

Effect of molecular hydrogen self-shielding modeling on galaxy formation in cosmological simulation

THỊNH HỮU NGUYỄN^{1,2} AND KIRK S. S. BARROW¹

¹Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

²Center for AstroPhysical Surveys, National Center for Supercomputing Applications
Urbana, IL 61801, USA

ABSTRACT

Keywords: Astronomical simulations(1857) — Radiative transfer simulations(1967) — Galaxy formation(595) — Reionization(1383)

1. INTRODUCTION

Molecular hydrogen is a building block for star formation. For a gas cloud to collapse into a star, the cloud's self-gravity must be sufficiently large to overcome its pressure gradient. In the absence of other environmental effects, to achieve the collapse, the gas cloud usually needs to have sufficient mass or a high enough cooling rate so that the cooling time is smaller than the free-fall time. Since the cooling time generally decreases when the gas density increases during the collapse, this leads to catastrophic cooling and the gas pressure being lost against gravity, resulting in a collapse. Depending on the temperature and density, there are multiple radiative cooling mechanisms for a cloud to release its internal energy. For primordial gas clouds with atomic hydrogen HI being the dominant component, Bremsstrahlung emission, atomic hydrogen's recombination, collisional ionization, and collisional excitation processes can help cool the gas down to at best the virial temperature of $10^4 K$, below which the cooling rate is too low (Thoul & Weinberg 1995). **Elaborate in what condition (temperature, density) molecular hydrogen forms from atomic hydrogen.** However, if molecular hydrogen H_2 is available, it can further stimulate cooling by opening up new cooling channels through its rotational and vibrational energy levels. This allows the gas to cool to ≈ 100 Kelvin degrees (Galli & Palla 1998), making it possible for the cloud to collapse into stars. This connection between star formation and H_2 can be reflected through the Kennicutt-Schmidt law (Schmidt 1959; Kennicutt 1998) and various observational evidence in nearby disk galaxies (Wong & Blitz 2002; Kennicutt et al. 2007; Leroy et al. 2013).

Ultraviolet (UV) radiation emitted by nearby active stars can photodissociate molecular hydrogen via the Solomon process (Solomon 1965 - private communica-

tion reported in Field et al. 1966, Stecher & Williams 1967). When absorbing a Lyman-Werner (LW) photon (11.2-13.6 eV), a molecular hydrogen molecule is excited from the ground state to an excited electronic state. Instead of radiatively decaying back to the bound state, about 15 percent of the molecules have their electrons decay into the vibrational continuum, which subsequently dissociates the molecule into atomic hydrogen. Dissociation rates through this mechanism can be inhibited through H_2 self-shielding, a phenomenon where the column density is large enough that the H_2 in the cloud's outer layer absorbs the LW radiation and hence shields the inner region. Previous studies point out the importance of including self-shielding in preserving the amount of molecular hydrogen required for star formation. When simulating a disc galaxy, (Christensen et al. 2012) notices an increase in the amount of cold gas and a clumpier interstellar medium when incorporating self-shielding in their simulation. Safranek-Shrader et al. (2017) finds that self-shielding is crucial in the development of H_2 in the disc's mid-plane, which accounts for 10-15 percent of the total gas mass. Therefore, the hydrogen self-shielding mechanism becomes pivotal in preserving the amount of molecular hydrogen and regulating star formation.

Through modeling a semi-infinite, static slab of gas irradiated on one surface, Draine & Bertoldi (1996) proposes an analytical expression to model H_2 self-shielding factor that includes the effect of line overlap (do we need to elaborate this? also do we need to show the equation here). Maybe show all the papers that use this expression? Look at Hartwig+2015 for the list As shown in Equation (37) of Draine & Bertoldi (1996), the self-shielding factor is a function of the H_2 column density. In the past, computing the column density for multiple 3D directions was

prohibitively computationally expensive (Shang et al. 2010; Wolcott-Green et al. 2011). Therefore, multiple methods have been suggested to approximate the self-shielding effect while trying to alleviate its demanding computational cost. A method is to define a local characteristic length scale across which the H_2 number density is constant. This characteristic length can be either the Jeans length (Shang et al. 2010; Johnson et al. 2011), the Sobolev length (Sobolev 1957; Yoshida et al. 2006), or the Sobolev-like length that is computed from the density-gradient (Gnedin et al. 2009; Gnedin & Kravtsov 2011). Rather than a single characteristic length, Hartwig et al. (2015a) approximates the 3D column densities by creating spherical maps of the column density around each Voronoi cell with 48 equal-area pixels and also taking into account the Doppler effect from the relative velocities of the infalling gas particles. Another non-local method is the six-ray approximation (Yoshida et al. 2003, 2007; Glover & Mac Low 2007a,b), where the H_2 column density is integrated along six directions along three Cartesian axes centered at each particle's position.

However, there are limitations to the applicability of approximate treatments in various test problems when compared to the accurate full ray-tracing calculation. In the context of photodissociation of H_2 in protogalaxies and direct-collapse black hole formation, Wolcott-Green et al. (2011) show that the Sobolev length, the density gradient, and the six-ray approximation methods overestimate the H_2 shielding degree by an order of magnitude in low-density regions of $n < 10^4 \text{ cm}^{-3}$. Also in the context of direct-collapse black hole, when comparing with their non-local approximation detailed in Hartwig et al. (2015a), Hartwig et al. (2015b) find that the Jeans approach returns a critical flux value J_{crit} - the lowest flux required for a halo with virial temperature above 10^4 K to collapse to a supermassive black hole seed - two times larger than the non-local approximation, leading to a disparity in the predicted number density of black hole seeds. Greif (2014) investigate the collapse of primordial star-forming clouds with an accurate H_2 line emission model and notice that the Sobolev method brings about a thermal instability for the collapsing cloud and an order-of-magnitude overestimation of the escape fraction for high optical depth. When evaluating the triggered Population III star formation at the limb of the HII region, Chiaki & Wise (2023) find that while the density-gradient method matches well with the ray-tracing method, the Jeans length approximation results in an overestimation of the number of the Population III stars at the front of the shock wave. The authors suggest that this is because the thickness of the

H_2 ring of the shock front is typically smaller than the Jeans length, and thus the local approximation is inappropriate. In their large-scale galactic discs simulations, Safranek-Shrader et al. (2017) notices that the Sobolev method underpredicts the H_2 abundance in the disc by a factor of 5, while the six-ray and the Jeans length methods perform better, which is contrary to the findings that Jeans length method is ineffective in previous studies. Thus, the validity of these approximation methods is highly problem-dependent and they cannot be used as a general substitution for the full ray-tracing calculation.

A simulation that uses approximated H2 self-shielding: COLDSDIM, Maio et al 2021

In this paper, we expand the investigation by comparing these approximation methods with the detailed ray-tracing methods, this time in the context of star formation during the Reionization period. This will help further examine and inform the community about the applicability of these approximations in modeling H_2 in different contexts. In Section ..., we will describe our simulation and the implemented H_2 self-shielding models. Section ... analyzes the effect of these models on star formation and galaxy evolution during Reionization. Last but not least, in Section ..., we discuss the implications of the findings and summarize the paper.

Maybe a paragraph at the beginning about the observational and theoretical evidence of the relation between H_2 and star formation (Kenicutt-Schmidt relation, for example)? I think this is a good idea. I'm happy to help with this part if you need some background.

2. METHOD

2.1. Cosmological simulations

We run and analyze outputs from a radiation-hydrodynamic adaptive mesh refinement simulation ENZO (Bryan et al. 2014). Two sets of cosmological simulations are generated to explore the self-shielding effect in the early stage and in the late stage of the Reionization period, which we will refer to as *EarlyRe* and *LateRe*, respectively. In the early stage of Reionization, the UV background radiation is still local. Thus, the *EarlyRe* set provides a control sample to study whether the self-shielding approximation affects the overall star formation activity during a low non-local UV-radiation scenario. On the other hand, the *LateRe* allows us to study how the resistance of molecular clouds against a strong photoionization background depends on the self-shielding treatments. Both sets assume a flat Λ CDM cosmology and are run with

the cosmological parameters obtained from Planck Collaboration et al. (2016): $\Omega_M = 0.3065$, $\Omega_\Lambda = 0.6935$, $\Omega_b = 0.0483$, $h = 0.679$, $\sigma_8 = 0.8154$, and $n = 0.9681$. The total moving volume of the EarlyRe and LateRe simulations are 13 (Mpc/h)^3 and 5 (Mpc/h)^3 , respectively. Both sets have the root grid dimension of 256^3 . The total number of dark matter particles is $\approx 10^{7.82}$ particles for the EarlyRe set and $\approx 10^8$ particles for the LateRe set. Inside the total volume of the EarlyRe (LateRe) simulation, we create a smaller zoom-in region using three (check the MUSIC file of LateRe) additional levels of nested refinement to create an effective grid size resolution of $2632.5 \times 2512.9 \times 2871.9 \text{ kpc}^3$ ($920.5 \times 1564.8 \times 1196.6 \text{ comoving kpc}^3$) at the last snapshot at redshift of 10.9 (6.4). In the zoom-in region, the most refined dark matter particle mass is 2.7×10^4 (1.5×10^3) M_\odot . As the density passes a certain threshold, the grids inside this zoom-in region can further be refined adaptively up to 2^{13} (2^{14}) times the root grid dimensions. This allows a maximum spatial resolution down to 3.07 pc (1.89 pc) at the last time step, which allows the tracking of large molecular clouds through initial collapse. We use radiative star cluster particle (Wise & Cen 2009) and Population III star particle (Abel et al. 2007) for our star formation prescription.

