KICKSTARTING REIONIZATION WITH THE FIRST BLACK HOLES: THE EFFECTS OF SECOND-ORDER PERTURBATION THEORY IN PRE-REIONIZATION VOLUMES

Kelly Holley-Bockelmann1,3, John H. Wise2, and Manodeep Sinha1

1 Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235, USA; kholley@vanderbilt.edu, manodeep.sinha@vanderbilt.edu
2 Center for Relativistic Astrophysics, Georgia Institute of Technology, Atlanta, GA 30332, USA; jwise@physics.gatech.edu

Received 2012 May 28; accepted 2012 November 6; published 2012 November 20

ABSTRACT

We explore structure formation in the dark ages (z ∼ 30–6) using two well-known methods for initializing cosmological N-body simulations. Overall, both the Zel’dovich approximation and second-order Lagrangian perturbation theory (2LPT) are known to produce accurate present-day dark matter halo mass functions. However, since the 2LPT method drives more rapid evolution of dense regions, it increases the occurrence of rare massive objects—an effect that is most pronounced at high redshift. We find that 2LPT produces more halos that could harbor Population III stars and their black hole remnants, and they produce them earlier. Although the differences between the 2LPT and Zel’dovich approximation mass functions are nearly erased by z = 6, this small boost to the number and mass of black holes more than doubles the reionized volume of the early universe. We discuss the implications for reionization and massive black hole growth.

Key words: black hole physics – cosmology: theory – galaxies: high-redshift – intergalactic medium – methods: numerical – radiative transfer

Online-only material: color figures

1. INTRODUCTION

The first generation of stars, so-called Population III stars (Pop III), forms from pristine gas within dark matter halos (DMHs) at very high redshifts (Couchman & Rees 1986; Tegmark et al. 1997). Both one-zone calculations and cosmological simulations suggest a top-heavy Pop III initial mass function (IMF) with an uncertain characteristic mass (e.g., Bromm et al. 1999; Omukai 2000; Abel et al. 2002; Smith et al. 2009). The uncertainty hinges on recent simulations that showed that the fragmentation and multiplicity may occur at protostellar densities (Turk et al. 2009; Clark et al. 2011; Greif et al. 2012), and on the idea that protostellar radiative feedback may limit mass accretion to ∼40 M⊙ (Hosokawa et al. 2011; Stacy et al. 2012). Massive Pop III stellar lifetimes are ∼2–20 Myr, after which they either enrich the intergalactic medium (IGM) with metals through supernovae (Heger et al. 2003) or collapse directly into black holes (BHs) that could grow into the supermassive ones at galactic centers (Madau & Rees 2001; Volonteri et al. 2003; Islam et al. 2003; Tanaka & Haiman 2009; Alvarez et al. 2009; Jeon et al. 2012; Micic et al. 2011). While the Pop III IMF is still uncertain, it is likely that DMHs do host Pop III stars, and a significant fraction produce seed BHs.

To form the first generation of stars, primordial gas collapses within the DMH center via two separate cooling channels—molecular hydrogen and atomic hydrogen. Trenti & Stiavelli (2009) found the minimum DMH mass to form cold dense core through molecular hydrogen cooling:

\[ M_{\text{min}} \simeq 6.44 \times 10^6 \, M_\odot \, J_{21}^{-0.457} \left( \frac{1 + z}{31} \right)^{-3.557} \],

(1)

where \( J_{21} \) relates to the Lyman–Werner flux and \( F_\star \) = 4πJ21 10^{-21} erg s^{-1} cm^{-2}. It is this flux that is responsible for photodissociating molecular hydrogen. To calculate \( F_\star \), we adopt the comoving Lyman–Werner photon density, \( n_{\text{LW}} \), from Trenti & Stiavelli (2009), which assumes stellar photon yields from Schaerer (2003):

\[ n_{\text{LW}}(z) = 7 \times 10^{66} \times 10^{3.3-3.3(1+z)/11} \, \text{Mpc}^{-3}, \]

and

\[ J_{21} = 1.6 \times 10^{-65} \left( \frac{n_{\text{LW}}}{1 \, \text{Mpc}^{-3}} \right) \left( \frac{1 + z}{31} \right)^3. \]

