

# NIHAO VII: Predictions for the galactic baryon budget in dwarf to Milky Way mass haloes

Liang Wang<sup>1,3\*</sup>, Aaron A. Dutton<sup>2,3</sup>, Gregory S. Stinson<sup>3</sup>, Andrea V. Macciò<sup>2,3</sup>,  
Thales Gutcke<sup>3</sup>, Xi Kang<sup>1</sup>

<sup>1</sup>*Purple Mountain Observatory, the Partner Group of MPI für Astronomie, 2 West Beijing Road, Nanjing 210008, China*

<sup>2</sup>*New York University Abu Dhabi, PO Box 129188, Abu Dhabi, UAE*

<sup>3</sup>*Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany*

submitted to MNRAS

## ABSTRACT

We use the NIHAO galaxy formation simulations to make predictions for the baryonic budget in present day galaxies ranging from dwarf ( $M_{200} \sim 10^{10} M_{\odot}$ ) to Milky Way ( $M_{200} \sim 10^{12} M_{\odot}$ ) masses. The sample is made of 88 independent high resolution cosmological zoom-in simulations. NIHAO galaxies reproduce key properties of observed galaxies, such as the stellar mass vs halo mass and cold gas vs stellar mass relations. Thus they make plausible predictions for the baryon budget. We present the mass fractions of stars, cold gas ( $T < 10^4 \text{K}$ ), cool gas ( $10^4 < T < 10^5 \text{K}$ ), warm-hot gas ( $10^5 < T < 5 \times 10^6 \text{K}$ ), and hot gas ( $T > 5 \times 10^6 \text{K}$ ), inside the virial radius,  $R_{200}$ . Compared to the predicted baryon mass, using the dark halo mass and the universal baryon fraction,  $f_b \equiv \Omega_b/\Omega_m = 0.15$ , we find that all of our haloes are missing baryons. The missing mass has been relocated past 2 virial radii, and cool gas dominates the corona at low mass ( $M_{200} \lesssim 3 \times 10^{11} M_{\odot}$ ) while the warm-hot gas dominates at high mass ( $M_{200} \gtrsim 3 \times 10^{11} M_{\odot}$ ). Haloes of mass  $M_{200} \sim 10^{10} M_{\odot}$  are missing  $\sim 90\%$  of their baryons. More massive haloes ( $M_{200} \sim 10^{12} M_{\odot}$ ) retain a higher fraction of their baryons, with  $\sim 30\%$  missing, consistent with recent observational estimates. Moreover, these more massive haloes reproduce the observed fraction of cold, warm-hot and hot gas. The fraction of cool gas we predict (0.11) is significantly lower than the observation from COS-HALOs (0.3-0.47), but agrees with the alternative analysis of Stern et al. (2016).

**Key words:** galaxies: evolution – galaxies: formation – galaxies: dwarf – galaxies: spiral – methods: numerical – cosmology: theory

## 1 INTRODUCTION

Cosmic structure formation has redistributed the baryons from a nearly uniformly distributed plasma into a variety of states, including stars, stellar remnants, cold (atomic and molecular) gas, and hot (ionized) gas. The theories of galaxy formation can predict the amount of mass in these different states, which can in turn be tested by observational constraints.

On cosmological scales, the ratio between the total baryonic and gravitating mass is measured to be  $f_b \equiv \Omega_b/\Omega_m \simeq 0.15$  (The Planck Collaboration 2014). However, the cold baryonic mass density implied by several galaxy baryon estimates (mainly stars and cold gas) is only 3-8% of the big bang nucleosynthesis expectation (Persic & Salucci

1992; Fukugita et al. 1998; Bell et al. 2003; McGaugh et al. 2010). The majority of the cosmic baryons are thought to be in the form of hot gas around or between galaxies (Cen & Ostriker 1999). Until recently only a fraction of these baryons had been detected (Bregman 2007; Shull et al. 2012). This discrepancy is referred to as the “missing baryon problem”. Several theoretical studies with cosmological simulations have constrained the phase of the potential reservoirs of the missing baryons in the intergalactic medium (IGM), and find a large fraction of the baryons with low density and high temperature resides between galaxies (Yoshida et al. 2005; He et al. 2005; Davé et al. 2010; Zhu et al. 2011; Haider et al. 2016).

As a part of the IGM, the circum galactic medium (CGM) is always treated as an major potential reservoir of the missing baryons. HI and metal absorption lines are expected to signpost such diffuse baryonic

\* wangliang@pmo.ac.cn

content. HI is mainly from gas with temperature  $T \sim 10^4$  K so that it is able to detect cold gas in the CGM. Meanwhile, theoretical work has predicted that a substantial portion of the CGM is in the warm or hot phase with temperature  $T > 10^{4.5}$  K. Gas enters this phase through photoionization, accretion shocks or shocks caused by galactic winds (van de Voort & Schaye 2012). Such dilute halo gas is at  $T \sim 10^{4.5-7}$  K, so the detection is dominated by metal lines, e.g. OVI. Recent advances in the detection of gas in the CGM have come from the COS survey (Tumlinson et al. 2011, 2013; Thom et al. 2012; Werk et al. 2012, 2013). On the scale of Milky Way mass haloes  $M_{200} \sim 10^{12} M_{\odot}$  a significant amount of warm ( $10^4 < T < 5 \times 10^6$  K) gas has been detected (Werk et al. 2014), accounting for 33-88% of the baryon budget. In the future such observations will be extended to a wider range of halo masses. A number of large volume cosmological simulations (Ford et al. 2013; ?; Suresh et al. 2015; Oppenheimer et al. 2016) and zoom-in cosmological simulations (Stinson et al. 2012; Hummels et al. 2013; Shull 2014) have given predictions for the HI and OVI absorption line. Gutcke et al. (2016) compared the column density profile of OVI and HI in the CGM of galaxies from the NIHAO (Wang et al. 2015) cosmological hydrodynamical simulation suite with observations, studied the covering fraction of dense HI, looked at the shape of the CGM and its chemical composition. The simulations reproduce the observational covering fraction and column density profile of cool HI well, and recover the observed trends of OVI column density with luminosity and impact parameter. In common with other simulations, however, the column density of OVI is lower and the extent of optically thick HI is smaller than observed.

The physical properties of the CGM has been shown to be able to test feedback models (Sharma et al. 2012; Marasco et al. 2013). Davé (2009) predicted galactic halo baryon fractions of galaxies with halo masses ranging from  $10^{11} M_{\odot}$  to  $10^{13} M_{\odot}$  using cosmological hydrodynamical simulations with a well-constrained model for galactic outflows. They found that, without the outflow model, the baryon fraction inside the virial radius is roughly the cosmic baryonic fraction, but with the outflow model, the baryon fraction is increasingly suppressed in lower mass haloes. By comparing results at  $z = 3$  and  $z = 0$ , they showed that large haloes remove their baryons at early times while small haloes lose baryons more recently due to the wind material taking longer to return to low-mass galaxies than high-mass galaxies. Muratov et al. (2015) showed similar results that the gas and baryon fractions are lower at lower redshift, after powerful outflows at intermediate redshift  $z \approx 0.5 - 2$  remove a large amount of gas from the halo. Several simulations (Crain et al. 2010; Christensen et al. 2016; van de Voort et al. 2016) found the baryon fraction and gas fraction are reduced compared to the cosmic baryon fraction, especially in low mass haloes. Sokolowska et al. (2016) studied the halo gas of three Milky way-sized galaxies using cosmological zoom-in simulations. They found that most of missing baryons actually resides in warm-hot and hot gas which contribute to 80% of the total gas reservoir. The recovered baryon fraction within 3 virial radii is 90%. The warm-hot medium is sensitive to the feedback model so that a reliable spatial mapping of the warm-hot medium will provide a stringent test for feedback models.

