PREDICTING THE NUMBER DENSITY OF HEAVY SEED MASSIVE BLACK HOLES DUE TO AN INTENSE LYMAN-WERNER FIELD

Hannah O'Brennan[1,*], John A. Regan[1], John Brennan[1], Joe McCaffrey[1], John H. Wise[2], Eli Visbal[3], Alessandro Trinca[4,5,6], and Michael L. Norman[7]

[1]Centre for Astrophysics and Space Sciences Maynooth, Department of Physics, Maynooth University, Maynooth, Ireland
[2]Center for Relativistic Astrophysics, Georgia Institute of Technology, 837 State Street, Atlanta, GA 30332, USA
[3]Department of Physics and Astronomy and Ritter Astrophysical Research Center, University of Toledo, 2801 W. Bancroft Street, Toledo, OH 43606, USA
[4]Como Lake Center for Astrophysics, DiSAT, Università degli Studi dell'Insubria, via Valleggio 11, I-22100, Como, Italy
[5]INAF/Osservatorio Astronomico di Roma, Via Frascati 33, 00040 Monte Porzio Catone, Italy
[6]INFN, Sezione Roma1, Dipartimento di Fisica, "Sapienza" Universit'a di Roma, Piazzale Aldo Moro 2, 00185, Roma, Italy and
[7]Center for Astrophysics and Space Sciences, University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093, USA
Version July 9, 2025

ABSTRACT

The recent detections of a large number of candidate active galactic nuclei at high redshift (i.e. $z \gtrsim 4$) has increased speculation that heavy seed massive black hole formation may be a required pathway. Here we re-implement the so-called Lyman-Werner (LW) channel model of Dijkstra, Ferrara & Mesinger (2014) to calculate the expected number density of massive black holes formed through this channel. We further enhance this model by extracting information relevant to the model from the `Renaissance` simulation suite. `Renaissance` is a high-resolution suite of simulations ideally positioned to probe the high-z Universe. Finally, we compare the LW-only channel against other models in the literature. We find that the LW-only channel results in a peak number density of massive black holes of approximately 10^{-4} cMpc^{-3} at $z \sim 10$. Given the growth requirements and the duty cycle of active galactic nuclei, this means that the LW-only is likely incompatible with recent JWST measurements and can, at most, be responsible for only a small subset of high-z active galactic nuclei. Other models from the literature (e.g. rapid assembly; relative velocities between baryons and dark matter) seem therefore better positioned, at present, to explain the high frequency of massive black holes at high z.

Subject headings: Early Universe, Galaxy Formation, Black Holes

1. INTRODUCTION

Observations of high-z (i.e. $z \gtrsim 6$) quasars, with masses in excess of 10^8 M$_\odot$, indicate that the number density of such massive objects is approximately 1 per cubic gigaparsec (Bañados et al. 2018; Inayoshi, Visbal & Haiman 2020). The origin of these high-z quasars poses a major challenge to our understanding of compact object formation and evolution in the early Universe. However, these extreme mass objects represent only the tip of the iceberg with smaller mass black holes likely being very much more abundant. Prior to the launch and subsequent observations of the James Webb Space Telescope (JWST), the number densities of this smaller mass population at high redshift could only be inferred from existing relations (Mortlock et al. 2011; Bañados et al. 2018; Matsuoka et al. 2019; Venemans et al. 2020; Wang et al. 2021; Yang et al. 2021; Izumi et al. 2021; Andika et al. 2022).

Recent observations of massive black holes (MBHs) with masses between 10^6 and 10^8 M$_\odot$ in the early Universe ($z \gtrsim 4$) have however started to indicate that the population of MBHs is potentially relatively high and certainly at the upper end of what theoretical models had previously suggested (e.g. Habouzit et al. 2016; Greene, Strader & Ho 2020; Kokorev et al. 2023; Larson et al. 2023; Maiolino et al. 2024; Greene et al. 2024; Harikane et al. 2023). Significant uncertainties in determining the

nature of a large population of so-called Little Red Dot (LRD) galaxies mean that it is too early, as of yet, to get clarity on the actual MBH population at high redshift (Lambrides et al. 2024; Li et al. 2025; Ma et al. 2025). Nonetheless, conservative estimates of the MBH population based on observations of LRDs indicate MBH number densities in excess of 10^{-4} cMpc^{-3} (e.g. Pérez-González et al. 2024) (as opposed to a number density of approximately 10^{-9} cMpc^{-3} for the MBH with masses in excess of 10^8 M$_\odot$)[1].

The question then arises as to what is the formation process that drives the existence of this large population of MBHs? In this work we focus on the so-called heavy seed pathway, where the initial black hole has a mass in excess of 10^3 M$_\odot$ (Regan & Volonteri 2024). While it is possible that the entire population of MBHs originates from light seeds with initial masses of less than 1000 M$_\odot$, we do not investigate that scenario here and instead direct interested readers to papers which investigate that channel (e.g. Madau & Rees 2001; Alvarez, Wise & Abel 2009; Madau, Haardt & Dotti 2014; Lupi et al. 2016; Smith et al. 2018; Shi, Kremer & Hopkins 2024; Mehta, Regan & Prole 2024).

For forming heavy seed MBHs, three mainstream (astrophysical) mechanisms have come to the fore which offer possible pathways to achieving the masses of this population within the required time-frame. All of the mech-

*E-mail:hannah.obrennan.2021@mumail.ie

---

[1] We use the letter 'c' to denote comoving here and so cMpc^{-3} refers to per comoving cubic megaparsec.

anisms we will discuss, except where we explicitly note, predict the formation of a super-massive star (SMS) as an intermediate stage and subsequently a transition into a MBH. This is often dubbed "Direct Collapse Black Hole (DCBH)" formation in the literature which is somewhat incorrect given the intermediate stage of stellar evolution which can, in theory, continue for ≥ 1 Myr. In fact, a more correct use of the DCBH terminology relates to the concept of the so-called "Dark Collapse" recently introduced into the literature by Zwick et al. (2023) and Mayer et al. (2024). We will therefore refrain from using the DCBH term here since we focus on the formation of massive stellar objects as precursors to MBH formation.

The first mechanism theorised to generate MBH seeds is through baryonic streaming velocities. Relative velocity differences between baryons and dark matter will arise following recombination (Tseliakhovich & Hirata 2010). While the mean offset will be zero there will, nonetheless, be regions of the Universe where variations from the mean will exist. It is within these regions that the relative velocities can impact the early formation of structure. In particular, streaming motions can act to suppress star formation in the lowest mass halos, pushing the onset of star formation to higher mass halos and perhaps all the way up to the atomic-cooling limit (Naoz, Yoshida & Gnedin 2012, 2013; Tanaka & Li 2014; Latif, Niemeyer & Schleicher 2014; Hirano et al. 2017; Schauer et al. 2017). In regions impacted by streaming velocities, the additional velocities of the baryons with respect to the dark matter means that baryons take additional time to settle in the halo centres, thus allowing the halo to grow in mass. Whether realistic streaming motions can truly allow halos to grow to the atomic-cooling limit without triggering star formation is unclear but it may be that the combination of streaming motions with a sufficiently intense Lyman-Werner (LW) flux (described below) may allow for this to occur (Kulkarni, Visbal & Bryan 2021; Schauer et al. 2021).

The second mechanism is through the rapid growth of structure itself, be it the early rapid assembly of galaxies (Yoshida et al. 2003) or MBH formation triggered through mergers (Mayer et al. 2010, 2015; Zwick et al. 2023; Regan 2023). In the case of early rapid assembly, halos below the atomic-cooling threshold experience rapid minor mergers and hence rapid growth. If the growth is particularly rapid then the heating caused by the dark matter inflows can offset the ability of H_2 to cool the gas effectively (Fernandez et al. 2014; Wise et al. 2019; Lupi, Haiman & Volonteri 2021). In this scenario normal metal-free (Population III; Pop III) star formation is delayed because the gas is unable to cool. Instead the halo continues to accumulate matter without forming stars. If the heating (i.e. rapid growth) is maintained then the halo can grow to the atomic-cooling limit where cooling by neutral hydrogen is triggered and star formation occurs regardless. However, having cooling commence at this mass scale offers the opportunity for SMS formation to occur as the gravitational potential is now sufficiently deep to enable this mechanism (e.g Regan et al. 2020; Latif et al. 2022; Regan 2023).

In addition to this pathway, the related pathway of major mergers can drive huge gas inflows directly into the centres of merging halos. The large gas inflows can in some cases initiate a dark collapse through the formation of a super-massive disk (Zwick et al. 2023) and ultimately the direct formation of a MBH (Mayer et al. 2010, 2024). While the investigations of Mayer et al. (2024) focused on very massive objects and the formation in particular of high-z quasars, it is likely that this mechanism also acts on smaller mass scales, leading to the formation of a population of high-z MBHs with masses closer to the expected MBH seed masses (Regan 2023).

The final mechanism and the one we will focus on in this paper is driven by local sources of LW radiation. LW radiation is emitted by stars and is composed of radiation below the hydrogen ionisation edge at photon energies between 11.2 and 13.6 eV. In order to form a Pop III star, the gas within a mini-halo[2] must cool down to $T_{gas} \approx 200$ K, allowing the gas to achieve the required densities and pressures to ignite star formation. Sufficiently intense LW radiation works against this process by dissociating H_2, thus removing (or at least suppressing) a critical coolant required for star formation to take place. In the primordial Universe, the formation of H_2 takes place via two possible routes (Galli & Palla 1998). The first is through the radiative association of H and H^+ below:

$$H + H^+ \rightarrow H_2 + \gamma. \tag{1}$$
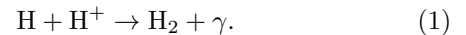
This reaction is only important at very high redshift ($z \gtrsim 365$) when $T_{CMB} > 10^3$ K and hence is less relevant for galaxy formation. The more relevant reaction is the associative detachment reaction of H and H^- which gives

$$H + H^- \rightarrow H_2 + e^-. \tag{2}$$
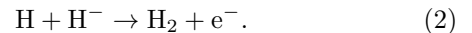
This is the critical reaction of H_2 formation in the high-z Universe relevant for galaxy formation and depends sensitively on the H^- abundances. In opposition to these formation pathways is LW radiation which can dissociate H_2. Radiation in the LW band excites electrons in the H_2 molecule, breaking the molecule down into its constituent atoms as such:

$$H_2 + \gamma_{LW} \rightarrow H_2^* \rightarrow H + H, \tag{3}$$
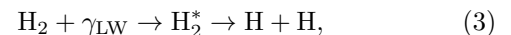
where initially the LW photons excite electrons in the H_2 molecule into an excited state from where they can decay with some probability into two H atoms. Crucially by removing the (inefficient) coolant H_2, the gas cannot cool and hence star formation is suppressed until the halo reaches the atomic-cooling limit, allowing the gas to cool and condense via neutral hydrogen transitions.