For \( J_{21} \sim 5 \) at \( z = 10 \), \( M_{\text{min}} \sim 5 \times 10^8 \, M_\odot. \)

On the other hand, all DMHs with virial temperatures \( \gtrsim 10^4 \, \text{K} \) trigger efficient atomic hydrogen cooling, independent of the Lyman–Werner background. A virial temperature of \( 10^4 \, \text{K} \) translates into a DMH mass of

\[ M_{\text{vir}} = 7.75 \times 10^6 \, M_\odot \left( \frac{1 + z}{31} \right)^{-1.5}. \]

Regardless of the cooling mechanism, the DMHs that host the first generation of stars are by far the most massive structures at this early epoch. For example, a DMH of \( \sim 10^9 \, M_\odot \) is a \( \sim 5 - \sigma \) peak at \( z = 20 \). These rare and massive peaks are the first collapsing structures in the universe, and are the most likely to be plagued by numerical transients and initial value issues in any cosmological simulation. The purpose of this Letter is to compare the effect on the Pop III era of two different techniques for initializing a cosmological N-body simulation: the Zel’dovich approximation and second-order Lagrangian perturbation theory (2LPT) on the Pop III era. Below, we merely outline the cosmological initialization techniques; see Scoccimarro (1998) for a comprehensive overview.

The particle positions in an N-body simulation are generated from a density spectrum consistent with the cosmic microwave background (CMB) power spectrum, and the particle positions are evolved using a tree code (Press et al. 1992; Springel 2005; Hernquist & Springel 2003). The code then advances the particle positions and velocities to a future redshift using the softened gravity of the N-body code. The code is run until the simulation is complete, and the results are analyzed using the high-redshift galaxy formation code (Hagemann et al. 2001; Wise & Holz 2006). The simulation code is run in parallel on a supercomputer using a large number of processors, and the results are validated using a set of benchmarks.

\(^{3}\) Adjunct at the Department of Physics, Fisk University, Nashville, TN 37208, USA.

\(^{4}\) Note that neutral H\(_1\) in the IGM can reduce the effective LW flux from Equation (2) by an order of magnitude (Haiman et al. 2000); this should not affect our results since we are concerned with the difference between two otherwise identical volumes.
background (CMB). Since $N$-body simulations start at much lower redshifts ($z_{\text{start}} \sim 50–250$), the particle evolution up to the starting redshift is estimated using a displacement field. In the Zel’dovich approximation, the displacement field assumes a linear evolution from recombination to $z_{\text{start}}$ so the particle displacement is determined by the linear overdensity and linear growth factor $D_1$. In 2LPT, the displacement field also includes a second-order growth factor $D_2 \sim -3D_1^2/7$ and accounts for some of the nonlinear density evolution.

Since Zel’dovich approximation is accurate only to the first order, any higher order growing modes will be incorrect. In addition, in Zel’dovich approximation the nonlinear decaying modes, called transients, damp away with $1/a^2$. This implies that the collapse time of true structures in an N-body simulation will be accurate for fewer $e$-folding times when using 2LPT (e.g., Scoccimarro 1998; Crocce et al. 2006; Jenkins 2010). Thus, for same starting redshift, simulations initialized with 2LPT will capture the formation of very high redshift halos more accurately compared to simulations initialized with Zel’dovich approximation. In general, high-$\sigma$ peaks are suppressed in Zel’dovich approximation, and the effect is larger at high redshift. Of course, given enough expansion factors, the transients eventually damp out for all but the most massive DMHs. The standard lore is that a Zel’dovich approximation simulation must only start earlier than 2LPT; indeed, by $z = 1$ there is only a $\sim 10\%$ difference in the DMH number density above $10^{15} M_\odot$ (Crocce et al. 2006). This small difference in the present-day DMH mass function is perhaps why the Zel’dovich approximation has been adopted so widely.