In this paper we make predictions for the baryonic bud-

get for stars, cold, warm and hot gas in and around the virial radius of haloes of mass ranging from  $M_{200} \sim 10^{10} M_{\odot}$  to  $10^{12} M_{\odot}$ . We use a sample of 88 zoom-in galaxy formation simulations from the NIHAO project. Reproducing the stellar mass content in dark matter haloes both today and in the past has been a formidable challenge for cosmological galaxy formation simulations (Weinmann et al. 2012; Hopkins et al. 2014). Even the latest state-of-the art simulations have trouble: the ILLUSTRIS simulation (Vogelsberger et al. 2014) strongly overpredicts the stellar masses in dwarf galaxy haloes ( $M_{200} \lesssim 10^{11} M_{\odot}$ ), while the EAGLE simulations (Schaye et al. 2015) underpredict the peak of the star formation efficiency in halos of mass  $M_{200} \sim 10^{12} M_{\odot}$ . In contrast, the NIHAO galaxies are consistent with the stellar mass vs halo mass relations from halo abundance matching since redshift  $z \sim 4$  (Wang et al. 2015), the galaxy star formation rate vs stellar mass relation since  $z \sim 4$  (Wang et al. 2015), and the cold gas mass vs stellar mass relation at  $z \sim 0$  (Stinson et al. 2015). Therefore, the simulations make plausible predictions for the mass fractions and physical locations of the warm and hot gas components. We find that all the haloes contain less baryons than expected according to the cosmic baryonic fraction, but the missing fraction is strongly mass dependent.

This paper is organized as follows: The cosmological hydrodynamical simulations including star formation and feedback are briefly described in §2; In §3 we present the results including the baryonic budget, baryon distribution, and a comparison with observations; §4 gives a summary of our results.

## 2 SIMULATIONS

In this study we use simulations from the NIHAO (Numerical Investigation of a Hundred Astrophysical Objects) project (Wang et al. 2015). The initial conditions are created to keep the same numerical resolution across the whole mass range with typically a million dark matter particles inside the virial radius of the target halo at redshift  $z = 0$ . The halos to be re-simulated at higher resolution with baryons have been extracted from 3 different pure N-body simulations with a box size of 60, 20 and 15  $h^{-1}$  Mpc respectively. We adopted the latest compilation of cosmological parameters from the Planck satellite (the Planck Collaboration et al. 2014). **Dark matter particle masses range from  $\sim 10^3 M_{\odot}$  in our lowest mass haloes to  $\sim 10^5 M_{\odot}$  in our most massive haloes, and their force softenings range from  $\sim 100$  pc in low mass range to  $\sim 900$  pc in high mass range. Gas particles are less massive in factor of 5 correspondingly, and the force softenings are 3 times less than dark matters.** More information on the collisionless parent simulations, the force softenings and particle masses for the highest refinement level for each simulation and sample selection can be found in Dutton & Macciò (2014) and Wang et al. (2015).

We use the SPH hydrodynamics code GASOLINE (Wadsley et al. 2004), with a revised treatment of hydrodynamics as described in Keller et al. (2014). The code includes a subgrid model for turbulent mixing of metal and energy (Wadsley et al. 2008), heating and cooling include photoelectric heating of dust grains, ultraviolet (UV) heating and

ionization and cooling due to hydrogen, helium and metals (Shen et al. 2010).

The star formation and feedback modeling follows what was used in the MaGICC simulations (Stinson et al. 2013). The gas is converted into stars according to the Kennicutt-Schmidt Law when it satisfies a temperature and density threshold. Stars feed both metals and energy back into the ISM gas surrounding the region where they formed. SN feedback is implemented using the blastwave formalism described in Stinson et al. (2006). Pre-SN feedback is an attempt to consider radiation energy from massive stars. Heating is introduced immediately after massive stars form based on how much star light is radiated. Our simulations use thermal feedback to provide pressure support and increase gas temperature above the star formation threshold, and thus to decrease star formation. There are two small changes in NIHAO simulations compared to MaGICC: The change in number of neighbors and the new combination of softening length and particle mass increases the threshold for star formation from 9.3 to  $10.3 \text{ cm}^{-3}$ , the increase of pre-SN feedback efficiency  $\epsilon_{\text{ESF}}$ , from 0.1 to 0.13. More details on the star formation and feedback modeling can be found in Wang et al. (2015).

### 3 BARYON BUDGET

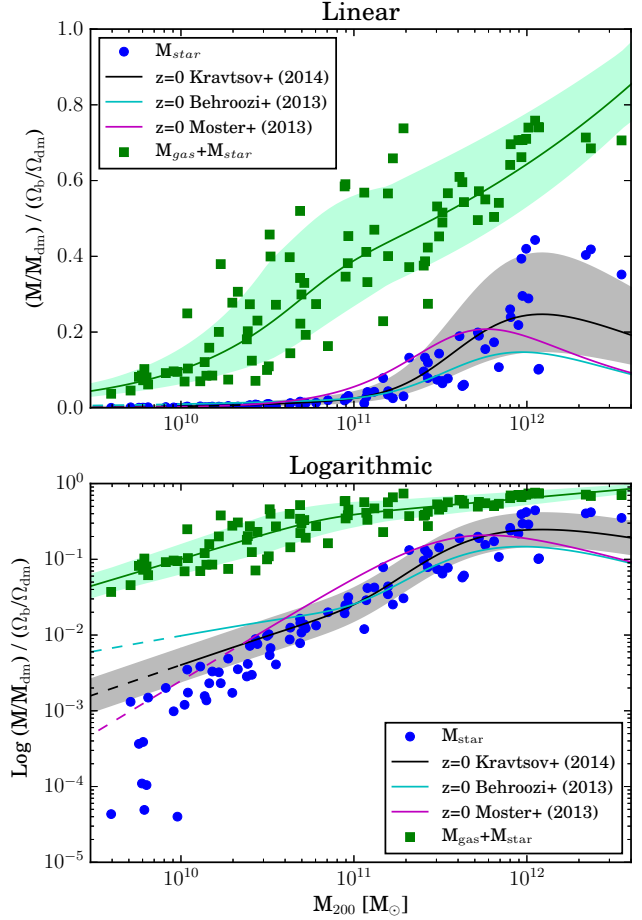
We define the fiducial baryonic mass as:

$$M_b \equiv M_b(R_{200}) = \frac{f_b}{1 - f_b} M_{\text{dm}}(R_{200}) \quad (1)$$

where the  $M_{\text{dm}}$  is the total dark matter mass of the halo, and the  $f_b = \Omega_b/\Omega_m \sim 0.15$  is the cosmic baryon fraction (the ratio between baryon density and mass density including baryonic mass plus dark matter), so that  $M_b$  would be the baryonic mass inside the virial radius if the baryons followed the dark matter closely.

Fig. 1 shows the ratio between the mass of each baryon component inside the virial radius to the fiducial baryonic mass for the most massive galaxy in each zoom-in region. We present the fractions of total stellar mass **within 20%**  $R_{\text{vir}}$  (blue points), and the total baryonic mass including stellar mass plus gas mass (green points). For the stellar mass fraction we also show the relations from the halo abundance matching (Moster et al. 2013; Behroozi et al. 2013; Kravtsov et al. 2014). The grey area is the one sigma scatter around the mean value from Kravtsov et al. (2014). Fig. 1 shows that all haloes in our study contain less than the universal fraction of baryons. The upper panel uses a linear y-axis, which highlights the large amount of baryons that are missing, especially in low mass haloes. The logarithmic scale in the lower panel highlights the power-law nature of the relations.