The intensity of LW radiation required to achieve full star formation suppression in mini-halos has been well-studied over the past decade. The consensus is that, while there is some spectral dependence on the critical flux of LW radiation required, values of at least 300 J_{21} are required for a $T \gtrsim 10^4$ K spectrum (e.g. Latif et al. 2014; Regan et al. 2017) while values of as high as 1000 J_{21} may be required for a spectrum composed of Pop III stars only (i.e. $T \sim 10^5$ K spectrum) (Agarwal & Khochfar 2015; Agarwal et al. 2016) (where $J_{21} = 10^{-21} \, erg \, s^{-1} \, Hz^{-1} \, sr^{-1} \, cm^{-2}$). The value of the LW flux required is then known as the critical flux, J_{crit}. The question then becomes under what circumstances can a LW flux greater than J_{crit} be achieved in

---

[2] We use the term mini-halo to refer to halos above $M_{halo} = 10^5$ $M_\odot$ and below the atomic-cooling threshold.

practice? The goal of this paper is to re-examine the methodology introduced by Dijkstra et al. (2008); Dijkstra, Ferrara & Mesinger (2014) (hereafter D08 & D14) while also augmenting the models with data from the Renaissance (Chen et al. 2014; O'Shea et al. 2015; Xu et al. 2016) simulation suite. Finally, we also compare the LW-only channel against more recent estimates of MBH heavy seed number densities through other channels (e.g. Trinca et al. 2022; McCaffrey et al. 2025). Using these combined results, we put tight limits on the number density of MBH seeds formed via LW feedback only and critically examine how viable the LW-only pathway is in the face of recent JWST observations. The structure of the paper is as follows: In §2, we outline the methodology including the model from D14 and the relevant output post-processed from Renaissance. In §3, we deliver the results of our analysis, showing an updated heavy seed number density plot. In §4, we summarize and discuss our results in light of recent JWST observations and other models in the literature.

2. METHODOLOGY

In this section, we detail our application of the heavy seed formation model used by D08 and D14. In §2.1, we describe the conditions a halo must meet to potentially host a heavy seed. In §2.2, we introduce the Renaissance simulations and how we post-process the data to supplement the fiducial (analytic-only) model. In §2.3, we define the halo mass function and how we use it to compute the halo number density. In §2.4, we discuss the impact of metal pollution, both from merger history and from neighbouring galaxy outflows. Finally, in §2.5, we show how the supercritical LW flux probability is computed, both in the fiducial model and one informed by Renaissance data.

2.1. Formation of Heavy Seeds

Here we refer to heavy seeds as MBH seeds that have initial masses $\geq 10^3$ $M_\odot$, as opposed to light seeds with initial masses $< 10^3$ $M_\odot$. We find the heavy seed number density $n_{\text{heavy seeds}}(z)$ (units: cMpc^{-3}) as a function of redshift z ($10 \leq z \leq 30$) by computing the number density of dark matter halos which meet the criteria to host heavy seeds (assuming one heavy seed forms per host halo). For the purpose of this study, the dominant criteria is the impact of a LW radiation field as discussed in §1 and in detail below.

Consider a potential host halo of mass M_{target} at redshift z (see Figure 1). It is surrounded by neighbouring dark matter halos of various masses, M, and physical separations, r. At this high-redshift range, the comoving distance from an observer on Earth to such a host halo is ≥ 9 cGpc (at $z = 10.0$, the proper distance would be $\gtrsim 800$ Mpc). We are considering physical separations between the host halo and its neighbours of $\lesssim 10$ Mpc. Since these separations are negligible compared to the proper distances from Earth, we approximate the neighbouring halos as having the same redshift as the potential host.

To form light seeds within a halo, baryonic gas would cool and fragment to form stars which would later collapse into stellar black holes[3]. Thus to reduce fragmen-

---

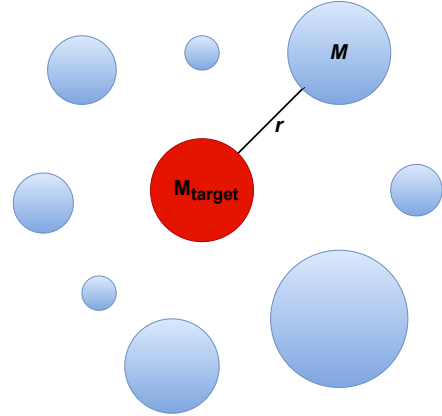[3] At time of writing, it is still not clear whether there is a mass

FIG. 1.— Target halo surrounded by neighbouring halos of various masses and physical separations. Here M refers to the mass of a given neighbouring halo, r is the physical separation between the neighbouring halo and the (central) target halo. The central target halo is later defined as having a mass $M_{\text{target}} = M_{\text{min}}(z)$ which is exactly equivalent to the virial mass of a halo at redshift z with $T_{\text{vir}} = 10^4$ K.

tation and the formation of light seeds (forming heavy seeds instead), the temperature of the gas within the target halo must remain high ($T_{\text{gas}} \gtrsim 8000$ K).

This can be achieved if the target halo meets the following criteria:

1. it is massive enough for the gas to cool via atomic hydrogen only;
2. it is chemically pristine i.e. free from metal enrichment;
3. it receives sufficient LW radiation from within its neighbours to suppress H_2 cooling.

The number density of dark matter halos (units: cMpc^{-3}) meeting these criteria (and thus the heavy seed number density) is described as:

$$n_{\text{heavy seeds}}(z) = \int_{M_{\text{min}}(z)}^{\infty} dM_{\text{target}} \frac{dn_{\text{SMT}}}{dM}(z, M_{\text{target}})$$

$$\times P_{\text{pristine}}(z) \, P_{\text{LW}}(z, M_{\text{target}}). \quad (4)$$

We now briefly define each term in the above integral. We integrate over M_{target}, the mass of a halo that is a potential heavy seed (or equivalently MBH) formation site. Here $M_{\text{min}}(z)$ is the minimum mass where the virial theorem is satisfied with $T_{\text{vir}} = 10^4$ K (Barkana & Loeb 2001) i.e. the atomic-cooling limit and is given as:

$$M_{\text{min}}(z) = 4 \times 10^7 \left(\frac{1+z}{11} \right)^{-3/2} M_\odot. \quad (5)$$

We integrate the halo mass function $\frac{dn_{\text{SMT}}}{dM}(z, M_{\text{target}})$ (units: $M_\odot^{-1}$ cMpc^{-3}) to find the number density of halos greater than or equal to the atomic-cooling limit. The

---

gap between the most massive neutron stars (≈ 2 $M_\odot$, Tolman-Oppenheimer-Volkoff limit (Heger et al. 2003)) and a minimum stellar black hole mass (≈ 5 $M_\odot$) (Farr et al. 2011; Kreidberg et al. 2012; Abbott et al. 2019).

halo mass function gives the number of halos per comoving volume per unit mass. It was originally derived analytically in a seminal paper by Press & Schechter (1974). In this work we use the semi-analytic form from Sheth, Mo & Tormen (2001) (SMT) which assumes ellipsoidal halo collapse (as opposed to the idealised spherical collapse assumed by Press & Schechter (1974)). We implement it using the Python package hmf, developed by Murray, Power & Robotham (2013).

The quantity $P_{\mathrm{pristine}}(z)$ refers to the probability of a target halo being pristine i.e. free from metal pollution. In §2.4, we discuss both the model used by D14 (Trenti & Stiavelli 2007, 2009) and one derived from Renaissance data and refer to the derived probabilities as $P_{\mathrm{pristine, fid.}}(z)$ and $P_{\mathrm{pristine, Ren.}}(z)$ respectively. $P_{\mathrm{pristine, fid.}}(z)$ accounts only for metal pollution inherited via merging episodes. $P_{\mathrm{pristine, Ren.}}(z)$ accounts for both inherited metal pollution and metal enrichment via supernovae outflows from neighbouring halos and does not distinguish between the two.

Finally, a given target halo receives LW radiation from its neighbours (which dissociates H_2) but also metal outflows from supernovae (which act as a coolant). If the target halo receives a total LW flux exceeding a threshold flux ($J > J_{\mathrm{crit}}$), it is a candidate for heavy seed formation due to the sufficient dissociation of H_2. The quantity $P_{\mathrm{LW, fid.}}(z, \mathrm{M_{target}})$ refers to the probability of a target halo at redshift z and of mass $\mathrm{M_{target}}$ receiving supercritical LW radiation while also avoiding metal outflows from its neighbouring halos in the fiducial model. The analogous quantity $P_{\mathrm{LW, Ren.}}(z, \mathrm{M_{target}})$ refers only to supercritical LW radiation as metal outflows are accounted for by $P_{\mathrm{pristine, Ren.}}(z)$ (see §2.4 for more detail). In this model, J_{crit} is a parameter and we evaluate the supercritical probability, and subsequently the heavy seed number density, at $J_{\mathrm{crit}} = 300$ and $1000 \, \mathrm{J_{21}}$. This probability is highly dependent on the form of the mean LW luminosity density $\langle L_{\mathrm{LW}}(z, M) \rangle$. We investigate both the form described by D14 and one derived from Renaissance data, referred to as $\langle L_{\mathrm{LW, fid.}}(z, M) \rangle$ and $\langle L_{\mathrm{LW, Ren.}}(z, M) \rangle$ respectively.

We describe each of the terms in Eq. 4 in greater detail in §2.3, §2.4 and §2.5 below. Our cosmology matches the cosmology used by D14: $n_{\mathrm{s}} = 0.9624$, h $= 0.6711$, $\Omega_{\mathrm{m,0}} = 0.3175$, $\Omega_{\mathrm{b,0}} \mathrm{h}^2 = 0.022068$, $\Omega_\Lambda = 0.6825$ and $\sigma_8 = 0.8344$. The impact of a specific cosmology on our model is however likely to be minimal. Note that we use units with factors of h when performing calculations. This is to be consistent with units used in hmf[4] and halomod[5], two Python packages that we use to compute the halo mass function, power spectrum and halo bias factor (Murray, Power & Robotham 2013; Murray et al. 2021). Our final results in Figure 8 have been converted to units without factors of h.

