more efficient in clusters than in groups. Notice that in groups of galaxies ($M_{\rm halo} \sim 10^{13}\,h^{-1}\,\mathrm{M}_\odot$) the fraction of subhaloes hosting H I is comparable to the fraction of subhaloes with stars.

[18] We have also identified some cases in which gas stripping is produced by gravitational collision.

In order to further check our gas stripping hypothesis, we plot in Fig. A2 the spatial distribution of matter, gas, stars and neutral hydrogen for a halo of mass $M = 1.1 \times 10^{14}\,h^{-1}\,\mathrm{M}_\odot$ from the AGN simulations at $z = 0$. We find that the total number of subhaloes identified by SUBFIND is equal to 67. Among those, 52 host stars but only 2 have neutral hydrogen (one of them being the cD galaxy). As can be seen, the positions of the neutral hydrogen blobs do not perfectly match, in the majority of the cases, the positions where stars are found. It is worth pointing out that overdensities in the matter field have correspondence in the distribution of stars, in the same way as overdensities in the gas distribution match the positions of the H I blobs, but in general there is a small displacement among matter/stars and gas/H I. This just points out that whereas stars follow the same dynamics as CDM, gas behaves differently.

In most of the cases, the positions of the H I blobs are just slightly displaced with respect to the positions where subhaloes are located. This small displacement is however enough for SUBFIND to fail at linking the gas to the subhaloes, and indicates that gas stripping has taken place recently. We notice that since most of the galaxies do not have neutral hydrogen, SUBFIND finds that the biggest subhalo of a given halo (that we refer to as the cD galaxy) hosts a very large amount of H I. The reason for this is that, since the spatial positions of the H I blobs and the galaxies are different, the neutral hydrogen blobs are accounted for in the largest subhalo, and not in the different galaxies. We find that in the 96 and 91 per cent of the haloes, the H I in their galaxies (mainly coming from the cD galaxy) account for more than 90 and 95 per cent of the overall halo H I content, respectively.

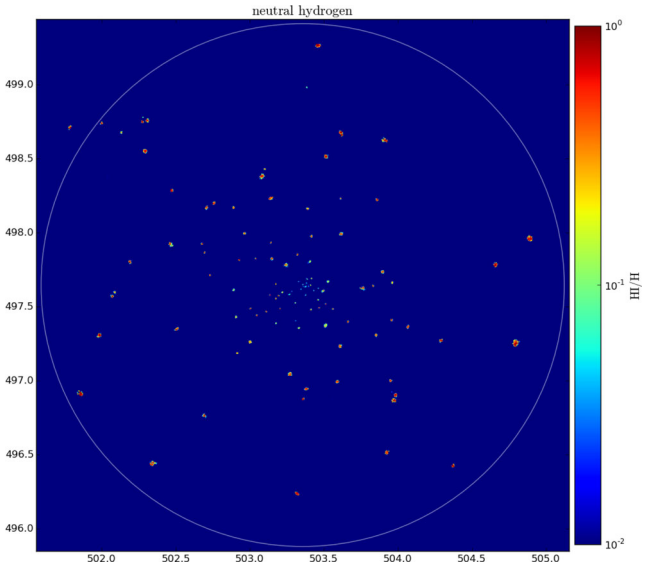

Figure A3. Distribution of neutral hydrogen in a very massive cluster of mass $M = 1.28 \times 10^{15}\,h^{-1}\,\mathrm{M}_{\odot}$ at $z = 0$. The neutral hydrogen fraction decreases quickly towards the centre of the halo.

Unfortunately, the problems discussed above do not allow us to study the dependence of the HIMF on halo mass. As a very rough approximation, we plot in the bottom panel of Fig. A1 the average H I mass per galaxy, computed by dividing the overall H I mass in a given halo by the total number of subhaloes within it, as a function of the halo mass. We warn the reader that the numbers quoted in the y-axis have to be taken with caution, since these results clearly depend on the resolution of the simulations: the higher the resolution, the larger the number of subhaloes it will be found, while the total H I mass is not expected to depend that strongly on resolution. In other words, the important thing in that plot is the trend rather than the absolute scale. Our results point out that the average H I mass in galaxies decreases with the mass of its host halo, although there is a rather large scatter in the trend. Therefore, H I-poor galaxies are more likely to be found in clusters of galaxies than in groups.

Even though in our simulations we cannot directly link H I to galaxies, we can still compare our results against some observational results. For instance, it is known that the fraction of galaxies with H I detected by surveys declines towards the centre of the haloes (see for instance Hess & Wilcots 2013; Yoon & Rosenberg 2014). In Fig. A3, we plot the spatial distribution of neutral hydrogen within a massive cluster of mass $M = 1.28 \times 10^{15}\,h^{-1}\,\mathrm{M}_{\odot}$ at $z = 0$ taken from the AGN simulations. It can be seen that towards the halo centre the H I clouds are smaller and the average neutral hydrogen within them quickly decreases, the same trend as seen in observations.

