FERMILAB-PUB-20-012-A

Draft version January 9, 2020
Typeset using LaTeX **twocolumn** style in AASTeX62

arXiv:2001.02233v1 [astro-ph.GA] 7 Jan 2020

Cosmic Reionization On Computers: The Galaxy-Halo Connection between $5 \leq z \leq 10$

Hanjue Zhu (朱涵珏),[1] Camille Avestruz,[2] and Nickolay Y. Gnedin[3,4,5]

[1] *The University of Chicago; Chicago, IL 60637, USA*
[2] *Leinweber Center for Theoretical Physics; Department of Physics; University of Michigan, Ann Arbor, MI 48109, USA*
[3] *Particle Astrophysics Center; Fermi National Accelerator Laboratory; Batavia, IL 60510, USA*
[4] *Kavli Institute for Cosmological Physics; The University of Chicago; Chicago, IL 60637, USA*
[5] *Department of Astronomy & Astrophysics; The University of Chicago; Chicago, IL 60637, USA*

(Received December 2, 2019)

Submitted to ApJ

ABSTRACT

We explore the connection between the stellar component of galaxies and their host halos during the epoch of reionization ($5 \leq z \leq 10$) using the CROC (Cosmic Reionization on Computers) simulations. We compare simulated galaxies with observations and find that CROC underpredicts the abundance of luminous galaxies when compared to observed UV luminosity functions, and analogously the most massive galaxies when compared to observed stellar mass functions. We can trace the deficit of star formation to high redshifts, where the slope of the star formation rate to stellar mass relation is consistent with observations, but the normalization is systematically low. This results in a star formation rate density and stellar mass density that is systematically offset from observations. However, the less luminous or lower stellar mass objects have luminosities and stellar masses that agree fairly well with observational data. We explore the stellar-to-halo mass ratio, a key quantity that is difficult to measure at high redshifts and that models do not consistently predict. In CROC, the stellar-to-halo mass ratio *decreases* with redshift, a trend opposite to some abundance matching studies. These discrepancies uncover where future effort should be focused in order to improve the fidelity of modeling cosmic reionization. We also compare the CROC galaxy bias with observational measurements using Lyman-Break Galaxy (LBG) samples. The good agreement of simulation and data shows that the clustering of dark matter halos is properly captured in CROC.

Keywords: galaxies — methods, numerical — cosmology

1. INTRODUCTION

Observations of high-redshift galaxies inform modeling of galaxy formation during the epoch of reionization (e.g. Bradley et al. 2012, Schmidt et al. 2014, Zitrin et al. 2014, Zitrin et al. 2015, Oesch et al. 2015, Oesch et al. 2016, Bernard et al. 2016, Livermore et al. 2017, Ishigaki et al. 2018, Morishita et al. 2018). Current observations now robustly constrain the faint-end slope of the UV luminosity functions at $z \gtrsim 6$. These observations come from gravitationally lensed images of the most distant galaxies, but lensing measurements have systematic uncertainties associated with them (Wyithe et al. 2011). Different measurements of the stellar mass function at these redshifts also have considerable discrepancies. Uncertainties in stellar mass measurements come from both the intrinsic scatter in empirical relations and uncertainties in model-dependent assumptions used to derive the stellar mass.

Future observational facilities, such as the *James Webb Space Telescope* (*JWST*, Gardner et al. (2006)) and next generation ground-based telescopes, will increase the sample of high redshift galaxies and push observational limits to fainter values. Data from these telescopes will help determine the faint end slopes of the UV luminosity function and the stellar mass function. To maximize the impact of observations, we need robust theoretical models to better predict high redshift

Corresponding author: Hanjue Zhu (朱涵珏)
hanjuezhu@uchicago.edu

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

galaxy properties and aid in the interpretation of upcoming data.

Galaxy formation is difficult to model because it incorporates non-linear interplay between a wide range of physical processes, such as gas cooling, star formation, and stellar feedback processes. On the other hand, the growth of large scale structure and the evolution of dark matter halos can be predicted robustly using numerical simulations in the ΛCDM model of cosmology. Dark matter halos grow through cosmic accretion and mergers with other halos. Galaxies form within dark matter halos, forming stars from gas reservoirs, and growing through both accretion of gas and stars and galaxy-galaxy mergers. Given the robust results from dark matter-only simulations and the complexity of galaxy formation processes, it is natural for us to explore the connections between observed galaxies and the underlying dark matter distribution (see Wechsler & Tinker (2018) for a recent review). Galaxy-halo connections encapsulate the physical processes driving galaxy formation. This connection generally illustrates empirical correlations between galaxy properties and the properties of the extended dark matter halos that host galaxies (e.g. Conroy et al. 2006, White et al. 2007, Zheng et al. 2007, Firmani & Avila-Reese 2010, Wang et al. 2013, Tinker et al. 2013, Birrer et al. 2014, Sun & Furlanetto 2016, Cohn 2017, Mitra et al. 2017).

Recent observations have also measured clustering properties of high redshift galaxies as a way to further connect observed galaxies with their host dark matter halos (Barone-Nugent et al. 2014, Sobacchi & Mesinger 2015, Harikane et al. 2018). The clustering of dark matter halos is easily obtained from numerical simulations. However, galaxy clustering depends on the detailed relationship between the galaxies and their dark matter host halos. Clustering statistics can therefore provide additional constraints to galaxy formation models.

There are two kinds of approaches to model the galaxy-halo connection - empirical modeling and physical modeling. Empirical modeling uses data to constrain parameters that describe the galaxy-halo connection; physical modeling either parameterizes (semi-analytical modeling) or directly simulates the physics of galaxy formation using cosmological hydrodynamical simulations. Simulations solve for gravity and hydrodynamics, while incorporating extensive physical prescriptions for galaxy formation processes.

The Cosmic Reionization on Computers (CROC) project produces such cosmological simulations that self-consistently model relevant physical processes in cosmological volumes during the Epoch of Reionization. CROC therefore connects the dynamics of dark matter and galaxy formation.

