MUFASA: Galaxy star formation, gas, and metal properties across cosmic time

Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 25 September 2018 (MN LATEX style file v2.2)

MUFASA: Galaxy star formation, gas, and metal properties across cosmic time

Romeel Davé[1,2,3], Mika H. Rafieferantsoa[1,5], Robert J. Thompson[4,1], Philip F. Hopkins[6]

[1] *University of the Western Cape, Bellville, Cape Town 7535, South Africa*
[2] *South African Astronomical Observatories, Observatory, Cape Town 7925, South Africa*
[3] *African Institute for Mathematical Sciences, Muizenberg, Cape Town 7945, South Africa*
[4] *National Center for Supercomputing Applications, Champaign-Urbana, IL 61801*
[5] *Max-Planck-Institüt für Astrophysik, Garching, Germany*
[6] *California Institute of Technology, Pasadena, CA 91125*

25 September 2018

ABSTRACT

We examine galaxy star formation rates (SFRs), metallicities, and gas contents predicted by the MUFASA cosmological hydrodynamic simulations, which employ meshless hydrodynamics and novel feedback prescriptions that yield a good match to observed galaxy stellar mass assembly. We combine $50, 25$, and $12.5h^{-1}$Mpc boxes with a quarter billion particles each to show that MUFASA broadly reproduces a wide range of relevant observations, including SFR and specific SFR functions, the mass-metallicity relation, H I and H_2 fractions, H I (21 cm) and CO luminosity functions, and cosmic gas density evolution. There are mild but significant discrepancies, such as too many high-SFR galaxies, overly metal-rich and H I-poor galaxies at $M_* \gtrsim 10^{10} M_\odot$, and sSFRs that are too low at $z \sim 1-2$. The H I mass function increases by $\times 2$ out to $z \sim 1$ then steepens to higher redshifts, while the CO luminosity function computed using the Narayanan et al. conversion factor shows a rapid increase of CO-bright galaxies out to $z \sim 2$ in accord with data. Ω_{HI} and Ω_{H2} both scale roughly as $\propto (1+z)^{0.7}$ out to $z \sim 3$, comparable to the rise in H I and H_2 fractions. MUFASA galaxies with high SFR at a given M_* have lower metallicities and higher H I and H_2 fractions, following observed trends; we make quantitative predictions for how fluctuations in the baryon cycle drive correlated scatter around galaxy scaling relations. Most of these trends are well converged with numerical resolution. These successes highlight MUFASA as a viable platform to study many facets of cosmological galaxy evolution.

Key words: galaxies: formation, galaxies: evolution, galaxies: star formation, galaxies: abundances, galaxies: ISM, methods: numerical

1 INTRODUCTION

Observations of galaxy properties from today back to the early universe are improving at a remarkable pace, thanks to advancing multi-wavelength photometric and spectroscopic galaxy surveys. Progress has been particularly impressive in the near-infrared and longer wavelengths, which provides more robust constraints on stellar and metal content at high redshifts and gas content across all redshifts. Models for galaxy formation thus find it increasingly challenging to be able to reproduce such observations within a physically-motivated concordance cosmology framework.

Recent cosmological hydrodynamic simulations have been impressively successful at broadly reproducing key galaxy demographic observables over cosmic time (see Somerville & Davé 2015, and references therein). A primary benchmark used to test galaxy formation models is the observed galaxy stellar mass function (GSMF). Many modern simulations can now match this to within a factor of several over the majority of cosmic time and mass (Davé et al. 2013; Genel et al. 2014; Crain et al. 2015; Khandai et al. 2015; Davé, Thompson, Hopkins 2016; Kaviraj et al. 2016), which is typically within the range of current systematic uncertainties in the data. To do so, all cosmological-scale simulations incorporate heuristic models for feedback processes associated with star formation that suppress galaxy formation at the low-mass end, combined with feedback often associated with active galactic nuclei (AGN) that suppresses massive galaxy growth. However, the precise physical mechanisms invoked for feedback can vary substantially amongst simulations, despite their predicted GSMFs being similar. To

further test and discriminate between models, and thereby constrain the physical mechanisms giving rise to feedback, it is thus important to move beyond the GSMF and consider other aspects of galaxy demographics.

Advancing multi-wavelength observations have made impressive progress at characterising the gas and metal content of galaxies across cosmic time. Metallicity measures at higher redshifts have been aided by new near-IR spectroscopic capabilities that have enabled the same optical emission line measures used at low redshifts to be applied to $z \sim 2-3$ galaxies (Steidel et al. 2014; Sanders et al. 2015). Molecular gas contents have now been measured out to similar redshifts thanks to deep millimetre-wave data that can detect redshifted carbon monoxide (CO) emission lines (Geach et al. 2011; Tacconi et al. 2013). Direct measures of atomic gas (H I) remain confined to low redshifts ($z \lesssim 0.5$) as of yet owing to the sensitivity of current instruments (Delhaize 2013; Fernández et al. 2016), but the Square Kilometre Array (SKA) and its prescursors such as MeerKAT aim to probe H I out to $z \sim 1$ and beyond (e.g. Holwerda, Blyth, Baker 2012). These observations provide a direct glimpse into the gaseous fuel for star formation, as well as products of massive star formation as traced by chemical enrichment, hence they can more directly probe the baryon cycle of gaseous inflows and outflows that are viewed as being the central driver of cosmological galaxy evolution.

Cosmological galaxy formation simulations have utilised these observations to provide additional constraints on feedback mechanisms and other physical processes of galaxy formation (e.g. Vogelsberger et al. 2014; Schaye et al. 2015; Davé, Thompson, Hopkins 2016). For instance, the slope of the mass-metallicity relation strongly suggests that low-mass galaxies preferentially eject more of their gas in outflows versus forming it into stars (e.g. Finlator & Davé 2008). The high gas fraction in low-mass galaxies is likewise a reflection of strong outflows that prevents the gas from forming into stars (e.g. Davé, Finlator, & Oppenheimer 2011). This broadly agrees with the notion that low-mass galaxies must have stronger feedback in order to suppress the faint end of the GSMF (e.g. Somerville et al. 2008; Davé, Finlator, & Oppenheimer 2011). While these trends generally point towards a qualitatively similar picture (Somerville & Davé 2015), it remains highly challenging for a single model to quantitively reproduce all the relevant observed relations across a wide range of mass scales and cosmic epochs.

In this paper we present a further analysis on the suite of cosmological hydrodynamic simulations of galaxy formation using Gizmo, called the Mufasa simulations, introduced in Davé, Thompson, Hopkins (2016, hereafter Paper I). Mufasa uses updated state of the art feedback modules, including two-phase kinetic outflows with scalings taken from the FIRE simulations (Muratov et al. 2015), an evolving halo mass-based quenching scheme (Gabor & Davé 2015; Mitra, Davé, & Finlator 2015), 11-element chemical evolution, and molecular gas-based star formation (Krumholz & Gnedin 2011; Thompson et al. 2014). We run three volumes, each with 512^3 dark matter particles and 512^3 gas elements, having box sizes of 50, 25, and $12.5h^{-1}$Mpc, in order to cover halo masses from $\sim 10^{10}-10^{14}M_\odot$ and stellar masses from $\sim 10^7-10^{12}M_\odot$.

In Paper I we showed that Mufasa does an excellent job at reproducing the observed evolution of the GSMF over most of cosmic time. Here we compare Mufasa to a wider suite of observations encompassing galaxy SFRs, gas, and metal content, in order to quantitatively examine whether a model that accurately reproduces stellar mass growth can also match these independent properties. One significant discrepancy seen in Paper I was that specific SFRs (sSFRs) at $z \sim 1-2$ were well below observations, even though galaxy growth rates as measured by GSMF evolution seemed to be in accord with data. Here we further investigate this issue using SFR and sSFR functions over cosmic time. Since Mufasa directly tracks H_2 within galaxies using a sub-grid prescription (Krumholz & Gnedin 2011), we investigate H I and H_2 contents separately, along with their evolution. Galaxy metallicities provide a crucial barometer for feedback, so we compare our predictions to emerging observations out to Cosmic Noon. Simulations naturally predict that deviations from the mean galaxy scaling relations are correlated, in that galaxies at a given stellar mass that are high in SFR are also low in metallicity (Davé, Finlator, & Oppenheimer 2011) and gas content (Rafieferantsoa et al. 2015). Here we generalize this analysis across all quantites considered, showing that deviations from the mean relations in SFR, metallicity, H I, and H_2 versus M_* are all correlated, and we quantify these correlations.

Taken together, these results extend the overall success of the Mufasa simulations as a reasonably faithful reproduction of the real universe, thereby highlighting Mufasa's utility as a platform to study of the physics of galaxy evolution across cosmic time. This paper is outlined as follows: In §2 we briefly recap the key ingredients of the Mufasa simulations. §3 discusses predicted SFRs and sSFRs, §4 presents the mass-metallicity relation, and §5 shows gas fractions and gas mass functions. In §6 we quantify the second-parameter dependences of the scatter around key scaling relations. We summarize our findings in §7.

2 SIMULATION DESCRIPTION

We employ a modified version of the gravity plus hydrodynamics solver Gizmo (Hopkins 2015), which uses the Gadget-3 gravity solver (Springel 2005), along with the meshless finite mass (MFM) hydrodynamics solver. We use adaptive gravitational softening throughout for all particles (Hopkins 2015), with a minimum (Plummer-equivalent) softening length set to 0.5% of the mean interparticle spacing. For more details on these aspects as well as the feedback choices summarised below, see Paper I.

We include radiative cooling from primordial (non-equilibrium ionisation) and heavy elements (equilibrium ionisation) using the Grackle-2.1 chemistry and cooling library (The Enzo Collaboration 2014; Kim et al. 2014). A spatially-uniform photo-ionising background is assumed, namely the 2011 update of the determination in Faucher-Giguere, Kereš, & Ma (2009). Gas above a threshold density is assumed to have an equation of state given by $T \propto \rho^{1/3}$ (Schaye & Dalla Vecchia 2008), and for the primary run employed in this paper the threshold density is taken to be 0.13 cm^{-3}. Stars are formed using a molecular gas-based prescription following Krumholz, McKee, & Tumlinson (2009), which approximates the H_2 fraction based on the local density, the Sobolev approximation in which the

optical depth is given by $\rho/|\nabla\rho|$ where ρ is the particle's density, and the particle's metallicity scaled to solar abundance based on Asplund et al. (2009). We vary the assumed clumping factor with resolution, as described in Paper I.

Young stellar feedback is modeled using decoupled, two-phase winds. Winds are ejected stochastically, with a probability that is η times the star formation rate probability. The formula for η is taken to be the best-fit relation from the Feedback In Realistic Environments (FIRE) suite of zoom simulations Muratov et al. (2015), namely

$$\eta = 3.55\Big(\frac{M_*}{10^{10}M_\odot}\Big)^{-0.351}, \quad (1)$$

where M_* is the galaxy stellar mass determined using an on-the-fly friends-of-friends galaxy finder. The ejection velocity v_w scaling is also taken to follow that predicted by FIRE, but with a somewhat higher amplitude:

$$v_w = 2\Big(\frac{v_c}{200\ \mathrm{km\,s^{-1}}}\Big)^{0.12} v_c + \Delta v_{0.25}. \quad (2)$$

where v_c is the galaxy circular velocity estimated from the friends-of-friends baryonic mass, and $\Delta v_{0.25}$ accounts for the potential difference between the launch location and one-quarter of the virial radius where Muratov et al. (2015) measured the scalings from FIRE. Winds are also ejected with a random 30% fraction being "hot", namely at a temperature set by the difference between the supernova energy and the wind launch energy (if this is positive), with the remaining 70% launched at $\ll 10^4$K. Wind fluid elements are allowed to travel without hydrodynamic forces or cooling until such time as its relative velocity versus surrounding (non-wind) gas is less than 50% of the local sound speed, or alternatively if it reaches limits in density of 0.01 times the critical density for star formation, or a time given by 2% of the Hubble time at launch. We further include energy Type Ia supernovae (SNIa) and asymptotic giant branch (AGB) stars, implemented as a delayed component using stellar evolution as tracked by Bruzual & Charlot (2003) models with a Chabrier (2003) initial mass function (IMF). See Paper I for full details.

Chemistry is tracked for hydrogen, helium, and 9 metals: C, N, O, Ne, Mg, Si, S, Ca, and Fe, comprising over 90% of metal mass in the universe. Type II SN yields are taken from Nomoto et al. (2006), parameterised as a function of metallicity, which we multiply by 0.5 in order to more closely match observed galaxy metallicities. Type II yields are added instantaneously to every star-forming gas particles at every timestep, based on its current star formation rate. For SNIa yields, we employ the yields from Iwamoto et al. (1999), assuming each SNIa yields $1.4M_\odot$ of metals. For AGB stars, we employ enrichment as a function of age and metallicity from various sources as described in Oppenheimer & Davé (2008), further assuming a 36% helium fraction and a nitrogen yield of 0.00118. The enrichment, like the energy, is added from stars to the nearest 16 gas particles, kernel-weighted, following the mass loss rate as computed assuming a Chabrier (2003) IMF.

We note that ISM gas ejected from our simulated galaxies is done so without any modification to its metallicity. We do not employ a separate "metal loading factor" parameter (i.e. the metallicity of the ejected gas relative to the ISM metallicity) which preferentially ejects enriched (or de-enriched, as in Illustris; Vogelsberger et al. 2014) ISM material; in other words, we assume a metal loading factor of unity. The physical justification for this is that, particularly in low mass galaxies where the mass loading factor η is high, direct supernovae ejectae represent only a very small portion of the total outflowing material, hence it makes sense that the outflow metallicity is dominated by ambient ISM gas (surrounding the launch site). In higher-mass galaxies where η is low, this assumption can break down, and it may be more appropriate to include a metal loading factor greater than unity. Without more detailed modeling, it is difficult to determine exactly what the appropriate metal loading factor is, so we eschew this complication for the present. Note that the FIRE simulations with self-consistently generated outflows find metal loading factors around unity for all galaxies (Ma et al. 2016), supporting our assumption.

To quench massive galaxies, we employ an on-the-fly halo mass-based quenching scheme that follows Gabor & Davé (2012, 2015). Above a halo quenching mass M_q, we maintain all halo gas at a temperature above the system virial temperature, by continuously adding heat. This is intended to mimic the effects of "radio mode" or "jet mode" quenching (Croton et al. 2006), where jets inflate superbubbles in surrounding hot gas which approximately spherically the jet energy and counteracts gas cooling (McNamara et al. 2007). We only add heat to gas that is not self-shielded, defined as having a neutral (atomic+molecular) fraction above 10% after applying a self-shielding correction following Rahmati et al. (2013). We take M_q as determined from the analytic "equilibrium model" constraints required to match the observed evolution of the galaxy population from $z = 0 - 2$ (Mitra, Davé, & Finlator 2015), namely:

$$M_q = (0.96 + 0.48z) \times 10^{12} M_\odot. \quad (3)$$

As demonstrated in Paper I this evolving quenching mass is nicely consistent with observations during early epochs ($z \sim 2$) and today, while providing a sharp turnover in the stellar mass function at late epochs that closely matches observations.