2.2. Renaissance Simulations

The Renaissance simulations (Xu, Wise & Norman 2013; Xu et al. 2014; Chen et al. 2014; O'Shea et al. 2015; Smith et al. 2018; Wise et al. 2019) were run using the massively parallel adaptive mesh refinement Enzo code (Bryan et al. 2014; Brummel-Smith et al. 2019). We

[4] hmf: https://pypi.org/project/hmf/1.6.2/
[5] halomod: https://github.com/halomod/halomod/

briefly describe the Renaissance suite here, but refer the interested reader to the previous papers for a more complete discussion. The Renaissance simulation suite is composed of three zoom-in regions extracted from a parent volume of $(40 \, \mathrm{cMpc})^3$. The three separate zoom-in regions were named the *Rarepeak* (RP) region, the *Normal* region and the *Void* region. Each volume was smoothed on a scale of 5 cMpc, with the RP region corresponding to a mean overdensity of $\langle \delta \rangle \equiv \frac{\langle \rho \rangle}{\Omega_{\mathrm{m}} \rho_c} - 1 \sim 0.68$ and the *Normal* region corresponding to $\langle \delta \rangle \equiv \frac{\langle \rho \rangle}{\Omega_{\mathrm{m}} \rho_c} - 1 \sim 0.09$. The RP and *Normal* subregions have volumes of 133.6 and 220.5 cMpc^3 respectively. They were resimulated with an effective initial resolution of 4096^3 grid cells and particles within the central, most refined regions, and have a particle mass resolution of 2.9×10^4 $\mathrm{M_\odot}$. In addition to the initial nesting procedure, which allows the RP and *Normal* regions to be selected, up to 8 further levels of refinement are allowed, giving a total maximum spatial resolution of ≈ 19 cpc. For this study we focus on the *Normal* region only since it is a representative volume of mean cosmic density. This is potentially quite conservative since it is less likely to find high values of $P_{\mathrm{LW}}(z, J_{\mathrm{crit}})$ within *Normal* regions. We make this choice since the alternatives (RP and *Void*) would be difficult to relate to observations.

From Renaissance we can compare some of the key characteristics of the D14 model against simulation datasets in order to test the veracity of some of the key underlying assumptions of the model. Considering again Eq. 4, we see that calculating the number density depends on the halo mass function, the probability of a halo receiving supercritical flux and the probability of a halo remaining pristine. From Renaissance the key quantities that can be extracted are the mean LW luminosity densities $\langle L_{\mathrm{LW, Ren.}}(z, M) \rangle$ as a function of redshift and neighbouring halo mass and the pristine fraction of halos $P_{\mathrm{LW, Ren.}}(z, \mathrm{M_{target}})$ as a function of redshift and target halo mass.

Note that while previous works (see e.g. Smith et al. (2018); McCaffrey et al. (2025)) showed that Renaissance post-processed black holes struggled to grow to heavy seed scales i.e. $M_{\mathrm{BH}} \sim 10^5 \, \mathrm{M_\odot}$, we are only interested in halos themselves as hosts whose environment fits the criteria to later form heavy-seed black holes. We do not investigate the later growth and dynamics of these black holes in this study.

However, we begin our analysis by examining the halo mass function, which is the main driver of the number of heavy seeds. We then move onto the pristine fraction (where Renaissance will play a role) and the probability of a halo receiving a super-critical flux (where again Renaissance will play a role).

2.3. Halo Mass Function

In Figure 2 we show how the halo mass function varies with halo mass ($10^7 \, \mathrm{h^{-1} \, M_\odot} \leq \mathrm{M_{target}} \leq 10^{10} \, \mathrm{h^{-1} \, M_\odot}$) at a number of fixed redshifts. More massive halos become significantly rarer at all redshifts, with the higher redshifts showing increased rarity of high mass halos as expected in a ΛCDM cosmology. The halo mass function is relatively well calibrated against N-body simulations at lower redshifts but can differ from N-body runs at high-z by up to an order of magnitude (for a discus-
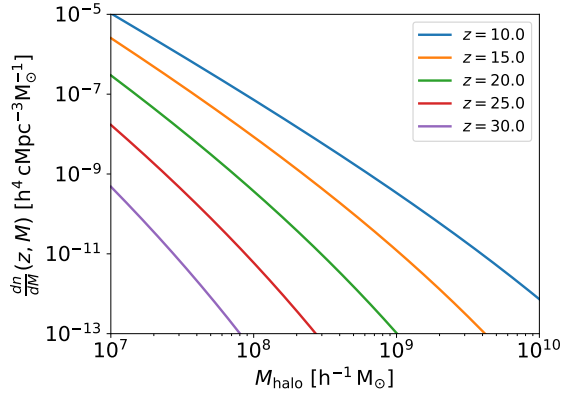
FIG. 2.— Halo mass function vs. halo mass at a number of redshifts. The lines shown here are generated using the `hmf` package developed by Murray, Power & Robotham (2013). We use the SMT halo mass function and the modified Planck13 cosmology (used also by D14). The rarity of halos in a given mass range increases with increasing redshift.

sion on this point see Yung et al. 2024; O'Brennan et al. 2024). However, for the halo masses which we are concerned with here (i.e. close to the atomic-cooling limit) the differences are approximately a factor of two and we do not investigate systematic differences due to the halo mass function here.

Since the halo mass function decreases by several orders of magnitude with increasing halo mass, we approximate Eq. 4 as:

$$n_{\text{heavy seeds}}(z) \approx \int_{\text{M}_{\text{min}}(z)}^{\infty} d\text{M}_{\text{target}} \frac{dn_{\text{SMT}}}{dM}(z, \text{M}_{\text{target}})$$
$$\times P_{\text{pristine}}(z) \, P_{\text{LW}}(z, \text{M}_{\text{min}}(z)),$$

$$n_{\text{heavy seeds}}(z) \approx n_{\text{halo}}(\text{M}_{\text{target}} > \text{M}_{\text{min}}(z))$$
$$\times P_{\text{pristine}}(z) \times P_{\text{LW}}(z, \text{M}_{\text{min}}(z)),$$

(6)

i.e. when computing P_{LW} in §2.5, we approximate all potential formation sites at some redshift z as having a fixed mass of $\text{M}_{\text{target}} = \text{M}_{\text{min}}(z)$.

2.4. *Metal Pollution*

Metal pollution can come in two forms. Firstly it can be inherited "genetically" through the hierarchical assembly process and secondly it can come from outflows from neighbouring halos. For the fiducial (analytic-only) model these two forms are computed separately, while for the model where we additionally use `Renaissance` data we simply calculate the halo metallicity (or more precisely the probability of a halo being metal-enriched via mergers or outflows from neighbours).

2.4.1. *Genetic Metal Pollution*

Starting with the analytic-only model: consider a target halo at redshift z with mass M_{target}. During previous episodes of halo merging, this target halo may have had a progenitor halo that became metal-enriched via supernova outflows from Pop III stars. If so, then the target halo would inherit this metal pollution i.e. genetic metal pollution. The quantity $P_{\text{pristine, fid.}}(z)$ refers to the probability of the target halo avoiding this genetic metal pollution i.e. it does not have a metal-enriched progenitor halo.

In our fiducial model, as with D14, we base the form of $P_{\text{pristine, fid.}}(z)$ on the model proposed by Trenti & Stiavelli (2007) (TS07) and Trenti & Stiavelli (2009) (TS09). TS07 investigated the formation of the first generation of Pop III stars within dark matter halos and tracked their descendant halos using a combination of `Gadget-2` N-body simulations and a Monte Carlo method based on linear theory. Using these simulation results, TS09 found how $\text{M}_{\text{H}_2}(z, J_{\text{bg}})$ (units: h^{-1} $\text{M}_\odot$) varied as a function of redshift z and LW background J_{bg}. M_{H_2} refers to the minimum mass where H_2 cooling may occur within a progenitor halo, and later Pop III star formation.[6] Figure 3 (a recreation of TS09 Figure 1 (upper right panel)) shows how this mass increases by a factor of ≈ 30 as redshift decreases from $z \approx 20$ to $z \approx 10$. Conversely, the atomic-cooling limit $\text{M}_{\text{min}}(z)$ increases only by a factor of ≈ 4 in the same redshift range and $\text{M}_{\text{min}}(z) < \text{M}_{\text{H}_2}(z, J_{\text{bg}})$ at $z \lesssim 13$. Thus a progenitor halo is more likely to cool via atomic hydrogen than H_2 at low redshifts.
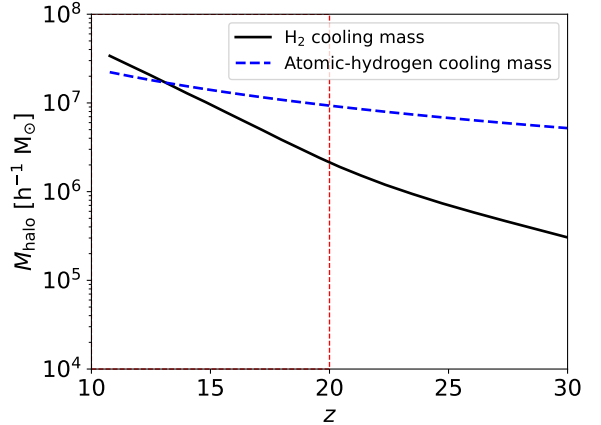

FIG. 3.— A recreation of Figure 1 (upper right panel) from Trenti & Stiavelli (2009). In this redshift range, the minimum halo mass for H_2 cooling to occur (solid black line) sharply increases as z decreases and is greater than the atomic-cooling limit (blue dashed line) at $z \lesssim 13$. The red dashed vertical line marks $z = 20$, the highest redshift considered by D14.

As a result, TS09 have shown that the probability of Pop III star formation is less likely in a progenitor halo as redshift decreases. Subsequently the probability, $P_{\text{gen}}(z)$, of a target halo inheriting metal pollution decreases as redshift decreases (see our Figure 4 for a recreation of TS09 Figure 1 (lower right panel)). TS09 found that $P_{\text{gen}}(z = 20) \approx 0.9$ and $P_{\text{gen}}(z = 10) \approx 0.1$.