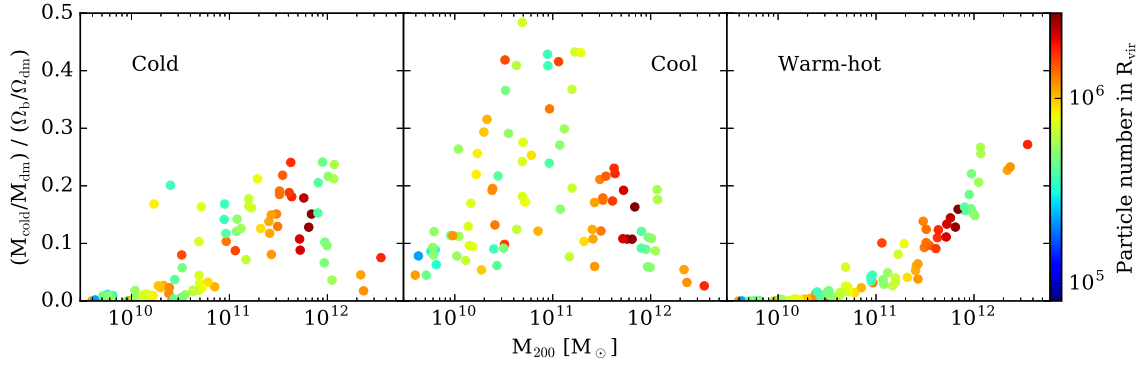
The trends of each component fraction are similar, in that the fractions are relatively low in low mass haloes, and increase as the halo mass increases. The main difference between the different components is the slope, with the baryonic mass fraction having a shallower slope than the stellar mass fraction. This is because in low mass haloes ( $M_{200} \sim 10^{10} M_\odot$ ) most of the baryons are in the form of gas, while in the highest mass haloes we study ( $M_{200} \sim 10^{12} M_\odot$ ) a substantial amount of gas has been turned into stars.



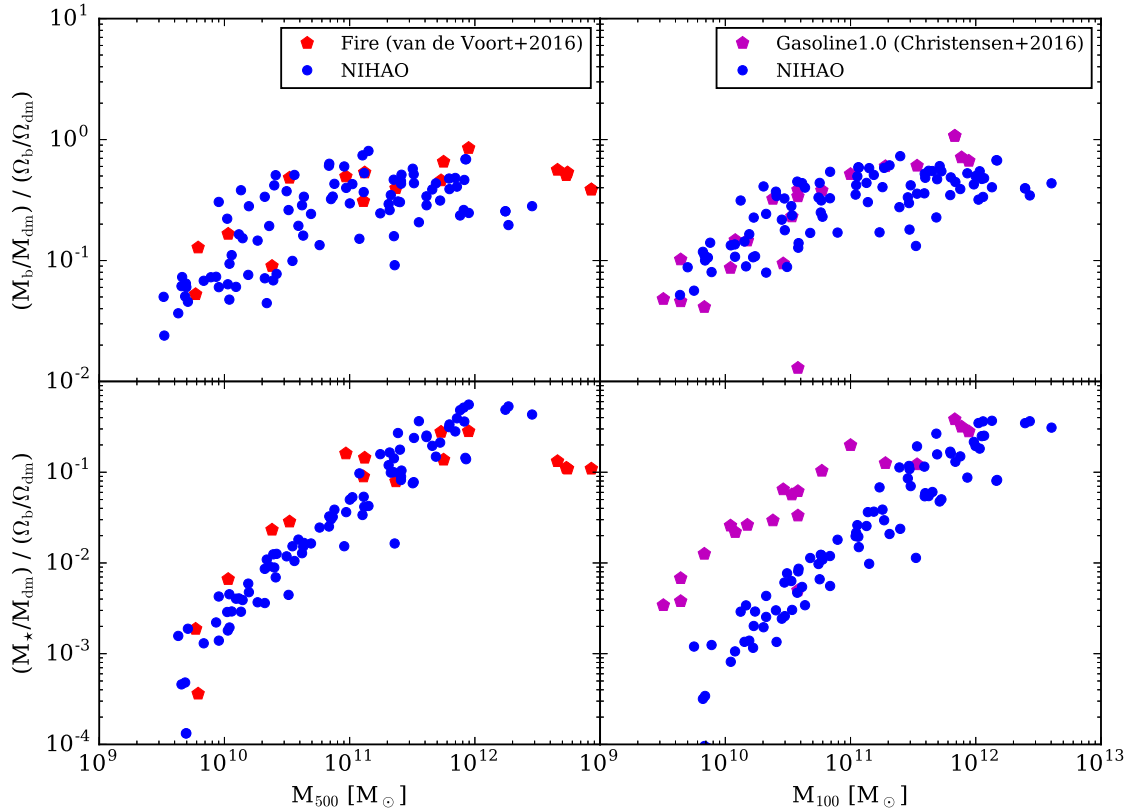
**Figure 1.** Fractional baryon content of our NIHAO simulations as a function of halo mass. The green points show the ratio between the baryonic mass (stars + gas) inside the virial radius and the total baryonic mass associated with the dark matter halo. The blue points show the corresponding fraction for the stars. The solid green line and shaded region shows a double power-law fit, together with the  $1\sigma$  scatter. For the stellar mass fraction we show several relations from halo abundance matching. The linear (upper panel) and logarithmic (lower panel) scales emphasize the large amount of “missing” baryons, and the power-law nature of the relations, respectively.

Since most of the haloes we study are above the mass where the cosmic UV background prevents gas from cooling, the missing baryons have most likely been ejected from the central galaxies in supernova/stellar feedback driven winds. Although the lower mass galaxies have converted a smaller fraction of their available baryons into stars, and hence there is proportionally less energy available to drive an outflow, they have expelled a larger fraction of their baryons. This is consistent with expectations from energy driven gas outflows, where the lower star formation efficiency is more than compensated by the shallower potential wells of lower mass halos according to the mass loading factor and circular velocity relation  $\eta \propto V^{-2}$  (e.g., Dutton 2012; Christensen et al. 2016).

The behavior of the baryonic mass fraction,  $f_{\text{bar}}$ , as a function of the halo mass is captured using a double power



**Figure 2.** Linear fractional baryon content of NIHAO simulations as a function of halo mass color coded by number of dark matter particles per halo.



**Figure 3.** Comparison between baryon (upper panels) and stellar fractions (lower panels) in the NIHAO simulations (blue circles) with simulations from van de Voort et al. (2016) (left, red pentagons) and Christensen et al. (2016) (right, purple pentagons).

law formula:

$$\frac{f}{f_0} = \left( \frac{M_{200}}{\mathcal{M}_0} \right)^\alpha \left\{ 0.5 \left[ 1 + \left( \frac{M_{200}}{\mathcal{M}_0} \right)^\gamma \right] \right\}^{\frac{\beta - \alpha}{\gamma}}. \quad (2)$$

In this formula, the lower and higher mass ends have logarithmic slope  $\alpha$  and  $\beta$ , respectively, while  $\gamma$  regulates how sharp the transition is from the lower to the higher ends. Giving all points equal weight, the best fit parameters are

as follows:

$$\begin{aligned} \mathcal{M}_0 &= 6.76 \times 10^{10} \\ f_0 &= 0.336 \\ \alpha &= 0.684 \\ \beta &= 0.205 \\ \gamma &= 3.40 \end{aligned} \quad (3)$$

The three most massive galaxies in the NIHAO suite have fairly high stellar masses compared to abundance matching results (see blue filled circles in Fig. 1), and are

thus possibly overcooled. However, they don't significantly bias the baryon fraction fitting formula since the high mass slope ( $M_{200} > 10^{11} M_{\odot}$ ) is constrained by  $\sim 40$  other haloes.