Paper I focused on the $50h^{-1}$Mpc MUFASA simulation using 512^3 gas fluid elements (i.e. mass-conserving cells), 512^3 dark matter particles, and $0.5h^{-1}$kpc minimum softening length. Table I of Paper I lists the details for two higher-resolution runs with the identical input physics and number of particles, having box sizes of $25h^{-1}$Mpc and $12.5h^{-1}$Mpc and proportionally smaller softening lengths. At that time, these simulations were only evolved to $z = 2$, but since then we have evolved the $25h^{-1}$Mpc volume to $z = 0$ and the $12.5h^{-1}$Mpc run to $z = 1$. We will use these to extend the dynamic range of our predictions and to test resolution convergence.

We generate initial conditions at $z = 249$ using MUSIC (Hahn & Abel 2011) assuming a cosmology consistent with Planck (2015) "full likelihood" constraints: $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_b = 0.048$, $H_0 = 68$ km s^{-1} Mpc^{-1}, $\sigma_8 = 0.82$, and $n_s = 0.97$. We output 135 snapshots down to $z = 0$ (105 to $z = 1$). We analyse the snapshots using SPHGR-yt[1] (Thompson 2015), which identifies galaxies using SKID and halos using ROCKSTAR (Behroozi et al. 2013), links them via

[1] `http://sphgr.readthedocs.org/en/latest/`

their positions, and outputs a catalog of properties required for all the analyses in this paper.

3 STAR FORMATION RATES

Paper I compared MUFASA to the evolution of the stellar mass function, showing general agreement with the growth of the stellar content of galaxies across much of cosmic time. However, it also reiterated a longstanding discrepancy in predictions of sSFRs at a given M_*, i.e. the main sequence, during the peak epoch of cosmic star formation, in which simulated galaxies have $\sim \times 2-3$ lower SFRs compared to observations at $z\sim 2$. Here we explore the distribution of SFRs in more detail, by comparing MUFASA to two other SFR observables, namely the star formation rate function and the specific star formation rate function.

3.1 Star formation rate function

Figure 1 shows SFR functions (SFRFs) at $z=0,1,2$ from our suite of MUFASA simulations. The red solid, green dashed, and blue dotted curves show the results from our $50h^{-1}$Mpc, $25h^{-1}$Mpc, and $12.5h^{-1}$Mpc (at $z\geqslant 1$) simulations. The hatched region shows cosmic variance as computed over the 8 sub-octants within each simulation volume. The vertical dotted line indicates the typical SFR at the steller mass resolution limit of 32 gas particle masses from a fit to the M_*-SFR relation; below this, the distribution of SFRs is expected to be significantly compromised by numerical resolution, and even above this SFR there may be some galaxies that are impacted by poor resolution owing to the scatter in the M_*-SFR relation. Hence this line should be regarded as an approximate rather than a strict resolution limit. Indeed, one can see from comparing the various simulations' SFR functions at the same SFR that the lack of resolution convergence seems to begin significantly above the dotted line.

Observations are shown in the various panels from $H\alpha$ luminosity functions, converted to SFR using the relation taken from Kennicutt (1998), adjusted for a Chabrier IMF. At $z\sim 0$, we show data from Bothwell et al. (2011, dotted black) and Gunawardhana et al. (2013, dashed black), at $z\sim 1$ from Colbert et al. (2013), and at $z\sim 2$ from Mehta et al. (2016). All these observations account for extinction based on considering Hβ and sometimes more, but there is still uncertainty in such corrections.

At $z=0$, the simulated SFRF are in good agreement with Bothwell et al. (2011), but overpredict by up to $\sim \times 3$ the more recent Gunawardhana et al. (2013) data from the Galaxy and Mass Assembly (GAMA) survey. This implies that there are several times more SFGs with SFR$\sim 1-10 M_\odot \mathrm{yr}^{-1}$ in MUFASA than in the real Universe. It is possible that Hα surveys miss the most highly star-forming galaxies since they are typically highly obscured. This could be mitigated by examining far-IR based SFR estimators, but that introduces the additional complexity of subtracting off the AGN contribution to the total flux, which is often substantial in luminous IR galaxies. This discrepancy is consistent with the finding in Paper I (see their Figure 3) that the cosmic SFR density is overpredicted by $\sim 50\%$ at $z=0$ in MUFASA. Hence MUFASA's predictions for the SFRF today are broadly consistent with data, but with a notable overprediction of galaxies with SFRs comparable to or exceeding that of the Milky Way.

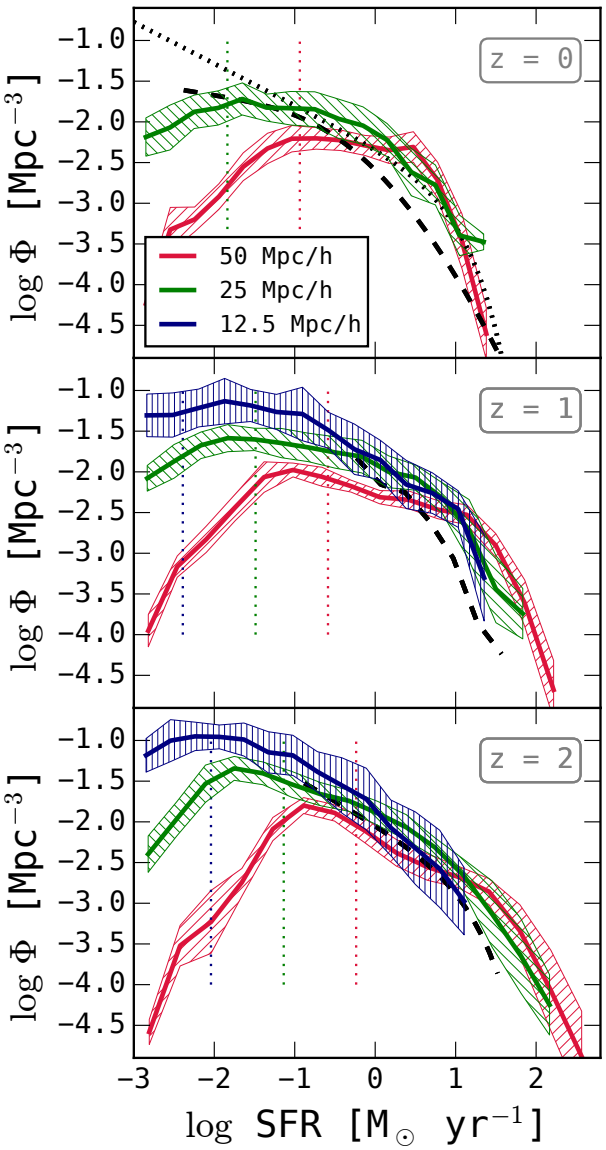

Figure 1. Star formation rate functions at $z=0,1,2$ (top to bottom), in our suite of MUFASA simulations. Results for the $50h^{-1}$Mpc, $25h^{-1}$Mpc, and $12.5h^{-1}$Mpc (at $z\geqslant 1$) are shown in red solid, green dashed, and blue dotted lines, respectively, with the hatched region showing the cosmic variance computed over the 8 sub-octants of the simulation volume. The vertical dotted line is an approximate resolution limit, taken as the mean SFR at our stellar mass resolution limit; there is likely some incompleteness even above this value, which can be seen by comparing the different volumes. Observations are shown as dashed black lines in each panel, using Hα luminosity functions converted to SFRFs from Gunawardhana et al. (2013, $z\approx 0$), Colbert et al. (2013, $z\approx 1$), and Mehta et al. (2016, $z\approx 2$).

At $z=1$, the predicted SFRF is similar to that at $z=0$ at low SFRs, but shows an excess at high SFRs, such that now we start to see galaxies with SFR$\gtrsim 100 M_\odot \mathrm{yr}^{-1}$ in our $50h^{-1}$Mpc volume. The SFRF does not show as strong a truncation at high-SFR as it does at low-z. Generally,

MUFASA exceeds observations at high SFRs, albeit with the same caveats regarding highly obscured galaxies that become more prevalent at high z.

At $z = 2$, the trend continues that the low-SFR end is mostly unevolving but the high-SFR end is more highly populated. Once again there is an excess in MUFASA relative to data, but it is fairly mild at this epoch. Interestingly, although MUFASA seems to reproduce the GSMF well at this epoch, and if anything the SFRF is overpredicted, it nonetheless yields an SFR$-M_*$ relation that is clearly too low (Paper I). Notably, the main sequence is typically derived from UV and/or rest-near infrared measures of SFR, not Hα. It is possible that various systematics operate differently at this epoch among the various observational SFR indicators. It is beyond the scope of this work to fully examine all the relevant systematics, but it highlights that, leaving aside the models, there appears to be some consistency issues purely among observational measures of SFRs during Cosmic Noon.

In summary, the SFRFs predicted by MUFASA generally show the observed shape from $z = 0 - 2$, though with an amplitude that is somewhat too high at low redshifts. The broad agreement is encouraging and may be within current systematic uncertainties in measuring a complete sample of star-forming galaxies across all these epochs. There is no obvious discrepancy in the SFRF at $z = 2$ that would explain the discrepancy in the SFR$-M_*$ relation. Resolution convergence in the SFRF between the various MUFASA volumes is reasonable, though not ideal.

3.2 Specific star formation rate function

A separate test of SFRs is whether our simulations reproduce the correct distribution of specific star formation rates at a given stellar mass. Qualitatively, at high redshifts the spread in sSFRs measures the fluctuations around the main sequence owing to inflow fluctuations (e.g. Mitra et al. 2016), while at lower redshifts a substantial low-sSFR population appears corresponding to quenched galaxies. Matching the amplitude and evolution of the distribution of sSFRs in stellar mass bins is thus a stringent test of whether the predicted MUFASA galaxy population is in accord with the rate at which galaxies are fluctuating around the main sequence, and eventually quenched (e.g. Tacchella et al. 2016).

Figure 2 shows the specific star formation rate function (sSFRF) in four bins of stellar mass from $10^{9.5} < M_* < 10^{11.5} M_\odot$ (left to right), at $z = 0.25, 1, 2$ (top to bottom). We only consider the $50h^{-1}$Mpc here volume for clarity, particularly since we want to well sample the rate of galaxy quenching for which we prefer our largest volume containing the most massive halos. Lines show the predicted sSFRF, while the hatched region shows the cosmic variance computed among 8 sub-octants. Observations are shown from a compilation by Ilbert et al. (2015) at $z = 0.2 - 0.4$ and $z = 0.8 - 1.2$, from various sources as described in the caption, generally from extinction-corrected UV measures or SED fitting. Note that the observations only consider galaxies which have a measurable SFR, which we mimic in our simulations by excluding galaxies with log sSFR< -3 (which would lie off this plot in any case).

At $z = 0.25$, the sSFRF shows a peak at the median sSFR within that M_* bin, a sharp truncation to higher sSFR, and a broader extension to low sSFR corresponding to green valley galaxies. MUFASA provides a remarkably good match (i.e. within cosmic variance) to the observed sSFRF in every stellar mass bin. This new test of models demonstrates that the scatter in sSFRs, and hence the fluctuations around the main sequence as well as the rate at which the green valley is being populated, is being well modeled in MUFASA. In particular, the amplitude and shape match in the most massive bin would suggest that MUFASA is not overproducing the number of galaxies with high sSFRs, even if Figure 1 suggested that it might be doing so. These can be reconciled if MUFASA is producing a few too many massive galaxies, which is indeed a trend noted in the $z = 0$ GSMF shown in Paper I, albeit with large cosmic variance.

At $z = 1$, the shape of the sSFRF is well reproduced, but there is clearly an offset in the distribution such that the predicted values are lower by $\sim \times 2$. This is simply reflecting the fact that the median sSFR is underproduced at this epoch, as shown in Paper I, continuing a trend generically seen in cosmological galaxy formation models. It appears that the discrepancy in the median sSFR is not reflective of the emergence of some new population of galaxies in observations that do not appear in the models, but rather an overall systematic shift in the measured sSFR values at that epoch. We would expect that these trends would continue on to $z = 2$, but we do not know of sSFRFs published at this epoch.

Overall, MUFASA does an excellent job of reproducing the low-z distribution of sSFRs, including the peak value, the sharp truncation to high sSFRs that highlights the rarity of starbursts locally, and the gradual decline towards low-sSFR that reflects the population of galaxies likely in the process of quenching. There are still a non-trivial number of SFGs even at the highest masses in MUFASA, which is in agreement with observations. This suggests that MUFASA does a good job reproducing the SFR fluctuations and quenching rate of galaxies, which provides some empirical support for the implemented subgrid models for star formation and quenching.

4 METALLICITY

Chemical enrichment provides a key tracer for star formation and feedback activity in and around galaxies. Within a simple equilibrium or bathtub-type model, the mass-metallicity relation directly reflects the mass loss rate in outflows together with the recycling of previously-ejected (enriched) material back into the ISM (e.g. Finlator & Davé 2008; Somerville & Davé 2015). Galaxy metallicities are thus a crucial test for how accurately a particular model is representing the baryon cycle.

The stellar mass–gas phase metallicity relation (MZR) is one of the tightest observed correlation between any two galaxy properties, with a scatter typically around 0.1 dex (Tremonti et al. 2004). Unfortunately, calibration issues may add significant systematic uncertainties (Kewley & Ellison 2008), but nonetheless the shape of the MZR is likely to be reasonably robust even if the amplitude is less certain. In this section we present predictions for the MZR from MUFASA, along with comparisons to key observations at the present epoch and in the early Universe.