In this paper, we first look at how CROC outputs compare to the observed luminosity and stellar mass properties - we show the CROC UV luminosity function in Section 3.1, the stellar mass-to-light ratio in Section 3.2, and the stellar mass function in 3.3. We additionally quantify the galaxy-halo connection in CROC by looking at the stellar-to-halo mass ratio (SHMR) in 3.4. Subsequently, we show the star formation histories in 3.5. Last, we use the relationship between the UV luminosity and halo mass in CROC to calculate the galaxy bias, or the excess clustering of galaxies compared with that of the underlying matter, and compare with observational results from Barone-Nugent et al. (2014) in Section 3.6. The galaxy bias complements the other tests of the galaxy-halo connection.

2. METHODOLOGY

2.1. *CROC Simulations*

All CROC simulations we performed with the Adaptive Refinement Tree (ART) code (Kravtsov 1999; Kravtsov et al. 2002; Rudd et al. 2008). They include a wide range of physical processes that are thought to be necessary for self-consistent modeling of cosmic reionization, such as gravity and gas dynamics, fully coupled radiative transfer, atomic cooling and heating processes, molecular hydrogen formation, star formation and stellar feedback. Full details of the simulations are described in the CROC method paper Gnedin (2014). For stellar population syntesis modeling, we use the Flexible Spectral Population Synthesis code (FSPS Conroy & Gunn 2010a).

In this paper we use two sets of simulations: a set of 6 independent random realizations in $L_{box} = 40h^{-1}$ Mpc comoving boxes with 1024^3 dark matter particles and a set of 3 independent random realizations in $L_{box} = 80h^{-1}$ Mpc comoving boxes with 2048^3 dark matter particles. Both simulation sets have the same spatial resolution of 100 pc in proper units (kept constant throughout the simulation) and effective mass resolution of $M_1 = 7 \times 10^6 \mathrm{M}_\odot$ (defined as the mass of a dark matter particle in an equivalent dark-matter-only simulation).

2.2. *Halo Finder*

We use the "yt" package to identify dark matter halos in the simulations (Turk et al. 2011). Yt supports two different halo finders: HOP (Eisenstein & Hut 1998) and Rockstar (Behroozi et al. 2013). We compared results using the $L_{box} = 40h^{-1}$ Mpc boxes, which have halo catalogs from both halo finders. Both halo finders produce virtually indistinguishable results for all rela-

tions discussed in future sections. HOP catalogs are available for all CROC simulation boxes; Rockstar catalogs are available for all $L_{box} = 40h^{-1}$ Mpc boxes but not for $L_{box} = 80h^{-1}$ Mpc due to numerical limitations of the Rockstar implementation on Blue Waters supercomputer. Unless specifically mentioned, by default we use HOP catalogs for the $L_{box} = 80h^{-1}$ Mpc dataset.

Note, the halo mass output by HOP we choose is the mass within a radius enclosing an average density that is 160 times the critical density of the universe. The halo mass selected from the *Rockstar* catalogs uses the mass definition from Bryan & Norman (1998): the mass within a radius enclosing an average density 360 times the background density at $z = 0$, including unbound particles.

3. RESULTS

3.1. *The UV Luminosity function*

We first compare the CROC galaxy UV luminosity function (UVLF) with observational measurements in the literature. Since the UVLF is relatively well measured all the way to $z \sim 10$, it is a natural choice for comparing theoretical predictions with observations. Galaxy luminosities at 1500Å are computed using the Flexible Stellar Population Synthesis (FSPS) code (Conroy et al. 2009, 2010; Conroy & Gunn 2010b), including an observationally calibrated dust correction, as described in the first CROC paper, Appendix B (Gnedin 2014).

Figure 1 shows the UVLF between $5 < z < 10$ for CROC galaxies from all $L_{box} = 80h^{-1}$ Mpc and $L_{box} = 40h^{-1}$ Mpc boxes. We overplot a compilation of observed values. There is general agreement within the observational measurements at low and moderate luminosities, but CROC simulations systematically underpredict luminosities of brighter galaxies with magnitudes $M_{1500} \lesssim -19$. Since the results from the $80h^{-1}$Mpc and $40h^{-1}$Mpc boxes agree fairly well, the source of the discrepancy is not in the limited volume of the CROC simulations.

The primary reason for the discrepancy, therefore, is the model for star formation and stellar feedback used in CROC. As discussed in Gnedin (2014), this model assumes a linear Kennicutt-Schmidt relation,

$$\dot{\Sigma}_* = \frac{\Sigma_{\mathrm{H}_2}}{\tau_{\mathrm{DEP}}}, \tag{1}$$

where the molecular gas depletion time $\tau_{\mathrm{DEP}} = 1.5\,\mathrm{Gyr}$ is taken as a universal constant, and Σ_{H_2} is computed using the fitting formulae from Gnedin & Draine (2014). Stellar feedback is modeled with a "blastwave" approximation (Stinson et al. 2009, 2013; Governato et al. 2010; Agertz et al. 2011, 2013; Brook et al. 2012).

The physical effects that are not modeled in CROC simulations, such as AGN feedback and cosmic ray feedback, are expected to reduce the abundance of luminous galaxies (Sijacki et al. 2015; Kaviraj et al. 2017). The inclusion of these additional processes would likely exacerbate, rather than resolve, the discrepancy with observations. The exact cause of the suppression of star formation in massive galaxies in CROC simulations is unclear. Such a cause may lie in the non-constancy of the molecular gas depletion time; this is less likely as the depletion time is a proxy to star formation efficiency, and it is well established that self-regulation makes star formation rates in the simulations independent of the assumed value of efficiency (Hopkins et al. 2018; Orr et al. 2018; Semenov et al. 2018); for the same reason the detailed model for molecular hydrogen production (which serves as another proxy for star formation efficiency) should not matter for massive galaxies that are expected to be in the self-regulation regime. Another possibility is the excessive strength of stellar feedback in the CROC model. The assumed feedback strength does indeed affect the overall star formation rates in simulated galaxies (Orr et al. 2018; Semenov et al. 2018) approximately proportionately; however, stronger feedback usually affects less massive galaxies more, while in our case it is in the most massive galaxies that the star formation rate is mis-modeled most significantly. These considerations serve as a motivation for further testing of the sub-grid models in CROC.