The green shaded region indicates the scatter about the best fit line, which is 0.151 dex for haloes with mass in the range of  $3 \times 10^9 M_{\odot} < M_{200} < 2 \times 10^{10} M_{\odot}$ , 0.236 for halo mass in  $2 \times 10^{10} M_{\odot} < M_{200} < 7 \times 10^{10} M_{\odot}$ , 0.125 for halo mass in  $7 \times 10^{10} M_{\odot} < M_{200} < 3 \times 10^{11} M_{\odot}$  and 0.052 for halo mass in  $3 \times 10^{11} M_{\odot} < M_{200} < 3.5 \times 10^{12} M_{\odot}$ .

As might be expected, haloes with the highest masses we study ( $\sim 10^{12} M_{\odot}$ ) have high baryon fractions ( $\sim 0.7$ ) with relatively small scatter (0.05 dex). As halo mass decreases, the baryon fraction decreases and the scatter increases reaching a maximum of  $\sim 0.24$  dex in haloes of mass  $\sim 4 \times 10^{10} M_{\odot}$ . However, at the lowest halo masses we study, below  $10^{10} M_{\odot}$ , the scatter starts to decrease. By contrast the scatter in the stellar mass fraction increases below this scale. Thus it seems unlikely that stellar feedback is primarily responsible for the low baryon fractions. Rather, we suggest an increased importance of the UV background, which heats gas to above the virial temperature, thus preventing it from collapsing into the low mass haloes. The baryon fractions **in low mass haloes** are thus controlled primarily by the halo masses, and thus are independent of the large scatter in the **stellar mass**.

Fig. 2 shows the budget fraction of cold, cool and warm-hot gas fraction. Since one can imagine that with increased resolution, higher densities can be achieved in the CGM, potential leading to more cooling and larger cold/cool gas fraction, all the points are color coded by the total number of dark matter particles within the haloes. In the plot, the number of dark matter particles per halo spans 2 magnitudes and has no correlation with cold and cool gas fraction. Therefore, the baryonic fractions predicted by NIHAO simulations are convergent.

### 3.0.1 Comparison with other zoom-in simulations

In Fig. 3 we compare the NIHAO results for the stellar and baryonic mass fractions with other recent state-of-the-art zoom-in simulations from van de Voort et al. (2016) and Christensen et al. (2016). Since these authors both use different halo mass definitions than we adopt here, we re-calculate the NIHAO results using a virial radius defined at  $500\times$  critical density (left panels) and  $100\times$  critical density (right panels). The simulations analyzed in van de Voort et al. (2016) are from the FIRE project (Hopkins et al. 2014), which uses a different hydrodynamical code and different sub-grid model for star formation and feedback. The simulations analyzed in Christensen et al. (2016) use a similar code as NIHAO, i.e., GASOLINE, but with important differences. NIHAO uses an upgraded version that improves mixing (Keller et al. 2014), and includes stronger feedback which was found to be necessary to delay star formation at early times (Stinson et al. 2013). Despite these different codes employed, the baryon fractions in the three sets of simulations are in remarkably good agreement.

For the massive haloes with  $M_{500} > 10^{12} M_{\odot}$ , there are three galaxies from NIHAO and four from FIRE. There is a hint that more massive haloes have lower baryon fractions. Of course the caveat here is that neither NIHAO or FIRE in-

cludes AGN feedback, which could have a significant impact on the gaseous content of these haloes.

For the stellar fractions, van de Voort et al. (2016) shows almost the same result as our work. Even at the low mass end, their two lowest mass galaxies may suggest the large scatter as found in NIHAO. However, Christensen et al. (2016) predicts one magnitude higher stellar fractions than what we find for halo masses below  $10^{11.5} M_{\odot}$ . This difference is likely due to the additional early stellar feedback included in the NIHAO version of gasoline.

In summary, while the stellar mass fractions are dependent on the sub-grid models, the baryon fractions (for haloes in the mass range  $10^{10} \lesssim M_{200} \lesssim 10^{12} M_{\odot}$ ) appear insensitive to the details of the simulation code. In particular all three codes predict a greater fraction of missing baryons in lower mass haloes. This suggests that the potential well of the halo **correlates the baryon fraction more closely** than the strength of the feedback.

### 3.1 Mass budget of the corona

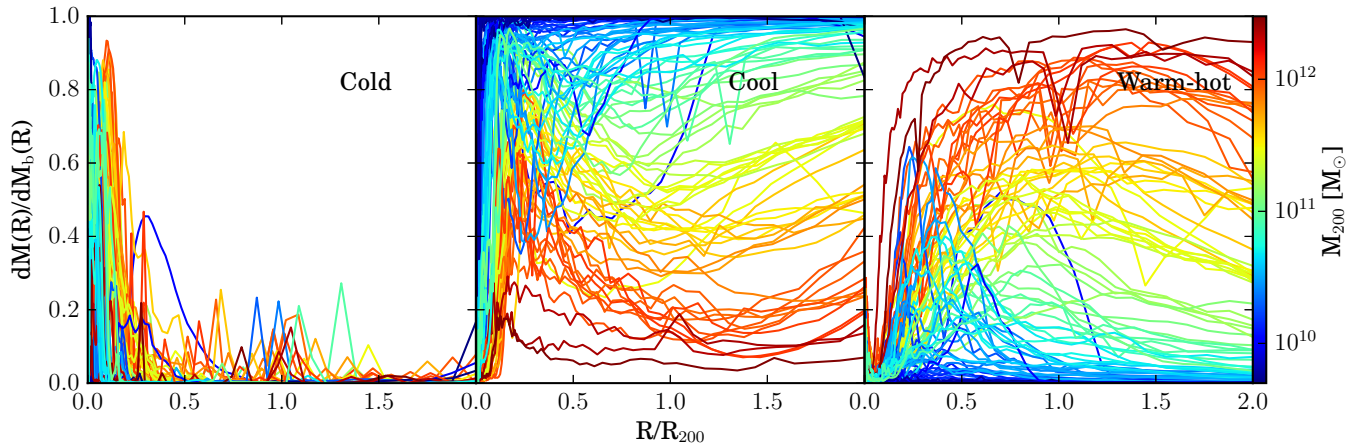
In Fig. 4, we present the radial distribution of gas in different phases at  $z = 0$ , normalized to the total baryonic mass profile, such that, in the region far away from the central galaxy where the stars are rare, for a given halo the four phases add up to unity, each line is colour coded by the virial mass. All simulations share a common attribute. The cold gas ( $T < 10^4 K$ ) is mostly located near the center ( $R < 0.2 R_{200}$ ) where most stars in galaxies form. In contrast, the cool ( $10^4 K < T < 10^5 K$ ) and warm-hot ( $10^5 K < T < 5 \times 10^6 K$ ) gas are located at large distances with roughly constant fractions up to 2 times  $R_{200}$ . The hot gas ( $T > 5 \times 10^6 K$ ) is a minority component for all galaxies in the NIHAO sample, with the maximum hot gas fraction at any radius being less than 5%.

Despite these similarities, we find a considerably higher proportion of cool gas in lower mass galaxies ( $M_{200} < 10^{11} M_{\odot}$ ) in the whole corona region. For higher mass galaxies, warm-hot gas dominates the corona which signals stronger virial shocks and higher efficiency of feedback. Even beyond the virial radius, the cool and warm-hot gas has similar features as the gas within virial radius which reveals the gas surrounding galaxies within large distances is the major reservoir of baryons.