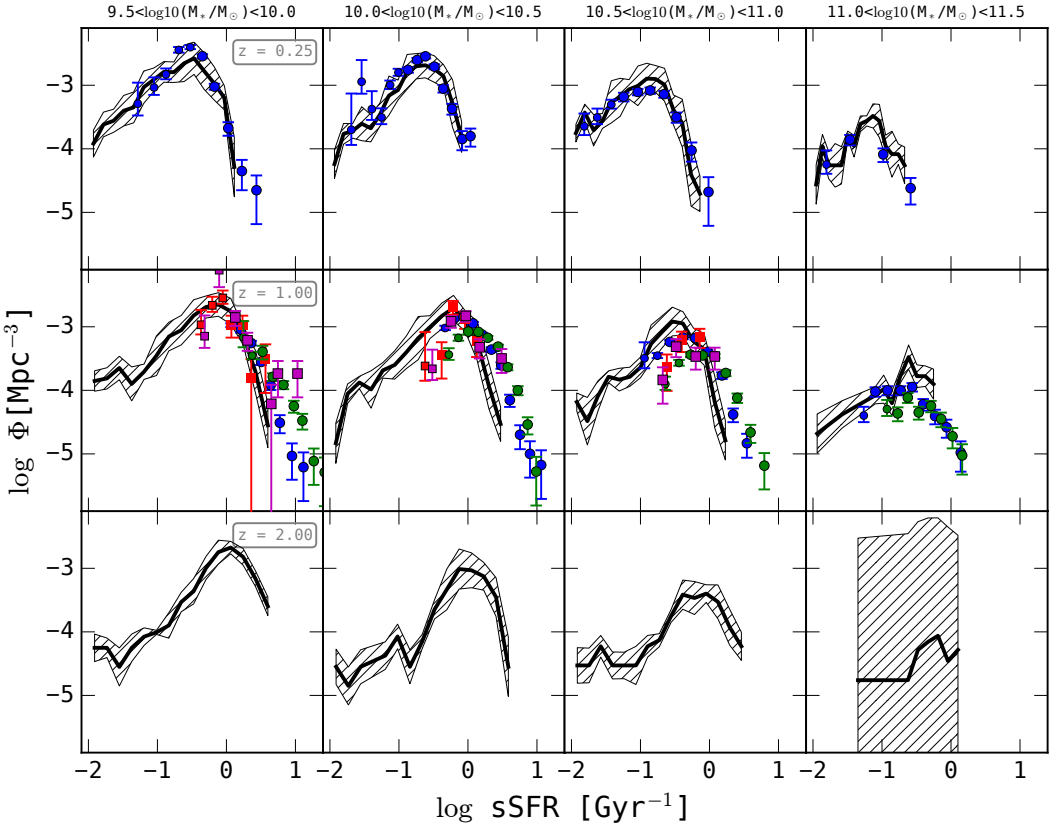

Figure 2. Specific star formation rate functions at $z = 0.25, 1, 2$ (top to bottom rows) in our $50h^{-1}$Mpc MUFASA simulation, in four bins of increasing stellar mass (left to right). Hatched regions show the cosmic variance computed over the 8 sub-octants in the volume. Observations from Ilbert et al. (2015) are shown, which includes data from COSMOS at $z = 0.2 - 0.4$ (blue in upper panels), and in the middle panels from COSMOS at $z = 0.8 - 1$ (blue), COSMOS at $z = 1 - 1.2$ (green), GOODS from $z = 0.8 - 1$ (red), and GOODS from $z = 1 - 1.2$ (magenta). The predicted sSFR functions match observations very well at $z \sim 0.25$, showing that MUFASA reproduces the distribution of sSFRs quite well at low-z. At $z = 1$ MUFASA matches well the shape of the distribution but is shifted to slightly lower sSFR. This indicates that MUFASA is properly capturing the physical causes of fluctuations around the main sequence, as well as the number of galaxies transitioning to quiescence.

Figure 3 shows the MZR at $z = 0, 2$ (left, right panels) in our MUFASA simulation suite. At $z = 2$, we have overplotted all three volumes down to each of their galaxy stellar mass resolution limit; these are the three "groupings" of points, with the $12.5h^{-1}$Mpc volume extending to the lowest masses, and the $50h^{-1}$Mpc volume dominating at high masses. At $z = 0$, we only have the $50h^{-1}$Mpc and $25h^{-1}$Mpc volumes. The thick red line shows a running median for the combined sample of simulated galaxies; while we do not show the individual volumes' medians separately, it is evident that the agreement between them is reasonable in the overlapping mass ranges, as there is no significant break in the median fit when crossing over a mass resolution threshold, though higher-resolution simulations tend to predict slightly higher metallicities at a given mass. The colour coding shows the deviation in log SFR for each galaxy off of the global M_*-SFR relation at that redshift (Paper I). Observations at $z = 0$ are shown from the Sloan Digital Sky Survey (SDSS), via nebular line fitting (Tremonti et al. 2004, black solid line) and "direct" abundance measures from stacked spectra (Andrews & Martini 2013, grey dashed line). At $z = 2$, we show observations from the Mosfire Deep Evolutionary Field (MOSDEF) survey (Kriek et al. 2015) using O3N2 abundances obtained from near-infrared Keck spectroscopy (Sanders et al. 2015, points with errorbars;).

Broadly, the agreement between MUFASA and observations is fairly good. The faint-end slope is generally consistent with data at both redshifts, and at high-z it can be seen that the simulated MZR slope extends unabated to much lower masses than can be observed prior to the *James Webb Space Telescope*. At the massive end, there is clearly a turnover at low redshifts above $M_* \gtrsim 10^{11} M_\odot$, and even at $z = 2$ there is a hint of a similar turnover though even the $50h^{-1}$Mpc volume does not adequately probe the very high-mass end at that epoch.

At low masses, there is $\sim 0.2 - 0.3$ dex increase in the metallicity at a fixed M_* from $z = 2 \rightarrow 0$. The evolution is slightly less at high masses, creating a more prominent flat portion of the MZR. This amount of evolution, and the trend of a more prominent turnover at low masses, is generally consistent with observations (Zahid et al. 2014; Steidel et al. 2014; Sanders et al. 2015).

A more careful comparison to MZR data reveals some notable discrepancies. Most obviously, there is a clear overprediction of the metallicity at $M_* \gtrsim 10^{10.3} M_\odot$ at $z = 0$. It appears that the high-mass flattening begins at a lower

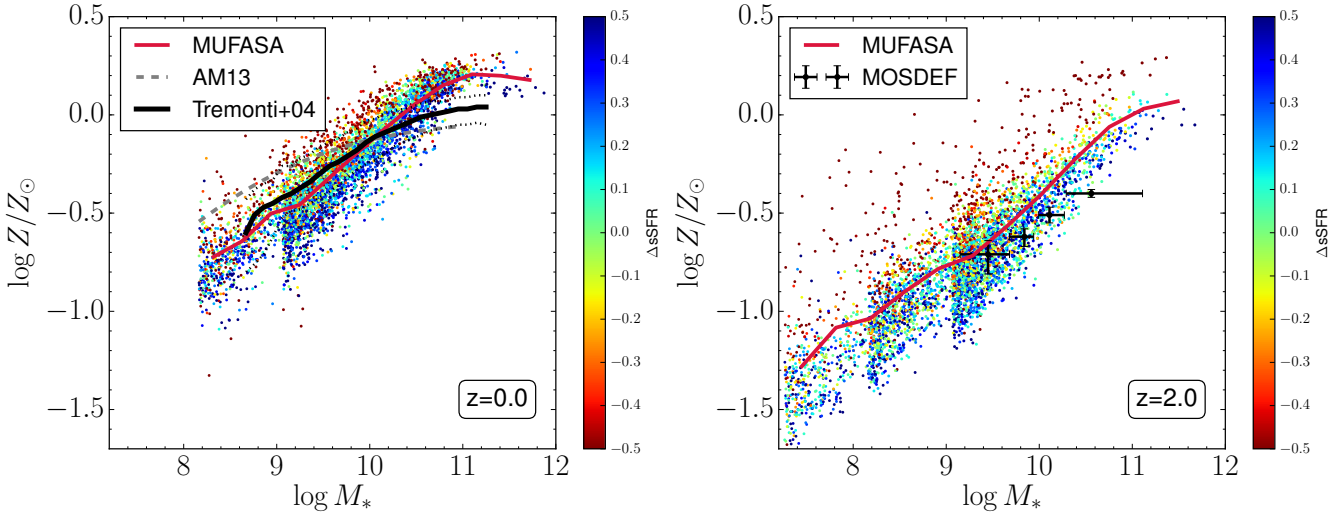

Figure 3. MUFASA galaxy mass-metallicity relations at $z = 0, 2$, computed from the predicted SFR-weighted oxygen abundance assuming the solar oxygen abundance from Vagnozzi et al. (2016). Displayed points are combined from the $50h^{-1}$Mpc, $25h^{-1}$Mpc, and $12.5h^{-1}$Mpc runs, and for each run every galaxy with gas is plotted down to the 64-particle stellar mass resolution limit where a break is evident. Points are colour-coded by their distance from the M_*-SFR relation; bluer points have higher SFR for their M_*, as indicated by the colour bar. Observations at $z = 0$ are shown from Tremonti et al. (2004, T04; solid black) and Andrews & Martini (2013, AM13; grey dashed), while $z \approx 2$ data is shown from Sanders et al. (2015).

mass scale in the data as compared to in MUFASA, which continues with an unabated power-law up to nearly $10^{11} M_\odot$ before flattening. There is even a hint of such an overproduction at $z = 2$; while the overall amplitude is slightly too large compared to these observations at all masses, this is particularly exacerbated for the highest mass bin. One possibility for reconciling this in the models would be that the metal loading factor at $M_* \gtrsim 10^{10.3} M_\odot$ should be greater than unity, which would preferentially eject a higher fraction of metals out of high mass galaxies. Alternatively, it could be that the models have excess wind recycling at high masses; we will examine mass flows and recycling in detail in future work.

One can also see that the low-mass end of the MZR is in better agreement with the Tremonti et al. (2004) nebular line MZR than the direct abundances measures by Andrews & Martini (2013). Such discrepancies between observational analyses highlight the difficulty in robustly calibrating metallicity indicators (Kewley & Ellison 2008). Moreover, at high redshifts it is possible that the typical stellar population in $z \sim 2$ star-forming galaxies may be substantially different than that at low redshifts (Steidel et al. 2016), which could alter the usual metallicity calibrations applied to nebular emission line measures. In light of this, the disagreements between MUFASA MZR predictions and observed may be regarded as preliminary.

Finally, the colours of the points show a clear trend that galaxies with low sSFR at a given mass will have high metallicity, and vice versa. This has been noted in data (Ellison et al. 2008; Lara-López et al. 2010; Mannucci et al. 2010; Salim et al. 2014; Telford et al. 2016), and Mannucci et al. (2010) dubbed this the fundamental metallicity relation (FMR) because they further argued that the SFR$-M_* - Z$ relation was also redshift-independent. More recent results have called into question whether the FMR is truly redshift independent (Salim et al. 2015; Brown et al. 2016; Grasshorn Gebhardt et al. 2016), and also whether it is even seen at high redshift (Steidel et al. 2014; Sanders et al. 2015). However, it appears that the samples at $z \sim 2$ may not be sufficient for such a trend to have been apparent, and moreover calibration issues can mask such subtle correlations (Salim et al. 2015). It is thus unclear whether the FMR exists at $z \sim 2$ observationally. We will discuss this second-parameter dependence of the MZR on the sSFR further in §6.

In MUFASA, the general trend of the SFR$-M_* - Z$ relation is apparent at both $z = 0$ and $z = 2$. However, the predicted MZR is notably tighter at $z = 0$ (typical variance of $\sigma \approx 0.1$ dex around the mean relation) than at $z = 2$ ($\sigma \approx 0.2$ dex). By $z = 2$, the most metal-rich galaxies already have metallicities comparable to the most metal-rich objects at $z = 0$, across all M_*, while the most metal-poor objects are much less enriched.

The physical explanation for the second-parameter correlation with SFR is that an increase in gas accretion will bring in metal-poor gas while fueling new star formation, and conversely a lull in accretion will result in an evolution more similar to a closed box that will raise the metallicity quickly by consuming its gas (e.g. Finlator & Davé 2008). As pointed out in Davé, Finlator, & Oppenheimer (2011), the lull is permanent for satellite galaxies, causing them to reach a slightly higher metallicity at a given mass before running out of fuel, as observed (Pasquali et al. 2010); though we don't show it here, this is true in MUFASA as well. Hence in the fluctuating "smooth accretion" scenario for galaxy fueling (Kereš et al. 2005; Dekel et al. 2009), the FMR is a natural outcome, and the scatter about the relation reflects the frequency and impact of accretion flucutations such as mergers. Confirming the reality of the FMR at $z \sim 2$ is thus a crucial test of this scenario. In §6 we will quantify pre-

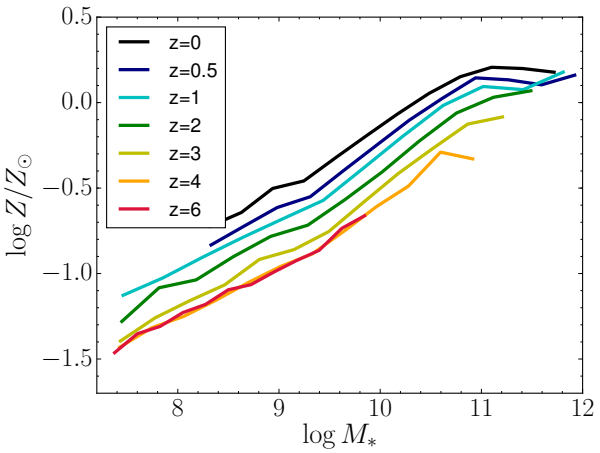

Figure 4. MUFASA median galaxy mass-metallicity relations at $z = 0, 0.5, 1, 2, 3, 4, 6$. Displayed relations combine galaxies from the $50h^{-1}$Mpc, $25h^{-1}$Mpc, and $12.5h^{-1}$Mpc runs down to each of their 64-particle resolution limit. There is steady upwards evolution of the MZR over time since $z \sim 4$, with $\sim \times 2$ increase in metallicity at a given M_* since $z = 2$.

dictions for this second-parameter correlation that can be tested against present and future observations.

Figure 4 shows the evolution of the MZR from $z = 6 \rightarrow 0$, computed as a running median from the combined sample of 3 runs. The MZR shows a constant low-M_* power-law slope of ≈ 0.5 at all redshifts. At the lowest redshifts, there is the onset of a flattening in the MZR at $M_* \gtrsim 10^{10.7} M_\odot$. The MZR rises steadily but slowly with time. At a given mass (below the flattening), the evolution is approximately 0.2 dex out to $z \sim 1$, and then 0.1 dex per unit redshift out to $z \sim 4$, and no further evolution to $z = 6$. In Finlator & Davé (2008) it was argued that, barring any evolution in η with M_* (MUFASA assumes none), then the evolution of the MZR must reflect the enrichment level of accreted material, i.e. wind recycling. It remains to be seen if such a scenario is consistent with a mass-independent increase in the metallicity down to quite low masses.

Qualitatively, the generally slow evolution and the onset of a high-mass flattened portion at lower redshifts is consistent with observations (e.g. Zahid et al. 2014), as well as data-constrained analytic models of galaxy evolution (Mitra, Davé, & Finlator 2015). However, the mass at which the flattening occurs is generally much higher in MUFASA than in such data, where the onset of flattening is typically below $M_* \lesssim 10^{10} M_\odot$. This again reflects the fact that MUFASA appears to produce too steep an MZR at $10^{10} \lesssim M_* \lesssim 10^{11} M_\odot$.