3.2. *The Stellar Mass-to-Light Ratio*

In this section, we examine the stellar mass to luminosity ratio, or the stellar mass-to-light ratio, as a function of galaxy stellar mass. We plot this instead of the $M_\star - L_\star$ relation to reduce the dynamic range and better illustrate modest trends. Note, the two relations are interchangeable.

In Figure 2 we show the stellar mass-to-light ratio of galaxies in the CROC simulations along with observational data from Song et al. (2016) and Duncan et al. (2014). The line colors correspond to a redshift range of $5 \leq z \leq 10$. We see that at fixed stellar mass, the mass-to-light ratio increases moderately with time for both the CROC sample and the observations, and the overall agreement is fair. This agreement means that the amount of light produced per unit mass of stars in CROC galaxies (i.e. the shapes of individual star formation histories) is captured faithfully. On the other hand, the lack of bright galaxies in the simulations is apparent in Figure 1, indicating that the overall level of star formation in the most massive halos in CROC is lower than that of observed galaxies. In other words, the star

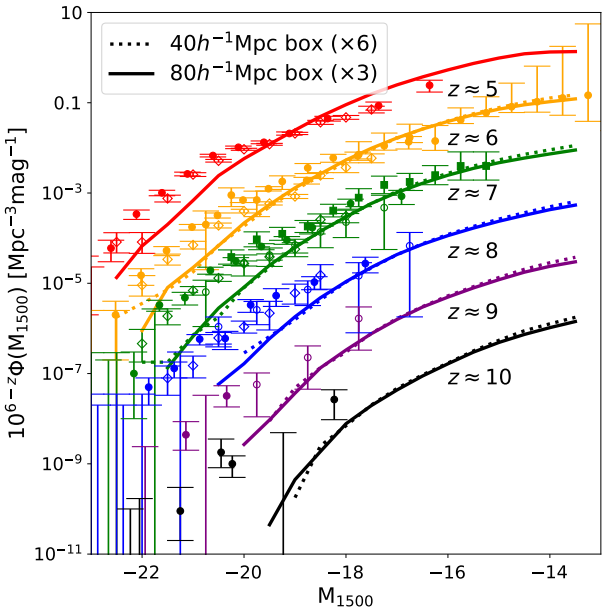

Figure 1. Comparison between the UV luminosity function (UVLF) from the CROC simulation suite boxes and observations. We manually apply an offset of 1 dex to each redshift to facilitate comparison between CROC and observational data at each redshift. The dotted lines correspond to the UVLF from the six $40h^{-1}$Mpc boxes in our suite. Solid lines correspond to the UVLF from the three $80h^{-1}$Mpc boxes. The two sets of boxes agree well at all redshifts shown. The observational data are from Atek et al. (2015) (filled squares), Bouwens et al. (2015, 2017) (filled circles), Finkelstein et al. (2015) (open diamonds), and Laporte et al. (2016) (open circles). While there is general agreement between the CROC predictions and observations, CROC systematically underpredicts current measurements at magnitudes $M_{1500} \lesssim -19$.

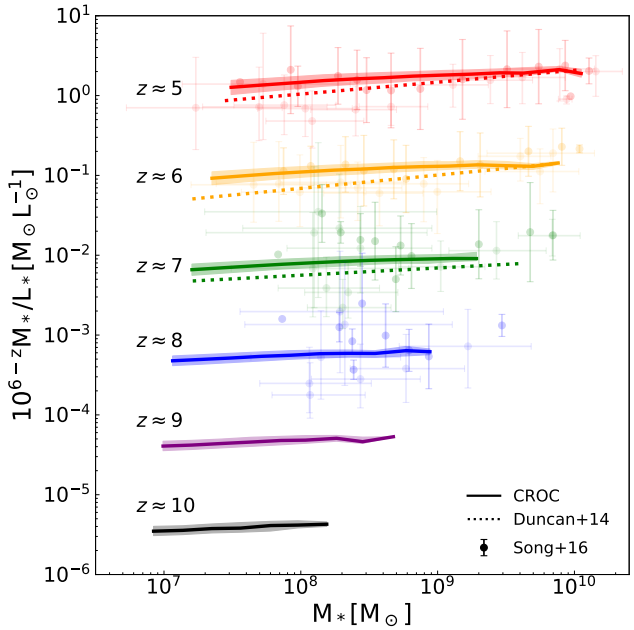

Figure 2. Mass to light ratio of CROC galaxies as a function of stellar mass. The solid line corresponds to the median CROC ratios in each stellar mass bin, and the transparent bands illustrate the 25th and 75th percentile values in each stellar mass bin. Observational data are from Song et al. (2016) and Duncan et al. (2014).

formation and stellar feedback model in the simulation works reasonably well in fainter galaxies ($M_{1500} \gtrsim -18$) ($L < L_\star$) but is underprucing stars in brighter, $L > L_\star$ ($M_{1500} \lesssim -18$), galaxies. The same deficiency should therefore manifest itself in the underprediction of stellar mass functions for the most massive galaxies.

The good agreement for stellar mass-to-light ratios also implies that any potential error due to stellar population synthesis modeling is insignificant: in the simulations we use Flexible Spectral Population Synthesis code (FSPS Conroy & Gunn 2010a), while observational measurements rely on Bruzual & Charlot models (Bruzual & Charlot 2003, 2011).