We now look at metal enrichment in `Renaissance`. For the `Renaissance` halos, we select only halos above the atomic-cooling threshold mass, M_{min}, and from the *Normal* region only. We show the probability of a halo being metal-enriched in `Renaissance` in Figure 4 as the orange line i.e. the fraction of halos found to be metal-enriched at a given redshift. Note that this metal enrichment is from both genetic pollution and external enrichment and hence direct comparisons should be treated with this in mind. We consider such halos to be pristine only if they are metal-free i.e. $Z < 10^{-16}$ $\text{Z}_\odot$. Halos with greater metallicity are counted as metal-enriched. At high red-

---

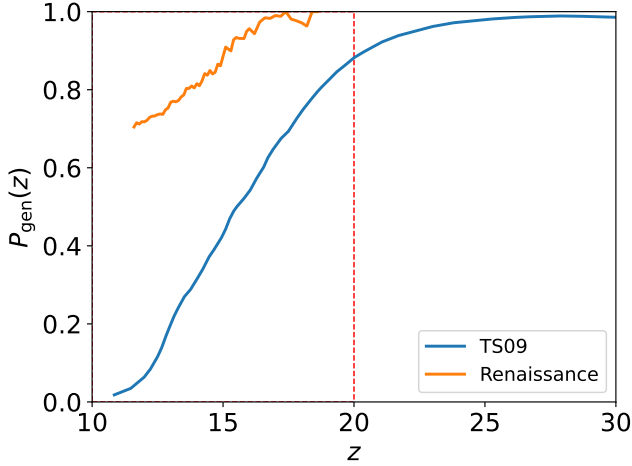[6] TS09 found $J_{\text{bg}}(z)$ as a function of z, so M_{H_2} is strictly a function of z only.

FIG. 4.— A recreation of Figure 1 (lower right panel) from Trenti & Stiavelli (2009). This is the probability for genetic metal pollution to occur within a halo with $T_{\rm vir} = 10^4$ K. From $z = 20$ to $z = 10$, $P_{\rm gen}(z)$ (solid blue line) sharply decreases i.e. a target halo at low redshift is less likely to inherit metal pollution from a progenitor halo. Overplotted as the solid orange line is the probability of a halo with mass in excess of the atomic-cooling mass being metal polluted taken from the Renaissance simulation suite. This metal pollution probability includes receiving metals from both genetic metal pollution and external metal pollution. The dashed red line marks $z = 20$.
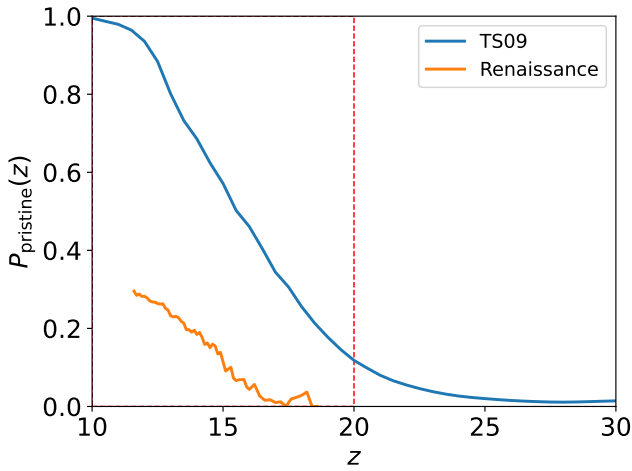

FIG. 5.— $P_{\rm pristine}(z) = 1 - P_{\rm gen}(z)$ as a function of redshift, based on the model used by TS07 and TS09 (solid blue line). From $z = 20$ to $z = 10$, $P_{\rm pristine, fid.}(z)$ sharply increases i.e. a target halo at low redshift is more likely to be free of genetic metal pollution. In orange we plot $P_{\rm pristine, Ren.}(z)$ from the Renaissance simulation suite (which again accounts for both genetic and external metal enrichment). The dashed red line marks $z = 20$.

shifts ($z \gtrsim 18$), the fraction is zero due to the fact that there are no halos with masses in excess of $M_{\rm min}$ at that point in the simulation (sub)volumes. As structure formation evolves and halos accrete sufficient mass, then the first atomic halos are predominantly star-forming and metal-enriched and hence the fraction of those halos which are pristine is exceptionally low to start (see Figure 5). As more and more halos form then the number (and fraction) of pristine halos increases substantially approaching 30% by the end of the simulation. No turnover is seen but this is expected to occur as metal diffusion becomes more widespread over cosmic time. Note that

the apparent sharp decrease in $P_{\rm pristine}$ from $z \approx 18$ to 17 in Figure 5 is likely noise due to small number statistics i.e. there are very few halos in this redshift range that are massive enough to be counted in this analysis.

The Renaissance results are clearly deviant from the analytical models - likely due to the self-consistent treatment of structure formation and metal enrichment available to the hydrodynamic simulations. However, as noted already, we need to be somewhat careful here. Metal enrichment in the Renaissance halos can come from either genetic enrichment as well as external enrichment (via outflows from neighbouring galaxies). While we see a similar trend to the analytic models of TS07 and TS09, we clearly see that the hydrodynamical simulations predict higher values of (genetic and external) metal pollution than the analytic models. Figure 5 shows the same result, albeit inverted since we show the pristine fraction (which is required by our model). For the analytic model we address the issue of external metal enrichment in §2.4.2 below.

2.4.2. *Metal Pollution from Neighbouring Halos*

Within a neighbouring halo of mass M (units: h^{-1} $M_\odot$), massive stars eject metals as supernovae at the end of their lives. We approximate this phenomenon by assuming a fraction of stars within the halo enter their supernova phase simultaneously and there is a single physical radius of metal pollution $r_{\rm s}(z, M, t)$ (units: h^{-1} Mpc) growing from the halo centre as t increases. It is given as:

$$r_{\rm s}(z, M, t) = \Big(\frac{E_0 \nu f_* \Omega_{\rm b,0} M}{\rho_{\rm gas}(z) \Omega_{\rm m,0}} \Big)^{1/5} t^{2/5}, \qquad (7)$$

where $E_0 = 10^{51}$ erg is the supernova explosion energy, $\nu = 0.01/h$ h $M_\odot^{-1}$ is the number of supernova per unit mass formed and $\rho_{\rm gas}(z) = \Delta \Omega_{\rm b,0} \rho_{\rm crit,0}(1 + z)^3$ is the density of the ambient gas. We set our over-density parameter to $\Delta = 60$. We assume that the beginning of the target halo collapse, the beginning of star formation within the neighbouring halo and the beginning of the neighbouring supernova phase occur simultaneously (at $t = 0$). The target halo must avoid metal pollution while $t < t_{\rm ff}(z)$ where $t_{\rm ff}(z)$ is the free-fall time of the target halo. In other words for a set redshift z and neighbouring halo mass M, the halo separation $r > r_{\rm s}(z, M, t = t_{\rm ff}(z))$ since $r_{\rm s}$ increases monotonically with t. When computing the heavy seed number density in the fiducial model, we include a Heaviside function $\Theta[r - r_{\rm s}(z, M, t_{\rm ff}(z))]$ to only include metal-free halos described above. This is only necessary for the analytic model. When modelling using the Renaissance data, we do not need to include the Heaviside function as this information is already contained as part of the pristine fraction of halos $P_{\rm pristine, Ren.}(z)$.

2.5. *Supercritical Flux*

Here we derive $P_{\rm LW}(z, M_{\rm target})$ i.e. the probability of a target halo receiving supercritical LW radiation. Since we use the approximation described in Eq. 6, this probability is now a function of z only i.e. $P_{\rm LW}(z, M_{\rm target}) = P_{\rm LW}(z, M_{\rm min}(z)) = P_{\rm LW}(z)$. We compute this analytically by approximating that the LW flux received by the

target halo is dominated by a single luminous nearby source.

We integrate a probability density over LW flux J:

$$P_{\text{LW, fid.}}(z) = \int_{\log_{10} J_{\text{crit}}}^{\infty} d\log_{10} J \frac{dP_{\text{fid.}}}{d\log_{10} J}(z, J),$$
$$P_{\text{LW, Ren.}}(z) = \int_{\log_{10} J_{\text{crit}}}^{\infty} d\log_{10} J \frac{dP_{\text{Ren.}}}{d\log_{10} J}(z, J). \tag{8}$$

Here J is in units of J_{21} and we integrate over the logarithmic value. The LW flux probability density is found by integrating over all possible masses (units: h^{-1} $M_\odot$) and physical separations (units: h^{-1} Mpc) of neighbouring halos:

$$\frac{dP_{\text{fid.}}}{d\log_{10} J}(z, J) = \int_{M_a}^{M_{b,\text{fid.}}} dM \int_{r_{\min}(z,M)}^{r_{\max}(z)} dr \frac{d^2 P}{dMdr}$$
$$\times \frac{dP}{d\log_{10} L}(\langle L_{\text{LW, fid.}}\rangle, J, r)$$
$$\times \Theta[r - r_s(z, M, t_{\text{ff}}(z))], \tag{9}$$

$$\frac{dP_{\text{Ren.}}}{d\log_{10} J}(z, J) = \int_{M_a}^{M_{b,\text{Ren.}}} dM \int_{r_{\min}(z,M)}^{r_{\max}(z)} dr \frac{d^2 P}{dMdr}$$
$$\times \frac{dP}{d\log_{10} L}(\langle L_{\text{LW, Ren.}}\rangle, J, r).$$

Here $\frac{d^2 P}{dMdr}(z, M, r)$ (units: h^2 $M_\odot^{-1}$ Mpc^{-1}) is the probability density used to count the number of neighbouring halos in a given mass-separation bin, $\frac{dP}{d\log_{10} L}(\langle L_{\text{LW}}(z, M)\rangle, J, r)$ is the probability density used to assign a LW luminosity density to each halo and we use the Heaviside function $\Theta[r - r_s(z, M, t_{\text{ff}}(z))]$ to account for metal pollution via supernova outflows from a neighbouring halo in the fiducial model. Now we define each term in the integral over M and r below.

2.5.1. *Counting the Neighbouring Halos*

If the neighbouring halos were uniformly distributed in space, then the probability of finding a neighbouring halo with halo mass $[M, M + dM/2]$ and physical separation $[r, r + dr/2]$ from the target halo would be:

$$\frac{d^2 P}{dMdr}(z, M, r)dMdr = 4\pi r^2(1 + z)^3 \tag{10}$$
$$\times \frac{dn_{\text{SMT}}}{dM}(z, M)\, dMdr.$$

Here $4\pi r^2 dr$ is the physical volume of the shell surrounding the target halo, $\frac{dn_{\text{SMT}}}{dM}(z, M)dM$ is the halo number per unit comoving volume and $(1 + z)^3$ converts from it from comoving to physical volume. But this only holds for a comoving separation $r_{\text{co, max}} \gtrsim 100\,\text{cMpc}$.[7] We must additionally account for dark matter clustering leading to deviations from the mean matter density $\bar{\rho}_m$ at close range. We denote this deviation $\delta(z, r_{\text{co}})$ by:

$$1 + \delta(z, r_{\text{co}}) = \frac{\rho_m(z, r_{\text{co}})}{\bar{\rho}_m(z)}. \tag{11}$$

[7] We computed this by finding $r_{\text{co, max}}$ such that $\xi_{\text{mm}} < 0$ for $r > r_{\text{co, max}}$ and halos are no longer correlated.