Overall, the slope and evolution of the MZR is in broad agreement with observations, showing mild evolution out to $z \sim 4$. However, there is a key discrepancy around L^* galaxies that bears further investigation. In future work we will examine the detailed origin for the evolution of the MZR, highlighting contributions from in situ enrichment versus pre-enriched accreted gas.

5 GAS CONTENT

The gas content of galaxies provides a measure of the fuel available for new star formation. Molecular gas (H_2) directly traces material that is forming into stars, while atomic gas (H I) typically resides in a more extended reservoir that connects the ionised IGM with the molecular ISM. Hence the gas content of galaxies represents a combination of the effects of how gas is converted into stars within the ISM, as well as the processes that fuel new star formation via gas from the IGM.

Observationally, it is generally believed that the atomic gas in galaxies evolves slowly out to high redshifts, while molecular gas evolves more rapidly upwards. The canonical explanation for this is that H I represents a transient reservoir which does not directly trace star formation, while H_2 traces star-forming gas much more closely and hence drops with time in a manner similar to what is seen for the cosmic star formation rate.

In actuality, the story is more subtle. In simple terms one can rewrite the ratio of star-forming gas to stars as

$$\frac{M_{\rm gas}}{M_*} = \frac{M_{\rm gas}}{SFR}\frac{SFR}{M_*} = t_{\rm dep}{\rm sSFR}, \quad (4)$$

where the first term is the depletion time and the second term is the specific SFR (e.g. Davé, Finlator, & Oppenheimer 2012). Given a fixed depletion time, one then expects the gas content of high redshift galaxies to be increased. However, one also expects the depletion time to be reduced to higher redshifts, since galaxies typically form a relatively fixed fraction of their gas into stars per dynamical time (Kennicutt 1998), and disk dynamical times are expected to scale approximately with the Hubble time (Mo, Mao, & White 1998). If sSFR$\propto (1+z)^{2.5}$, and $t_{\rm dyn} \propto H^{-1}(z)$, then one gets approximately $f_{H2} \propto (1+z)$. Hence galaxies are expected to have higher star-forming gas fractions at earlier epochs.

Meanwhile, the evolution of atomic hydrogen is not so straightforward to predict. In the simplest model where the timescale to pass through the atomic phase also scales with the halo (or, equivalently, disk) dynamical time, H I should follow H_2. But physically, atomic gas occurs when gas can self-shield against ionising radiation, yet is not dense enough to be molecular (i.e. to self-shield against H_2 dissociating radiation). At high redshifts, gas is physically denser and accretion is more filamentary (Dekel et al. 2009), but the ionising background is stronger. Which effect wins will depend on the detailed interplay of how gas is accreted around galaxies.

In this section we examine how the atomic, molecular, and total neutral (atomic+molecular) gas evolves within galaxies, as a function of stellar mass, in terms of mass functions, and globally as a cosmic mass density.

5.1 Gas fractions

MUFASA, like many recent simulations of galaxy formation, tracks the amount of molecular gas formed in galaxies. Owing to limitations of resolution, this is done via a subresolution prescription as described in Paper I, broadly following Krumholz, McKee, & Tumlinson (2009) with minor additions.

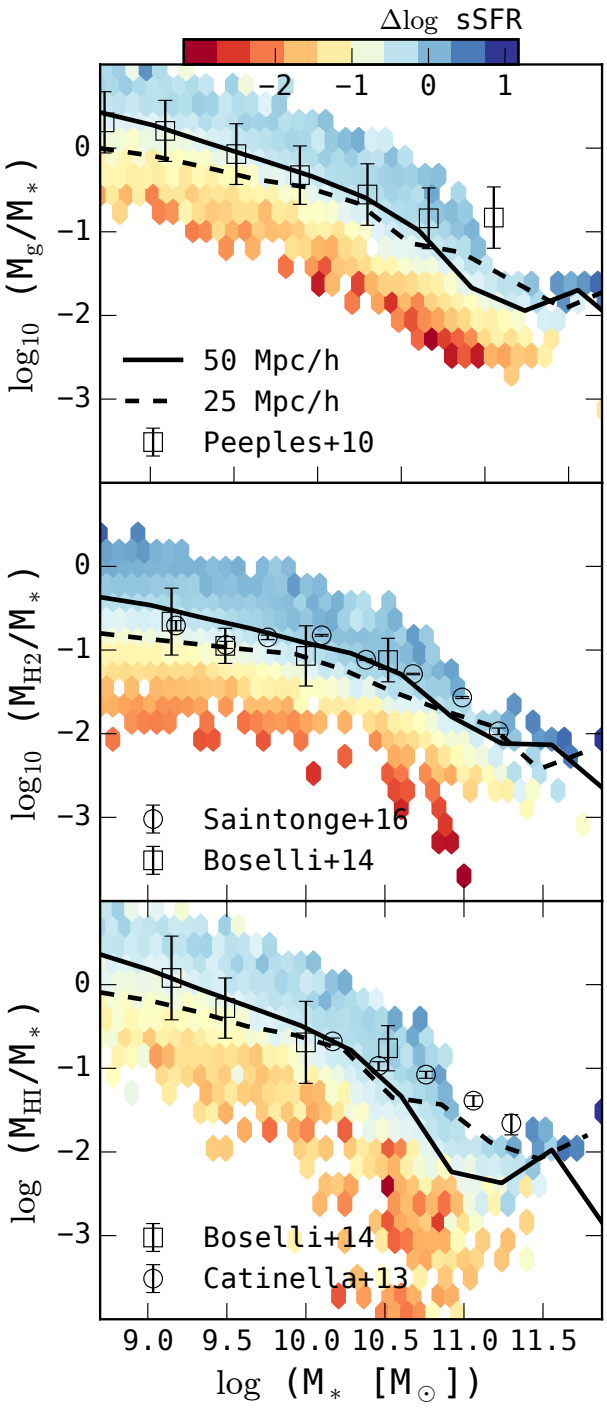

Figure 5. Total neutral (H I+H_2; top panel), molecular (middle), and atomic (bottom) gas fractions as a function of stellar mass predicted from the MUFASA $50h^{-1}$Mpc simulation at $z = 0$. Running medians are shown as the solid black lines. Colour-coding shows the mean sSFR deviation from the main sequence (Δ log sSFR) in each hexbin. Dashed line shows the running median from the $25h^{-1}$Mpc run to assess resolution convergence. Data is shown in the top panel from the compilation by Peeples & Shankar (2011), in the middle panel from COLDGASS (Saintonge et al. 2016) and HRS (Boselli et al. 2014), and in the bottom panel from HRS and GASS (Catinella et al. 2013). The $50h^{-1}$Mpc box shows good agreement with observations over the mass range probed by the data, but the $25h^{-1}$Mpc run tends to show lower gas fractions at low masses. At a given M_*, galaxies with higher gas content have higher SFR, and the trend appears tighter for H_2.

Meanwhile, the atomic gas fraction is typically significant only in regions that are able to self-shield against the cosmic metagalactic flux (ignoring, as we do here, ionising radiation emitted locally by the galaxy itself). Hence we must account for self-shielding in order to separate the neutral gas from the ionised gas.

We follow the prescription in Rahmati et al. (2013) for determining the self-shielded fraction. They provide a fitting formula to the attenuation in the cosmic metagalactic flux as a function of local density, based on full radiative transfer simulations. Given the attenuated ionising flux impinging on each gas particle, we then compute the rate balance equations to determine the equilibrium atomic fraction following Popping et al. (2009). For particles at low densities ($n_H \lesssim 10^{-3}$cm^{-3}) the gas is generally optically thin, but above this density one quickly gets more self-shielded gas, increasing the fraction to unity typically above $n_H \gtrsim 10^{-2}$cm^{-3}. From this self-shielded gas, we then subtract the molecular fraction as tracked directly in the simulation, which yields the atomic fraction. We compute a galaxy's H I content by summing all atomic gas that is more gravitationally bound to that galaxy relative to any other galaxy, using the total baryonic mass to compute the gravitational binding. In practice, we do not consider gas with $n_H < 10^{-4}$ cm^{-3} since this is never self-shielded and thus contributes negligibly to the total H I content.

Figure 5 shows the total (H I+H_2) (top panel), molecular (middle), and atomic (bottom) gas fractions as a function of stellar mass at $z = 0$. The solid black line shows a running median for the fiducial $50h^{-1}$Mpc volume. The overlaid hexbins are colour-coded by the average sSFR at that gas fraction relative to the global average sSFR at the given M_*. The dashed line shows a similar running median for the $25h^{-1}$Mpc run, to illustrate the level of resolution convergence.

In the top panel, the total gas fraction as a function of M_* in the $50h^{-1}$Mpc run is in excellent agreement with a compilation of observations by Peeples & Shankar (2011) over most of the mass range. At the highest masses, the observations lie above the model predictions. While these data only include galaxies where gas was detected, and many of the simulated galaxies have such low gas fractions that they would likely evade detection, since there are no predicted galaxies at all at the median total gas fraction, it appears at face value that the discrepancy is real. For $M_* \lesssim 10^{10.5} M_\odot$, however, galaxy samples are quite complete, and hence the agreement is a robust success of the models.

The middle panel shows that the molecular gas fractions are likewise in good agreement with observations from the COLDGASS survey (Saintonge et al. 2016), as well as the Herschel Reference Survey (HRS; Boselli et al. 2014). COLDGASS (Saintonge et al. 2011) is an M_*-complete survey and hence is quite directly comparable to our simulated galaxies. MUFASA even traces the slight turn-down in f_{H2} at $M_* \gtrsim 10^{10.5}$ relative to an extrapolated trend from lower masses, which is indicative of a typical mass scale at which quenching kicks in.

The atomic gas fractions are compared to data from the GASS survey (Catinella et al. 2010), which is the parent survey of COLDGASS and hence also a M_*-selected sample of SDSS galaxies down to very low H I fractions. At low masses ($M* \lesssim 10^{10.5} M_\odot$) there is quite good agreement

with the GASS data, which again is a non-trivial success. However, our $50h^{-1}$Mpc volume predicts a sharp drop in f_{HI} above this mass, whereas the data show a more gradual trend. This is likely the origin of the discrepancy in the total gas fraction at these masses, since f_{H2} shows good agreement in this mass range.

The $25h^{-1}$Mpc volume (dashed lines) consistently shows lower gas content at $M_* \lesssim 10^{10} M_\odot$, and thus a shallower trend with M_* that results in a $\sim \times 2$ deficit with respect to the $50h^{-1}$Mpc volume at the lowest probed masses. The deficit is essentially identical in both H I and H_2, which suggests that gas consumption is more rapid in the $25h^{-1}$Mpc volume, likely owing to its higher resolution that achieves higher densities where more rapid star formation can occur. Interestingly, this volume shows no "dip" in the H I, and hence total gas, content at $M_* \sim 10^{11} M_\odot$, indicating that the disagreement in the $50h^{-1}$Mpc may be a peculiarity in that simulation or else some issue with resolution convergence in terms of the way it interacts with the quenching model. One possibility is that the $25h^{-1}$Mpc is able to self-shield gas in massive halos more effectively owing to its ability to resolve clumpier structures, and thus the quenching model is less impactful here since by construction it only operates on non-self-shielded gas. In any case, at high and low masses it appears that resolution convergence is not ideal for predicting gas fractions, and the resulting systematic uncertainties are of the order of a factor of two.

The coloured hexbins show that at a given M_*, both molecular and atomic gas content are highly correlated with ongoing star formation. In both cases, galaxies with enhanced gas content for their M_* also have higher sSFR. The trend appears to be qualitatively stronger in the molecular case, which is unsurprising since stars form out of molecular gas in our simulations. Nonetheless it is also clearly present in the atomic gas, indicating that the H I reservoir plays a role in regulating star formation even if it is not directly forming stars. This is qualitatively consistent with observations that show more low-metallicity gas in the outskirts of bluer (i.e. higher sSFR) galaxies (Moran et al. 2012).

As with the metallicity, the qualitative explanation of this is that a temporary enhancement (lull) of accretion results in both increased (decreased) gas content and star formation, along with the aforementioned reduction (increase) in metallicity. Since it takes some time for the inflowing gas to first turn into atomic gas and then molecular and finally stars, the molecular gas content is expected to be more highly correlated with the SFR. Hence as with the FMR, the second parameter dependence of gas content on SFR most directly reflects fluctuations in the inflow rate (Mitra et al. 2016); we will quantify this in §6.

Other models that track molecular and atomic gas generally predict H I and H_2 fractions broadly in accord with observations, be they semi-analytic (Lagos et al. 2011; Popping, Somerville, & Trager 2014) or state-of-the-art hydrodynamic models such as EAGLE (Lagos et al. 2015; Crain et al. 2016). Together with MUFASA's success, this suggests that the overall gas content is a fairly robustly predictable quantity in models, at least at $z = 0$. We note that all these models (including ours) have been tuned at various levels in order to match the present-day stellar mass function. It may be that predicting this correctly, plus having a molecular gas-based prescription for converting gas into stars, generically leaves the proper amount of gas in galaxies. If so, this represents a non-trivial success for current models of galaxy formation.

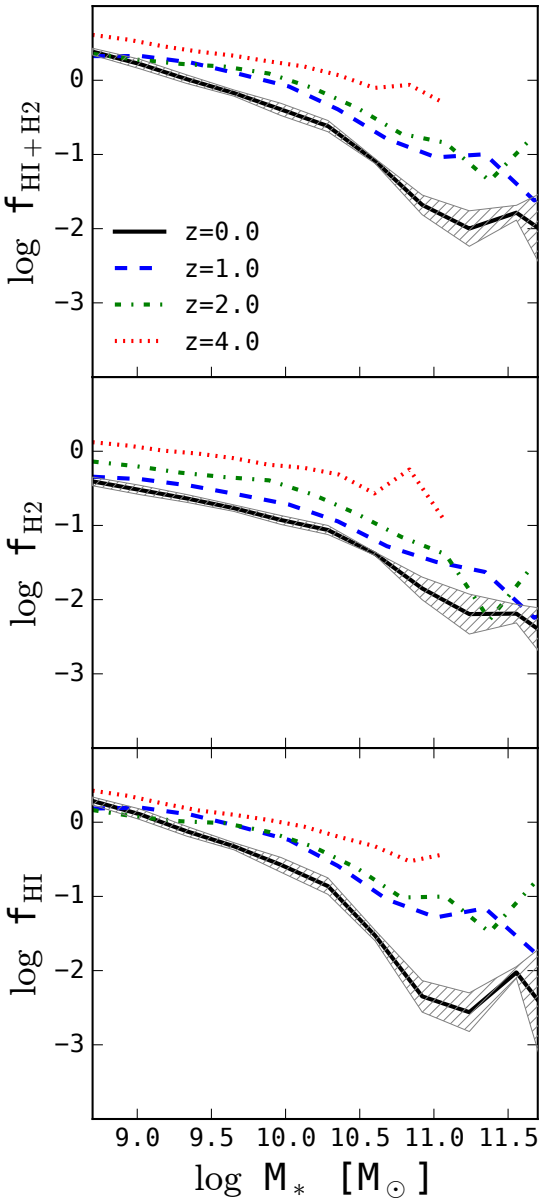

Figure 6. MUFASA median gas fractions as a function of M_* at $z = 0$ (solid black), $z = 1$ (dashed blue), $z = 2$ (dot-dashed green), and $z = 4$ (dotted red) from the $50h^{-1}$Mpc run. Cosmic variance over 8 sub-octants is shown as the hatched region around the $z = 0$ curve. The H I fraction (bottom panel) decreases with time for massive galaxies, while low-mass galaxies are always H I-dominated. The molecular gas fraction (middle panel) increases steadily across all masses, but is typically sub-dominant to H I. The total molecular plus atomic content (top panel) is thus driven by the H I evolution.