Each set contains a version running with detailed ray-tracing calculation (EarlyRe-ray and LateRe-ray) and a version running with a local approximation of the column density (EarlyRe-apx and LateRe-apx) when computing the H_2 self-shielding factor. The self-shielding model is set by the RadiativeTransferOpticallyThinH2 ENZO radiation parameter, where a value of 1 enables an optically thin H_2 dissociating Lyman-Werner radiation field and a value of 0 uses ray-tracing for the calculation. For the EarlyRe simulation set, the EarlyRe-ray and the EarlyRe-apx are initialized with an identical list of parameters and initial conditions except for the RadiativeTransferOpticallyThinH2 parameter. We evolve the two simulations from $z = 100$ to $z \approx 10.9$, each output is 1 Myr apart. For the LateRe set, to save computational time, we run the two simulations without ray-tracing until $z \approx 7.3$, then we switch the RadiativeTransferOpticallyThinH2 parameter to 0 for one of the simulations (the LateRe-ray) and run both of them until $z \approx 6.4$. This allows us to examine the effect of self-shielding near the end of the reionization epoch.

2.2. Halo tracking and merger tree algorithms

To identify the halos in the simulations, we employ SHINBAD (Barrow et al 2025, in preparation), an all-in-one algorithm that both finds halos and builds merger trees based on robust energy definition and particle

tracking of dark matter particles. (Note that this is version 1088 of SHINBAD, ask Kirk how to expand this part). We exclude all the halos that are outside the refined region of the simulation box - the region containing only the highest and the second-most highest resolution dark matter particles - to ensure that the halos are physical.

2.3. Numerical H_2 self-shielding models

A H_2 self-shielding process happens when a molecular cloud has a sufficiently high column density that its outer layer absorbs all the Lyman-Werner UV photons, effectively shielding the molecules within it against photodissociation. Via modeling a semi-infinite, static slab of gas that is irradiated on one surface, Draine & Bertoldi (1996) express a self-shielding-accounted photodissociation rate as

$$k_{\text{diss}}(N_{H_2}, T) = f_{sh}(N_{H_2}, T) \cdot k_{\text{diss}}(N_{H_2} = 0, T), \quad (1)$$

where $f_{sh}(N_{H_2}, T)$ is a self-shielding factor or shielding function, and $k_{\text{diss}}(N_{H_2} = 0, T)$ is the dissociation rate in a optically-thin regime. At a low temperature (hundreds of K), the self-shielding factor can be approximately computed from only the H_2 column density,

$$f_{sh} = \min \left[1, \left(\frac{N_{H_2}}{10^{14} \text{ cm}^{-2}} \right)^{-3/4} \right]. \quad (2)$$

This is controlled by setting the ENZO radiation parameter RadiativeTransferH2ShieldType = 0, which is also its default value. We use this expression to calculate the self-shielding factor in the EarlyRe-apx simulation. It is important to note that the convention of f_{sh} value is slightly counter-intuitive: the f_{sh} value of 0 means full self-shielding is available, and the f_{sh} value of 1 means there is no self-shielding. We also acknowledge that Draine & Bertoldi (1996) lists a more accurate formula to calculate f_{sh} that takes temperature as a parameter. However, using this formula does require additional computational resources that we save to run the ray-tracing scheme. Thus, we choose to use the approximated expression at low temperatures. Check Gnedin and Draine 2014, where they do a comparison between this equation with a more accurate one (including line overlap). Summarize their results here.

Maybe talking about the limitations of Draine and Bertoldi 1996, as on page 5, right column of Wolcott-Green.

To avoid the high computational expense of finding the H_2 column density in multiple 3D directions, previous studies employ different techniques to approximate

it (see Section 1). In this paper, we compare the results from two methods: a direct integration from ray tracing and a Sobolev-like length approximation (Gnedin et al. 2009; Wolcott-Green et al. 2011). Even though demanding more computational time, the ray tracing treatment ensures a more accurate radiative transfer calculation. ENZO uses the adaptive ray tracing technique (Wise & Abel 2011) that is based on the HEALPix (Hierarchical Equal Area isoLatitude Pixelation, Gorski et al. 2005) scheme. The column density is integrated along a HEALPix rays,

$$N_{H_2} = \int n(H_2) ds. \quad (3)$$

The Sobolev-like treatment is a local approximation in which the number density n_{H_2} is assumed to be constant within a characteristic length scale L_{char} . In the Sobolev-like approximation, the characteristic length scale is defined as a distance over which the cell's gas density ρ diminishes. The gas beyond this distance is sparse enough that its influence on shielding is negligible, and thus molecular hydrogen is dissociated in an optically thin fashion. The column density is then computed by

$$N_{H_2} \approx n_{H_2} L_{\text{char}} = n_{H_2} \frac{\rho}{|\nabla \rho|} \quad (4)$$

The approximation is enabled by setting the ENZO parameters `RadiativeTransferOpticallyThinH2 = 1` and `RadiationShield = 2`.

Figure 1 compares the self-shielding factor f_{sh} between the ray-tracing model and the Sobolev-like approximation model. At the time step corresponding to $z = 10.9$, we select 31 halos that contain stars or overlap with another halo that contains stars. For each halo, we select two gas cells to evaluate the self-shielding factor: (1) the gas cell with the highest gas density, and (2) the gas cell with the highest gas density gradient. Each scatter point in Figure 1 represents a selected gas cell. In the approximation model, the self-shielding factor of a gas cell is calculated using Equations 2 and 3. For the ray-tracing model, since the simulation output does not store the photodissociation rate calculation, we need to re-compute the self-shielding factor using Equation 2 with an attenuated LW flux-weighted average column density. The LW flux from each star to a selected gas cell is calculated as follows. Firstly, the intrinsic spectrum of a star particle is generated using the Flexible Stellar Population Synthesis (fsps) with the age and metallicity of the star particle as inputs. We assume an initial mass function from Davé (2008) evaluated at $z = 0$, and we use the MIST isochrone library and MILES spectral library to generate the spectra. **may cite Susie's paper**

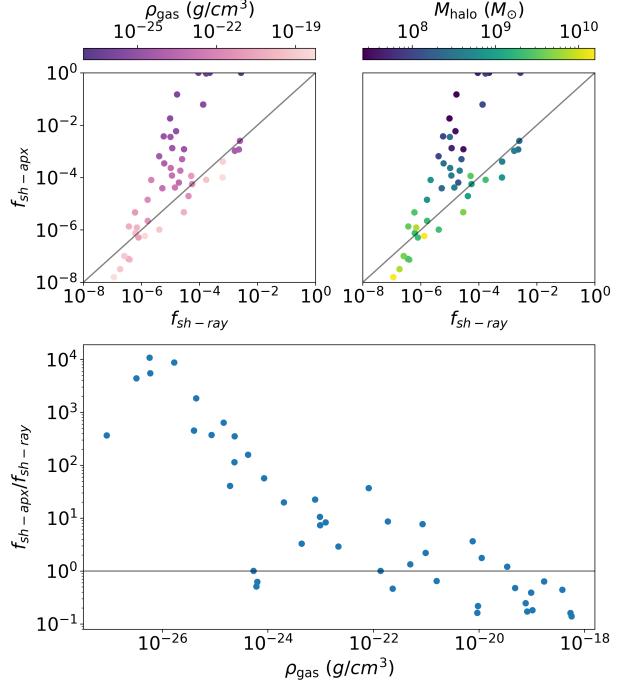


Figure 1. The comparison between the self-shielding factor between the *EarlyRe-ray* and the *EarlyRe-apx* simulation, evaluated at $z = 10.9$. Two gas cells with the highest gas density and highest gas density gradients are chosen in 31 halos in different mass range to evaluate the self-shielding factors. (Top) The relationship between the self-shielding factors calculated by the Sobolev-like approximation model and the ray-tracing model, colored by the cell's gas density (top left) and the host halo's mass (top right). (Bottom) The ratio between the approximated and the ray-traced self-shielding factors is plotted as a function of gas density. The Sobolev-like approximation tends to overestimate the amount of H_2 self-shielding in the high gas density regime and underestimate it in the low gas density regime.

here or ask her for more information. Then, we calculate the attenuation by including both scattering and absorption due to H/He and metals **Ask Kirk to include how the absorption and scatterings are computed.** With the scattering and absorption coefficients, we can compute the total optical depth and then the attenuation of the stellar spectrum from the source to the target gas cell. Lastly, we calculate the attenuated flux of each source to the gas cell, compute the LW flux-weighted average H_2 column density, and get the self-shielding factor from the ray-tracing model. It is important to note that we only include stars within the halo's virial radius to estimate the flux-weighted average H_2 density because stars outside the radius are too far from the selected gas cell to contribute considerably to the total flux at the selected gas cell.