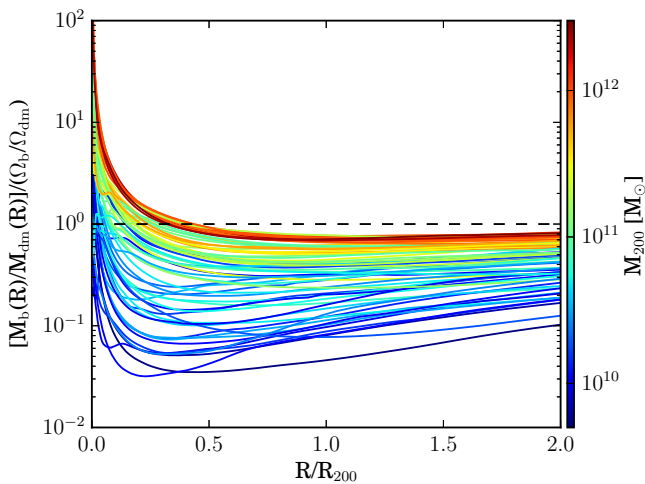
### 3.2 Where are the missing baryons?

Fig. 5 shows the cumulative fraction of total baryons for each simulation. Here the y-axis is the ratio between the baryonic to dark matter mass,  $M_b(< R)/M_{dm}(< R)$ , enclosed within a sphere of radius,  $R$ , normalized by the cosmic baryon-to-dark matter ratio,  $\Omega_b/\Omega_{dm}$ .

Each solid curve represents a halo, and the curves are coloured by their halo mass (red for high masses to blue for low masses). Broadly speaking, the curves have a similar shape, with a normalization that depends on halo mass. They have a cusp in the central region where the stars and cold gas dominate, then become flat in the outer region. More massive haloes have higher baryon fractions at all radii. At small radii, the baryon to dark matter ratio is higher than the cosmic value due to gas dissipation. Beyond



**Figure 4.** Radial profile of the mass fraction of the gas in each phase to total baryonic mass in each radial bin at  $z = 0$  for all galaxies in NIHAO sample. Each solid line is from one galaxy and colour coded with the halo mass.



**Figure 5.** Baryon distribution of each galaxy from NIHAO simulations. The lines are colour coded by their halo mass, which shows a clear trend that the more massive haloes preserve more baryons at all radii.

0.5 virial radii, all haloes are missing baryons. Even beyond the virial radius, there is little change in the baryon fraction up to 2 virial radii. We thus conclude that the missing baryons are well outside of the virial radius.

Ford et al. (2016) compare results for the cumulative baryon profile from two cosmological hydrodynamic simulations which employed different prescriptions for galactic outflow models. They constructed samples of simulated galaxies with a similar distribution of stellar masses to that of COS-Halos.

In the hybrid energy/momentum driven winds model (“ezw”) 65% of all available baryons are inside the halo. This is broadly consistent with estimates of baryonic mass derived from Werk et al. (2014) and our finding in Fig. 1. Comparing the cumulative baryon fraction profile in detail, Fig. 10 in Ford et al. (2016) shows the baryon fraction within  $0.1R_{\text{vir}}$  is only  $\sim 30\%$  and the fraction gradually increases to 65% at  $R_{\text{vir}}$ . In the NIHAO simulations the profile of galaxies with

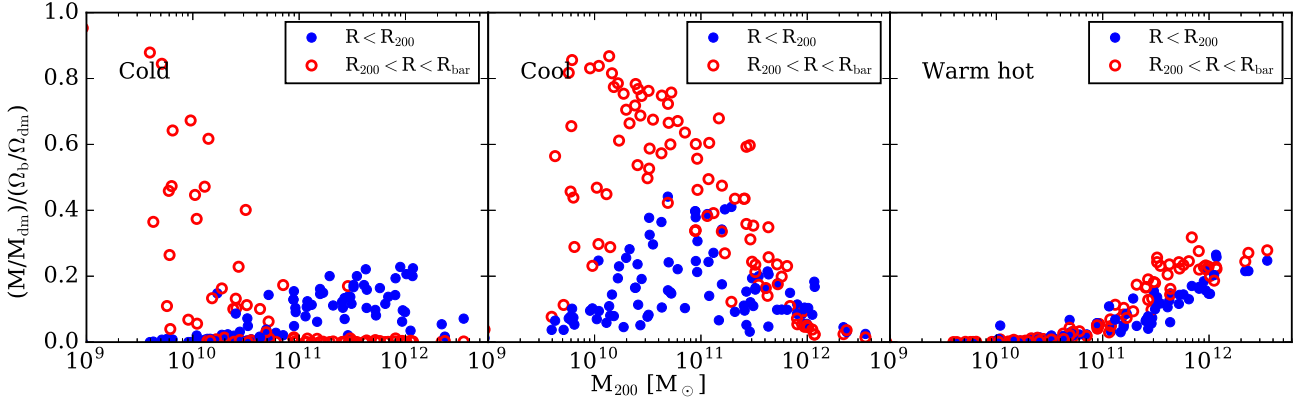
halo masses below  $10^{11}M_{\odot}$  in Fig. 5 have similar features, the more massive galaxies all have roughly flat slopes.

The simplified constant wind outflow model (“cw”) shows a lower fraction at all radii inside  $R_{\text{vir}}$ , even though the “ezw” and “cw” models generally gives similar observational absorption line properties. This suggests that the cumulative baryon fraction profile is complementary to the total amount of CGM gas for distinguishing between competing outflow models.

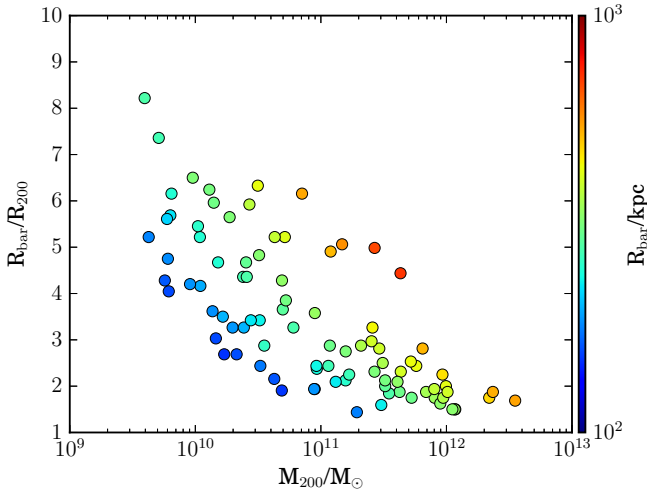
To estimate how far the baryons escape, we measured the radius,  $R_{\text{bar}}$ , within which the total baryon mass equals the fiducial baryonic mass defined by Eq. 1. This is a lower limit to the true extent of the missing baryons since the baryon mass includes gas and stars that belong to nearby lower mass haloes, **and is a radius within the effective volume of the simulations since contamination fraction in a shell between 0.9 and 1.0  $R_{\text{bar}}$  shows most NIHAO galaxies have a fraction close to 0, with largest only around 10 - 20.** Fig. 6 shows the baryon radius of each galaxy as function of the virial mass. In physical units, we find that the baryon radius generally increases with virial mass. When normalized by the virial radius, the distance baryons are ejected gradually decreases as halo mass increases, varying from  $R_{\text{bar}}/R_{200} \sim 5$  at a halo mass of  $M_{200} \sim 10^{10}M_{\odot}$  to  $\sim 2$  at a halo mass of  $M_{200} \sim 10^{12}M_{\odot}$ .