5.2 Gas fraction evolution

Galaxies at a given mass are observed to be more molecular gas-rich at earlier epochs (e.g. Geach et al. 2011; Tacconi et al. 2013). The amount of evolution is subject to some uncer-

tainties regarding the conversion between the observed CO intensity and the molecular gas mass (see e.g. Bothwell et al. 2013), but this is unlikely to erase the qualitative trend. Far-IR dust continuum measures can also be used to probe gas content evolution, though likewise subject to some uncertainties regarding the conversion of dust to gas mass (e.g. Scoville et al. 2016); in general, such studies tend to find less strong evolution than CO-based studies.

Neutral gas above $z \gtrsim 0.3$ is currently only observable in absorption line studies such as with Mg II absorbers (Rao et al. 2006) or Damped Lyman Alpha (DLA) systems (Prochaska & Wolfe 2009); it is not obvious how these systems trace galaxies, as it is usually challenging to identify the individual galaxy giving rise to such absorbers, though clustering measures offer some general guide that they typically live in $10^{11-12} M_{\odot}$ halos (e.g. Bouché et al. 2005; Font-Ribera et al. 2012). Measuring H I fractions requires having a measure of the stellar mass from optical or near-IR data for individual galaxies, for which 21 cm emission can be observed. While current H I-21 cm surveys only probe to $z \sim 0.4$ (Fernández et al. 2016), upcoming radio telescopes promise to push direct H I-21 cm gas content measures in optically-selected samples out to $z \gtrsim 1$, for example using the new MeerKAT telescope (Holwerda, Blyth, Baker 2012), and will be further advanced with the SKA. Here we make testable predictions for gas fractions which can guide such efforts.

Figure 6 shows the evolution at $z = 0, 1, 2, 4$ of the median total (H I+H_2) gas fraction versus stellar mass from the $50h^{-1}$Mpc MUFASA run in the top panel, and the next two panels show this subdivided into molecular and atomic gas fractions. The $z = 0$ curve is identical to that in Figure 5, but here we also show show with shading the cosmic variance estimated via jackknife resampling over 8 volume sub-octants. Here we do not show the second-parameter dependence on SFR as we did in Figure 5, but a similar trend persists at all redshifts. We do not explicitly show any observations on this plot, since molecular gas observations span some range depending on the type of data, while atomic gas fraction measures do not yet exist at $z \gtrsim 0.4$.

The total gas content of galaxies at a given M_* is higher at earlier epochs. There appears to be some mild mass dependence to this statement, as high-mass galaxies lose their gas more quickly than low-mass galaxies, with an overall effect of steepening the $f_{\rm gas} - M_*$ relation. Much of the evolution occurs from $z \sim 1 \rightarrow 0$, prior to which the evolution was somewhat slower.

Neutral hydrogen (bottom panel) represents the majority of the cold gas content of galaxies at almost all epochs and masses, except at high masses today. Hence the evolutionary trends in H I fraction tend to drive those of the total gas content. The strong evolution particularly at high masses out to $z \sim 1$ is good news for upcoming H I surveys designed to measure 21 cm emission from galaxies out to this epoch such as LADUMA, and will figure prominently in the evolution of the H I mass function discussed in §5.3.

Although we don't show it, the $25h^{-1}$Mpc box actually shows quite good resolution convergence with the $50h^{-1}$Mpc box shown here for all redshifts *except* $z = 0$. At $z = 0$, the $25h^{-1}$Mpc volume shows a flatter relation (as seen in Figure 5), but at higher-z the relations are similar, which implies less mass dependence to the evolution. Hence one should regard the detailed mass dependence of the evolution as a less robust prediction.

The trend to earlier epochs for the molecular gas (middle panel) is broadly similar to that for the atomic gas, in that it is increasing at all masses. There is a steady decrease in f_{H2} with time across all M_* of about $0.2 - 0.3$ dex between $z = 2 \rightarrow 0$, with only a very slight trend for more evolution at the highest masses. Predicted gas fractions continue to increase at a given M_* out to $z = 4$, so we expect even more gas-rich galaxies at high masses, but unfortunately even ALMA will have difficulty measuring the molecular content at these epochs except in the very largest systems (Decarli et al. 2016).

Comparing to observations, it appears that MUFASA predicts H_2 fractions that are too low versus data at $z \sim 1 - 2$. CO-based gas fractions from Tacconi et al. (2013) show a large scatter but generally lie between $20 - 40\%$ for the most massive galaxies, and 50% for moderate-mass galaxies. To lower masses, fractions up to 90% are inferred for the smallest $z \sim 2$ galaxies by inverting the Kennicutt (1998) relation (e.g. Erb et al. 2006). The dust continuum-based measures from Scoville et al. (2016) also show typically molecular fractions of $20 - 40\%$ for main sequence galaxies at $z \sim 1 - 2$, and even higher for starbursts. In contrast, MUFASA predicts $z = 2$ gas fractions of $\sim 10\%$ for massive ($M_* \sim 10^{11}$) galaxies, and only up to $\sim 40\%$ (relative to the molecular+stellar mass) for the smallest galaxies that are well below the what can be probed directly with observations. Hence in general it appears that MUFASA high-z gas fractions are too low by $\sim \times 2$. Given the uncertainties this is not a gross failure, but it is notable.

Such low high-z molecular gas fractions are predicted in other simulations and SAMs as well (Popping et al. 2015; Lagos et al. 2015). This may be partially but certainly not completely explained by selection effects in which targeted CO observations tend to select highly star-forming (and thus gas-rich) galaxies; Tacconi et al. (2013) accounted for this and still found generally higher f_{H2} than predicted here. Another possibility is that locally-calibrated CO-to-H_2 conversion factors may not be correct at high-z; we will explore this issue in more depth in §5.4. Nonetheless, at face value it appears that many models including MUFASA struggle to reproduce quite as high gas fractions as inferred for massive high-z galaxies.

5.3 H I mass function

The H I mass function (HIMF) combines information from the galaxy mass function and H I fractions to provide a complementary constraint on models. Observations of the HIMF extend to quite low masses locally thanks to deep surveys with the Arecibo telescope such as the H I-selected The Arecibo Legacy Fast Alfa survey (ALFALFA; Haynes et al. 2011) and the stellar mass-selected Galex Arecibo SDSS Survey (GASS; Catinella et al. 2010). However, the sensitivity of current instrumentation precludes characterisation of the HIMF at significantly higher redshifts. The SKA and its precursors aim to improve on this, and hence predictions for the evolution of the HIMF are useful for quantifying expectations for upcoming surveys such as the Looking At the Distant Universe with the MeerKAT Array (LADUMA; Holwerda, Blyth, Baker 2012) survey.

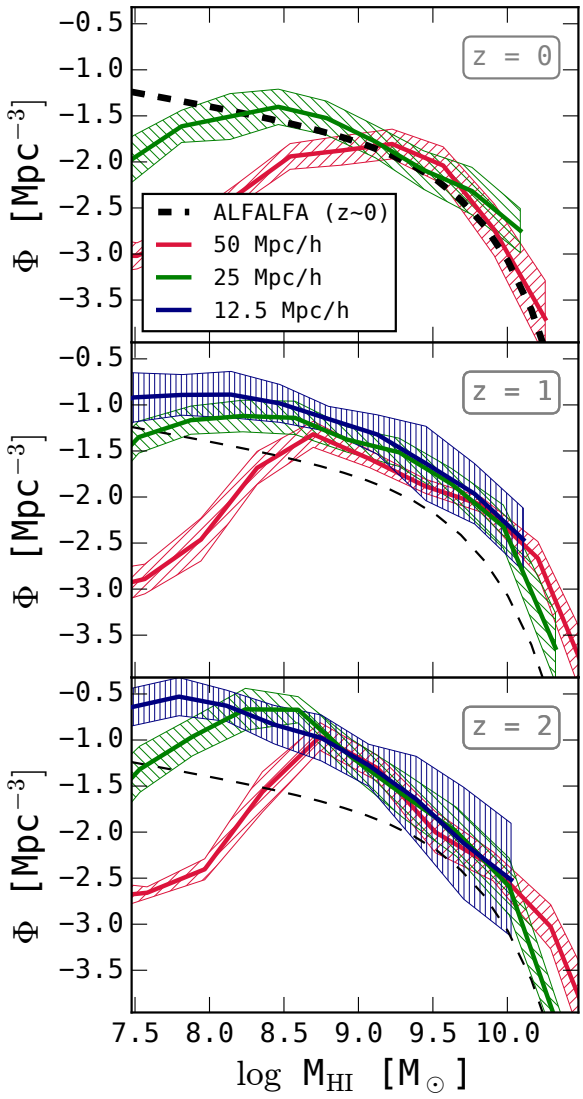

Figure 7. H I mass functions in the $50h^{-1}$Mpc (red solid), $25h^{-1}$Mpc (green dashed), and $12.5h^{-1}$Mpc (blue dotted, for $z \geqslant 1$) MUFASA simulations. Hatched region shows the cosmic variance computed over 8 sub-octants in each volume. Black dashed line shows $z \approx 0$ observations from the ALFALFA survey, reproduced at $z = 1, 2$ in order to better depict the evolution. MUFASA does reasonably well reproducing the HIMF over the M_{HI} range of convergence; at $z \sim 0$, the $50h^{-1}$Mpc box begins to deviate from the $25h^{-1}$Mpc at $M_{HI} \lesssim 10^9 M_\odot$, while at higher redshifts this occurs at $M_{HI} \sim 10^{8.7} M_\odot$. The HIMF generally increases in amplitude to higher redshifts, and also steepens noticeably at $z = 2$. There is fair resolution convergence down to $M_{HI} \sim 10^9 M_\odot$ in the $50h^{-1}$Mpc volume, and to $\approx 8\times$ lower in the $25h^{-1}$Mpc volume.

Figure 7 shows the predicted HIMF from MUFASA, showing the three volumes in different colours with cosmic variance (shading) estimated as before, from the variance of the HIMF in each of the 8 sub-octants within each simulation volume. As discussed in Paper I, this is likely to somewhat underestimate the true cosmic variance. The three panels show the HIMF at $z = 0, 1, 2$ from top to bottom. The ALFALFA mass function at $z \approx 0$ is shown as the thick black dashed line, and this is reproduced in the other redshift panels to better visualise the amount of evolution in the models; however, direct comparison to MUFASA should only be made at $z = 0$.

The top panel shows that MUFASA provides a reasonable match to the observed HIMF, in the resolved range. At $M_{HI} \lesssim 10^9 M_\odot$, the $50h^{-1}$Mpc volume shows a departure from the data, but the higher resolution run continues unabated, suggesting that the turnover at low masses is an artifact of numerical resolution. Indeed, if one combines the stellar mass resolution limit of $10^{8.7} M_\odot$ with the fact that galaxies at that M_* have an H I fraction of around two (Figure 5), this suggests that galaxies with $M_{HI} \lesssim 10^9 M_\odot$ will suffer from incompleteness in our simulations. The $25h^{-1}$Mpc volume extends another factor of almost 8 lower in mass before turning over, as expected from its $8\times$ higher mass resolution.

The agreement of MUFASA with both the stellar mass function (Paper I) and the HIMF is an important success. Previous simulations by Davé et al. (2013) also showed good agreement with data for both, even when subdivided into stellar mass bins (Rafieferantsoa et al. 2015). The EAGLE simulation likewise shows good agreement with both (Crain et al. 2016), including subdivided into M_* bins (Bahé et al. 2016). However, semi-analytic models constrained to match the stellar mass function don't necessarily agree well with the HIMF (e.g. Benson 2014). The SAMs of Popping, Somerville, & Trager (2014) do fairly well at $M_{HI} \gtrsim 10^9 M_\odot$, but predict a significant upturn to lower masses that is not observed. Such an upturn is also seen in the older Overwhelmingly Large Simulations (OWLS) HIMF as well (Duffy et al. 2008). The semi-empirical model of Popping, Behroozi, & Peeples (2015) likewise produces a steep faint end of the HIMF, deviating strongly at $M_{HI} \lesssim 10^9 M_\odot$. In simulations such as MUFASA, H I represents a transient reservoir of cold gas infalling into a galaxy, as demonstrated in Crain et al. (2016); such a dynamic origin suggests that fully dynamical models are best suited to make predictions for the nature and evolution of H I in galaxies. It appears that the low-mass ($M_{HI} \lesssim 10^9 M_\odot$) HIMF may be a key discriminant for the dynamics of gas infall.

Looking at the evolution to $z = 1$, we see that the HIMF is best described by an overall increase in the mass of H I in each galaxy by a factor of $\sim \times 2-3$, particularly at the massive end. This is consistent with the evolution seen in Figure 6. This is good news for surveys such as LADUMA that will probe the bright end of the HIMF at these redshifts; in future work we will make more specific predictions for LADUMA. Interestingly, this is somewhat contrary to the trend predicted by our previous simulations in Davé et al. (2013), which showed a steepening of the HIMF to higher redshifts, but the massive end was generally unchanged or lowered. This is because the H I fraction in the Davé et al. (2013) simulations was invariant with redshift, whereas in MUFASA galaxies are substantially more H I-rich, particularly at high masses.

From $z = 1$ to 2, the main trend is that the HIMF is steeper at low masses, while the massive end does not evolve significantly. This is driven by the steepening of the stellar mass function, since the H I fraction if anything has a shallower trend with stellar mass at higher redshifts (Figure 6). This general trend agrees better with that in Davé et al. (2013).

Overall, the HIMF in MUFASA at $z = 0$ is a reasonable

match to observations, even though the dynamic range is limited compared to other simulations such as EAGLE. By combining various box sizes, we can span a similar dynamic range, and the HIMF shows good resolution convergence in the overlapping H I mass range. MUFASA predicts a noticeably higher HIMF at $z \sim 1$, and then a steepening trend to $z \sim 2$, which must await future SKA and pathfinder telescope data for testing.