3.3. *Evolution of the Stellar Mass Function*

The stellar mass function (SMF) provides another illustrative comparison between CROC and observations; in the simulations the stellar mass is a primary, directly simulated quantity, while galactic UV luminosity is computed in post-processing. However, this is with the caveat that the observational inference of the (SMF) is more assumption dependent than the UVLF. Figure 3 shows the stellar mass function of CROC galaxies and the observed stellar mass functions from Duncan et al. (2014) and Song et al. (2016).

The stellar mass function exhibits the deficit of massive galaxies, $M_\star \gtrsim 3 \times 10^8 M_\odot$, which is consistent with the deficit of bright galaxies discussed above. Since the stellar mass-to-light ratios of massive galaxies in CROC are realistic, the discrepancy we see in the UVLF in Figure 1 is driven by the inaccuracy of the star formation and stellar feedback model and not, for example, by the errors in computing stellar luminosities. In other words, the galaxies that live in the most massive CROC halos are less massive and less luminous than those actually observed. In order to quantify this further, we consider below the relation between stellar and halo masses in the simulations.

For some applications, this deficiency of the simulations can be "corrected" in post-processing by re-scaling the simulated stellar masses as

$$M_\star \to M_\star \left(1 + A_z \log_{10}\left(1 + \frac{M_\star}{3 \times 10^8 M_\odot}\right)\right), \quad (2)$$

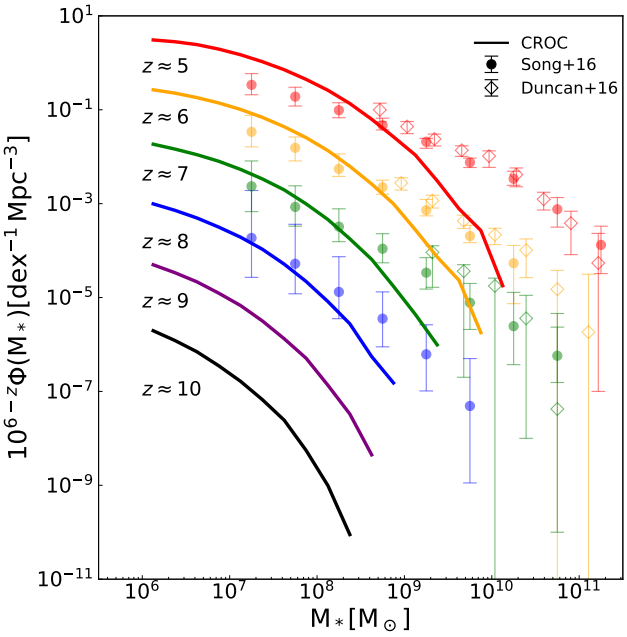

Figure 3. Evolution of stellar mass function of CROC galaxies. We overplot observational data from Song et al. (2016) in circles and data from Duncan et al. (2014) in diamonds. Relative to observations, CROC has a deficit of high mass objects at $z < 8$.

where $A_z \approx 6, 10, 30$ at $z = 5, 6, 7$. We do not use such re-scaling in this paper, as our goal is to emphasize the regimes where the simulations fail and to identify what needs to be fixed in future simulations, not to simply match all the existing observations.

3.4. *Galaxy Properties and Halo Mass*

In order to quantify the galaxy-halo connection in CROC, we first examine the stellar-to-halo mass ratio (SHMR), which quantifies the integrated efficiency of star formation and merger-driven growth.

Figure 4 shows the stellar-to-halo mass ratio as a function of halo mass between $5 \leq z \leq 10$. At fixed redshift, the relationship flattens for halos with $M_h \gtrsim 10^{10}\mathrm{M}_\odot$. The flattening is consistent with the deficiency in massive galaxies apparent in Figs. 1 and 3. As host halos grow through hierarchical formation to form the largest objects, galaxies do not seem to be forming more stars. We explore the evolution of star formation separately in subsequent sections.

Separate from the flattening of the SHMR is the redshift evolution, which is somewhat of an open question given discrepancies between models and observations. At fixed halo mass, the stellar-to-halo mass ratio of CROC galaxies decreases with increasing redshift, indicating the star formation efficiency in galaxies is lower in the past. This is consistent with the results from Tacchella et al. (2013), where they assume a time delay for star formation that has a larger impact at higher redshifts.

This evolutionary trend, however, is contrary to the observational findings of Finkelstein et al. (2015), Harikane et al. (2016) and Stefanon et al. (2017). Both Finkelstein et al. (2015) and Harikane et al. (2016) found evidence that the SHMR *increases* from $z = 4$ to $z = 7$. On the other hand, Stefanon et al. (2017) found no evidence of evolution between $4 \leq z \leq 7$, consistent with model predictions from the *BLUETIDES simulation* (Wilkins et al. 2017; Bhowmick et al. 2018) and from the FIRE-2 simulations (Ma et al. 2018).

The observational methods for constraining the SHMR vary; Harikane et al. (2016) combined a halo occupation distribution model for angular correlation functions to constrain the SHMR with measurements of stellar mass, and Stefanon et al. (2017) used an abundance matching method. Additionally, the uncertainties associated with the observed SHMR at high redshifts have the same sources as those in the SMF, where stellar masses are derived from empirical relations with large intrinsic scatter. These factors could explain the disparity across observations in constraining the SHMR, and much improved constraints can be expected from the forthcoming JWST observations.

The comparison with observational constraints from Finkelstein et al. (2015) and Harikane et al. (2016) needs to be done with care as the observations only sample halo masses in excess of $10^{11}\mathrm{M}_\odot$ ($M_* > 3 \times 10^8\mathrm{M}_\odot$), where CROC simulations fail. As an illustration, however, we can apply the correction from Equation (2) to extend simulation results to higher stellar masses. Surprisingly, we find that we are unable to match Harikane et al. (2016) with corrections (2) that are chosen to reproduce Duncan et al. (2014) measurement. In fact, factors A_z in Equation (2) would have to be 3 times smaller to match Harikane et al. (2016). In other words, we are unable to populate CROC dark matter halos with stars with *any* $M_*(M_h)$ in such a way as to match simultaneously Harikane et al. (2016) and Duncan et al. (2014).