In Fig. 7, we show the mass fractions of gas (in three phases) inside the virial radius (filled blue points) and between the virial and baryon radius (open red points) **which are relative to the fiducial baryonic mass within the virial radius.** The fractions of cold gas,  $f_{\text{cold}}$ , are shown in the left panel. Inside the virial radius  $f_{\text{cold}}$  increases gradually from zero at a halo mass of  $10^{10}M_{\odot}$  to  $\sim 20\%$  at a halo mass of  $10^{12}M_{\odot}$ . The gas outside the virial radius has the opposite and much stronger trend:  $f_{\text{cold}} \sim 50\%$  in haloes of mass  $10^{10}M_{\odot}$  and decreases to zero by halo masses of  $10^{11}M_{\odot}$ . The fractions of cool gas,  $f_{\text{cool}}$ , are shown in the middle panel. Inside the virial radius,  $f_{\text{cool}}$ , has a maximum of 40% at a halo mass of  $10^{11}M_{\odot}$ , and declines to less than 10% below and above halo masses of  $10^{10}M_{\odot}$  and  $10^{12}M_{\odot}$ , respectively. For most haloes there is more cool gas outside than inside the virial radius. In haloes of mass  $10^{10}M_{\odot}$ ,





**Figure 7.** Mass fraction of gas in three phases (relative to the fiducial baryonic mass within the virial radius) inside the virial radius (filled blue points) and between virial and baryon radius (open red points), respectively. Cool gas is the dominant component of the fiducial baryonic mass for most galaxies ( $10^{10} < M_{200}/M_{\odot} < 10^{12}$ ). Cold and warm-hot gas is the majority only for galaxies at the low and high mass ends, respectively.



**Figure 6.** Normalized baryon radius as function of total virial mass. The points are colour coded by baryon radius of each galaxy.

$f_{\text{cool}} \sim 80\%$ , and decreases to less than 10% by a halo mass of  $10^{12}M_{\odot}$ . The fractions for warm-hot gas,  $f_{\text{warm}}$ , are shown in the right panel. The trends of the gas inside and outside the virial radius are quite similar,  $f_{\text{warm}}$  increases monotonically with halo mass with maximum values of  $\sim 30\%$ . The hot gas isn't shown since it is negligible both inside and outside the virial radius across the whole mass range we study. We thus conclude that, for galaxies with halo masses in the range  $10^{10} \lesssim M_{200} \lesssim 10^{11}M_{\odot}$ , the majority of baryons associated with the dark matter halo are in the cold and cool phases, and are located well outside of virial radius. For haloes in the mass range  $10^{11} \lesssim M_{200} \lesssim 10^{12}M_{\odot}$ , the fractions of cold gas, cool gas and warm-hot gas are comparable.

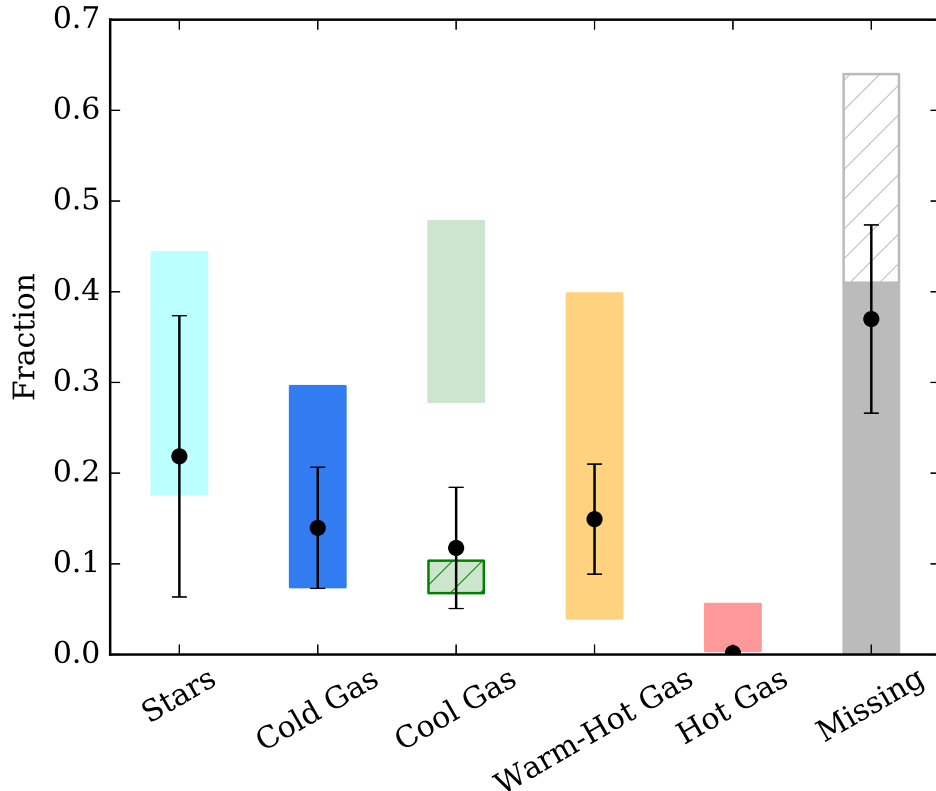
### 3.3 Comparison with observations of Milky Way mass haloes

Observations can gain information of CGM from absorption and emission lines. Although emission lines allow us

to directly obtain a 3D picture of the distribution of gas in the CGM, emission line studies preferentially probe the dense gas closer to galaxies, since gas emissivity scales with the square of density. While the situation is improving with new facilities, e.g. Hayes et al. (2016), absorption lines are the most common observational constraints on the physical state of the CGM.

The COS-HALOs survey is filling in details about the CGM at low redshift (Peeples et al. 2014; Tumlinson et al. 2011, 2013; Werk et al. 2012, 2013, 2014). For the CGM of low-redshift  $\sim L^*$  galaxies ( $M_{\text{star}} \sim 10^{10.5}M_{\odot}$ ), Tumlinson et al. (2013) and Peeples et al. (2014) constrain the mass of the warm-hot CGM ( $T \sim 10^{5-6.7}\text{K}$ ), Werk et al. (2014) provides a strict lower limit to the mass of cool material ( $T \sim 10^{4-5}\text{K}$ ) in the CGM of these galaxies. In a study using X-rays, Anderson et al. (2013) place a constraints on the mass of hot gas ( $T > 5 \times 10^6\text{K}$ ) residing in the extended hot halos. These observational constraints are shown in Fig. 8 with the same colour scheme as in Fig.11 of Werk et al. (2014). The stellar mass fractions are based on halo abundance matching as described in Kravtsov et al. (2014), while the cold disk gas mass comes from Dutton et al. (2011). In addition, since Crain et al. (2010) and van de Voort & Schaye (2013) show that soft X-ray emission traces gas with temperatures a few times the virial temperature ( $T_{\text{vir}} = 10^6\text{K}$  for  $M_{200} \sim 10^{12}M_{\odot}$ ), we also estimate the hot gas fraction with lower boundaries ( $3 \times 10^6\text{K}$  and  $10^6\text{K}$ ) between the hot gas and warm-hot gas, the fractions can increase to 0.00511 ( $3 \times 10^6\text{K}$ ) and 0.04168 ( $10^6\text{K}$ ). The mean value of missing fraction is the residual fraction of the sum of mean values of all components fractions in unity and the lower and upper limits are calculated in the same manner by the upper and lower limits of the all fractions.

The blue points and error bars show the mean values and standard deviation of the mass fraction of stars and different components of gas in our most massive galaxies ( $3.49 \times 10^{11}M_{\odot} < M_{200} < 3.53 \times 10^{12}M_{\odot}$ ). The gas is assigned to a range of temperature bins: cold gas ( $T < 10^4\text{K}$ ), cool gas ( $10^4\text{K} < T < 10^5\text{K}$ ), warm gas ( $10^5\text{K} < T < 5 \times 10^6\text{K}$ ) and hot gas ( $T > 5 \times 10^6\text{K}$ ). The observa-



**Figure 8.** Baryonic budget of NIHAO haloes of mass  $3.5 \times 10^{11} < M_{200}/M_{\odot} < 3.5 \times 10^{12}$  (blue points with  $1\sigma$  error bars) compared with observations of  $M_{200} \sim 10^{12} M_{\odot}$  haloes (shaded regions). There is good agreement, except for the cool gas which has two conflicting measurements: Werk et al. (2014, upper) upper, Stern et al. (2016, lower).

tions and the simulations match well in every phase except the cool CGM gas, where the observations from Werk et al. (2014) find  $3\times$  the mass that NIHAO simulations predict.