Figure 1 demonstrates a clear disparity in the f_{sh} value between EarlyRe-apx and EarlyRe-ray. In particular, the approximation model tends to overestimate the amount of shielding for cells with a gas density larger than $10^{-20} g/cm^3$ by up to 1 dex. On the other hand, diffuse gas is subjected to underestimated H_2 self-shielding when the Sobolev-like approximation model is used. For highly diffuse gas, this underestimation can be up to 3–4 dex. This connection between gas density and f_{sh} implies a mass-dependent effect of the self-shielding approximation model on halos. The top right subfigure of Figure 1 shows that larger halos ($M_{\text{halo}} > 10^{9.5} M_{\odot}$) have their molecular cloud region over-shielded in the approximation model because the gas is denser. On the other hand, the molecular clouds in small halos ($M_{\text{halo}} < 10^{8.5} M_{\odot}$) are under-shielded against LW radiation when using the Sobolev-like approximation. In some cases, the approximation returns no shielding effect even though those gas cells are still weakly shielded ($f_{sh} \approx 10^{-3}$ when ray-tracing is utilized). Only molecular gas clouds in halos whose mass is about $10^9 M_{\odot}$ are well modeled in the approximation model. The significance of this effect on star formation will be explored in Section 3.1.

Maybe a comparison in terms of the time run between the two methods

2.4. Cross-matching halos

To examine how each individual halo is affected by the implementation scheme of H_2 self-shielding, we perform cross-matching between halos in the simulations with ray-tracing (*EarlyRe-ray* and *LateRe-ray* sets) and their counterparts in the simulations with the Sobolev-like approximation (*EarlyRe-apx* and *LateRe-apx* sets). The matching requirements are

$$\frac{2}{3} < \frac{M_{vir_{ray}}}{M_{vir_{apx}}} < \frac{3}{2}, \quad \frac{d_{COM}}{R_{vir_{ray}}} < 0.25, \quad \frac{d_{COM}}{R_{vir_{apx}}} < 0.25 \quad (5)$$

The virial mass and radius in Equation 5 come from ROCKSTAR and are calculated using only dark matter. We cross-match all the halos at the last timestep of the simulations and then trace back their merger tree using the consistent-tree outputs (details in Subsection 2.2).

2.5. Star assignment

We uniquely assign each star particle in the simulation to one dark matter halo in a two-step fashion. For the first step, a star particle is assigned to a halo where it is

first created. If a star particle is created inside the intersection of multiple halos, we will calculate the orbital energy of that star particle with respect to each halo's center to determine which halo the particle belongs to. The assigned halo is the one with lower relative total orbital energy. We assume that a star particle never leaves its assigned halo unless that halo merges with another one, in which case the star particle will become a member of the descendant halo of the merger. This assumption allows a quick assignment of stars to halos without the need to calculate the orbital energy of each star to multiple halos, which is computationally expensive. The assignment's first step starts from the first snapshot to the last snapshot of the simulation.

Even though stars do not generally leave a halo's potential well for isolated halos, the behavior becomes more complex during halo-halo interaction as stripping can occur and stars can be lost from one halo to another (Kannan et al. 2015). Therefore, the second step of our process is to address this shortcoming of the first step's assumption and to refine the star assignment result. In this step, we validate the output from the first step and check whether a star remains in its assigned halo's virial radius throughout its lifetime. If a star escapes the virial radius of its originally assigned halo, we assign that star to a new halo under two conditions: (1) the star's position must be within the new halo's virial radius and (2) the star's total orbital energy with respect to the new halo must be negative. Similar to the first step, if a star is located inside multiple potential new halos, we assign that star to the halo with the lowest negative orbital energy. If either Condition (1) or (2) fails at a certain timestep, this suggests that the star is stripped out of the potential well of all halos and thus we remove that star from the halo assignment process of that time step.

Once each star particle is uniquely associated with one halo, we calculate a halo's stellar mass and SFR exclusively based on its member star particles. This gravitational unbinding of star particles guarantees that each halo's stellar mass profile are not overlap and independent of each other.

2.6. Defining ISM, CGM, and IGM

The boundary of the interstellar medium (ISM), the circumgalactic medium (CGM), and the intergalactic medium (IGM) region in our analysis is defined as follows. The ISM region (or the boundary of a galaxy) starts from the galaxy's center of mass to $R_{\text{bary},2000}$, which is the radius enclosing a baryonic (gas and stars) mass density 2000 times the universe's critical mass density at a certain redshift. Even though there are multiple other definition of the boundary of galaxy in literature,

such as using a fixed fraction of the virial radius, a fixed physical radius, or a radius that is scaled with mass or surface brightness (see Stevens et al. 2014 for a summary of these techniques), our over-density approach helps define the galaxy more robustly and more physically because it is redshift and density dependent. It also helps better define the galaxy during a close galaxy interaction when a half-mass radius or half-light radius approach may fail as another galaxy is in the target galaxy’s virial radius. Depending on the redshift and the galaxy’s compactness, $R_{\text{bary},2000}$ typically ranges from 0.15 to 0.25 times the halo’s virial radius, which falls in the same range as other galaxy’s aperture definitions in cosmological simulations (Stevens et al. 2014). The CGM starts from $R_{\text{bary},2000}$ to $R_{\text{halo,vir}}$, and the IGM is the region outside of the virial radius of all halos.

Also, it is important to note that the baryonic center of mass does not always coincide with the halo’s center of mass. Thus, in some cases, we shift the virial region to be centered on the gas and stellar mass instead of the dark matter mass; however, this shift is not large and does not affect our interpretation.

3. RESULTS

3.1. Effect on the molecular hydrogen content and star formation

Because the column density approximation model directly concerns the calculation of the molecular hydrogen density, we first investigate the amount of H_2 in the ray tracing and approximation treatment. Figure 2 shows the H_2 fraction (f_{H_2}), which is the ratio between the H_2 mass with the total gas mass, as a function of the total halo mass in three snapshots, one at the time where star particles first appear in the simulations ($z \approx 17.49$), one at the time step in the middle of the simulation run ($z \approx 13.32$), and one being the last time step of the run ($z \approx 10.9$). Each halo in the snapshot is represented by a scatter point, the line represents the running average of each total mass bin, and the shaded region around the line is the standard deviation of the scatter points within each mass bin. The contour plots represent the distribution of the halos in the $M_{\text{gas}} - M_{\text{total}}$ space. According to Figure 2, both EarlyRe-ray and Early-apx simulations have similar gas mass distribution across all timesteps, showing f_{H_2} represents the amount of H_2 in the halos. However, f_{H_2} shows a notable difference between the two radiative transfer treatments, especially at the later time step and in the lower-halo-mass regime. For halo mass smaller than $10^8 M_\odot$, the f_{H_2} in the ray-tracing treatment is larger than that of the Sobolev-like approximation treatment by 3 to 6 dex. This difference grows larger as the simulation evolves,

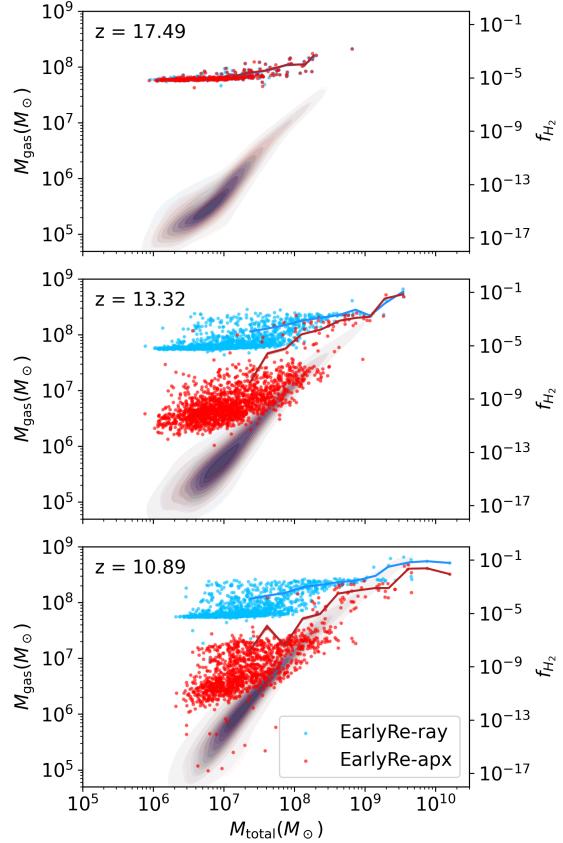


Figure 2. The relationship between the H_2 fraction and the total mass in the timestep where stars first appear, in the intermediate, and in the last timestep of the *EarlyRe* dataset. The contour plots show the M_{gas} distribution and the scatter points show the H_2 fraction of all halos in the simulation box.

reaching up to 12 dex at the last time step for small halos. The EarlyRe-ray simulation also develops a lower limit for f_{H_2} of around 10^{-5} , whereas the f_{H_2} is allowed to reach much further value in the EarlyRe-apx simulation. This shows that the ray-tracing treatment models the self-shielding effect much better and thus helps preserve more molecular hydrogen in the halos. As pointed out in Figure 1, the approximation treatment considerably underestimates H_2 self-shielding in a low gas density regime and molecular hydrogen to get photodissociated more easily, inferring that smaller halos are the most susceptible to the choice of the model. At a higher mass range, the difference in f_{H_2} between the two models gradually diminishes. This is reasonable because larger halos have deeper gravitational potential well and have more dense cool gas regions where H_2 exists. In these regions, the approximation model matches well with the ray-tracing model (as shown in Figure 1), leading to an agreement of f_{H_2} between the two runs.