Rather than computing $\delta(z, r_{\text{co}})$ directly, we find a related quantity: the dimensionless two-point halo-halo correlation function $\xi_{\text{hh}}(z, M_{\min}(z), M, r_{\text{co}})$. This accounts for the excess probability of finding a neighbouring halo of mass M at a comoving separation of r_{co} from our target halo of mass $M_{\min}(z)$, both at redshift z. We can separate the M-dependence using the dimensionless halo bias terms $b(z, M)$ and the r_{co}-dependence using the dimensionless two-point matter-matter correlation function $\xi_{\text{mm}}(z, r_{\text{co}})$ (van den Bosch et al. 2013):

$$\xi_{\text{hh}}(z, M_{\min}(z), M, r_{\text{co}}) \approx b(z, M_{\min}(z))b(z, M)$$
$$\times \xi_{\text{mm}}(z, r_{\text{co}}). \tag{12}$$

The redshift and mass dependence of the halo bias terms $b(z, M)$ are determined by a fitting function developed by Sheth, Mo & Tormen (2001). We can later relate physical and comoving separations by:

$$r = \frac{1}{1 + z}r_{\text{co}}. \tag{13}$$

The quantity $\xi_{\text{mm}}(z, r_{\text{co}})$ is found by taking the inverse Fourier transform of the nonlinear power spectrum $P(z, k)$ in spherical coordinates where we assume the power spectrum is spherically symmetric:

$$\xi_{\text{mm}}(z, r_{\text{co}}) = \frac{1}{2\pi^2} \int_0^{\infty} dk\, k^2 P(z, k) \frac{\sin(kr_{\text{co}})}{kr_{\text{co}}}. \tag{14}$$

We note that the linear power spectrum $P(z, k)$ varies with z as:

$$P(z, k) = d(a(z))^2 P(z = 10, k), \tag{15}$$

where $a(z) = 1/(1 + z)$ and $d(a)$ is the normalised linear growth factor at $z = 10$ (Lukić et al. (2007)):

$$d(a) = \frac{D^+(a)}{D^+(a = 1/(1 + 10))}, \tag{16}$$

$$D^+(a) = \frac{5\Omega_{m,0}}{2} \frac{H(a)}{H_0} \int_0^a \frac{da'}{[a'H(a')/H_0]^3}. \tag{17}$$

Thus we can approximate the redshift dependence of the matter-matter correlation function as:

$$\xi_{\text{mm}}(z, r_{\text{co}}) = d(a(z))^2 \xi_{\text{mm}}(z = 10, r_{\text{co}}). \tag{18}$$

We create a fitting function based on arrays of r_{co} and $\xi_{\text{mm}}(z = 10, r_{\text{co}})$ values. This allows us to find $\xi_{\text{mm}}(z, r_{\text{co}})$ using Eq. 18, rather than computing it via integration as in Eq. 14 which is much more computationally heavy.

For brevity, we shall refer to $\xi_{\text{hh}}(z, M_{\min}(z), M, r_{\text{co}})$ as ξ_{hh}. Finally, the probability of finding a neighbouring halo with halo mass $[M, M + dM/2]$ and physical separation $[r, r + dr/2]$ is given by:

$$\frac{d^2 P}{dMdr}(z, M, r)\, dMdr = 4\pi r^2(1 + z)^3 \tag{19}$$
$$\times \frac{dn_{\text{SMT}}}{dM}(z, M)[1 + \xi_{\text{hh}}]\, dMdr.$$

For the fiducial model, we integrate over the halo mass range $[M_a, M_{b,\text{fid.}}] = [M_{\min}(z), 10^{15}\, M_\odot]$. For the model informed by `Renaissance` data, we integrate over the range $[M_a, M_{b,\text{Ren.}}] = [M_{\min}(z), 10^9\, M_\odot]$ since no halos

were found above this mass range for the redshifts considered ($11.6 \leq z \leq 18.6$). We define the minimum physical separation, $r_{\min}(z, M)$, as $r_{\min}(z, M) = 2r_{\mathrm{vir}}(z, M)$ where r_{vir} is the virial radius of the neighbouring halo (Johnson (2012)). This avoids a neighbouring halo overlapping with the target halo.

$$r_{\mathrm{vir}}(z, M) = (7.84 \times 10^{-4})\Big(\frac{M}{10^8\,\mathrm{h}^{-1}\mathrm{M}_\odot}\Big)^{1/3} \Omega_{\mathrm{m},0}^{-1/3}$$
$$\times \Big(\frac{1+z}{10}\Big)^{-1} \mathrm{Mpc/h}. \tag{20}$$

We set $r_{\max}(z) = r_{\mathrm{co,\,max}}/(1+z)$ as the maximum physical separation. We found $r_{\mathrm{co,\,max}} \approx 119.448$ cMpc/h, which is the maximum comoving separation where $\xi_{\mathrm{mm}}(z = 10.0, r_{\mathrm{co}}) > 0$. Beyond this value, a distant halo would no longer be correlated with the target halo.

2.5.2. *Assigning Lyman-Werner Luminosity Density*

The probability density that a neighbouring halo of mass M (units: $\mathrm{h}^{-1}\,\mathrm{M}_\odot$ and physical separation r (units: Mpc/h) has a LW luminosity density L is given as:

$$\frac{dP}{d\log_{10} L}(\langle L_{\mathrm{LW}}\rangle, J, r) = \frac{1}{\sigma_{\mathrm{LW}}\sqrt{2\pi}} \tag{21}$$
$$\times \exp\left[\frac{-(x-\mu)^2}{2\sigma_{\mathrm{LW}}^2}\right],$$

where $\sigma_{\mathrm{LW}} = 0.4$, $x = \log_{10} L$ (L in units of 10^{26} erg s^{-1} Hz^{-1}), $L = 16\pi^2 r^2 J$ and $\mu = \log_{10}\langle L_{\mathrm{LW}}(z, M)\rangle$ (adapted from Eq. 3 of Dijkstra et al. (2008)). A LW luminosity density (and LW flux) is assigned to each halo such that they follow a lognormal distribution in r with a mean LW luminosity density $\langle L_{\mathrm{LW}}(z, M)\rangle$. In our fiducial model, we adapt the model used by D08 and D14.

To compute $\langle L_{\mathrm{LW,\,fid.}}(z, M)\rangle$, we must first consider the mean LW photon production rate $\langle Q(t)\rangle$ (units: h s^{-1} M$_\odot^{-1}$):

$$\langle Q(t)\rangle = (Q_0)[1 + (t_6/4)]^{-3/2} e^{-t_6/300}, \tag{22}$$

where $Q_0 = (10^{47}/\mathrm{h})$ photons h s^{-1} M$_\odot^{-1}$ and $t = (t_6)(10^6\,\mathrm{yr})$. This quantity is derived from the Starburst99 population synthesis model developed by Leitherer et al. (1999) while assuming that star formation occurs with a Salpeter initial mass function.

The mean LW luminosity density $\langle L_{\mathrm{LW,\,fid.}}(t, M)\rangle$ (units: erg s^{-1} Hz^{-1}) of a halo of mass M at a time t after star formation begins is given as:

$$\langle L_{\mathrm{LW,\,fid.}}(t, M)\rangle = \frac{h_{\mathrm{P}}\langle \nu\rangle}{\Delta\nu}\langle Q(t)\rangle f_{\mathrm{esc}} f_* \frac{\Omega_{\mathrm{b},0}}{\Omega_{\mathrm{m},0}} M, \tag{23}$$

where h_{P} is Planck's constant, $\langle \nu\rangle$ is the mean LW frequency, $\Delta\nu$ is the LW frequency range, $f_{\mathrm{esc}} = 1$ is the LW photon escape fraction and $f_* = 0.05$ is the star formation efficiency.

For possible heavy seed formation, it is imperative that the gas within the target halo fully collapses before it may cool and fragment to form stars. It must receive supercritical LW flux from its neighbouring halos for the
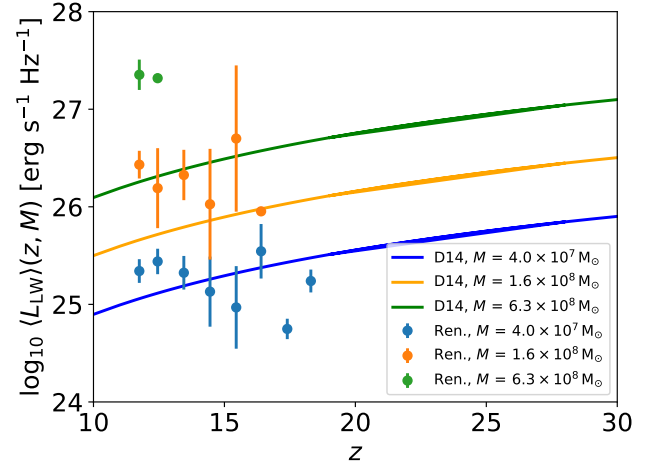

FIG. 6.— Mean LW luminosity density vs. redshift for a number of halo masses. This is the luminosity density, in units of erg s^{-1} Hz^{-1}, emitted by halos with the masses shown in the legend. The mean LW luminosity density in the fiducial model is given by Eq. 23. This quantity increases as time decreases and it is evaluated at the free-fall time. The free-fall time decreases as redshift increases, thus the mean luminosity increases as redshift increases. We are evaluating the luminosity when the stars are younger (and hence more luminous) at higher redshift. We also plot data points from the Renaissance simulation suite with the mean LW luminosity emitted from halos of a given mass as a function of z. In this case the mean LW luminosity is almost flat with a small decrease (i.e. lower star formation efficiencies) seen for the lowest mass halos. The Renaissance data points shown are the median values within a redshift bin, with the error bars being the interquartile range.

duration of its free-fall time $t_{\mathrm{ff}}(z)$. The free-fall time is given as:

$$t_{\mathrm{ff}}(z) = \sqrt{\frac{3\pi}{32G\rho(z)}} \sim 83\left[\frac{1+z}{11}\right]^{-\frac{3}{2}} \mathrm{Myr}, \tag{24}$$

where the density of a halo $\rho(z) \approx 200\rho_{\mathrm{m}}(z)$ and we assume that star formation begins in all neighbouring halos simultaneously. Since $\langle L_{\mathrm{LW,\,fid.}}(t, M)\rangle$ monotonically decreases as t increases, if the target halo receives supercritical LW flux at $t = t_{\mathrm{ff}}(z)$, then it has received supercritical LW flux at $t < t_{\mathrm{ff}}(z)$.