If the observations are correct, the simulations have either ejected cool gas too far, or they have created a CGM with the wrong mix of gas temperatures. The total gas fractions (0.39 in COS-HALOs, 0.41 in NIHAO) suggest the latter option. However, Stern et al. (2016) developed a new method to constrain the physical conditions in the cool CGM from measurements of ionic columns densities. This new method combines the information available from different sight-lines during the photoionization modeling, and was applied to the COS-HALOs data yielding a total cool CGM mass within the virial radius of  $1.3 \times 10^{10} M_{\odot}$ . The corresponding cool gas fraction is shown by the green hashed bar in Fig 1 and is in good agreement with our prediction. We should note that Stern et al. (2016) does not follow a same assumption on CGM ionization that assume all OVI is produced through photoionization and no collisional ionization, however, Gutcke et al. (2016) finds the CGM in  $L^*$  galaxies in NIHAO is mostly ionized by collisional ionization.

As the Fig. 4 and Fig. 7 show, the cool gas is the most important component for most haloes in the NIHAO simulations ( $10^{10} < M_{200}/M_{\odot} < 10^{12}$ ). The cold and warm-hot gas only dominate for galaxies at the lower and higher mass ends. As the CGM of lower mass galaxies will soon be observed, Table 1 lists information about CGM mass fractions of the different components of gas in haloes down to a halo mass of  $\sim 10^{10} M_{\odot}$ .

A recent estimation for the baryonic fractions from a set of eight Milky Way-sized zoom-in cosmological simulations (Colin et al. 2016) is generally **consistent** with our results. The fraction of cool gas from their simulations is even lower ( $f_{\text{cool}} = 0.034$ ) than ours and the result from Stern et al. (2016). On the other hand, Peebles et al. (2014) gave a conservative observational estimate about warm-hot gas and showed that the fraction of gas in this phase is only 5%, which is the lower limit of the estimates from Werk et al. (2014) and less than the average values from Colin et al. (2016) ( $f_{\text{warm-hot}} \approx 24\%$ ) and our finding ( $f_{\text{warm-hot}} \approx 17\%$ ). Therefore, the mix of temperatures of the CGM has large uncertainties. More accurate knowledge of the physical properties of the CGM are necessary to better constrain feedback models and to understand the role of the CGM in galaxy formation.

#### 4 SUMMARY

We have used the NIHAO galaxy simulation suite (Wang et al. 2015) to study the statistical features of the baryonic budget and distribution spanning halo masses of  $\sim 10^{10}$  to  $\sim 10^{12} M_{\odot}$ . NIHAO is a large (currently 88) set of high resolution cosmological zoom-in hydrodynamical galaxy formation simulations. As shown in previous papers the NIHAO galaxies reproduce several key observed scaling relations, and thus they make plausible predictions for the baryon



budget in and around galaxies. We summarize our results as follows:

- All of the NIHAO haloes have a lower baryon to dark matter ratio, inside the virial radius, than the cosmic baryon fraction (Fig. 1). We refer to the cosmic baryon fraction associated with each dark matter halo as the fiducial baryons.
- Lower mass haloes are missing a larger fraction of their fiducial baryons, even though they convert a much lower fraction of the baryons into stars (Fig. 1). Similar trends are found by other recent simulations (Christensen et al. 2016; van de Voort et al. 2016) using different codes (Fig. 3).
- The missing baryons have been expelled well beyond the virial radius,  $R_{200}$ , (Fig. 5). Relative to the virial radius, the baryons are expelled to smaller radii in more massive haloes:  $R_{\text{bar}} \sim 5R_{200}$  for  $M_{200} = 10^{10} M_{\odot}$  and  $R_{\text{bar}} \sim 2R_{200}$  for  $M_{200} = 10^{12} M_{\odot}$  (Fig. 6).
- Cold gas ( $T < 10^4 \text{K}$ ) is mostly restricted to be within 0.2 virial radii (Fig. 4). Cool gas ( $10^4 < T < 10^5 \text{K}$ ) dominates the baryonic mass outside the virial radius, **as well as outside the 20%  $R_{\text{vir}}$** , at low masses ( $M_{200} \lesssim 3 \times 10^{11} M_{\odot}$ ) while the warm-hot gas ( $10^5 < T < 5 \times 10^6 \text{K}$ ) dominates at high masses ( $M_{200} \gtrsim 3 \times 10^{11} M_{\odot}$ ) (Figs. 4 & 7).
- For the highest mass haloes in our study  $\sim 10^{12} M_{\odot}$  our simulations are consistent with the observed fractions (e.g. Werk et al. 2014) of stars, cold gas, warm and hot gas inside the virial radius (Fig. 8).
- For the cool gas we predict  $f_{\text{cool}} = 0.11 \pm 0.06$  which is significantly lower than the observations from COS-HALOs ( $f_{\text{cool}} = 0.28 - 0.48$ ), but is in excellent agreement with the analysis of Stern et al. (2016).

We hope our results will motivate observers to obtain more accurate measurements of the mass fractions in different phases of the CGM over a wide range of galaxy masses, and simulators to make the corresponding predictions.

## ACKNOWLEDGMENTS

We thank the two anonymous referees whose suggestions greatly improve the paper. We thank Freeke van de Voort for kindly sharing data of the FIRE simulation. GASOLINE was written by Tom Quinn and James Wadsley. Without their contribution, this paper would have been impossible. The simulations were performed on the THEO cluster of the Max-Planck-Institut für Astronomie and the HYDRA cluster at the Rechenzentrum in Garching; and the Milky Way supercomputer, funded by the Deutsche Forschungsgemeinschaft (DFG) through Collaborative Research Center (SFB 881) "The Milky Way System" (subproject Z2), hosted and co-funded by the Jülich Supercomputing Center (JSC). We greatly appreciate the contributions of all these computing allocations. AAD, GSS and AVM acknowledge support through the Sonderforschungsbereich SFB 881 The Milky Way System (subproject A1) of the German Research Foundation (DFG). The analysis made use of the pynbody package (Pontzen et al. 2013). The authors acknowledge support from the MPG-CAS through the partnership programme between the MPIA group lead by AVM and the PMO group lead by XK. LW acknowledges support of the MPG-CAS student programme. XK acknowledge the support from 973 program (No. 2015CB857003, 2013CB834900), NSFC

project No.11333008 and the "Strategic Priority Research Program the Emergence of Cosmological Structures" of the CAS(No.XD09010000).