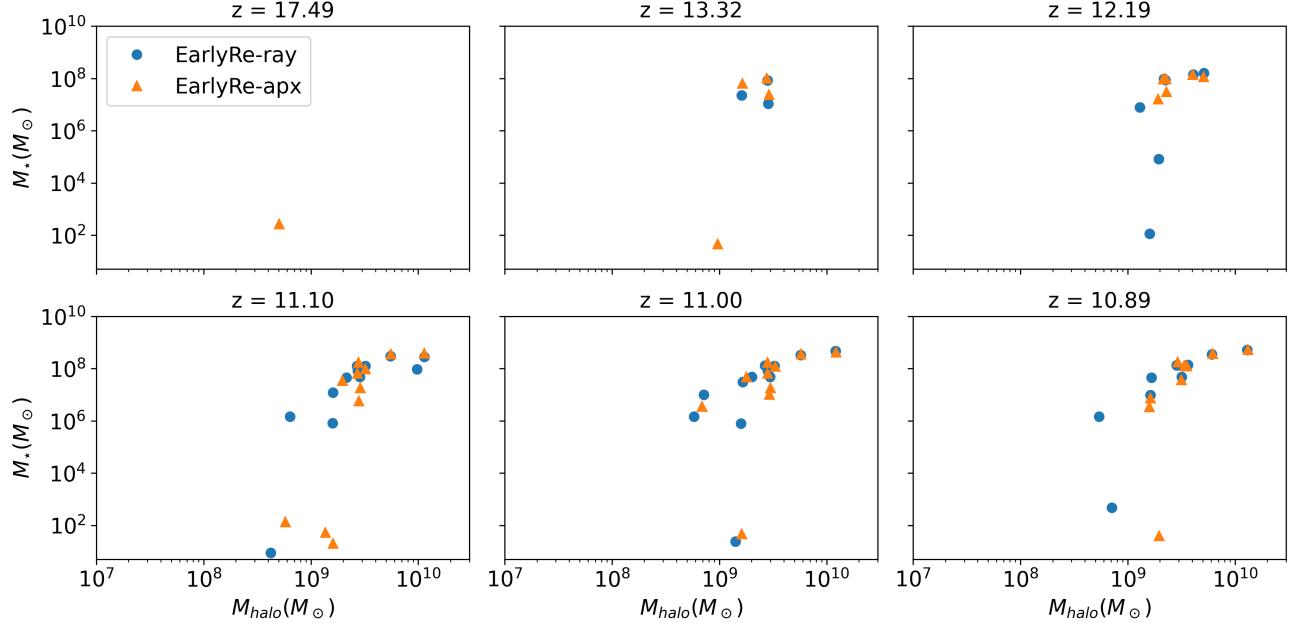


Figure 3. The relationships between stellar mass and total mass of all the halos in the timestep where stars first appear, in the intermediate timesteps, and in the last timestep of the *EarlyRe* dataset.

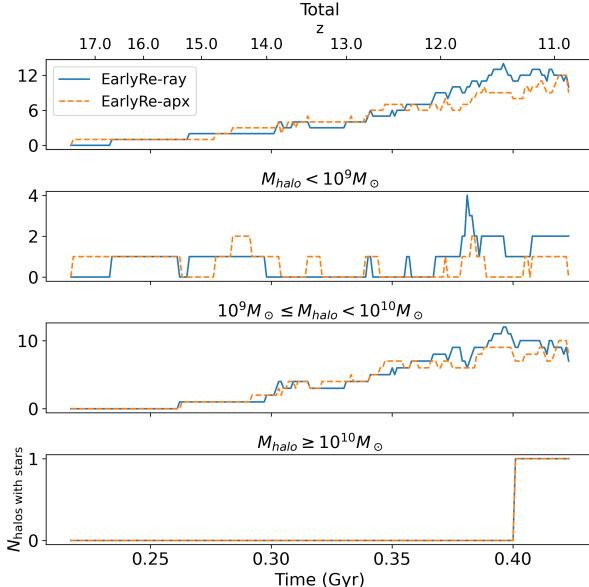


Figure 4. The comparison between the number of halos with stars as a function of time for each halo mass bin between the *EarlyRe-ray* and the *EarlyRe-apx* simulations.

We proceed to investigate whether the self-shielding treatment affects the stellar mass of galaxies in the simulation. Figure 3 shows the relationship between stellar mass and halo mass between the *Early-Ray* and *Early-apx* in the time step where stars first appear (top left panel), in the last time step of the dataset (bottom right panel), and in the intermediate time steps. Figure 4 displays the number of halos that host stars as a function

of time for different halo mass bins. The stellar component of the halos is unbound by using the star assignment procedure described in Section 2.5. Even though the number of halos with stars is even between *EarlyRe-apx* and *EarlyRe-ray* in the beginning, there are more halos with stars in *EarlyRe-ray* at the end of the simulation, particularly halos with $M_{halo} < 10^9 M_\odot$. For ease of reference and lack of a better term, we will use the term "galaxy" to refer to a halo with stars, as opposed with halos with only gas and dark matter particles. From 375 million years after the Big Bang to near the end of the simulation ($z \approx 11.88 - 11$), there are two to four galaxies more in the simulation with the ray-tracing treatment. Furthermore, at a similar halo mass, smaller halos in *EarlyRe-ray* tend to have more stars than their counterparts in *EarlyRe-apx* by several dex in solar mass. This indicates that the use of Sobolev-like approximation for H_2 self-shielding treatment inhibits star formation activity in small halos and dwarf galaxies. On the other hand, halos whose dark matter mass is larger than $10^9 M_\odot$ show good agreement in stellar mass, suggesting that the stellar components of large halos are unaffected by our choice of H_2 self-shielding model. This observation aligns with Figure 2, where we see that H_2 is more accurately shielded in higher halo mass in *EarlyRe-apx*.

Another difference between *EarlyRe-ray* and *EarlyRe-apx* is the time when star particles first appear in the simulation. According to Figure 4, the formation of the first galaxy in *EarlyRe-apx* occurs 16 million years be-

fore that in *EarlyRe-ray*. The first galaxies in the two simulations both form in a halo with a dark matter mass smaller than $10^9 M_\odot$. This is rather intuitive, given that previous evidence shows that for small halos, stars form easier in the simulation using a ray-tracing model. However, further investigation suggests that the timing discrepancy in the formation of first stars is likely due to a stochastic nature of our star formation prescription rather than a physical reason. Before the first stars appear in a halo in *EarlyRe-apx*, the temperature and density distribution of gas and molecular hydrogen of the counterpart halo in *EarlyRe-ray* are similar to the halo in *EarlyRe-apx*. More discussion and plots regarding the first galaxies in the two simulation will be directed to the Appendix. **Reminder to include the appendix**

567 3.2. Case study

568 In this section, we analyze the properties of two halos, one being the largest halo in the simulation and one having a halo mass of $\approx 1.7 \times 10^9 M_\odot$, which is in the mass range where the approximation in H_2 self-shielding fails to accurately model star formation, as suggested in Figure 3. We match the halos between *EarlyRe-ray* and *EarlyRe-apx* using the procedure in Section 2.4. These two pairs of halos serve as examples to how the H_2 self-shielding modeling affects the galaxy properties.

577 Figures 5 and 6 show the projection of various properties at the last time step and the time evolution of those properties for the two halos in our case study. As hinted earlier by Figure 2, gas mass stays consistent between *EarlyRe-apx* and *EarlyRe-ray* throughout the halo's evolution and thus is unaffected by the choice of H_2 self-shielding modeling. The gas density projection plot also looks relatively similar between our two runs. However, even though there is an agreement on total gas content, the amount of H_2 is considerably different in two simulations. For the largest halo, the mass of H_2 starts to diverge around $z \approx 14$, and *EarlyRe-ray* has about 10 times more H_2 in that halo than *EarlyRe-apx*. For the smaller halo, the discrepancy in the H_2 content is much more severe. Early in the simulation, the ray-tracing model consistently returns about 0.5 dex more mass than the approximation model. At $z \approx 11.5$, while *EarlyRe-ray* witnesses an quick rise in molecular gas, *EarlyRe-apx* maintains the same level of H_2 , intensifying the difference to be up to 2 dex in M_{H_2} between the two simulations at the last snapshot. This difference in H_2 matches our observation in Figure 2. In addition to H_2 mass, the distribution of molecular hydrogen is also different. In both halos that we use as examples, molecular hydrogen exists in clumps of high density ($\approx 10^{-21} g/cm^3$) with respect to the surrounding. These

603 clumps are located inside the galaxy (indicated by the 604 white circle in the figures) and are concentrated in the 605 galactic center. Outside of these clumps, H_2 gas den- 606 sity is quite low, reaching down to $10^{-36} g/cm^3$, proving 607 that H_2 is completely photo-dissociated. In contrast, 608 the halos in the *EarlyRe-ray* set displays a much more 609 uniform distribution of molecular hydrogen. While H_2 610 density is still highest at the halo's center, H_2 density 611 in the surrounding region is only 2-3 dex lower, show- 612 ing that H_2 is able to exist and form in the outer part 613 of the galaxy and even in the CGM. Indeed, the exis- 614 tence of H_2 outside the galactic center is what causes 615 the discrepancy in the H_2 fraction between *EarlyRe-apx* 616 and *EarlyRe-ray*. More details on the topic of ISM and 617 CGM will be discussed in Section 3.3.