It is not obvious whether the gas blobs we find represent a physical situation (e.g. gas stripping by the ICM; see for instance Sun, Donahue & Voit 2007; Jáchym et al. 2014), a numerical artefact of our hydrodynamic solver, where residual errors in our SPH scheme prevent the disruption of those H I blobs, or a mix of the two. In order to address this issue, one could run simulations with different SPH methods or with grid methods, and investigate how results change. Besides, the stability of our result should be verified against resolution. This is however beyond the scope of this paper and we leave it for a future work.

Two important questions naturally then show up: (1) How important is the H I mass outside galaxies? and (2) How important is the H I mass arising from numerical artefacts on the overall H I content of haloes?

In order to address the first question, we have computed, for each halo of the AGN simulations at $z = 0$, the H I mass within its R_{200} which is located within galaxies. We have taken special care of the cD galaxy, since the output of SUBFIND will assign to it most of the H I blobs. In order to circumvent this problem, we have computed the H I mass of the cD galaxy as the H I mass contained in a sphere of radius of $50\,h^{-1}$ kpc[19]. In the left-hand panel of Fig. A4, we show the results using this procedure together with the results taking all H I within R_{200}. We find that the H I mass hosted by galaxies in clusters is much lower than its overall H I mass, while for groups the difference is much smaller. The right-hand panel of Fig. A4 shows the ratio, for each halo, between the H I mass in galaxies and the overall H I mass. It can be seen that for groups, the H I mass in galaxies is very similar to the overall H I mass while for galaxy clusters the galaxy content in H I is a rather small fraction of the overall halo neutral hydrogen mass. In the left-hand panel of Fig. A4, we also display the prediction of the Bagla et al. (2010) model for the $M_{\mathrm{H\,I}}(M, z)$ function. Our results point out that the Bagla et al. (2010) model reproduces pretty well our measurements of the $M_{\mathrm{H\,I}}(M, z)$ function when the contribution of H I outside galaxies is removed, although the scatter is very large. We note however again that the Bagla model is inconsistent with the lower limit on the H I mass from the Sausage and Virgo clusters.

In order to address, at least partially, the question of how important is the H I mass in blobs whose nature is not physical but simply numerical, we have run SUBFIND, on top of all of our resimulated regions at $z = 0$ for the simulation with AGN feedback, allowing it to identify gas clouds as subhaloes. We have then identified blobs of gas as subhaloes with no CDM and star particles in them. The reason for doing this is that we expect the numerical blobs to show up as isolated and concentrated clouds of gas. Therefore, SUBFIND is likely to identify these structures as purely gaseous subhaloes, while stripped gas will be more diffuse and thus SUBFIND will not recognize it as subhaloes. In Fig. A5, we show, for each halo, the overall H I mass within R_{200} together with the H I mass within galaxies and within the blobs.

We find that for groups, the H I mass in numerical blobs is just a small fraction of the overall H I mass, while for galaxy clusters the H I mass in those blobs is much larger and surpasses the H I mass within galaxies. In the right-hand panel of Fig. A5, we plot, for each halo, the ratio between the H I mass in numerical blobs over the overall H I mass. We find that the contribution to the overall H I mass from numerical blobs (notice that we are assuming that all purely gaseous subhaloes are numerical, which may not be true) is below 50 per cent in the majority of the cases. Thus, we conclude that the results of this paper are robust against numerical problems and the H I masses we quote can be trusted within a factor of 2 for galaxy clusters while the H I mass within groups is much less sensitive to these issues.

We however emphasize that a more detailed study is required in order to more closely address the nature of the isolated H I blobs we find in our simulations. Notice that H I clouds without optical counterparts have been found in Davies et al. (2004), Minchin et al. (2005), Kent et al. (2007), Koopmann et al. (2008), Taylor et al. (2012, 2013) and Janowiecki et al. (2015). We plan to address this in a future paper.

[19] We notice that our results are robust against reasonable variations in the size of the cD galaxy.