5.4 CO luminosity function

The mass function of molecular gas is more complicated to determine than that of atomic gas, since observations typically do not directly trace H_2 but rather some proxy such as CO. For ordinary (non-starburst) galaxies, canonically the best proxy for H_2 is the $J = 1 - 0$ rotational transition of CO. Nonetheless, this still requires a conversion factor (X_{CO}) to obtain the H_2 mass, and the dependence of X_{CO} on the intrinsic properties of galaxies such as star formation rate and metallicity is uncertain. This becomes particularly problematic at high redshifts, where the ISM conditions in typical main sequence galaxies vary substantially from that today.

A typical assumption is that galaxies that are near the main sequence have "Milky Way-like" $X_{CO} \approx 4$, whereas starbursts have $X_{CO} \approx 0.8$ (Tacconi et al. 2013). However, substantial work has gone into predicting X_{CO} based on galaxy properties from detailed simulations, yielding a continuous rather than bimodal trend. In particular, Narayanan et al. (2012) used zoom simulations together with a CO line radiative transfer code to develop an approximate fitting function for X_{CO} as a function of H_2 surface density and metallicity:

$$X_{CO} = \frac{1.3 \times 10^{21}}{Z' \times \Sigma_{H2}^{0.5}} \quad (5)$$

where Z' is the metallicity in solar units.

Here, we use this formula to compute X_{CO} individually for each galaxy, obtaining Σ_{H2} by dividing the H_2 half-mass of each galaxy by the area computed from the H_2 half-mass radius. Using this X_{CO}, we then convert our simulated H_2 masses into CO luminosities (L_{CO}), which can be compared more directly against observations. In this way, we specifically account for the metal and gas content evolution in CO-to-H_2 conversions when comparing to observations. This is analogous to the approach in Narayanan, Bothwell, & Davé (2012), except that here we convert simulated galaxies to get L_{CO}, while they took the converse approach of converting observations into M_{H2} to compare with models. However, we will see that our conclusions are similar.

Figure 8 shows the CO luminosity function (COLF) from our MUFASA simulations, showing once again our available simulation volumes at each redshift $z = 0, 1, 2$ (top to bottom). At $z = 0$, it is possible to directly observe CO 1-0 down to very low L_{CO}, and such observations by Kereš, Yun, & Young (2003) are shown as the data points. To higher redshifts, blind CO surveys where the survey volume can be robustly estimated are difficult, so one typically uses another proxy for this. The dashed lines show observations from Vallini et al. (2016), which used far-infrared luminosity as a proxy for CO luminosity; at $z = 0$, they agree with the

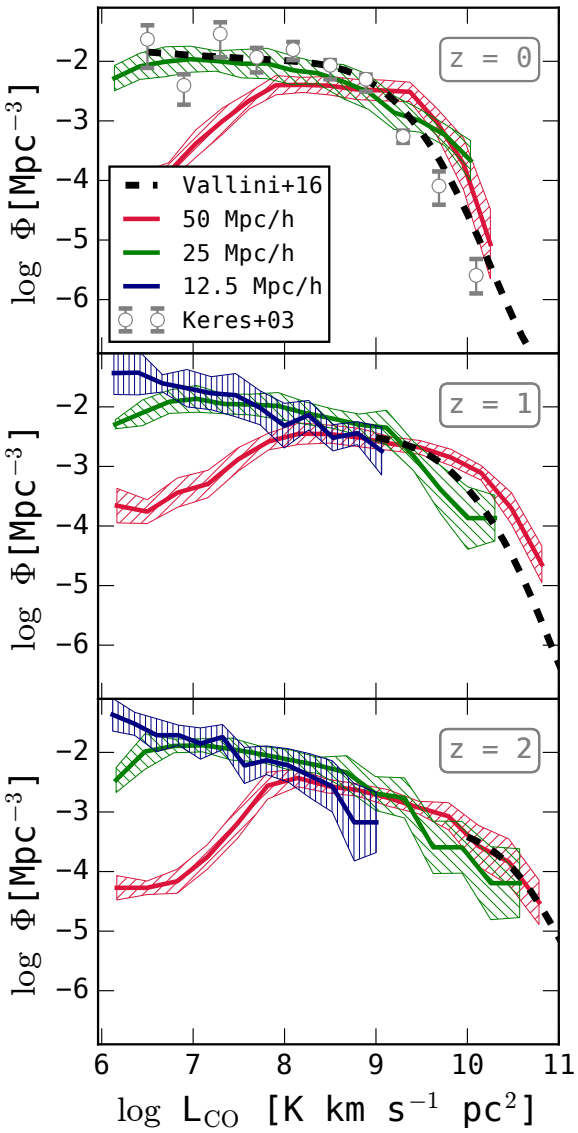

Figure 8. CO luminosity functions in the $50h^{-1}$Mpc (red solid), $25h^{-1}$Mpc (green dashed), and $12.5h^{-1}$Mpc (blue dotted, for $z \geqslant 1$) MUFASA simulations. Hatched region shows the cosmic variance computed over 8 sub-octants in each volume. We compute the CO1-0 luminosity from our predicted H_2 mass based on the prescription in Narayanan et al. (2012) derived from zoom simulations coupled with CO radiative transfer. We show observations from Vallini et al. (2016) at $z = 0, 1, 2$ as the black dashed lines, down to their approximate completeness limit at each redshift; note that these are based on an L_{IR}-to-L_{CO} conversions. Resolution convergence amongst the volumes is generally quite good.

Kereš, Yun, & Young (2003) data. At higher redshifts, we plot their observations down to their approximate completeness limit. We note that recent direct CO measures from ALMA by Decarli et al. (2016) indicate a somewhat higher number of high-L_{CO} objects at $z \sim 2$ than Vallini et al. (2016), but the statistics are small and the cosmic variance is large, so the discrepancy is only marginally significant.

At $z = 0$, MUFASA generally predicts a reasonable COLF, with a hint of an excess at high L_{CO}. The $50h^{-1}$Mpc volume shows a turnover at low-L_{CO} owing to numerical

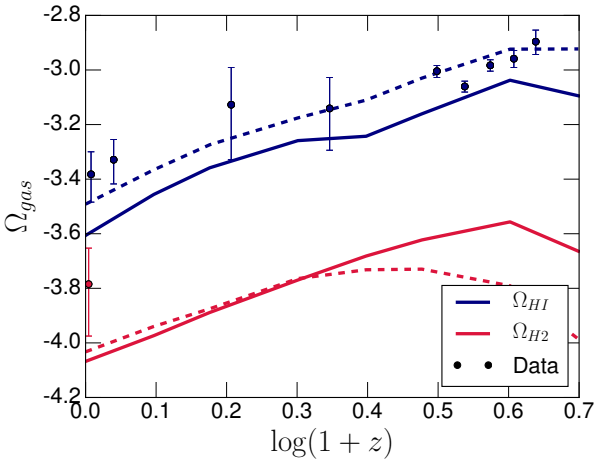

Figure 9. Cosmic mass density in H I (blue) and H_2 (red) as a function of redshift, in our $50h^{-1}$Mpc (solid) and $25h^{-1}$Mpc (dashed) MUFASA simulations. Ω_{HI} observations are shown from Delhaize (2013, $z < 0.1$), Rao et al. (2006, $0.5 < z < 1.3$), and Noterdaeme et al. (2012, $2.1 < z < 3.35$). The predicted trend of approximately $\Omega_{HI} \propto (1+z)^{0.74}$ is a good match to the compiled observations, as is the normalization although it is somewhat sensitive to numerical resolution. Ω_{H2} shows a similar redshift trend as Ω_{HI}, which is substantially slower than the evolution of the cosmic SFR density as shown in Figure 3 of Paper I.

resolution, while the $25h^{-1}$Mpc continues to agree well with the observations down to the lowest probed L_{CO}. The observations of Vallini et al. (2016) generally find an increase in the number of high-L_{CO} galaxies with redshift, and the simulations follow this trend, generally agreeing with data with still a hint of a high-L_{CO} excess. By $z = 2$ the observations only probe the brightest CO galaxies, where only the $50h^{-1}$Mpc volume has comparably bright systems, but these are in very good agreement with the data.

It is interesting that despite the relatively mild evolution of H_2 fractions in Figure 6 and a putative underprediction of f_{H2} at $z \sim 2$, MUFASA reproduces well the evolution of the COLF out to $z = 2$, and shows significantly more high-L_{CO} objects at high redshifts. This suggests that using a physically-motivated prescription for converting CO into H_2 (or vice versa) can lead to inferring a different amount of evolution in the gas fractions, and in general could potentially reconcile the relatively low amount of evolution in simulations versus the stronger evolution inferred using standard assumptions regarding X_{CO}; this broadly echoes the conclusions of Narayanan, Bothwell, & Davé (2012). Generically, metallicity-dependent X_{CO} prescriptions such as Narayanan et al. (2012) and Feldmann, Gnedin, & Kravtsov (2012) tend to predict more H_2 at high masses and less at low masses owing to enhanced H_2 production at high metallicities, which serves to increase the bright end of the COLF and flatten the faint end (Popping, Somerville, & Trager 2014) thus yielding better agreement with the COLF. Empirical luminosity-dependent X_{CO} calibrations have a qualitatively similar effect (Boselli et al. 2014). Hence MUFASA may be plausibly reproducing the evolution of the molecular gas content in galaxies in spite of its modest evolution of $\sim \times 2$ in the H_2 content at a fixed M_* out to $z \sim 2$.

5.5 Cosmic gas mass evolution

A synthesis of all the above evolutionary measures is provided in the evolution of the global cosmic gas density, typically parameterised in units of the critical density (i.e. as $\Omega_{\rm gas}$). The slow evolution of Ω_{HI} relative to the overall cosmic star formation rate density has been noted as evidence that H I is not directly physically associated with star formation, while the more rapid evolution of H_2 fractions can explain at least part of the rapid evolution in the cosmic SFRD. However, such interpretations are complicated by detailed assumptions regarding X_{CO} as discussed in the previous section, and how H I gas traces galaxies. Here we examine predictions for the evolution of the cosmic H I and H_2 mass densities, in the context of the evolutionary trends we have discussed above.

Figure 9 shows the evolution of the cosmic density in atomic gas (blue) and molecular gas (red) as a function of $\log(1+z)$. Solid and dashed lines show the results from our $50h^{-1}$Mpc and $25h^{-1}$Mpc, respectively. This is obtained by summing over all SKID-identified galaxies; using instead the sum over all H I or H_2 in the volume (which includes the IGM) makes a negligible difference.

Data points with the blue error bars correspond to various observational measures of Ω_{HI}: From 21cm emission (Delhaize 2013, $z < 0.1$), using Mg II absorbers as a proxy for DLAs (Rao et al. 2006, $0.5 < z < 1.3$), and DLA absorbers selected from the Sloan Digital Sky Survey (Noterdaeme et al. 2012, $2.1 < z < 3.35$). For H_2, no data is shown; at $z = 0$, Kereš, Yun, & Young (2003) inferred $\Omega_{H2} \approx 2 \times 10^{-4}$ which is above the predictions, but given the good agreement shown versus the $z \approx 0$ COLF from the previous section, this could be subject to uncertainties regarding X_{CO}.

Ω_{HI} roughly follows a power law in $(1+z)$; a best-fit relation to the $50h^{-1}$Mpc run is given by $\Omega_{HI} = 10^{-3.53}(1+z)^{0.74}$, and is higher in amplitude by 20% for the $25h^{-1}$Mpc volume. Generally, this provides a good fit to the trend seen in the compilation of observations from various sources and techniques, particularly for the higher-resolution volume. The difference between the volumes, while only about 0.1 dex, nonetheless suggests that there is suboptimal resolution convergence in this quantity, likely driven by the fact that the $50h^{-1}$Mpc volume does not resolve many low-M_{HI} galaxies as seein in Figure 7. Semi-analytic models tend to predict that Ω_{HI} rises somewhat out to intermediate redshifts, but then falls at $z \gtrsim 1-2$ (Obreschkow et al. 2009; Popping, Behroozi, & Peeples 2015), in clear disagreement with a continued rise in Ω_{HI} out to $z \sim 3.5$. Hence the broad agreement in the redshift evolution of Ω_{HI} is highly encouraging, and suggests that H I in and around galaxies is being viably modeled by MUFASA across a range of epochs.

In contrast, the evolution of Ω_{H2} is predicted to be substantially slower than often believed. MUFASA predicts essentially the same redshift evolution for Ω_{H2} as for Ω_{HI}, with an increase of a factor of $\times 3$ from $z = 0 \rightarrow 3$. The SAMs of Lagos et al. (2011) predict almost no evolution for Ω_{HI}, but a $\times 7$ increase from $z = 0 \rightarrow 3$ for Ω_{H2}. More recently, Lagos et al. (2015) found slower evolution of Ω_{H2} in the EAGLE simulation, more similar to MUFASA. Observations cannot yet clearly distinguish between these predictions.

In summary, MUFASA predicts mild evolution in both the total H I and H_2 cosmic mass densities, scaling approxi-

mately as $(1+z)^{0.8}$. Such a scaling roughly follows from the simple equilibrium model arguments outlined at the start of this section. The evolution of Ω_{HI} is in good agreement with observations, but the predictions for Ω_{H2} are not currently robustly testable. As CO and far-infrared surveys improve with ALMA and other facilities, such constraints will provide important tests of these and other models.

6 FLUCTUATIONS AROUND SCALING RELATIONS

In the prevalent baryon cycling paradigm, quasi-continuous gas inflows drive galaxy growth, modulated by feedback (Somerville & Davé 2015). The net result is that galaxies live on fairly tight scaling relations between stellar mass, star formation rate, metallicity, and gas content (e.g. Finlator & Davé 2008; Davé, Finlator, & Oppenheimer 2012; Lilly et al. 2013; Lagos et al. 2015). Fluctuations in the inflow rate owing to e.g. mergers can cause fluctuations around these scaling relations. Indeed, inflow fluctuations owing to stochastic dark matter infall alone yield a scatter that is in good agreement with the observed scatter in the SFR$-M_*$ relation (Forbes et al. 2014; Mitra et al. 2016).

In addition to scatter in SFR, such fluctuations also give rise to correlated scatter in the metallicity and gas content. For instance, a boost in inflow will enhance the gas content, lower the metallicity, while boosting the SFR owing to the abundance of fresh fuel. Hence one expects that, at a given M_*, high gas content should correlate with low metallicity and high SFR. This results in second-parameter dependences with SFR in the scatter around these scaling relations. In this section we quantify these second-parameter dependences in MUFASA, which provides predictions for baryon cycling that are testable with current and future observations.

Figures 3 and 5 already showed clear second-parameter dependences on the SFR in MUFASA galaxies: Galaxies that have higher SFR for their M_* also have lower metallicities and higher gas fractions in both H I and H_2. To quantify this, we use "deviation plots," i.e. we plot the deviation away from the mean scaling relation in two quantities versus M_* against each other. This isolates the second-order aspects of baryon cycling-driven galaxy evolution by directly quantifying how fluctuations drive correlated scatter, while removing the dependence on the overall inflow rate that sets the first-order (mean) scaling relationship between quantities.