Therefore we evaluate $\langle L_{\mathrm{LW,\,fid.}}(t, M)\rangle$ at $t = t_{\mathrm{ff}}(z)$, making it strictly a function of z and M only. Figure 6 depicts how $\langle L_{\mathrm{LW,\,fid.}}\rangle$ varies with z for three different values of M. The different values of M are marked as solid lines and vary from 4×10^7 M$_\odot$ up to 6×10^8 M$_\odot$. What we see is that for the analytic model the mean luminosity increases with redshift - primarily driven by the dependence on the free-fall times which depends on the redshift z.

To test the physicality of this model we again appeal to Renaissance and plot the mean LW luminosity $\langle L_{\mathrm{LW,\,Ren.}}\rangle$ as a function of halo mass and redshift from Renaissance. When calculating the LW flux from Renaissance halos we first determine the stellar mass of that halo and from that calculate the mean LW flux that is produced by that stellar mass according to the Renaissance model. While the data for the largest halo masses is relatively sparse (green data points), the data for the smaller halos is well sampled. In this case we see a relatively flat (slightly decreasing) LW luminosity as a function of redshift. For the lowest mass halos we
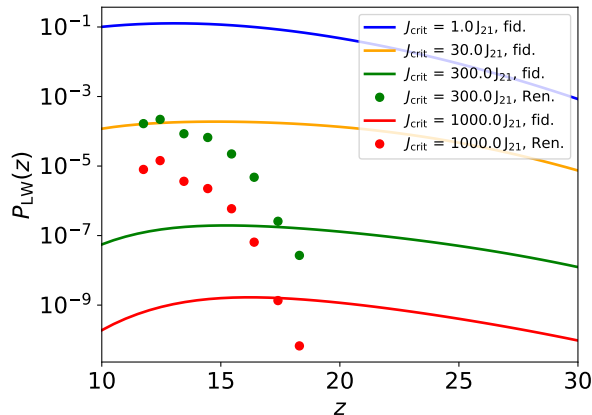
FIG. 7.— The supercritical probability vs. redshift. The probability of a halo receiving a super-critical LW flux is given by the y-axis. Line colours refer to values of $J_{\rm crit}$. As expected, the probability of a halo receiving a high flux (e.g. $J_{\rm crit} \geq 300$ J$_{21}$) is low. The solid lines are from the analytic model. The points are from the Renaissance-informed models. For these models, the probability of a halo receiving a super-critical flux drops sharply with redshift - more in line with expectations.

would expect a slightly lower star formation efficiency and hence a lower mean LW luminosity. However, what we do not see, and is a limitation of the analytic model, is a mean LW luminosity which increases with redshift over this range. This divergence of the analytic and hydrodynamical models will feed into our results. Having now introduced the methodology behind our analysis we now present our results.

3. RESULTS

As previously stated, the aim of this work is firstly to reproduce and verify the results from D14 and secondly to compare the results of this analysis against the results of other numerical experiments from the literature. We augment this goal by also taking advantage of the Renaissance suite of simulations and use some of the relevant Renaissance data in the analytic models. Although the Renaissance suite cannot capture the rare halos that experience super-critical LW radiation at the values thought necessary to produce heavy seeds (i.e. $J_{\rm crit} \gtrsim 300$ J$_{21}$), it is nonetheless an important check on the self-consistency of the analytic model, particularly the probability of finding pristine halos $P_{\rm pristine}(z)$ as a function of redshift and the mean LW luminosity density $\langle L_{\rm LW}(z, M) \rangle$ as a function of halo mass and redshift.

3.1. *Supercritical Probability*

In previous works, $J_{\rm crit}$ was chosen to be $30 - 300$ J$_{21}$ for a T $= 10^4$ K blackbody spectrum (see Shang, Bryan & Haiman (2010)) and 1000 J$_{21}$ for a T $= 10^5$ K spectrum (see Wolcott-Green, Haiman & Bryan (2011)). D14 chose $J_{\rm crit} = 300$ J$_{21}$ as an intermediate value between these two spectra. In this work we follow D14 and choose $J_{\rm crit} = 300$ J$_{21}$ as a critical threshold for the LW pathway. We also investigate solutions with $J_{\rm crit} = 1.0, 30.0$ and 1000.0 J$_{21}$.

In Figure 7, we plot the probability of a target halo receiving a supercritical flux at different redshifts. The probability is plotted for a range of different values of the critical flux, $J_{\rm crit}$, from $J_{\rm crit} = 1.0$ J$_{21}$ up to $J_{\rm crit} = 1000.0$ J$_{21}$. Solid lines are from the analytic model. As

expected the probability of a target halo receiving a flux in excess of $J = 1.0$ J$_{21}$ is very high. In this case by a redshift of $z = 10$, approximately 1 in 10 halos meeting the target halo criteria will receive a flux greater than or equal to $J = 1.0$ J$_{21}$. However, the probability of a target halo receiving significantly higher fluxes is much less. For example the probability of a target halo receiving a flux in excess of $J = 1000.0$ J$_{21}$ at $z = 10$ is less than 1 in 10^9. Note that in this model, the peak in terms of probabilities occurs at $z \sim 15$ and decreases at higher and lower redshifts.

It should also be noted that the probability of a halo receiving a supercritical flux does not decrease as rapidly as perhaps expected towards very high redshift (i.e. $z \geq 20$). This is an inherent characteristic of the fiducial model. The model requires that a target halo is illuminated by a super-critical flux for a free-fall time (see §2.5.2). While the number of star-forming halos decreases as per the halo mass function, the free-fall time shrinks dramatically with redshift. The two effects cancel each other out somewhat and hence the probability of receiving a super-critical flux does not decrease as redshift increases as rapidly as expected.

Overplotted in Figure 7 are the results when applying information from the Renaissance datasets (median values of each redshift bin). In this case the mean LW values $\langle L_{\rm LW,\ Ren.}(z, M) \rangle$ are used instead of the analytic model values $\langle L_{\rm LW,\ fid.}(z, M) \rangle$ and we exclude the Heaviside function from the first line of Eq. 9, i.e. we do not account for the influence of metal pollution in this plot but we do in Figure 8. For the simulation values, we see a much steeper decline in the probability of a halo receiving a super-critical flux (driven mainly by the mean LW luminosity density). Hence, for redshifts approaching $z \sim 20$ the probability of a halo receiving a supercritical flux is negligible. This is in comparison to the analytic model where the probabilities only decline slowly.

3.2. *Heavy Seed Number Density*

Putting everything together, we now plot the number density of halos acting as heavy seed formation sites as a function of redshift in Figure 8. In total we plot 10 datasets. With the exception of the black dots, all dotted datasets are the median values of a redshift bin. These datasets include: our fiducial model with $J_{\rm crit} = 300$ J$_{21}$ (green line) and $J_{\rm crit} = 1000$ J$_{21}$ (red line); our Renaissance-augmented model with $J_{\rm crit} = 300$ J$_{21}$ (green dots) and $J_{\rm crit} = 1000$ J$_{21}$ (red dots); the heavy seed number density from Trinca et al. (2022) (solid blue line with circular markers); the DCBH number density from Chiaki et al. (2023) (solid magenta line with circular dots); the heavy seed formation site number density from McCaffrey et al. (2025) (orange dots); and the heavy seed formation site number density from D14 with $J_{\rm crit} = 300$ J$_{21}$ (black dots).

We also include an AGN number density estimate from Inayoshi (2025). This model is based on the observed and inferred abundance of LRDs, suggesting that LRDs are a phase in galaxy evolution. The brown dashed line in Figure 8 shows this fit for $6 \leq z \leq 10$. The brown dots are the inferred AGN number densities based on recent observations for $6.5 \leq z \leq 8.5$ from Kokorev et al. (2024) and Kocevski et al. (2024). The fit from
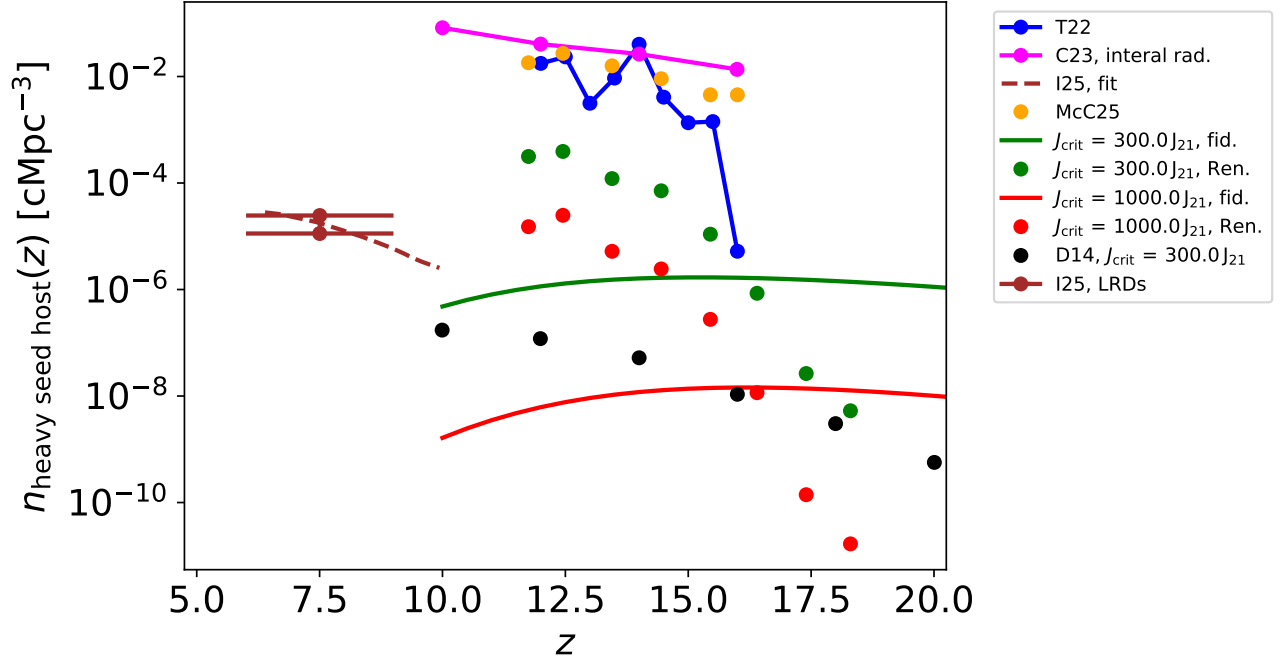
FIG. 8.— Number density of heavy seeds vs. redshift. Black dots are original points taken from D14, green and red lines show our use of the same analytic model methodology as outlined in D14, the green and red dots use `Renaissance`-informed data as part of the analytic model, orange dots are from McCaffrey et al. (2025), the solid blue line with circular markers is from Trinca et al. (2022), the solid magenta line with circular markers is from Chiaki et al. (2023) (with internal radiation), and the brown lines are from Inayoshi (2025) (dashed line for the AGN number density fit; dots for the LRD number density estimate based on observations). The LW-only channels models (solid lines, black dots, green and red dots) are unable or only very marginally able to reproduce recent JWST high-z AGN number density of $\gtrsim 10^{-4}$ cMpc^{-3}. On the other hand, the models of McCaffrey et al. (2025) and Trinca et al. (2022) produce higher number densities and are thus more compatible with recent JWST observations.