For comparison, in Figure 4 we show model predictions from abundance matching (Behroozi et al. 2019) with dashed lines. It is interesting to note that for halo masses below $M_h \lesssim 5 \times 10^{11}\mathrm{M}_\odot$, the Behroozi et al. (2019) model predicts that the SHMR *decreases* with from $z = 5$ to $z = 8$, but above this mass range, the redshift trend reverses. Despite the fact that both approaches reproduce the observed galaxy UVLF and SMF in a given mass range (for CROC only for $M_\star < 10^8\mathrm{M}_\odot$), CROC and Behroozi et al. (2019) display opposite evo-

lution trends in the SHMR. This implies that the observational errors in the UVLF and SMF are still large enough to allow *opposite trends* in the SHMR to be consistent with the data.

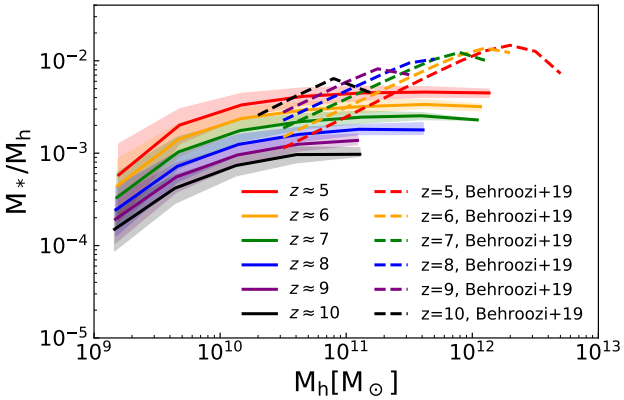

Figure 4. Stellar-to-halos-mass ratio as a function of halo mass for CROC simulations (solid lines with bands) from $z = 5$ to $z = 10$ together with the same relation from Behroozi et al. (2019) obtained from abundance matching. The SHMR flattens at halo masses above $M_h \gtrsim 10^{10}\mathrm{M}_\odot$, reflecting the deficiency in massive galaxies apparent in Figs. 1 and 3. The bands illustrate the 25th and 75th percentile values in each mass bin.

3.5. *Star formation over cosmic time*

The deficiency at high stellar masses can be better understood by looking at how the star formation rate at fixed stellar mass varies with redshift, otherwise known as the evolution of the "main sequence of galaxy formation." The solid lines in Figure 5 show that the star formation rate at a fixed stellar mass *decreases* with redshift, consistent with the trends generally found in observational studies (c.f. Santini et al. 2017). At a given redshift, the slope of the SFR-$M_\star$ relation is sub-linear and is slightly shallower than the slope of the observed relation at $2 \leq z < 3$.

The star formation rate in galaxies has been relatively well measured by several teams over the last few years, such as Salim et al. (2007); Pannella et al. (2015); Catalán-Torrecilla et al. (2015); Zahid et al. (2017) for low redshift galaxies and Whitaker et al. (2012) for galaxies up to $z = 2.5$. At even higher redshifts the measurements are limited by a number of systematic effects, such as those present in gravitationally lensed systems that magnify and enable measurements of high redshift objects. Therefore, the robustness of the relationship between star formation rate and stellar mass at high redshifts is not yet well established. In general our galaxies form stars at rates consistent with, but perhaps somewhat lower than, observations. Note, however, that any difference between CROC and observed galaxy star formation rates is much smaller than the correction (Eq. 2) required to bring CROC stellar mass functions in agreement with observational data.

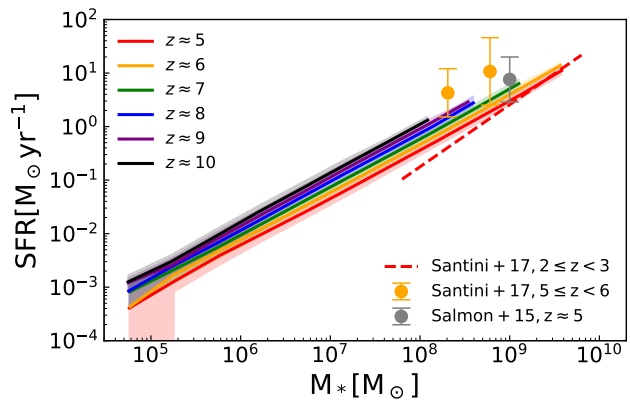

Figure 5. Star formation rate as a function of stellar mass from $z = 5$ to $z = 10$. The bands illustrate the 25th and 75th percentile values in each mass bin. We also show the relation from observations in Santini et al. (2017), where the slope is similar and a similar evolutionary trend exists. The $z = 5$ observed slope matches those in CROC, but the observed normalization is higher. Also, note that the observed slope steepens with time.

In Figure 6 we show the evolution of the global star formation rate density (SFRD). First, we see that the SFRD decreases as redshift increases. Next, as we include fainter galaxies into our calculation of SFRD, the SFRD increases as expected. In order to compare with observations, we integrate the UV luminosity functions above the magnitude cutoff of $M_{1500} \approx -17$. Compared to simulations, the observational results give higher SFRD value at all redshifts. This is another illustration that that CROC is underproducing stars in massive galaxies at all cosmic times.

The corresponding evolution of the global stellar mass density is shown in Figure 7. The global stellar mass density is harder to measure than the global SFR. The observational uncertainties are therefore larger. Still, the underprediction of stellar mass in CROC simulations in galaxies above $10^8\mathrm{M}_\odot$ (thick dashed red line) is very apparent.