2. SIMULATIONS AND BLACK HOLE MODELING

2.1. Simulations

In any cosmological simulation, the goal is to displace the particles only slightly from their initial positions. The rms displacement, $\Delta_{\text{rms}}$ is

$$\Delta_{\text{rms}}^2 =\frac{4\pi}{3} \int_{k_f}^{k_{\text{Nyq}}} P(k, z_{\text{start}}) \, dk,$$

where $k_f = 2\pi/L_{\text{box}}$ is the fundamental mode, $L_{\text{box}}$ is the simulation box size, $k_{\text{Nyq}} = N/2 \times k_f$ is the Nyquist frequency for an $N^3$ simulation, and $P(k, z_{\text{start}})$ is the power spectrum at the starting redshift, $z_{\text{start}}$. If $\Delta_{\text{rms}}$ is larger than the mean interparticle separation $\Delta_p = L_{\text{box}}/N$, then “orbit crossing” occurs and invalidates the accuracy of the initial conditions. Strictly speaking, imposing $\Delta_{\text{rms}}/\Delta_p = 0.1$ would be ideal in that it makes orbit crossing a $\sim 10\sigma$ event. The challenge in small volume, high-resolution simulations is that $\Delta_p$ is small and the starting redshift must be very high to satisfy this $\Delta_{\text{rms}}/\Delta_p$ constraint. For instance, to satisfy $\Delta_{\text{rms}}/\Delta_p \sim 0.1$, a $10^{-1}$ Mpc, $512^3$ simulation would need $z_{\text{start}} \approx 799$. However, at very high redshifts the matter distribution is quite smooth, and the net force may be dominated by numerical round-off errors (Lukić et al. 2007)—this will suppress small-scale structure. Therefore, we adopt $\Delta_{\text{rms}}/\Delta_p = 0.25$, which sets $z_{\text{start}} = 300$. We note that this criterion is rarely mentioned in the literature. Several small volume, high-resolution simulations appear to have $\Delta_{\text{rms}}/\Delta_p > 1.0$, a clear sign that the initial conditions are not valid. Since canonical Pop III-hosting halos are expected to appear at $z \sim 30$, this $z_{\text{start}}$ allows 10 expansion factors to occur, which reduces transients from the initial conditions in 2LPT by a factor of 100.

We initialize six $512^3$, $10^{-1}$ Mpc, DM-only cosmological volumes of a $\Lambda$CDM universe and evolved them with Gadget-2 (Springel et al. 2001b; Springel 2005) from $z_{\text{start}} = 300$ to $z = 6$. The six simulations directly contrast the initialization technique, because each pair samples the volume identically from the CMB transfer function, and only displaces these identical initial positions to the same starting redshift using either 2LPT or Zel’dovich approximation. Each volume is initialized using WMAP-5 parameters (Komatsu et al. 2009).

Given that we are concerned with the evolution of our volumes at very high redshift, care was taken to integrate the positions and velocities with high accuracy. We adopted many of the tree code parameters of the Coyote Universe (Heitmann et al. 2010). PMGRID, which defines the Fourier grid, is $1024^3$, ASMTH, which defines the split between long- and short-range forces, is $1.5$ times the mesh cell size, RCUT, which controls the maximum radius for short-range forces, is $6.0$ times the mesh cell size, the force accuracy is $0.002$, the integration accuracy is $0.00125$, and the softening is $1/40\Delta_p = 0.5 h^{-1}$ kpc. The particle mass for all simulations is $5.3 \times 10^5 h^{-1} M_\odot$.

2.2. Black Hole Growth and Accretion Luminosity

In post-processing, we identified halos with at least 20 particles using SUBFIND (Springel et al. 2001a) and constructed a halo merger tree (Sinha & Holley-Bockelmann 2012). We seeded $100 h^{-1} M_\odot$ BHs in halos with the mass criteria defined in the previous section. Once a BH seed is sown, we do not allow another BH to form within that halo. The merger tree now allows us to track the growth and merger of BHs as well. We assumed a simple BH growth prescription: Eddington-limited accretion triggered by a major merger, where the BHs merge promptly after DMHs merge, and there is no gravitational wave recoil (see Micic et al. 2006 for more details). While the IMF of Pop III stars and BH growth are still a matter of debate, the goal of this Letter is simply to examine the differences between 2LPT and Zel’dovich approximation-seeded volumes. A different IMF or growth prescription will change absolute BH number densities and masses, and more (or less) aggressive BH growth will result in a difference in the ionized volumes. Hence, we chose the simplest Pop III IMF and BH growth prescription to contrast the two initialization techniques.\footnote{Note that our particle mass does not allow us to resolve all Pop III-hosting halos at the highest redshifts.}