Inayoshi (2025) was normalised such that the AGN number density matches the observed LRD number density of $\approx 3 \times 10^{-5}$ cMpc^{-3} for $z \sim 4 - 7$. We include these estimates to provide context for our heavy seed datasets but we must note that a heavy seed number density cannot be directly compared to these AGN number density estimates. These AGN number densities which are informed by recent observations likely account for only the most active of AGN, while in reality many AGN will be quiescent and thus making the total AGN number density larger. We also note that only a fraction of heavy seeds will undergo efficient growth to reach SMBH masses by $z \sim 6$. Thus in Figure 8, if a heavy seed number density estimate at $z \sim 10$ is equal to or greater than this AGN number density estimate, that pathway may be viable to account for SMBH number densities. If a heavy seed number density estimate is less than the AGN number density estimate, that pathway can at best only account for a fraction of SMBHs.

For our fiducial models (green and red lines), we are unable to match results from D14 (black dots) exactly despite following their methodology and our results deviate from theirs, particularly at high redshift. Without access to their code base we cannot determine where the discrepancy arises. Our codebase and pipeline will be publicly available on GitHub but until then may be accessed upon request.

Both the results from our fiducial model and the results of D14 do however agree that the number density of heavy seed black holes (or indeed heavy seed hosting halos) is less that 10^{-6} cMpc^{-3} at $z \gtrsim 10$. With the

number density of MBH hosting galaxies at least two orders of magnitude greater than this (and potentially likely much higher) at $z \gtrsim 4$ (e.g. Pérez-González et al. 2024; Inayoshi 2025), the LW channel is likely unable to explain the high abundances of MBHs in the early Universe based on our fiducial model. This is the first takeaway from our analysis.

We plot the results using values obtained from the `Renaissance`-informed analytic models as green ($J_{\rm crit} = 300\,{\rm J_{21}}$) and red dots ($J_{\rm crit} = 1000\,{\rm J_{21}}$). In this case the results are more encouraging - albeit with a steeper decline. Note that we are not using the `Renaissance` output to make predictions but use their mean LW luminosity density values $\langle L_{\rm LW, \, Ren.}(z, M) \rangle$ and the impact of genetic and external metal pollution is informed by their $P_{\rm pristine, \, Ren.}(z)$ function (see Figure 6 and Figure 5). We see that the number density of heavy seed hosting halos peaks at approximately 10^{-4} cMpc^{-3} at $z \sim 12.5$ for $J_{\rm crit} = 300\,{\rm J_{21}}$. These numbers are on the face of it consistent, albeit marginally, with the recent results from JWST (see AGN number density estimates from Inayoshi (2025) (brown dots and brown dashed line). However, our models are for the seeds and not the candidate AGN detected by JWST. Given the growth requirements of the seeds combined with the expected duty cycle of AGN, these number densities are still likely incompatible with current JWST observations.

We also plot more recent results by Trinca et al. (2022) (solid blue line with circular markers), Chiaki et al. (2023) (solid magenta line with circular dots) and McCaffrey et al. (2025) (orange dots). McCaffrey et al. (2025)

used the `Renaissance` simulations to analyse the formation and later the growth of heavy seed black holes while accounting for rapid assembly and merger history of halos. We do not account for growth in this comparison and only consider their number density of heavy seed formation sites. They used the following criteria when identifying halos that may host massive black holes:

- $T_{vir} \gtrsim 10^4$ K,

- $Z < 10^{-3} Z_{\odot}$,

- $M(r < 0.5R_{vir})/M(r < R_{vir}) > 0.5$,

- $\dot{M} > 0.1 M_{\odot} yr^{-1}$.

The results from McCaffrey et al. (2025) show that the formation of heavy seed MBHs via the so-called rapid assembly channel results in significantly higher number densities with values of approximately 10^{-2} cMpc^{-3} or higher. These results from McCaffrey et al. (2025) are consistent with recent JWST results (again with the caveat that the growth of seeds and the duty cycle of AGN will push these number densities downwards).

Trinca et al. (2022) performed a similar analysis to McCaffrey et al. (2025) using the COSMIC ARCHAEOLOGY TOOL (CAT), a semi-analytic model which allowed them to follow the formation of the first stars and black holes while accounting for accretion and mergers. They also account for how star formation in mini-halos can be affected by LW flux from nearby highly star-forming galaxies. This inhibits H_2 cooling and potentially sterilises these halos until they reach the atomic-cooling regime. In their model, if a dark matter halo meets the following conditions:

1. $T_{vir} \gtrsim 10^4$ K,

2. $Z < 10^{-3.8} Z_{\odot}$,

3. $J_{LW} > 300 J_{21}$,

then a heavy seed black hole with $M_{BH} = 10^5$ M$_{\odot}$ is set at the centre of the galaxy within the halo. Their data in Figure 8 indicates the number density of newly-formed heavy seeds at a given redshift. Their results show similar number densities to McCaffrey et al. (2025) with peak values 10^{-2} cMpc^{-3}. Their model shows number densities significantly beyond the LW-only channel, again compatible with more recent results from JWST.

When computing the number density of DCBHs, Chiaki et al. (2023) consider both external and internal LW radiation, rather than only external sources like we do. They make a direct comparison to Trinca et al. (2022) by considering halos with $T_{vir} > 10^4$ K, $J_{crit} = 300$ J$_{21}$ and $Z < 10^{-3.8}$ Z$_{\odot}$. Internal radiation from a single halo depends on its LW photon emissivity (from both Pop III and Pop II stars) and the radius of the galaxy hosted by that halo (where the galaxy radius is 10% of the halo virial radius). External radiation received by a single halo is computed from the total LW flux from all other halos and is dependent on LW photon emissivity and the distance between the target and source halos. Metal pollution is accounted for via chemical evolution equations of Pop I/II stars and SN winds. Their highest number density value is almost 10^{-1} cMpc^{-3} at $z = 10$.

Additionally, Bhowmick et al. (2021) explored a similar LW channel with the inclusion of low gas angular momentum to track SMBH seed formation and Bhowmick et al. (2024) used the `BRAHMA` simulations to track DCBH formation. Both works found that more efficient heavy seeding channels may be necessary to account for the most massive black holes at high redshift.

In summary, Figure 8 tells us that both D14 and our re-implementation of the D14 model show results which are incompatible with recent JWST data. Other pathways investigated by McCaffrey et al. (2025), Trinca et al. (2022) and Chiaki et al. (2023) appear more promising to explain the overall MBH population. A similar result was shown by Bhowmick et al. (2021).

4. DISCUSSION AND CONCLUSIONS

In this paper, we have reviewed the analytic model of Dijkstra, Ferrara & Mesinger (2014) in terms of calculating the number density of MBHs that can be formed through the so-called Lyman-Werner (LW) channel. In this framework, a super-critical flux of LW irradiates a target halo. The target halo must have a mass exceeding the atomic-cooling threshold and must be metal-free. The super-critical flux required can vary from halo to halo but is likely to be excess of 300 J$_{21}$. Such a high value can only be produced by a nearby neighbouring halo. A weakness of this model is assuming that there is no correlation between a halo being metal-free and receiving a super-critical LW flux i.e. we multiply the probabilities $P_{pristine}(z)$ and $P_{LW}(z, J_{crit})$. More realistically, a super-critical LW flux may suppress star formation in a progenitor halo and thus the target halo has an increased probability of being metal-free. Such an analysis of the merger history of the target halos is beyond the scope of this study. In agreement with D14, we find that the number density of target halos receiving a critical flux in excess of 300 J$_{21}$ is approximately 10^{-6} cMpc^{-3} at $z = 10$ (see Figure 8). The number densities drop, as expected, towards higher redshifts.

Despite considerable effort, we were unable to reproduce the exact results of D14 and our fiducial model results differ from the D14 results at $z \gtrsim 20$. We varied the cosmology implemented, the integration limits for neighbouring halo mass M and separation r, how the halo-halo correlation was computed and the definition of the mean LW luminosity density. Our analysis tools and pipeline will be available on GitHub. Our analysis, following the methodology of D14, shows that the number density of heavy seeds is almost constant out to very high redshift ($z \gtrsim 20$) - this is primarily due to how the mean LW luminosity $\langle L_{LW, fid.}(z, M) \rangle$ produced by a halo is calculated in the D14 model. Nonetheless, our analysis agrees very well with the D14 as we approach $z = 10$. However, in both cases the number densities remain close to or below 10^{-6} cMpc^{-3}.

To check the physical consistency of the D14 model, we augment the analytic model with information taken directly from the `Renaissance` simulation suite. Specifically, we take data of the mean LW luminosity density $\langle L_{LW, Ren.}(z, M) \rangle$ produced as a function of redshift and halo mass and the pristine fraction $P_{pristine, Ren.}(z)$ of halos as a function of redshift (see Figures 6 and 5). Using this augmented model, we find that number density of heavy seed hosting halos increases steeply

between $z \sim 18$ and $z \sim 10$. The peak number density of heavy seed hosting halos reaching values close to 10^{-4} cMpc^{-3} at $z \sim 10$. The Renaissance-informed models are nonetheless still likely incompatible with the recent results on AGN fractions at high-redshift (e.g. Pérez-González et al. 2024; Greene et al. 2024; Inayoshi 2025) given subsequent growth requirements of the seeds combined with the expected duty cycle of AGN.