618 With regards to stellar mass, the different in self- 619 shielding modeling results in greater differences in 620 smaller halos than in bigger halos, as pointed out in Fig- 621 ure 3. For the case of the biggest halo in the simulation, 622 Figure 5 shows that the stellar mass of the halo converge 623 after $z \approx 13$, even though *EarlyRe-ray* forms its first 624 stellar particles a few million years earlier. On the other 625 hand, for the smaller halo in Figure 6, the halo modeled 626 with ray-tracing forms stars approximately 20 million 627 years prior to the *EarlyRe-apx*'s counterpart. Further- 628 more, throughout most of this halo's evolution, the stel- 629 lar mass in *EarlyRe-ray* stays about several dex higher 630 in the *EarlyRe-apx* version. Thus, this proves the choice 631 of H_2 self-shielding model affect stellar mass and the 632 timing of star formation in halos.

633 In addition to the difference in stellar mass, the star 634 formation history shows that the halo in the ray-tracing 635 model experience more prolonged period of star forma- 636 tion while the halo in the approximation model experi- 637 ence shorter but stronger star formation episode (**more 638 detailed explanation maybe?**). This behavior also 639 happens in other large halos in the simulation box (for 640 the brevity purpose, they are not shown in the paper). 641 The prolonged star formation period leads to prolonged 642 supernova rate, thus affecting the metal mixing and 643 metal distribution of the galaxy. As supernovas go off 644 in longer time period, more metals are produced at once 645 and thus it takes a longer time for the metals to mixed 646 with the surrounding gas. This leads to what we see 647 in the metallicity projection plot, where metals are less 648 mixed and more uniformly distributed in the *EarlyRe- 649 ray* set. More metals also results in more efficient cooling 650 of gas, which is reflected in the temperature projection 651 plot. These behaviors are more prominent in smaller 652 halos, showing again that smaller halos are more sus- 653 ceptible by H_2 self-shielding modeling than of bigger 654 halos.

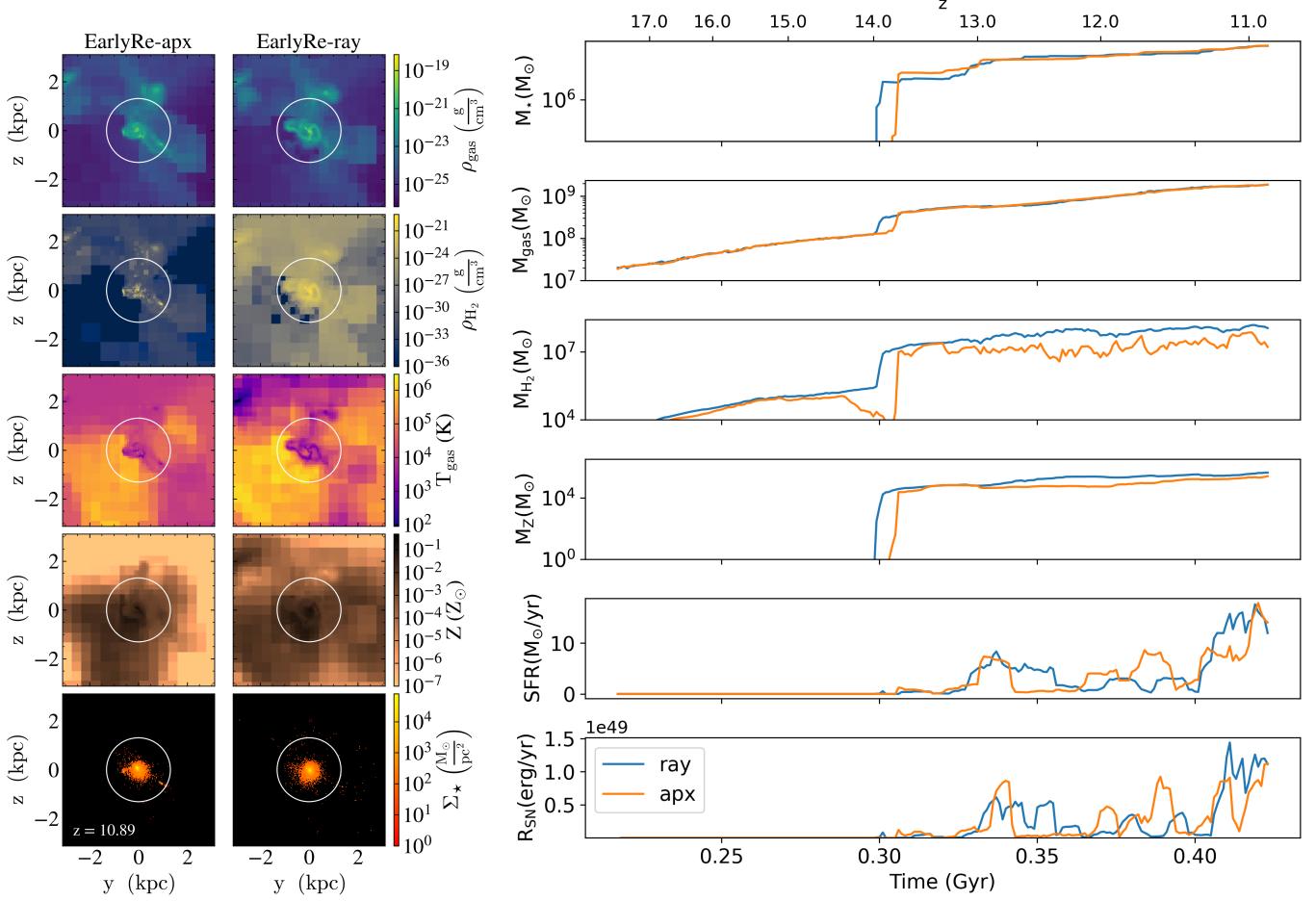


Figure 5. (Left) The gas density, H_2 density, temperature, metallicity, and stellar surface density projection plot of Halo 0 ($M_{\text{halo}} \approx 1.3 \times 10^{10} M_\odot$ at $z = 10.89$) within one virial radius. The white circle represents the boundary of the galaxy, which is defined as the radius enclosing the baryonic density 2000 times larger than the critical density of the universe. (Right) The time evolution of different properties in Halo 0 between the *EarlyRe-ray* and *EarlyRe-apx* simulations.

3.3. Effect on the galaxy and CGM

Sub-section 3.1 explores how two treatments of modulating H_2 affect star formation in a halo as a whole. In this section, we more closely examine the effect of H_2 self-shielding modeling on the ISM, the CGM, and the IGM to help further explain the previous observations we have on star formation.

For the radiative star cluster model and the Population III star model used in our simulation, their overdensity thresholds for a star particle to form are 10^7 and 10^6 , respectively. At the time step where we evaluate the f_{sh} ($z = 10.9$), these overdensity thresholds correspond to a gas density threshold of 4.5×10^{-20} and 4.5×10^{-21} . According to Figure 1, these thresholds lie in the regime where f_{sh} is typically over-estimated in the approximation model. On the other hand, H_2 gas in the interstellar medium (ISM) of EarlyRe-apx can be under-shielded.

Make a big figure showing all the halo pairs projection plot at the last snapshot.

3.4. Effect on the large-scale universe and reionization

We extend our analysis to the whole simulation box to investigate the effect of H_2 self-shielding modeling on a large-scale universe. Figure 9 shows the projection plots at different snapshots within the simulation's refined region, which is about 1/10000 volume of the whole simulation box. While the large-scale structure of gas is similar between two simulations, the distribution of other properties differ greatly. The amount of molecular hydrogen in the ISM is much lower when using the Sobolev-like approximation method

As suggested by Figure 9, the universe in EarlyRe-apx is much hotter at $z \approx 10.9$ than its counterpart in EarlyRe-ray. Because cosmic reionization is linked