## REFERENCES

- Anderson, M. E., Bregman, J. N., Dai, X. 2013, ApJ, 762, 106
- Agertz, O., Moore, B., Stadel, J. 2007, MNRAS, 380, 963
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57
- Bell, E. F., McIntosh, D. H., Katz, N., Weinberg, M. D., 2003, ApJ, 585, 117
- Bregman, J. N. 2007, ARAA, 45, 221
- Cen, R. Y., Ostriker, J. P. 1999, ApJ, 514, 1
- Christensen, C. R., Davé, R., Governato, F., et al. 2016, ApJ, 824, 57
- Colin, P., Avila-Reese, V., Roca-Fabrega, S., et al. 2016, ApJ, in press
- Crain, R. A., McCarthy, I. G., Frenk, C. S., et al. 2010, MNRAS, 470, 1403
- Davé, R. 2009, ASPC, 419, 347D
- Davé, R., Oppenheimer, B. D., Katz, N., et al. 2010, MNRAS, 408, 2051
- Dutton, A. A., Conroy, C., van den Bosch, F. C., et al. 2011, MNRAS, 416, 322
- Dutton, A. A. 2012, MNRAS, 424, 3123
- Dutton, A. A., & Macciò, A. V. 2014, MNRAS, 441, 3359
- Ford, A. B., Oppenheimer, B. D., Davé, R., et al. 2013, MNRAS, 432, 89
- Ford, A. B., Werk, J. W., Davé, R., et al. 2016, MNRAS, 459, 1745
- Fukugita, M., Hogan, C. J., Peebles, P. J. F. 1998, ApJ, 503, 518
- Gutcke, T. A., Stinson, G. S., Macciò, A. V., et al. 2016, arXiv:1602.06956
- Haider, M., Steinhauser, D., Vogelsberger, M., et al. 2016, MNRAS, 457, 3024
- Hayes, M., Melinder, J., Östlin, G., et al. 2016, arXiv:1606.04536
- He, P., Feng, L. L., Fang, L. Z. 2005, ApJ, 623, 601
- Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581
- Hummels, C. B., Bryan, G. L., Smith, B. D., et al. 2013, MNRAS, 430, 1548
- Keller, B. W., Wadsley, J., Benincasa, S. M., & Couchman, H. M. P. 2014, MNRAS, 442, 3013
- Kravtsov, A., Vikhlinin, A., & Meshcheryakov, A. 2014, arXiv:1401.7329
- Marasco, A., Marinacci, F., Fraternali, F. 2013, MNRAS, 433, 1634
- McGaugh, S. S., Schombert, J. M., de Blok, W. J. G., Zargursky, M. J. 2010, MNRAS, 408, 14
- Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121
- Muratov, A. L., Kerés, D., Faucher-Giguère, C., et al. 2015, MNRAS, 454, 2691
- Oppenheimer, B. D., Crain, R. A., Schaye, J., et al. 2016, arXiv:1603.05984
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, AA16

**Table 1.** The baryonic budget parameters for NIHAO galaxies in different halo mass bins. We refer to gas in the temperature range  $T < 10^4$  K as cold;  $10^4$  K  $\leq T < 10^5$  K as cool;  $10^5$  K  $\leq T < 5 \times 10^6$  K as warm; and  $T \geq 5 \times 10^6$  K as hot.

$\langle \log_{10}(M_{200}/M_{\odot}) \rangle$	$9.992 \pm 0.223$	$10.588 \pm 0.143$	$11.267 \pm 0.209$	$11.963 \pm 0.278$
$\langle \log_{10}(M_b/M_{\odot}) \rangle$	$9.243 \pm 0.215$	$9.829 \pm 0.139$	$10.493 \pm 0.206$	$11.176 \pm 0.273$
$\langle M_{\star}/M_b \rangle$	$1.61 \times 10^{-3}$	$9.57 \times 10^{-3}$	$6.43 \times 10^{-2}$	0.240
$\sigma_{\star}$	$1.63 \times 10^{-3}$	$4.24 \times 10^{-3}$	$5.02 \times 10^{-2}$	0.156
$\langle M_{\text{cold}}/M_b \rangle$	$1.37 \times 10^{-2}$	$3.92 \times 10^{-2}$	0.124	0.130
$\sigma_{\text{cold}}$	$8.24 \times 10^{-3}$	$4.08 \times 10^{-2}$	$3.58 \times 10^{-2}$	$7.78 \times 10^{-2}$
$\langle M_{\text{cool}}/M_b \rangle$	0.103	0.192	0.203	0.109
$\sigma_{\text{cool}}$	$7.56 \times 10^{-2}$	0.135	0.138	$6.37 \times 10^{-2}$
$\langle M_{\text{warm}}/M_b \rangle$	$7.92 \times 10^{-4}$	$1.48 \times 10^{-2}$	$5.83 \times 10^{-2}$	0.167
$\sigma_{\text{warm}}$	$1.25 \times 10^{-3}$	$6.53 \times 10^{-3}$	$3.19 \times 10^{-2}$	$6.05 \times 10^{-2}$
$\langle M_{\text{hot}}/M_b \rangle$	0.000	0.000	0.000	$1.82 \times 10^{-4}$
$\sigma_{\text{hot}}$	0.000	0.000	0.000	$2.33 \times 10^{-4}$
$\langle M_{\text{missing}}/M_b \rangle$	0.880	0.744	0.548	0.350
$\sigma_{\text{missing}}$	$5.65 \times 10^{-2}$	0.137	0.110	$9.66 \times 10^{-2}$

Peebles, M. S., Werk, J. K., Tumlinson, J., et al. 2014, *ApJ*, 786, 54

Persic, M., Salucci, P. 1992, *MNRAS*, 258, 14

Pontzen, A., Roškar, R., Stinson, G., & Woods, R. 2013, *Astrophysics Source Code Library*, 1305.002

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, 446, 521

Shen, S., Wadsley, J., & Stinson, G. 2010, *MNRAS*, 407, 1581

Sharma, P., McCourt, M., Parrish, I. J., Quataert, E. 2012, *MNRAS*, 427, 1219

Shull, J. M., Smith, B. D., Danforth, C. W. 2012, *ApJ*, 759, 23

Shull, J. M. 2014, *ApJ*, 784, 142

Sokolowska, A., Mayer, L., Babul, A., Madau, P., Shen, S. 2016, *ApJ*, 819, 21

Stern, J., Hennawi, J. F., Prochaska, J. X., Werk, J. K. 2016, *arXiv:1604.02168*

Stinson, G. S., Seth, A., Katz, N., et al. 2006, *MNRAS*, 373, 1074

Stinson, G. S., Brook, C., Prochaska, J. X., et al. 2012, *MNRAS*, 425, 129

Stinson, G. S., Brook, C., Macciò, A. V., et al. 2013, *MNRAS*, 428, 129

Stinson, G. S., Dutton, A. A., Wang, L., et al. 2015, *MNRAS*, 454, 1105

Suresh, J., Rubin, K. H. R., Kannan, R., et al. 2015, *arXiv:1511.00687*

Thom, C., Tumlinson, J., Werk, J. K. 2012, *ApJL*, 758, L41

Tumlinson, J., Thom, C., Werk, J., et al. 2011, *Science*, 334, 948

Tumlinson, J., Thom, C., Werk, J., et al. 2013, *ApJ*, 777, 59

van de Voort, F., Schaye, J. 2012, *MNRAS*, 423, 2991

van de Voort, F., Schaye, J. 2013, *MNRAS*, 430, 2688

van de Voort, F., Quataert, E., Hopkins, P. F., et al. 2016, *arXiv:1604.01397*

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *MNRAS*, 444, 1518

Wadsley, J. W., Stadel, J., & Quinn, T. 2004, *NewA*, 9, 137

Wadsley, J. W., Veeravalli, G., & Couchman, H. M. P. 2008, *MNRAS*, 387, 427

Wang, L., Dutton, A. A., Stinson, G. S., et al. 2015, *MNRAS*, 454, 83

Weinmann, S. M., Pasquali, A., Oppenheimer, B. D., et al. 2012, *MNRAS*, 426, 2797

Werk, J. k., Prochaska, J. X., Thom, C., et al. 2012, *ApJS*, 198, 3

Werk, J. k., Prochaska, J. X., Thom, C., et al. 2013, *ApJS*, 204, 17

Werk, J. k., Prochaska, J. X., Thom, C., et al. 2014, *ApJ*, 792, 8

Yoshida, N., Furlanetto, S. R., Hernquist, L. 2005, *ApJ*, 618L, 91

Zhu, W., Feng, L. L., Fang, L. Z. 2011, *MNRAS*, 415, 1093