A slight weakness of this model is the lack of incorporation of baryonic matter streaming velocities (Tseliakhovich & Hirata 2010). This would suppress star formation in halos with masses $\lesssim 10^6$ M$_\odot$ (Tseliakhovich, Barkana & Hirata 2011; O'Leary & McQuinn 2012; Xu et al. 2014). This is an order of magnitude below the M$_{\rm min}$ values we consider, so this may not affect our results greatly. It could push $P_{\rm pristine, Ren.}(z)$ to higher values since star formation would be suppressed in more progenitor halos, leading to more target halos avoiding genetic metal pollution.

We have shown through our heavy seed number density predictions that the LW-only channel is likely subdominant compared to other channels when accounting for AGN number density estimates based on recent observations. While these observations are still hotly debated and the exact make-up of the JWST galaxies unclear, even if some fraction of the galaxies host AGN (as is strongly suspected) then the LW-only channel cannot be responsible - the predicted number densities are simply too low.

REFERENCES

Abbott B. P. et al., 2019, ApJ, 882, L24

Agarwal B., Khochfar S., 2015, MNRAS, 446, 160

Agarwal B., Smith B., Glover S., Natarajan P., Khochfar S., 2016, MNRAS, 459, 4209

Alvarez M. A., Wise J. H., Abel T., 2009, ApJ, 701, L133

Andika I. T. et al., 2022, AJ, 163, 251

Bañados E. et al., 2018, Nature, 553, 473

Barkana R., Loeb A., 2001, Physics Reports, 349, 125

Bhowmick A. K., Blecha L., Torrey P., Kelley L. Z., Vogelsberger M., Nelson D., Weinberger R., Hernquist L., 2021, Monthly Notices of the Royal Astronomical Society, 510, 177–196

Bhowmick A. K. et al., 2024, MNRAS, 533, 1907

Brummel-Smith C. et al., 2019, The Journal of Open Source Software, 4, 1636

Bryan G. L., Norman M. L., O'Shea B. W., Abel T., Wise J. H., Turk M. J., The Enzo Collaboration, 2014, ApJS, 211, 19

Chen P., Wise J. H., Norman M. L., Xu H., O'Shea B. W., 2014, ApJ, 795, 144

Chiaki G., Chon S., Omukai K., Trinca A., Schneider R., Valiante R., 2023, MNRAS, 521, 2845

Dijkstra M., Ferrara A., Mesinger A., 2014, MNRAS, 442, 2036

Dijkstra M., Haiman Z., Mesinger A., Wyithe J. S. B., 2008, Monthly Notices of the Royal Astronomical Society, 391, 1961

Farr W. M., Sravan N., Cantrell A., Kreidberg L., Bailyn C. D., Mandel I., Kalogera V., 2011, ApJ, 741, 103

Fernandez R., Bryan G. L., Haiman Z., Li M., 2014, MNRAS, 439, 3798

Galli D., Palla F., 1998, A&A, 335, 403

Greene J. E. et al., 2024, ApJ, 964, 39

Greene J. E., Strader J., Ho L. C., 2020, ARA&A, 58, 257

Habouzit M., Volonteri M., Latif M., Dubois Y., Peirani S., 2016, MNRAS, 463, 529

Harikane Y. et al., 2023, ApJ, 959, 39

Heger A., Fryer C. L., Woosley S. E., Langer N., Hartmann D. H., 2003, ApJ, 591, 288

Hirano S., Hosokawa T., Yoshida N., Kuiper R., 2017, Science, 357, 1375

Inayoshi K., 2025, arXiv e-prints, arXiv:2503.05537

Inayoshi K., Visbal E., Haiman Z., 2020, ARA&A, 58, 27

Izumi T. et al., 2021, ApJ, 914, 36

Johnson J. L., 2012, in The First Galaxies, Springer Berlin Heidelberg, pp. 177–222

Kocevski D. D. et al., 2024, arXiv e-prints, arXiv:2404.03576

Kokorev V. et al., 2024, ApJ, 968, 38

Kokorev V. et al., 2023, ApJ, 957, L7

Kreidberg L., Bailyn C. D., Farr W. M., Kalogera V., 2012, ApJ, 757, 36

Kulkarni M., Visbal E., Bryan G. L., 2021, ApJ, 917, 40

Lambrides E. et al., 2024, arXiv e-prints, arXiv:2409.13047

Larson R. L. et al., 2023, ApJ, 953, L29

Latif M. A., Bovino S., Van Borm C., Grassi T., Schleicher D. R. G., Spaans M., 2014, MNRAS, 443, 1979

Latif M. A., Niemeyer J. C., Schleicher D. R. G., 2014, MNRAS, 440, 2969

Latif M. A., Whalen D. J., Khochfar S., Herrington N. P., Woods T. E., 2022, Nature, 607, 48

Leitherer C. et al., 1999, ApJS, 123, 3

Li J. et al., 2025, ApJ, 981, 19

Lukić Z., Heitmann K., Habib S., Bashinsky S., Ricker P. M., 2007, ApJ, 671, 1160

Lupi A., Haardt F., Dotti M., Fiacconi D., Mayer L., Madau P., 2016, MNRAS, 456, 2993

Lupi A., Haiman Z., Volonteri M., 2021, MNRAS, 503, 5046

Ma Y. et al., 2025, ApJ, 981, 191

Madau P., Haardt F., Dotti M., 2014, ApJ, 784, L38

Madau P., Rees M. J., 2001, ApJ, 551, L27

Maiolino R. et al., 2024, A&A, 691, A145

Matsuoka Y. et al., 2019, ApJ, 872, L2

Mayer L., Capelo P. R., Zwick L., Di Matteo T., 2024, ApJ, 961, 76

Mayer L., Fiacconi D., Bonoli S., Quinn T., Roškar R., Shen S., Wadsley J., 2015, ApJ, 810, 51

Mayer L., Kazantzidis S., Escala A., Callegari S., 2010, Nature, 466, 1082

McCaffrey J., Regan J., Smith B., Wise J., O'Shea B., Norman M., 2025, The Open Journal of Astrophysics, 8, 11

Mehta D., Regan J. A., Prole L., 2024, The Open Journal of Astrophysics, 7, 107

Mortlock A., Conselice C. J., Bluck A. F. L., Bauer A. E., Grützbauch R., Buitrago F., Ownsworth J., 2011, MNRAS, 413, 2845

Murray S., Diemer B., Chen Z., Neuhold A., Schnapp M., Peruzzi T., Blevins D., Engelman T., 2021, Astronomy and Computing, 36, 100487

Murray S. G., Power C., Robotham A. S. G., 2013, Astronomy and Computing, 3, 23

Naoz S., Yoshida N., Gnedin N. Y., 2012, ApJ, 747, 128

Naoz S., Yoshida N., Gnedin N. Y., 2013, ApJ, 763, 27

O'Leary R. M., McQuinn M., 2012, The Astrophysical Journal, 760, 4

O'Shea B. W., Wise J. H., Xu H., Norman M. L., 2015, ApJ, 807, L12

O'Brennan H., Regan J. A., Power C., Ward S., Brennan J., McCaffrey J., 2024, The Open Journal of Astrophysics, 7

Pérez-González P. G. et al., 2024, ApJ, 968, 4

Press W. H., Schechter P., 1974, ApJ, 187, 425

Regan J., 2023, The Open Journal of Astrophysics, 6, 12

Regan J., Volonteri M., 2024, The Open Journal of Astrophysics, 7, 72

Regan J. A., Visbal E., Wise J. H., Haiman Z., Johansson P. H., Bryan G. L., 2017, Nature Astronomy, 1, 0075

Regan J. A., Wise J. H., Woods T. E., Downes T. P., O'Shea B. W., Norman M. L., 2020, The Open Journal of Astrophysics, 3, 15

Schauer A. T. P., Glover S. C. O., Klessen R. S., Clark P., 2021, MNRAS, 507, 1775

Schauer A. T. P., Regan J., Glover S. C. O., Klessen R. S., 2017, MNRAS, 471, 4878

Shang C., Bryan G. L., Haiman Z., 2010, Monthly Notices of the Royal Astronomical Society, 402, 1249

Sheth R. K., Mo H. J., Tormen G., 2001, Monthly Notices of the Royal Astronomical Society, 323, 1

Shi Y., Kremer K., Hopkins P. F., 2024, A&A, 691, A24

Smith B. D., Regan J. A., Downes T. P., Norman M. L., O'Shea B. W., Wise J. H., 2018, MNRAS, 480, 3762

Tanaka T. L., Li M., 2014, MNRAS, 439, 1092

Trenti M., Stiavelli M., 2007, ApJ, 667, 38

Trenti M., Stiavelli M., 2009, ApJ, 694, 879

Trinca A., Schneider R., Valiante R., Graziani L., Zappacosta L., Shankar F., 2022, MNRAS, 511, 616

Tseliakhovich D., Barkana R., Hirata C. M., 2011, MNRAS, 418, 906

Tseliakhovich D., Hirata C., 2010, Phys. Rev. D, 82, 083520

van den Bosch F. C., More S., Cacciato M., Mo H., Yang X., 2013, MNRAS, 430, 725

Venemans B. P. et al., 2020, ApJ, 904, 130

Wang F. et al., 2021, ApJ, 907, L1

Wise J. H., Regan J. A., O'Shea B. W., Norman M. L., Downes T. P., Xu H., 2019, Nature, 566, 85

Wolcott-Green J., Haiman Z., Bryan G. L., 2011, MNRAS, 418, 838

Xu H., Ahn K., Wise J. H., Norman M. L., O'Shea B. W., 2014, ApJ, 791, 110

Xu H., Norman M. L., O'Shea B. W., Wise J. H., 2016, ApJ, 823, 140

Xu H., Wise J. H., Norman M. L., 2013, ApJ, 773, 83

Yang J. et al., 2021, ApJ, 923, 262

Yoshida N., Abel T., Hernquist L., Sugiyama N., 2003, ApJ, 592, 645

Yung L. Y. A., Somerville R. S., Finkelstein S. L., Wilkins S. M., Gardner J. P., 2024, MNRAS, 527, 5929

Zwick L., Mayer L., Haemmerlé L., Klessen R. S., 2023, MNRAS, 518, 2076

This paper was built using the Open Journal of Astrophysics LATEX template. The OJA is a journal which provides fast and easy peer review for new papers in the `astro-ph` section of the arXiv, making the reviewing process simpler for authors and referees alike. Learn more at http://astro.theoj.org.