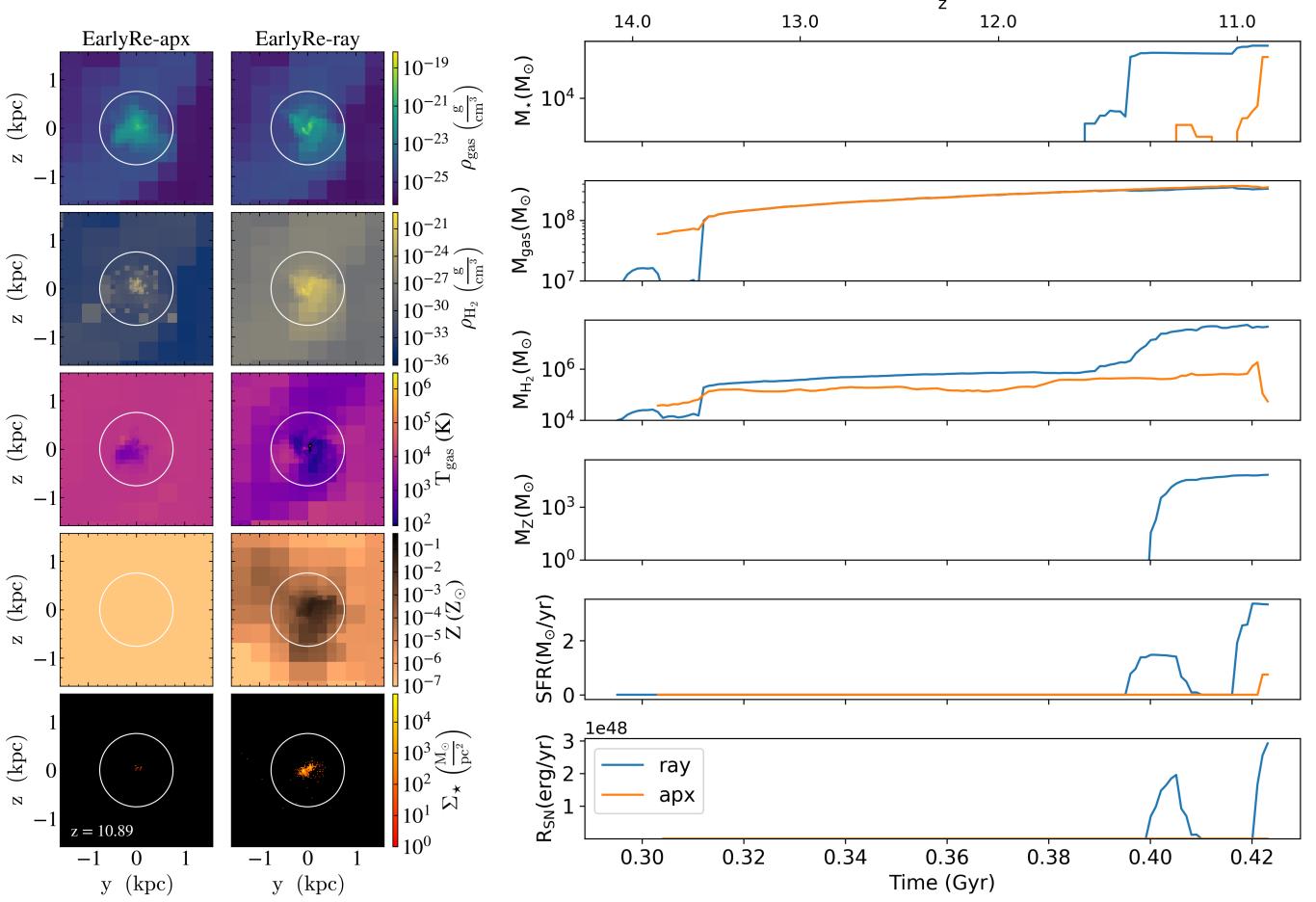


Figure 6. Same as Figure 5, but for another halo whose $M_{\text{halo}} \approx 1.7 \times 10^9 M_{\odot}$ at $z = 10.89$.

to the heating of the IGM (D’Aloisio et al. 2019), we expect that reionization is affected by the choice of self-shielding modeling. Figure 10 shows the distribution of neutral hydrogen fraction χ_{HI} within the refined reionization region between our two simulations. The χ_{HI} value of 0 corresponds to complete ionization of atomic hydrogen. High-energy radiation from galaxies ionizes hydrogen, decreasing χ_{HI} in the surrounding IGM regions. χ_{HI} can go up again during recombination between protons and electrons to form neutral hydrogen, which is efficient at temperature lower than 10^4 K (Wise 2019). In all time steps shown, *EarlyRe-apx* displays more extended and prominent region of ionized hydrogen, expanding out to the further IGM region. On the other hand, hydrogen is less ionized in *EarlyRe-ray*, and ionized hydrogen is limited around the galaxy filament (shown in the top two rows of Figure 9). This observation is quantified in Figure 11 where we plot the volume-weighted average of χ_{HI} within the simulation’s refined region as a function of time. At the start of both simulations, the universe

is dominated with neutral hydrogen ($\chi_{\text{HI}} \approx 1$). However, about 0.32 Gyr after the Big Bang ($z \approx 13.5$), the volume-weighted average of the neutral hydrogen fraction starts to drop considerably in the approximation model when the ray-tracing model only shows a slight decline. By the end of the simulation, while more than 90% of hydrogen is still neutral in *EarlyRe-ray*, only about 65% of hydrogen remains as such in *EarlyRe-apx*. While both simulations are not run long enough to cover the whole reionization epoch, Figure 11 suggests that the reionization may be complete earlier in *EarlyRe-apx* than in *EarlyRe-ray*, even though the reionization’s start time is similar. This illustrates that the use of Sobolev-like approximation to model H_2 self-shielding speeds up the reionization process and thus affect our theoretical prediction and modeling of the reionization epoch.

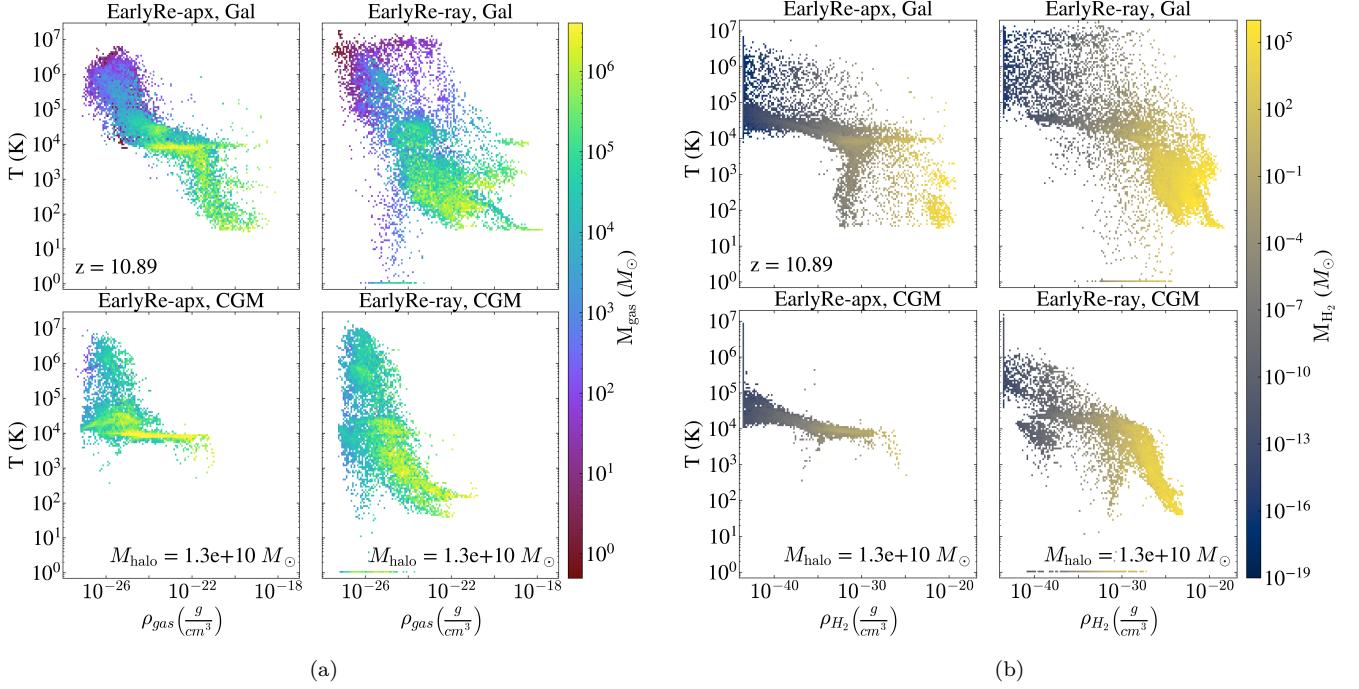


Figure 7. The gas phase plot (Sub-figure (a)) and the molecular hydrogen phase plot (Sub-figure (b)) of the largest halo ($M_{\text{halo}} = 1.3 \times 10^{10} M_{\odot}$) in our simulation set, evaluated at the last time step $z = 10.89$. For each sub-figure, the left column shows the halo in *EarlyRe-apx*, and the right column shows its counterpart in *EarlyRe-ray*. The top row in each subfigure shows the phase plot using the gas cells within the galaxy region ($r \leq R_{\text{bary},2000}$), and the bottom row shows the phase plot using the gas cells in the CGM region ($R_{\text{bary},2000} < r \leq R_{\text{vir}}$).