As an example, in order to make a deviation plot for sSFR vs. H I, we begin with the sSFR$-M_*$ and $f_{HI}-M_*$ relations. For each galaxy, we then compute the difference between log sSFR of that galaxy and the median of all galaxies' log sSFR at that M_*; we call this $\Delta\log$ sSFR. Similarly, we compute the difference between $\log f_{HI}$ for that galaxy and the median $\log f_{HI}$ at that galaxy's M_*; this is $\Delta\log f_{HI}$. We can analogously compute $\Delta\log f_{H2}$ and $\Delta\log Z$ for the molecular gas and metallicity, respectively. Note that here we are always using M_* as our independent variable, because this quantity is stable on the (relatively) short timescales over which deviations are occuring; in principle, it is possible to use any property as the independent variable, but we leave such explorations for future work.

Figure 10 shows 2-D histograms of the deviations $\Delta\log f_{HI}$, $\Delta\log f_{H2}$, $\Delta\log$sSFR, and $\Delta\log Z$ plotted against each other. Only star-forming galaxies are included, and for simplicity we only show the $50h^{-1}$Mpc volume at $z=0$ but the trends are similar in the other volumes. The panels along the diagonal show the histograms of deviation values for each quantity, which illustrate the shape of the scatter around the median scaling relation versus M_*. The solid line in each panel shows the best-fit power law to the deviations shown, and the number in the upper right corner is the best-fit slope.

Figure 10 at its most basic level shows that deviations in the SFR, H I, and H_2 all correlate positively with each other, while metallicity deviations (bottom row) anti-correlate with all of the others. This quantifies the amount by which galaxies that lie above the mean MZR also tend to lie below the mean relations in sSFR, f_{HI}, and f_{H2} vs. M_*. Such trends arise naturally in a "gas regulator" type model, which is an ISM mass-balance formalism in which the gas content is allowed to vary (Lilly et al. 2013).

The slope of the best-fit line contains information about how well quantities track each other. For instance, consider $\Delta\log f_{H2}$ vs. $\Delta\log$ sSFR: The slope is close to linear, which means that fluctuations in H_2 are directly tracking fluctuations in SFR. This is unsurprising, since in our simulations it is assumed that the star formation rate of any given gas element is proportional to its f_{H2}; nonetheless, it is not trivial that this translates into a similar trend in galaxy-integrated quantities. The slope versus $\Delta\log f_{HI}$, in contrast, is somewhat sublinear for $\Delta\log f_{H2}$ and $\Delta\log$sSFR, indicating that fluctuations in H_2 and sSFR do not perfectly reflect fluctuations in H I.

The deviations in metallicity versus the MZR, i.e. the plots along the lowermost row, have garnered much attention in the literature. For instance, the panel showing $\Delta\log$ sSFR vs. $\Delta\log Z$ (lower rightmost) corresponds to the FMR, showing that galaxies with higher SFR at a given M_* have lower Z. The best-fit line has a slope of -0.16, which represents a higher-order testable prediction of MUFASA's ability to depict the fluctuations in baryon cycle that give rise to scatter around the scaling relations.

The bottom leftmost panel corresponds to the observational trend noted by Hughes et al. (2013), Lara-López et al. (2013), and Bothwell et al. (2013), which the latter dubbed the H I-FMR: Galaxies with higher H I content at a given M_* are seen to have lower metallicities. Resolved spectroscopy by Moran et al. (2012) indicated that the excess in H I tends to be accompanied by a drop in the outer metallicity, strongly suggesting that this trend is driven by accretion in the outskirts of galaxies. MUFASA predicts a slope for $\Delta\log f_{HI}$ vs. $\Delta\log Z$ of -0.18, similar to but slightly stronger than that vs. $\Delta\log$sSFR. Robertson et al. (2013) measured this deviation slope to be -0.41 ± 0.14 for field galaxies (-0.31 for cluster galaxies). This is steeper than our current predictions, but this was done at a fixed sSFR rather than M_*, which likely accounts for some of the difference. Metallicity is formally most strongly tied to H_2; the slope of this deviation relation is -0.22.

One can also examine the spread of points around the best-fit linear relation within the deviation plots. This is another measure of how tightly any given two quantities fluctuate. One can quantify this by measuring the mean deviation in, say, metallicity, from the best-fit relations involving

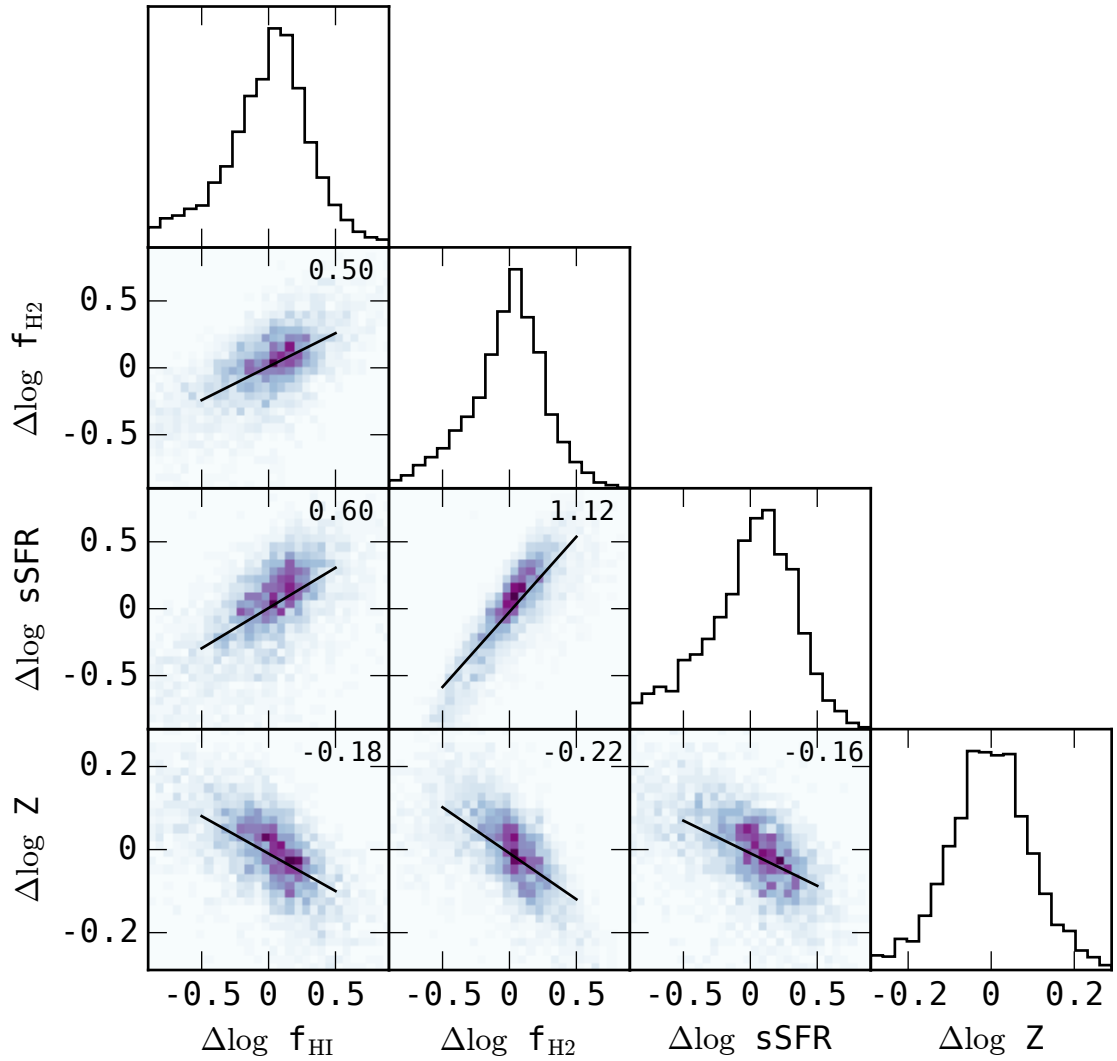

Figure 10. Plots showing the deviation from mean scaling relations versus M_* in our $50h^{-1}$Mpc volume at $z = 0$ for four quantities (in logarithm): H I fraction, H_2 fraction, specific SFR, and metallicity. Scatter plots show these deviations plotted against each other, depicting how fluctuations in these quantities are correlated. The trends can be reasonably represented by the best-fit power laws shown as the solid lines, with the slope indicated in the upper right of each panel. Overall, galaxies at a given M_* with higher sSFR have higher H I and H_2 fractions and lower metallicity. The panels along the diagonal show histograms of the scatter around each scaling relation. While generally Gaussian, there is a tail to low sSFR and gas content arising from green valley galaxies.

$\Delta \log Z$. For the metallicity relations, the mean departure in $\Delta \log Z$ is $0.062, 0.061, 0.064$ with respect to $\Delta \log f_{HI}$, $\Delta \log f_{H2}$, and Δ logsSFR, respectively. This again suggests, at a very marginal level, that metallicity more strongly follows H I than sSFR, which is a conclusion also reached in observational analysis by Bothwell et al. (2013). Still, metallicity tracks H_2 slightly better than either of these quantities (as also found by Lagos et al. 2015).

Finally, the diagonal panels show the scatter of each quantity around the mean scaling relation versus M_*. The shape is generally Gaussian, with a spread that is slightly smaller in H_2 relative to sSFR and H I. Metallicity has quite small scatter, consistent with ~ 0.1 dex as observed Tremonti et al. (2004). In detail there is a longer tail to low-ΔsSFR and correspondingly low gas fraction deviations, which arises from galaxies on their way to quenching.

This deviation plot represents the global view over all galaxies down to the resolution limit of our $50h^{-1}$Mpc volume at $z = 0$. Clearly, it is instructive to examine this plot using galaxies binned by mass, or colour, or at different redshifts. We do not show this here, but we have checked that the trends depicted in Figure 10 are generally well-converged with resolution in the overlapping mass range, and they are qualitatively similar at higher redshifts. One can also examine trends by fixing other quantities besides M_*, such as SFR. In future work, we will explore the implications of these deviation plots in terms of baryon cycling, and present more detailed comparisons to relevant observations.

In summary, deviations plots quantify how galaxies respond to fluctuations in the baryon cycle. By examining only the departures around the mean relations (with respect to M_*), we remove the "first-order" component of galaxy growth (along with many associated systematics) and isolate the impact of "second-order" fluctuations on observ-

able quantities. We thus quantify the correlation in scatter among these various quantities, thereby presenting a new and higher-order test of galaxy formation models. While there currently exist various forms of observational characterisations for these trends, we plan to conduct a more thorough and direct comparison to data regarding second-parameter trends for both gas fractions and metallicities in future work.

7 SUMMARY

We have presented predictions of the MUFASA simulations and compared to observations of the star formation rate, metal, and gas content of galaxies. MUFASA uses state of the art feedback modules and hydrodynamics methodology taken from high-resolution zoom simulations and analytic models. To further extend our dynamic range we employ several simulations using identical input physics but varying in volume (box sizes of $50, 25, 12.5h^{-1}$Mpc), and check that generally simulations at different numerical resolution make similar predictions in their overlapping mass ranges.

Following on Paper I where we showed that MUFASA performed creditably at reproducing the observed stellar mass function over a range of cosmic epochs, here we further show that it also fares well against a number of other key barometers, including several that have not been examined extensively versus previous models such as the specific star formation rate function. We also make novel and testable predictions for the correlations in the fluctuations around mean scaling relations in SFR, metallicity, and gas content, as a direct means to quantify how galaxies respond to fluctuations in the baryon cycle.

Our main results are summarised as follows:

• The star formation rate function in MUFASA shows a Schechter shape with a relatively shallow faint end, in broad agreement with observations out to $z \sim 2$, albeit with a hint that MUFASA overpredicts high-SFR galaxies. This is curious given that in Paper I we demonstrated that MUFASA matches the stellar mass function well but strongly underpredicts the $z \sim 2$ sSFR$-M_*$ relation, which implies that, if anything, MUFASA should *under*predict the SFR function. This highlights the continued difficulty in reconciling current SFR measures during Cosmic Noon with models and, in some cases, among the various data sets themselves.

• The specific SFR function provides a more detailed test of how well models reproduce the scatter around the main sequence. MUFASA reproduces the observed sSFR function at low-z quite well, indicating that this simulation is nicely reproducing the number of galaxies in the green valley, is correctly capturing the spread around the main sequence, and is not missing a large population of starbursts. At $z \sim 1$, the entire sSFR function is shifted by $\sim \times 2$ with respect to observations although the shape matches well, reiterating the result from Paper I showing that the mean sSFR at that epoch is underpredicted by a similar factor.

• The mass-metallicity relation shows a reasonable low-mass slope and amplitude versus observations at both $z \approx 0$ and 2. In contrast, $M_* \gtrsim 2 \times 10^{10} M_\odot$ star-forming galaxies at $z = 0$ continue to show a strong rise in the MZR that does not agree well with observations, and then abruptly flattens at roughly the appropriate metallicity. We conjecture that wind recycling, which plays a key role in setting the MZR at these masses, may be too vigorous in our simulations at these masses, or else these galaxies should have a metal loading factor above our assumed value of unity. Finally, the MZR clearly shows a second-parameter trend such that galaxies with high SFR at a given M_* have lower metallicity.

• MUFASA directly tracks molecular gas, hence we can separate the gas content into atomic, molecular, and ionised. MUFASA well reproduces observations of the total cold gas fraction (H I+H_2) as a function of M_*, and provides a fair match to the H I fraction individually, with a notable deficit at high masses in only our lowest-resolution run. Like with the metallicity, H I and H_2 content also show a second parameter trend that galaxies with high SFR tend to have higher gas fractions.

• Gas fractions are broadly predicted to increase with redshift, which at least qualitatively agrees with observations. However, the predicted rate of evolution for H_2 ($\sim \times 2-3$ out to $z \sim 2$) is slower than canonically observed for molecular gas. H_2 evolves similarly across all masses, while H I evolves slightly faster at higher masses; there is an order of magnitude more H I in a $M_* = 10^{11} M_\odot$ galaxy at $z \sim 1-2$ versus today.

• As a result of the rapid H I evolution at high masses, the bright end of the H I mass function evolves fairly rapidly as well. The predicted HIMF agrees well with observations at $z \sim 0$, and evolves upwards at all masses by $\sim \times 2-3$ by $z = 1$. At $z = 2$ we predict a steeper faint end, although this may not be accessibly observationally in 21 cm prior to the full SKA.