3.6. *Galaxy Bias*

The galaxy bias measurement provides an additional constraint on the galaxy-halo connection, measuring the excess clustering of galaxies compared with the clustering of matter. The bias can be calculated in terms of the two-point correlation function, $\langle b_{\mathrm{gal}}^2 \rangle = \xi_{gg}/\xi_{mm}$,

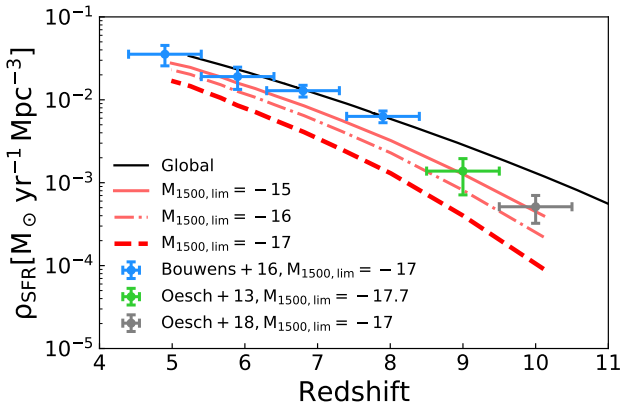

Figure 6. Evolution of star formation rate density (SFRD) in CROC compared with data from Bouwens et al. (2016) in blue errorbars, Oesch et al. (2013) in green errorbar, and Oesch et al. (2018) in grey errorbar. The solid black line is derived from the global SFR, while the red solid, dash-dot, and dashed lines show the results obtained by having different M_{1500} cuts.

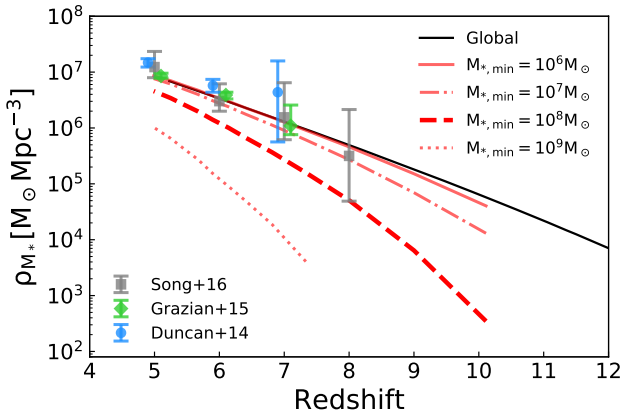

Figure 7. Evolution of stellar mass density in CROC with data from Song et al. (2016) in grey errorbars, Grazian et al. (2015) in green errorbars, and Duncan et al. (2014) in blue errorbars. These observational results are derived by integrating their best-fit stellar mass functions above $M_\star > 10^8 M_\odot$. We show the CROC results obtained by including halos with different lower stellar mass limits, including the $10^8 M_\odot$ limit in dot-dash line.

where ξ is the two-point correlation function of either galaxies (gg) or matter (mm). When compared to observed galaxies, we can constrain whether or not our modeled galaxies have similar spatial clustering as the observed galaxies. Given consistent cosmologies, we can assume that our simulated halo bias matches that in the real universe, $\langle b^2_{\rm halo}\rangle = \xi_{hh}/\xi_{mm}$. We can then use the galaxy bias comparison to determine whether or not CROC galaxies live in the right dark matter halos.

In order to calculate the galaxy bias across mass bins including the bins of the most massive galaxies, whose number counts are small, we can use the empirical form of the bias dependence on the halo mass, $b(M_h)$, and compute the average galaxy bias for a given sample as an integral over the halo mass function, instead of actually computing the ratio of two power spectra. For the explicit calculation, we use the Tinker et al. (2010) halo bias model, $b(M_h)$, from the cosmology package *Colossus* (Diemer 2018). The simulated halo mass function of CROC well-matches the halo mass function from Tinker et al. (2010), the corresponding bias model is therefore the appropriate model to use for CROC halos. As an example, for a halo sample with halo masses between $M_{\rm min}$ and $M_{\rm max}$ the average bias is,

$$\langle b^2\rangle = \frac{\int_{M_{\rm min}}^{M_{\rm max}} b^2(M_h)\frac{dn}{dM_h}dM_h}{\int_{M_{\rm min}}^{M_{\rm max}} \frac{dn}{dM_h}dM_h}. \quad (3)$$

For a galaxy sample containing N galaxies (with an arbitrary sample selection criterion) we can analogously compute the average bias as the quadratic average of bias factors of individual galaxies,

$$\langle b^2\rangle = \frac{1}{N}\sum_{i=1}^{N} b^2(M_{h,i}), \quad (4)$$

where $M_{h,i}$ is the known halo mass for galaxy i from the sample.

Figure 8 shows the so computed average *galaxy* bias from all 6 boxes in the CROC suite. The blue and red squares respectively show the average galaxy bias for bright ($M_{1500} < -19.4$) and faint ($M_{1500} > -19.4$) galaxies in CROC as a function of redshift. The black squares show the average galaxy bias for all galaxies in the simulation as a function of redshift. For comparison, the points with errorbars show observational inferences of the bias from Barone-Nugent et al. (2014) with the same UV magnitude selection. There is good agreement between the simulations and the data, especially at $z \sim 7$, which indicates two things.

Given the assumption that the halo bias in simulations reflects that of the real universe, and the fact that the faint (low mass) end of the CROC luminosity (stellar mass) function matches the observed function, we can first say that the low luminosity sample of CROC galaxies below $M_{1500} = -19.4$ live in the correct halos and well reproduce the observed behavior.

The minimum stellar mass of galaxies brighter that $M_{1500} = -19.4$ corresponds to $M_\star \approx 3\times10^8 M_\odot$. These brighter galaxies above the luminosity cutoff also seem to well reproduce the observed galaxy bias, illustrating that the bias factor from Equation 4 has been modeled correctly. However, this population of galaxies have

UVLFs that are discrepant with the observed UVLF. In other words, the high mass halos have galaxies in the correct luminosity bin brighter than $M_{1500} = -19.4$, but there are not enough stars in these galaxies to match the UVLF at the bright end.