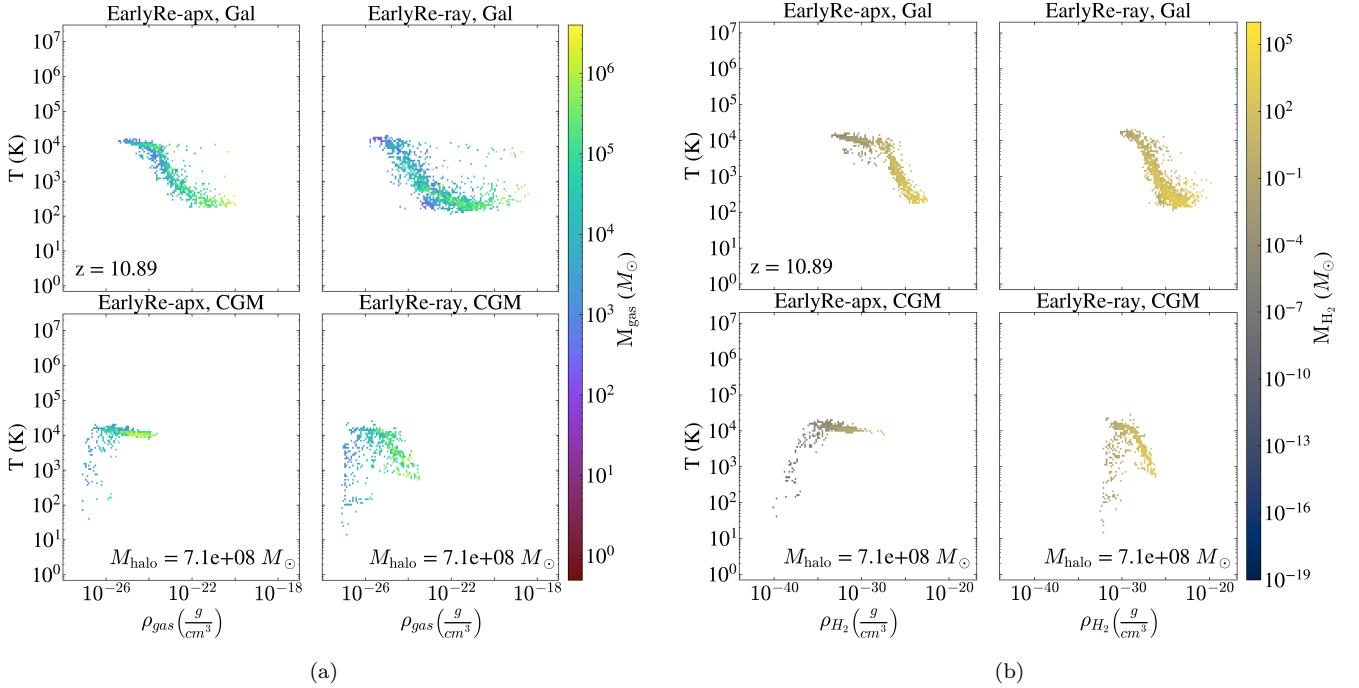


Figure 8. Similar to Figure 7, for a halo with mass $M_{\text{halo}} = 7.1 \times 10^8 M_{\odot}$.

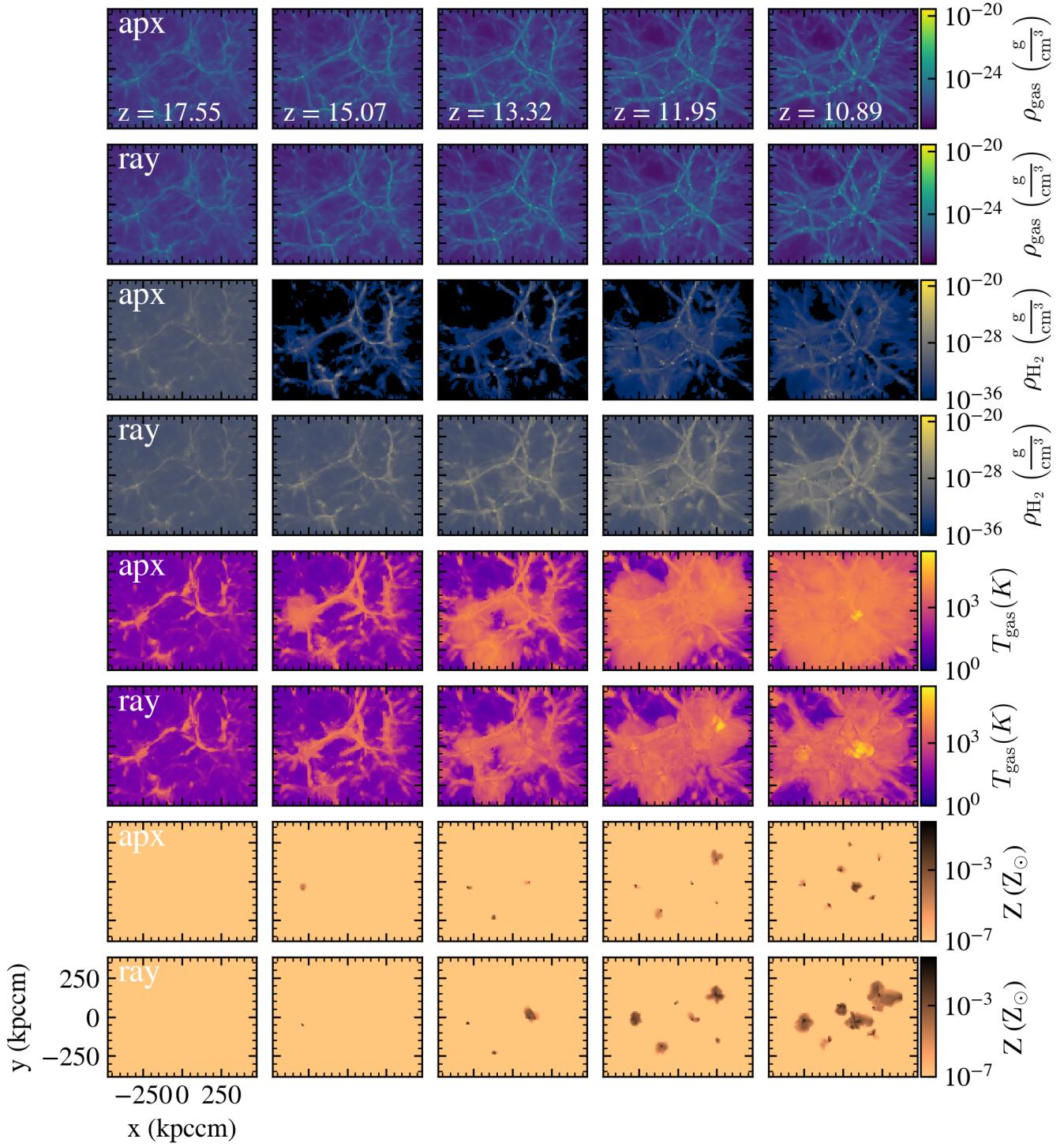


Figure 9. The gas density (1st and 2nd row), H_2 density (3rd and 4th row), gas temperature (5th and 6th row), and gas metallicity (7th and 8th row) within the simulation's refined region between the *EarlyRe-apx* and *EarlyRe-ray* across five time steps.

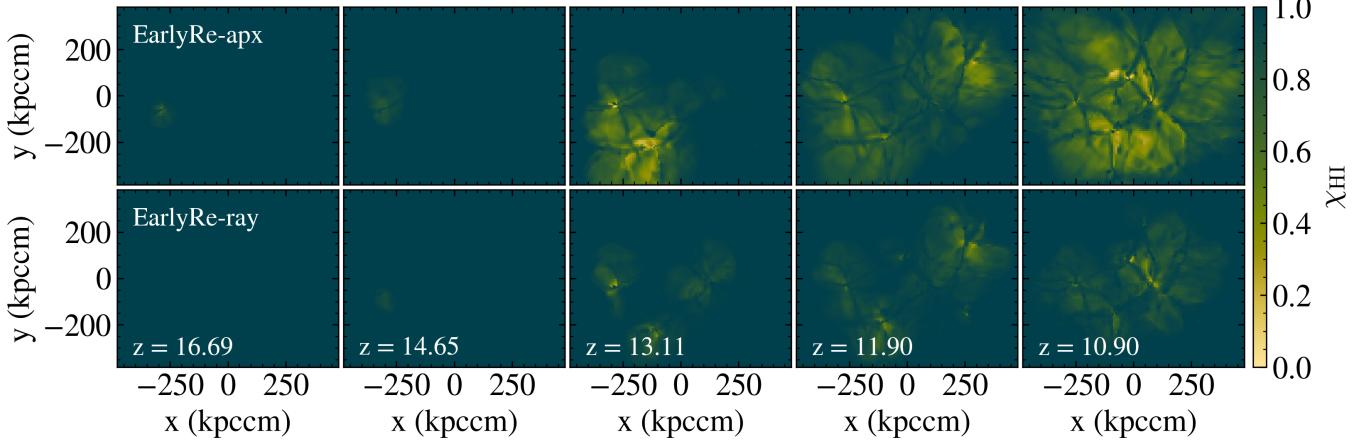


Figure 10. The neutral hydrogen fraction in the refined region box between the *EarlyRe-ray* and *EarlyRe-apx* simulations across five time steps.

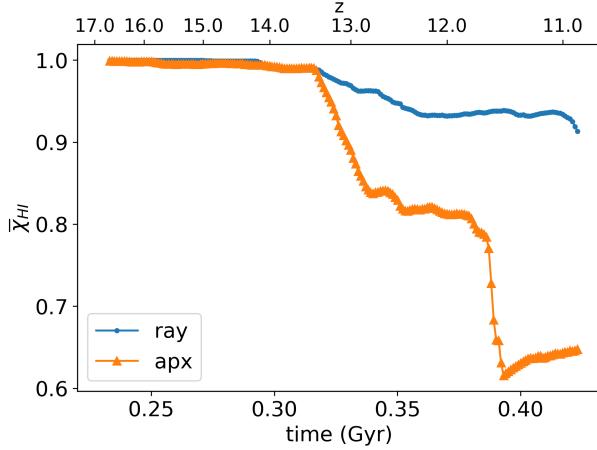


Figure 11. The comparison between the volume-weighted average of the neutral hydrogen fraction as a function of time between the *EarlyRe-ray* and the *EarlyRe-apx* simulations.

725 Maybe mention that the Renaissance Simulation (and
726 maybe others) doesn't implement self-shielding, so can
727 affect the results.

728 The strong radiation in late reionization destroys H₂
729 in both *ray* and *apx* sets. This is reflected by the fact
730 that the separation between the *ray* and the *apx* is not
731 as prominent as in the EarlyRe set. In the EarlyRe set,
732 the radiation is still weak, thus it only affects the non-
733 shielding version. The halos of *ray* version are shielded
734 with this weak radiation, but not with the strong radi-
735 ation as in LateRe.