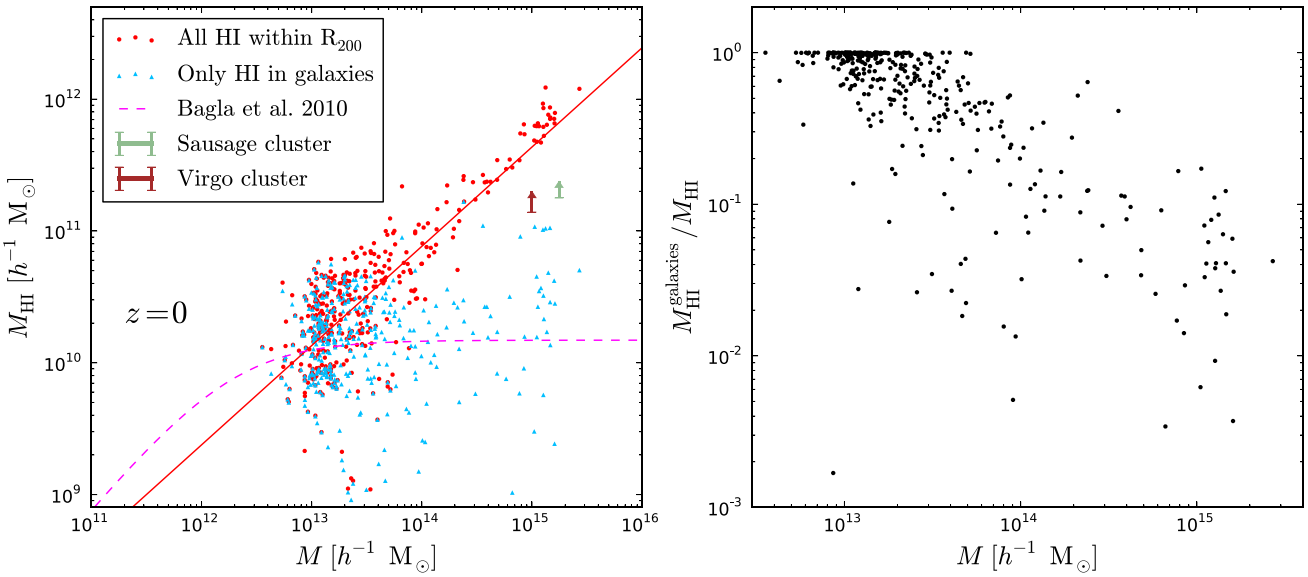

Figure A4. Neutral hydrogen mass outside galaxies. For each halo in the AGN simulations, we have computed the H I mass within R_{200} at $z = 0$. The red points in the left-hand panel show the results as a function of the halo mass together with the best fit obtained by using the expression $M_{\rm HI} = {\rm e}^{\gamma} M^{\alpha}$. For each halo, we have also computed the H I mass residing only in galaxies, and the blue triangles in the left-hand panel display the results. The green and brown arrows represent a lower limit on the neutral hydrogen mass hosted by the Sausage and Virgo clusters, respectively. The right-hand panel shows the ratio, for each halo, between the H I mass in galaxies over the whole neutral hydrogen mass.

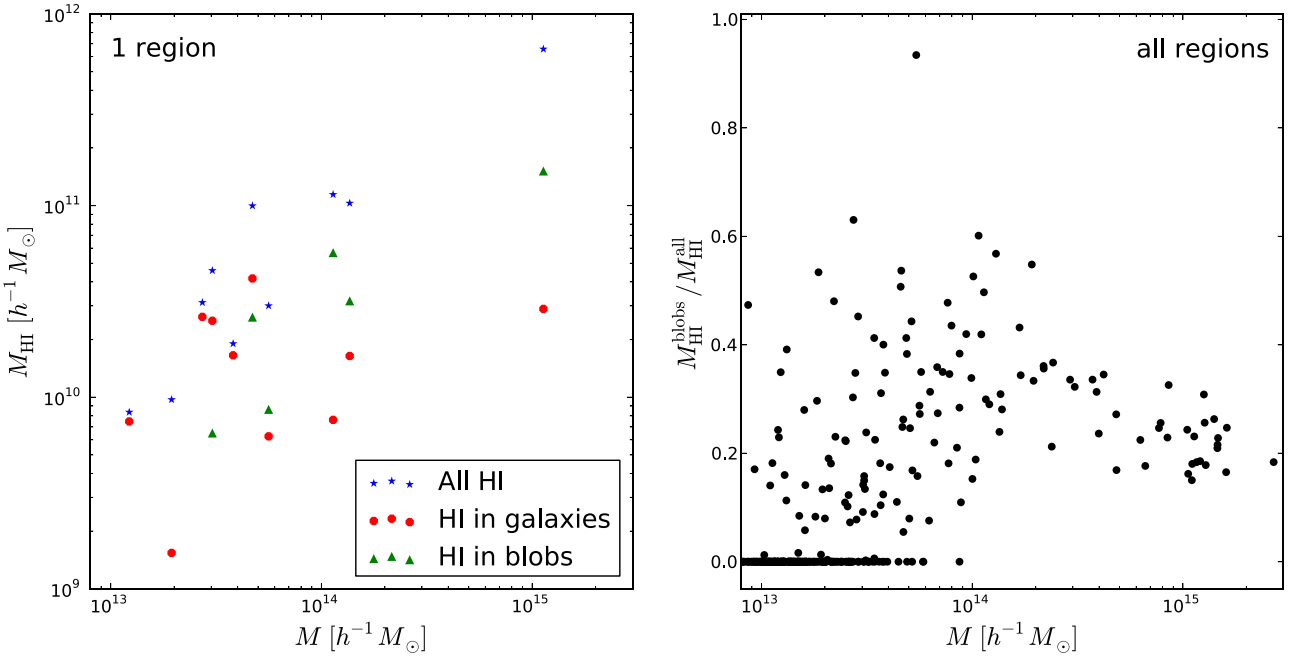

Figure A5. H I mass in numerical blobs. For each halo of the AGN simulations at $z = 0$, we have computed all H I mass within R_{200} (blue stars), the H I mass in galaxies (red points) and the H I in blobs (subhaloes with no stars and DM; green triangles), and we show the results in the left-hand panel. For clearness, we only display the results for one single region. The right-hand panel shows the ratio between the H I mass in blobs and the overall H I mass for all haloes.

This paper has been typeset from a TEX/LATEX file prepared by the author.