Note, the galaxy bias of the fainter galaxy bin may appear to be on the high side. However, we use the same number of CROC galaxies whose luminosities are below the cutoff as the number of CROC galaxies that lie above the cutoff, selecting the brightest galaxies when rank ordered by luminosity. The resulting faint galaxy sample is comprised of galaxies whose total masses are above $M_{\rm tot} \gtrsim 10^{11}{\rm M}_\odot$. The resulting galaxy bias is therefore larger than if we included lower mass galaxies.

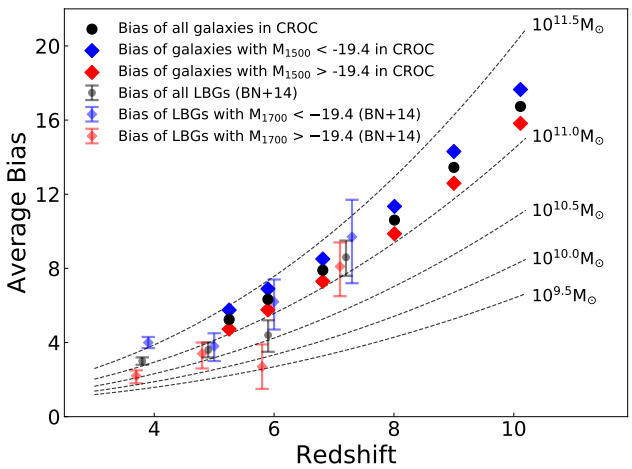

Figure 8. Galaxy bias in the CROC simulations compared with measurements from Barone-Nugent et al. (2014) and dark matter halo bias from Tinker et al. (2010). We illustrate the bias in two magnitude bins, brighter (blue) and fainter (red) than $M_{1500} = -19.4$. As expected, the brighter galaxies exhibit a higher average bias in each redshift snapshot. The CROC measurements are consistent with observations. However, the faint bin of CROC galaxies exhibit a somewhat higher bias than the observed value at z≤6.

4. SUMMARY AND DISCUSSION

In this paper, we examine the galaxy-halo connection of halos in the mass range $\rm M_{halo} \approx 10^9$ to $10^{12}{\rm M}_\odot$ at $5 \leq z \leq 10$ in CROC simulations. CROC simulations include a wide range of physical processes necessary to model cosmic reionization, including gravity, gas dynamics, fully coupled radiative transfer, radiative cooling, star formation, and stellar feedback. We look at the UV luminosity function, the stellar mass function, the stellar-to-halo mass ratio, star formation histories and the galaxy bias in CROC.

Our main findings include:

- CROC galaxies match the UV luminosity function and the stellar mass function at the faint and low mass end respectively.
- Our simulations underproduce stellar masses of the most massive halos. Physical processes that are not modeled in CROC, such as AGN feedback and cosmic ray feedback, can reduce star formation rates in massive halos, further exacerbating the discrepancy. Hence the most likely reason for the discrepancy is that the CROC model for star formation and feedback mis-models star formation in high mass halos. Another, admittedly much less likely, possibility is that the discrepancy between CROC and observations in the high mass end is due to a bias in observational results from, for example, AGN contributions to galactic UV luminosities (Adams et al. 2019).
- The stellar-to-halo mass ratio in CROC decreases with redshift, consistent with some other models, but inconsistent with abundance matching predictions from Behroozi et al. (2019) and some of the observed trends. In fact, we are unable to populate CROC dark matter halos with stars with *any* function form of $M_*(M_h)$ in such a way as to match simultaneously Harikane et al. (2016) and Duncan et al. (2014). Since all models and observational constraints match the observed luminosity functions and use the same cosmology, the lack of consistency even in the sign of the redshift evolution implies that the observed luminosity functions are not accurate enough to constrain the stellar-to-halo-mass ratio at $z > 5$.

These results will help focus efforts on improving the accuracy and physical fidelity of future simulations of cosmic reionization.

The authors are grateful to Mimi Song and Kenneth Duncan, who enabled the comparison of the CROC simulations with observational data. The authors thank Huanqing Chen and Hsiao-Wen Chen for helpful discussions that guided the analysis and interpretation of results, and Andrey Kravtsov and Phil Mansfield for extensive and comprehensive comments that allowed us to significantly improve the original manuscript. HZ acknowledges support from the Kavli Institute for Cosmological Physics and the Jeff Metcalf Fellowship Program. CA acknowledges support from both the Enrico Fermi Institute the Kavli Institute for Cosmological Physics at the University of Chicago, and both the Leinweber Center for Theoretical Physics and the LSA Collegiate Fel-

lowship at the University of Michigan. Fermilab is operated by Fermi Research Alliance, LLC, under Contract No. DE-AC02-07CH11359 with the United States Department of Energy. This work was partly supported by a NASA ATP grant NNX17AK65G, and used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. This research is also part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications.

REFERENCES

Adams, N. J., Bowler, R. A. A., Jarvis, M. J., et al. 2019, arXiv e-prints, arXiv:1912.01626

Agertz, O., Kravtsov, A. V., Leitner, S. N., & Gnedin, N. Y. 2013, ApJ, 770, 25

Agertz, O., Teyssier, R., & Moore, B. 2011, MNRAS, 410, 1391

Atek, H., Richard, J., Kneib, J.-P., et al. 2015, ApJ, 800, 18

Barone-Nugent, R. L., Trenti, M., Wyithe, J. S. B., et al. 2014, ApJ, 793, 17

Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143

Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013, ApJ, 762, 109

Bernard, S. R., Carrasco, D., Trenti, M., et al. 2016, ApJ, 827, 76

Bhowmick, A. K., Campbell, D., Di Matteo, T., & Feng, Y. 2018, MNRAS, 480, 3177

Birrer, S., Lilly, S., Amara, A., Paranjape, A., & Refregier, A. 2014, ApJ, 793, 12

Bouwens, R. J., Oesch, P. A., Illingworth, G. D., Ellis, R. S., & Stefanon, M. 2017, ApJ, 843, 129