736 **Remember to use bullet points for each main**
737 **result and specifically state the figure associated**
738 **with each main result.**

739 Acknowledgement: —

REFERENCES

- 740 Abel, T., Wise, J. H., & Bryan, G. L. 2007, The
741 Astrophysical Journal, 659, L87, doi: [10.1086/516820](https://doi.org/10.1086/516820)
- 742 Bryan, G. L., Norman, M. L., O'Shea, B. W., et al. 2014,
743 The Astrophysical Journal Supplement Series, 211, 19,
744 doi: [10.1088/0067-0049/211/2/19](https://doi.org/10.1088/0067-0049/211/2/19)
- 745 Chiaki, G., & Wise, J. H. 2023, Monthly Notices of the
746 Royal Astronomical Society, 520, 5077,
747 doi: [10.1093/mnras/stad433](https://doi.org/10.1093/mnras/stad433)
- 748 Christensen, C., Quinn, T., Governato, F., et al. 2012,
749 Monthly Notices of the Royal Astronomical Society, 425,
750 3058, doi: [10.1111/j.1365-2966.2012.21628.x](https://doi.org/10.1111/j.1365-2966.2012.21628.x)
- 751 D'Aloisio, A., McQuinn, M., Maupin, O., et al. 2019, The
752 Astrophysical Journal, 874, 154,
753 doi: [10.3847/1538-4357/ab0d83](https://doi.org/10.3847/1538-4357/ab0d83)
- 754 Davé, R. 2008, Monthly Notices of the Royal Astronomical
755 Society, 385, 147, doi: [10.1111/j.1365-2966.2008.12866.x](https://doi.org/10.1111/j.1365-2966.2008.12866.x)
- 756 Draine, B. T., & Bertoldi, F. 1996, The Astrophysical
757 Journal, 468, 269, doi: [10.1086/177689](https://doi.org/10.1086/177689)
- 758 Field, G. B., Somerville, W. B., & Dressler, K. 1966,
759 Annual Review of Astronomy and Astrophysics, 4, 207,
760 doi: [10.1146/annurev.aa.04.090166.001231](https://doi.org/10.1146/annurev.aa.04.090166.001231)
- 761 Galli, D., & Palla, F. 1998, Astronomy and Astrophysics,
762 335, 403, doi: [10.48550/arXiv.astro-ph/9803315](https://doi.org/10.48550/arXiv.astro-ph/9803315)
- 763 Glover, S. C. O., & Mac Low, M.-M. 2007a, The
764 Astrophysical Journal Supplement Series, 169, 239,
765 doi: [10.1086/512238](https://doi.org/10.1086/512238)
- 766 —. 2007b, The Astrophysical Journal, 659, 1317,
767 doi: [10.1086/512227](https://doi.org/10.1086/512227)

- 768 Gnedin, N. Y., & Kravtsov, A. V. 2011, The Astrophysical
 769 Journal, 728, 88, doi: [10.1088/0004-637X/728/2/88](https://doi.org/10.1088/0004-637X/728/2/88)
- 770 Gnedin, N. Y., Tassis, K., & Kravtsov, A. V. 2009, The
 771 Astrophysical Journal, 697, 55,
 772 doi: [10.1088/0004-637X/697/1/55](https://doi.org/10.1088/0004-637X/697/1/55)
- 773 Gorski, K. M., Hivon, E., Banday, A. J., et al. 2005, The
 774 Astrophysical Journal, 622, 759, doi: [10.1086/427976](https://doi.org/10.1086/427976)
- 775 Greif, T. H. 2014, Monthly Notices of the Royal
 776 Astronomical Society, 444, 1566,
 777 doi: [10.1093/mnras/stu1532](https://doi.org/10.1093/mnras/stu1532)
- 778 Hartwig, T., Clark, P. C., Glover, S. C. O., Klessen, R. S.,
 779 & Sasaki, M. 2015a, The Astrophysical Journal, 799, 114,
 780 doi: [10.1088/0004-637X/799/2/114](https://doi.org/10.1088/0004-637X/799/2/114)
- 781 Hartwig, T., Glover, S. C. O., Klessen, R. S., Latif, M. A.,
 782 & Volonteri, M. 2015b, Monthly Notices of the Royal
 783 Astronomical Society, 452, 1233,
 784 doi: [10.1093/mnras/stv1368](https://doi.org/10.1093/mnras/stv1368)
- 785 Johnson, J. L., Khochfar, S., Greif, T. H., & Durier, F.
 786 2011, Monthly Notices of the Royal Astronomical Society,
 787 410, 919, doi: [10.1111/j.1365-2966.2010.17491.x](https://doi.org/10.1111/j.1365-2966.2010.17491.x)
- 788 Kannan, R., Macciò, A. V., Fontanot, F., et al. 2015,
 789 Monthly Notices of the Royal Astronomical Society, 452,
 790 4347, doi: [10.1093/mnras/stv1633](https://doi.org/10.1093/mnras/stv1633)
- 791 Kennicutt, R. C. 1998, Annual Review of Astronomy &
 792 Astrophysics, 36, 189,
 793 doi: [10.1146/annurev.astro.36.1.189](https://doi.org/10.1146/annurev.astro.36.1.189)
- 794 Kennicutt, R. C. J., Calzetti, D., Walter, F., et al. 2007,
 795 The Astrophysical Journal, 671, 333, doi: [10.1086/522300](https://doi.org/10.1086/522300)
- 796 Leroy, A. K., Walter, F., Sandstrom, K., et al. 2013, The
 797 Astronomical Journal, 146, 19,
 798 doi: [10.1088/0004-6256/146/2/19](https://doi.org/10.1088/0004-6256/146/2/19)
- 799 Planck Collaboration, Ade, P. A. R., Aghanim, N., et al.
 800 2016, Astronomy and Astrophysics, 594, A13,
 801 doi: [10.1051/0004-6361/201525830](https://doi.org/10.1051/0004-6361/201525830)
- 802 Safranek-Shrader, C., Krumholz, M. R., Kim, C.-G., et al.
 803 2017, Monthly Notices of the Royal Astronomical Society,
 804 465, 885, doi: [10.1093/mnras/stw2647](https://doi.org/10.1093/mnras/stw2647)
- 805 Schmidt, M. 1959, The Astrophysical Journal, 129, 243,
 806 doi: [10.1086/146614](https://doi.org/10.1086/146614)
- 807 Shang, C., Bryan, G. L., & Haiman, Z. 2010, Monthly
 808 Notices of the Royal Astronomical Society, 402, 1249,
 809 doi: [10.1111/j.1365-2966.2009.15960.x](https://doi.org/10.1111/j.1365-2966.2009.15960.x)
- 810 Sobolev, V. V. 1957, Soviet Astronomy, 1, 678
- 811 Stecher, T. P., & Williams, D. A. 1967, The Astrophysical
 812 Journal, 149, L29, doi: [10.1086/180047](https://doi.org/10.1086/180047)
- 813 Stevens, A. R. H., Martig, M., Croton, D. J., & Feng, Y.
 814 2014, Monthly Notices of the Royal Astronomical Society,
 815 445, 239, doi: [10.1093/mnras/stu1724](https://doi.org/10.1093/mnras/stu1724)
- 816 Thoul, A. A., & Weinberg, D. H. 1995, The Astrophysical
 817 Journal, 442, 480, doi: [10.1086/175455](https://doi.org/10.1086/175455)
- 818 Wise, J. H. 2019, Contemporary Physics, 60, 145,
 819 doi: [10.1080/00107514.2019.1631548](https://doi.org/10.1080/00107514.2019.1631548)
- 820 Wise, J. H., & Abel, T. 2011, Monthly Notices of the Royal
 821 Astronomical Society, 414, 3458,
 822 doi: [10.1111/j.1365-2966.2011.18646.x](https://doi.org/10.1111/j.1365-2966.2011.18646.x)
- 823 Wise, J. H., & Cen, R. 2009, The Astrophysical Journal,
 824 693, 984, doi: [10.1088/0004-637X/693/1/984](https://doi.org/10.1088/0004-637X/693/1/984)
- 825 Wolcott-Green, J., Haiman, Z., & Bryan, G. L. 2011,
 826 Monthly Notices of the Royal Astronomical Society, 418,
 827 838, doi: [10.1111/j.1365-2966.2011.19538.x](https://doi.org/10.1111/j.1365-2966.2011.19538.x)
- 828 Wong, T., & Blitz, L. 2002, The Astrophysical Journal,
 829 569, 157, doi: [10.1086/339287](https://doi.org/10.1086/339287)
- 830 Yoshida, N., Abel, T., Hernquist, L., & Sugiyama, N. 2003,
 831 The Astrophysical Journal, 592, 645, doi: [10.1086/375810](https://doi.org/10.1086/375810)
- 832 Yoshida, N., Oh, S. P., Kitayama, T., & Hernquist, L. 2007,
 833 The Astrophysical Journal, 663, 687, doi: [10.1086/518227](https://doi.org/10.1086/518227)
- 834 Yoshida, N., Omukai, K., Hernquist, L., & Abel, T. 2006,
 835 The Astrophysical Journal, 652, 6, doi: [10.1086/507978](https://doi.org/10.1086/507978)