• In order to explore the potential discrepancy in molecular gas evolution further, we use the simulation-based prescription from Narayanan et al. (2012) to convert MUFASA molecular gas masses to a CO luminosity based on the metallicity and molecular gas content, and compare to inferred CO luminosity functions observed out to $z \sim 2$. We find surprisingly good agreement at all masses for the COLF, despite the mild evolution in f_{H2}. This highlights that systematic uncertainties in X_{CO} can be a overriding factor in making robust comparisons to molecular gas content data at intermediate redshifts.

• The cosmic mass density in H I is predicted to evolve mildly upwards out to high-z, such that $\Omega_{HI} \propto (1+z)^{0.7-0.8}$. This evolution is in good agreement with observations from a variety of techniques. The amplitude is somewhat sensitive to resolution, and our $25h^{-1}$Mpc volume has 10–20% higher Ω_{HI} than our $50h^{-1}$Mpc cube, which agrees slightly better with data.

• In accord with our predicted mild evolution for f_{H2}, we also predict mild evolution for Ω_{H2}, with a similar redshift scaling as Ω_{HI} but lower by $\sim \times 3$. We note that this is much less steep than the evolution of the cosmic SFR density; it has been suggested that the drop in cosmic SFR owes directly to the drop in molecular gas mass, but MUFASA does not support this intepretation, as the cosmic SFRD (see Paper I) evolves significantly more rapidly than Ω_{H2}.

• An independent and higher-order test of models is whether they reproduce the observed scatter around the mean scaling relations. We make predictions for this using deviation plots, where we correlate the deviations for each galaxy relative to scaling relations in sSFR, metallicity, f_{H2}, and f_{HI} versus M_*. We show that MUFASA qualitatively re-

produces observed trends that indicate that at a given M_*, galaxies with high SFR have low metallicity and high gas content. We make predictions for the power-law slopes between deviations in sSFR, metallicity, f_{H2}, and f_{HI} that can be tested against observations.

As in Paper I, MUFASA continues to demonstrate good agreement with now a wider range of galaxy observables across cosmic time, indicating that it provides a viable platform to study the physics of galaxy evolution in a cosmological context on $\gtrsim$kpc scales. This implies, among other things, that employing scalings taken from the FIRE simulations into cosmological-scale runs satisfyingly reproduces some of the same data-concordant trends as individual FIRE zoom runs. Our current heuristic quenching model seems to populate the green valley and lower their gas contents relative to blue cloud galaxies approximately as observed, though we are working towards a more self-consistent black hole growth and feedback model that may substantially impact these predictions particularly at the massive end. MUFASA's successes showcase the emerging promise of cosmologically-situated galaxy formation simulations in helping to understand the Universe as mapped through large-scale multi-wavelength galaxy surveys probing the various constituents of galaxies back to early epochs.

ACKNOWLEDGEMENTS

The authors thank D. Anglés-Alcázar, F. Durier, K. Finlator, S. Huang, N. Katz, D. Narayanan, and R. Somerville for helpful conversations and comments. RD, MR, and RJT acknowledge support from the South African Research Chairs Initiative and the South African National Research Foundation. Support for MR was also provided by the Square Kilometre Array post-graduate bursary program. Support for RD was provided by NASA ATP grant NNX12AH86G to the University of Arizona. Support for PFH was provided by an Alfred P. Sloan Research Fellowship, NASA ATP Grant NNX14AH35G, and NSF Collaborative Research Grant #1411920 and CAREER grant #1455342. The MUFASA simulations were run on the Pumbaa astrophysics computing cluster hosted at the University of the Western Cape, which was generously funded by UWC's Office of the Deputy Vice Chancellor. These simulations were run with revision e77f814 of GIZMO hosted at `https://bitbucket.org/rthompson/gizmo`.

REFERENCES

Andrews, B. & Martini, P. 2013, ApJ, 765, 140
Asplund, M., Grevesse, N., Sauval, A. J., Scott, P. 2009, ARA&A, 47, 481
Bahé, Y. M. et al. 2016, MNRAS, 456, 1115
Behroozi, P. S., Wechsler, R. H., Conroy, C. 2013, ApJ, 770, 57
Benson, A. J. 2014, MNRAS, 444, 2599
Boselli, A., Cortese, L., Boquien, M., Boissier, S., Catinella, B., Lagos, C., Saintonge, A. 2014, A&A, 564, 66
Bothwell, M. S. et al. 2011, MNRAS, 415, 1815
Bothwell, M. S., Maiolino, R., Kennicutt, R., Cresci, G., Mannucci, F., Marconi, A., Cicone, C. 2013, MNRAS, 433, 1425
Bouché, N., Gardner, J. P., Katz, N., Weinberg, D. H., Davé, R., Lowenthal, J. D. 2005, ApJ, 628, 89
Brown, J. S., Martini, P., & Andrews, B. H. 2016, MNRAS, 458, 1529
Bruzual, G., Charlot, S. 2003, MNRAS, 344, 1000
Catinella, B. et al. 2010, MNRAS, 403, 683
Catinella, B. et al. 2013, MNRAS, 436, 34
Chabrier G., 2003, PASP, 115, 763
Colbert, J. W. 2013, ApJ, 779, 34
Crain, R. A. 2015, MNRAS, 450, 1937
Crain, R. A. 2016, MNRAS, submitted, arXiv:1604.06803
Croton, D. J., et al. 2006, MNRAS, 365, 11
Davé, R. 2008, MNRAS, 385, 147
Davé, R., Finlator, K. M., Oppenheimer, B. D. 2011, MNRAS, 416, 1354
Davé, R., Finlator, K. M., Oppenheimer, B. D. 2012, MNRAS, 421, 98
Davé, R., Katz, N., Oppenheimer, B. D., Kollmeier, J. A., Weinberg, D. H. 2013, MNRAS, 434, 2645
Davé, R., Thompson, R. J., Hopkins, P. F. 2016, MNRAS, accepted
Decarli, R. et al. 2016, ApJ, submitted, arXiv:1607.06771
Delhaize, J., Meyer, M. J., Staveley-Smith, L., Boyle, B. J. 2013, MNRAS, 433, 1398
Dekel, A. et al. 2009, Nature, 457, 451
Duffy, A.R., Schaye, J., Kay, S.T., Dalla Vecchia, C. 2008, MNRAS, 390, L64
Ellison, S. L., Patton, D. R., Simard, L., & McConnachie, A. W. 2008, ApJL, 672, L107
The Enzo Collaboration, Bryan, G. L., Norman, M. L., et al. 2014, ApJS, 211, 19
Erb, D. K., Shapley, A. E., Pettini, M., Steidel, C. C., Reddy, N. A., Adelberger, K. L. 2006, ApJ, 644, 813
Faucher-Giguere, C. A., Lidz, A., Zaldarriaga, M., Hernquist, L. 2009, ApJ, 703, 1416
Feldmann R., Gnedin N. Y., Kravtsov A. V., 2012, ApJ, 747, 124
Fernández, X. et al. 2016, ApJ, 824, 1
Finlator, K. & Davé, R. 2008, MNRAS, 385, 2181
Font-Ribera, A. et al. 2012, JCAP, 11, 59
Forbes, J. C., Krumholz, M. R., Burkert, A., Dekel, A. 2014, MNRAS, 443, 168
Gabor, J. M., Davé, R., Oppenheimer, B. D., Finlator, K. M. 2011, MNRAS, 417, 2676
Gabor, J. M. & Davé, R. 2012, MNRAS, 427, 1816
Gabor, J. M. & Davé, R. 2015, MNRAS, 447, 374
Geach, J. E., Smail, I., Moran, S. M., MacArthur, L. A., Lagos, C. P., Edge, A. C. 2011, ApJ, 730, 19
Genel, S. et al. 2014, MNRAS, 445, 175
Grasshorn Gebhardt, H. S., Zeimann, G. R., Ciardullo, R., et al. 2016, ApJ, 817, 10
Gunawardhana, M. L. P. et al. 2013, MNRAS, 433, 2764
Hahn, O., Abel, T. 2011, MNRAS, 415, 2101
Haynes, M. P. et al. 2011, AJ, 142, 170
Holwerda, B. W., Blyth, S.-L., Baker, A. J. 2012, IAUS v.284, p.496
Hopkins, P.F. 2015, MNRAS, 450, 53
Hughes, T. M., Cortese, L., Boselli, A., Gavazzi, G., & Davies, J. I. 2013, A&A, 550, A115
Ilbert, O. et al. 2015, A&A, 579, 2

Iwamoto, K., Brachwitz, F., Nomoto, K., Kishimoto, N., Umeda, H., Hix, W. R., Thielemann, F.-K. 1999, ApJS, 125, 439
Kaviraj, S., Laigle, C., Kimm, T., Devriendt, J. E. G., Dubois, Y., Pichon, C., Slyz, A., Chisari, E., Peirani, S. 2016, MNRAS, submitted, arXiv:1605.09379
Kennicutt, R. C. 1998, ApJ, 498, 541
Kereš, D., Yun, M. S., Young, J. S. 2003, ApJ, 582, 659
Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2
Kewley, L. J. & Ellison, S. L. 2008, ApJ, 681, 1183
Khandai, N., Di Matteo, T., Croft, R., Wilkins, S., Feng, Y., Tucker, E., DeGraf, C., Liu, M.-S. 2015, MNRAS, 450, 1349
Kim, J.-h., Abel, T., Agertz, O., et al. 2014, ApJS, 210, 14
Kriek, M. et al. 2015, ApJS, 218, 15
Krumholz, M. R., McKee, C. F., Tumlinson, J. T. 2008, ApJ, 693, 216
Krumholz, M. R., McKee, C. F., Tumlinson, J. T. 2011, ApJ, 729, 36
Lagos, C. Del P., Baugh, C. M., Lacey, C. G., Benson, A. J., Kim, H.-S., Power, C. 2011, MNRAS, 418, 1649
Lagos, C. Del P., et al. 2015, MNRAS, 459, 2632
Lara-López, M. A. et al. 2010, A& 521, L53
Lara-López, M. A., Hopkins, A. M., López-Sánchez, A. R., et al. 2013, MNRAS, 433, L35
Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., Peng, Y. 2013, ApJ, 772, 119
Ma, X., Hopkins, P. F., Faucher-Giguère, C.-A., Zolman, N., Muratov, A. L., Kereš, D., Quataert, E. 2016, MNRAS, 456, 2140
Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., Gnerucci, A. 2010, MNRAS, 408, 2115
McNamara, B. R. & Nulsen, P. E. J. 2007, ARA&A, 45, 117
Mehta, V. et al. 2016, ApJ, accepted, arXiv:1505.07843
Mitra, S., Davé, R., Finlator, K. 2015, MNRAS, 452, 1184
Mitra, S., Davé, R., Simha, V., Finlator, K. 2016, MNRAS, submitted, arXiv:1606.07436
Mo, H.J., Mao, S., White, S.D.M. 1998, MNRAS, 295, 319
Moran, S.M. et al. 2012, ApJ, 745, 66
Muratov, A.L., Kereš, D., Faucher-Giguere, C. A., Hopkins, P. F., Quataert, E., Murray, N. 2015, MNRAS, 454, 2691
Narayanan, D., Krumholz, M. R., Ostriker, E. C., Hernquist, L. 2012, MNRAS, 421, 3127
Narayanan, D., Bothwell, M., Davé, R. 2012, MNRAS, 426, 1178
Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., Maeda, K. 2006, NuPhA, 777, 424
Noterdaeme, P. et al. 2012, A&A, 547, L1
Obreschkow, D., Croton, D., De Lucia, G., Khochfar, S., Rawlings, S. 2009, ApJ, 698, 1467
Oppenheimer, B. D. & Davé, R. 2008, MNRAS, 387, 577
Pasquali, A., Gallazzi, A., Fontanot, F., van den Bosch, F. C., De Lucia, G., Mo, H. J., Yang, X. 2010, MNRAS, 407, 937
Peeples, M. S., Shankar, F. 2011, MNRAS, 417, 2962
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, arXiv:1502.01589
Popping, A., Davé, R., Braun, R., Oppenheimer, B. D. 2009, A&A, 504, 15
Popping, G., Somerville, R. S., Trager, S. C. 2014, MNRAS, 442, 2398
Popping, G., Behroozi, P. S., Peeples, M. S. 2015, MNRAS, 449, 477
Popping, G. et al. 2015, MNRAS, 454, 2258
Prochaska, J.X., Wolfe, A.M. 2009, ApJ, 696, 1543
Rafieferantsoa, M., Davé, R., Anglés-Alcàzar, D., Katz, N., Kollmeier, J. A., Oppenheimer, B. D. 2015, MNRAS, 453, 3980
Rahmati, A., Pawlik, A. H., Raičević, M., Schaye, J.
Rao, S.M., Turnshek, D.A., Nestor, D.B. 2006, ApJ, 636, 610
Robertson, P., Shields, G. A., Davé, R., Blanc, G. A., Wright, A. 2013, ApJ, 773, 4
Salim, S. et al. 2007, ApJS, 207, 173, 267
Salim, S. et al. 2014, ApJ, 797, 126
Salim, S., Lee, J. C., Davé, R., Dickinson, M. 2015, ApJ, 808, 25
Saintonge, R. L. et al. 2011, MNRAS, 415, 32
Saintonge, R. L. et al. 2016, MNRAS, 462, 1749
Sanders, R. L. et al. 2015, ApJ, 799, 138
Sanders, R. L. et al. 2016, ApJ, 816, 23
Schaye, J. & Dalla Vecchia, C. 2008, MNRAS, 383, 1210
Schaye, J. et al. 2015, MNRAS, 446, 521
Scoville, N. et al. 2016, ApJ, 820, 83
Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., Hernquist, L. 2008, MNRAS, 391, 481
Somerville, R. S., Davé, R. 2015, ARA&A, 53, 51
Springel, V. 2005, MNRAS, 364, 1105
Steidel, C. C. et al. 2014, ApJ, 795, 165
Steidel, C. C., Strom, A. L., Pettini, M., Rudie, G. C., Reddy, N. A., Trainor, R. F. 2016, ApJ, 826, 159
Tacchella, L. J. et al. 2016, MNRAS, 458, 242
Tacconi, L. J. et al. 2013, ApJ, 768, 74
Telford, O. G., Dalcanton, J. J., Skillman, E. D., & Conroy, C. 2016, ApJ, 827, 35
Thompson, R., Nagamine, K., Jaacks, J., Choi, J.-H. 2014, ApJ, 780, 145
Thompson, R. 2015, ASCL, 1502, 12
Tremonti, C. A. et al. 2004, ApJ, 613, 898
Vagnozzi, S., Freese, K., Zurbuchen, T. H. 2016, arXiv:1603.05960
Vallini, L., Gruppioni, C., Pozzi, F., Vignali, C., Zamorani, G. 2016, MNRAS, 456, L40
Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Nelson, D., Hernquist, L. 2014, MNRAS, 444, 1518
Zahid, H. J. et al. 2014, ApJ, 792, 75