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34

Bouwens, R. J., Aravena, M., Decarli, R., et al. 2016, ApJ, 833, 72

Bradley, L. D., Trenti, M., Oesch, P. A., et al. 2012, ApJ, 760, 108

Brook, C. B., Stinson, G. S., Gibson, B. K., et al. 2012, MNRAS, 426, 690

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000

—. 2011, GALAXEV: Evolutionary Stellar Population Synthesis Models, , , ascl:1104.005

Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80

Catalán-Torrecilla, C., Gil de Paz, A., Castillo-Morales, A., et al. 2015, A&A, 584, A87

Cohn, J. D. 2017, MNRAS, 466, 2718

Conroy, C., & Gunn, J. E. 2010a, FSPS: Flexible Stellar Population Synthesis, , , ascl:1010.043

—. 2010b, ApJ, 712, 833

Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486

Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, ApJ, 647, 201

Conroy, C., White, M., & Gunn, J. E. 2010, ApJ, 708, 58

Diemer, B. 2018, ApJS, 239, 35

Duncan, K., Conselice, C. J., Mortlock, A., et al. 2014, MNRAS, 444, 2960

Eisenstein, D. J., & Hut, P. 1998, ApJ, 498, 137

Finkelstein, S. L., Ryan, Jr., R. E., Papovich, C., et al. 2015, ApJ, 810, 71

Firmani, C., & Avila-Reese, V. 2010, ApJ, 723, 755

Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, SSRv, 123, 485

Gnedin, N. Y. 2014, ApJ, 793, 29

Gnedin, N. Y., & Draine, B. T. 2014, ApJ, 795, 37

Governato, F., Brook, C., Mayer, L., et al. 2010, Nature, 463, 203

Grazian, A., Fontana, A., Santini, P., et al. 2015, A&A, 575, A96

Harikane, Y., Ouchi, M., Ono, Y., et al. 2016, ApJ, 821, 123

—. 2018, PASJ, 70, S11

Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 480, 800

Ishigaki, M., Kawamata, R., Ouchi, M., et al. 2018, ApJ, 854, 73

Kaviraj, S., Laigle, C., Kimm, T., et al. 2017, MNRAS, 467, 4739

Kravtsov, A. V. 1999, PhD thesis, NEW MEXICO STATE UNIVERSITY

Kravtsov, A. V., Klypin, A., & Hoffman, Y. 2002, ApJ, 571, 563

Laporte, N., Infante, L., Troncoso Iribarren, P., et al. 2016, ApJ, 820, 98

Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, ApJ, 835, 113

Ma, X., Hopkins, P. F., Garrison-Kimmel, S., et al. 2018, MNRAS, 478, 1694

Mitra, S., Davé, R., Simha, V., & Finlator, K. 2017, MNRAS, 464, 2766

Morishita, T., Trenti, M., Stiavelli, M., et al. 2018, ApJ, 867, 150
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labbé, I., & Stefanon, M. 2018, ApJ, 855, 105
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2013, ApJ, 773, 75
Oesch, P. A., van Dokkum, P. G., Illingworth, G. D., et al. 2015, ApJL, 804, L30
Oesch, P. A., Brammer, G., van Dokkum, P. G., et al. 2016, ApJ, 819, 129
Orr, M. E., Hayward, C. C., Hopkins, P. F., et al. 2018, MNRAS, 478, 3653
Pannella, M., Elbaz, D., Daddi, E., et al. 2015, ApJ, 807, 141
Rudd, D. H., Zentner, A. R., & Kravtsov, A. V. 2008, ApJ, 672, 19
Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267
Santini, P., Fontana, A., Castellano, M., et al. 2017, ApJ, 847, 76
Schmidt, K. B., Treu, T., Trenti, M., et al. 2014, ApJ, 786, 57
Semenov, V. A., Kravtsov, A. V., & Gnedin, N. Y. 2018, ApJ, 861, 4
Sijacki, D., Vogelsberger, M., Genel, S., et al. 2015, MNRAS, 452, 575
Sobacchi, E., & Mesinger, A. 2015, MNRAS, 453, 1843
Song, M., Finkelstein, S. L., Ashby, M. L. N., et al. 2016, ApJ, 825, 5
Stefanon, M., Bouwens, R. J., Labbé, I., et al. 2017, ApJ, 843, 36
Stinson, G. S., Brook, C., Macciò, A. V., et al. 2013, MNRAS, 428, 129
Stinson, G. S., Dalcanton, J. J., Quinn, T., et al. 2009, MNRAS, 395, 1455
Sun, G., & Furlanetto, S. R. 2016, MNRAS, 460, 417
Tacchella, S., Trenti, M., & Carollo, C. M. 2013, ApJL, 768, L37
Tinker, J. L., Leauthaud, A., Bundy, K., et al. 2013, ApJ, 778, 93
Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878
Turk, M. J., Smith, B. D., Oishi, J. S., et al. 2011, The Astrophysical Journal Supplement Series, 192, 9
Wang, L., Farrah, D., Oliver, S. J., et al. 2013, MNRAS, 431, 648
Wechsler, R. H., & Tinker, J. L. 2018, ARA&A, 56, 435
Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, ApJL, 754, L29
White, M., Zheng, Z., Brown, M. J. I., Dey, A., & Jannuzi, B. T. 2007, ApJL, 655, L69
Wilkins, S. M., Feng, Y., Di Matteo, T., et al. 2017, MNRAS, 469, 2517
Wyithe, J. S. B., Yan, H., Windhorst, R. A., & Mao, S. 2011, Nature, 469, 181
Zahid, H. J., Kudritzki, R.-P., Conroy, C., Andrews, B., & Ho, I. T. 2017, ApJ, 847, 18
Zheng, Z., Coil, A. L., & Zehavi, I. 2007, ApJ, 667, 760
Zitrin, A., Zheng, W., Broadhurst, T., et al. 2014, ApJL, 793, L12
Zitrin, A., Labbé, I., Belli, S., et al. 2015, ApJL, 810, L12