Once the BHs are seeded and accreting, we estimate the ionizing photon luminosity from the growing BH, as well as its effect on the neutral gas in the neighbourhood through radiative transfer calculations. We assume that the BH is fed at the Eddington rate by major mergers, and this presupposes a rich supply of cold gas in these halos—an assumption that has not been observationally verified at such a high redshift. In fact, massive first stars can expel most of the gas in the host halo, starving the remnant BH (Kitayama et al. 2004; Whalen et al. 2004; Alvarez et al. 2009) until a merger with a gas-rich DMH can restock it. In particular, early simulations with $\sim 100 M_\odot$ Pop III stars found that gas expulsion delayed accretion for $10–50$ Myr while the host DMH restocked its gas reservoir (Johnson et al. 2007; Wise & Abel 2008; Wise et al. 2012). Recent simulations predict that lower Pop III stellar masses\footnote{We assume that the radiative feedback from the accreting BH affects its effect on the neutral gas in the neighbourhood through radiative transfer calculations.}
and their radiative feedback can still drive the majority of gas from their host halos. However, most of the gas infall at high redshifts and in rare peaks happen through cold streams—these cold streams present a small solid angle to the stellar photon flux, which should make it more difficult to photoevaporate. Thus, it may be plausible for Pop III DMHs, especially in more massive ones suppressed by LW radiation, to retain a reservoir of gas for future consumption.

Our simulation is collisionless, so we calculate the impact of enhanced BH growth on the neutral gas numerically in post-processing. For an accreting BH, the temperature profile of the accretion disk is

\[
T(r) = T_0 \left( \frac{r_{\text{in}}}{r} \right)^{3/4} \left( 1 - \sqrt{\frac{r_{\text{in}}}{r}} \right)^{1/4},
\]

\[
T_0 = \left( \frac{3GM_{\text{BH}}M_{\text{gas}}}{8\pi\sigma\sigma_{\text{in}} r_{\text{in}}^2} \right)^{1/4}, \tag{6}
\]

where \(M_{\text{BH}}\) is BH mass, \(M_{\text{gas}}\) is gas accretion rate, \(\sigma\) is Stefan–Boltzmann’s constant, and \(r_{\text{in}}\) is the innermost accretion disk radius. The accretion disk luminosity at a given frequency, \(\nu\), can then be computed as

\[
L_{\nu} = 4\pi^2 \int_{r_{\text{in}}}^{\infty} 2\frac{h\nu^3}{c^2} \left[ \exp\left( \frac{h\nu}{kT(r)} \right) - 1 \right]^{-1} r \, dr, \tag{7}
\]

where \(L_{\nu}\) is the emitted luminosity per frequency band, \(c\) is the speed of light, \(k\) is Boltzmann’s constant, and \(h\) is Planck’s constant. The accretion efficiency, \(\eta\), converts the accreted mass \((M_{\text{acc}})\) to energy: \(M_{\text{acc}} c^2 = 1/2 r_s/r_{\text{in}}\), where \(r_s = GM_{\text{BH}}/c^2\). For a Schwarzschild BH \(r_s \approx 6r_g\) with \(\eta \sim 8\%\). In practice, we used \(\eta = 10\%\) and adjusted \(r_{\text{in}}\) accordingly.

To quantify the effect of these BHs on reionization, we conduct radiative transfer calculations with ENZO+MORAY (Wise & Abel 2011). After every snapshot, we accumulate a fraction \(\Omega_{\text{bh}}/\Omega_{M}\) of the mass into a fixed 512³ grid, assuming that all gas is hydrogen. At these scales and resolution, the gas should follow the DM. The density distribution is fixed during the radiative transfer calculation. We treat each accreting BH as a point source, with luminosities described by Equation (7) divided into bins of 13.6–40, 40–100, 100–10³, and 10³–10⁴ eV. Photon packages have the average photon energy in each bin, ~28, 70, 400, and 1200 eV, and they obey periodic boundary conditions. We use a nonequilibrium chemistry model (Anninos et al. 1997) with hydrogen only. We consider Compton cooling and free electron heating by the CMB, radiative losses from atomic cooling in the optically thin limit, and secondary ionizations from high-energy radiation, using the fitting formulae from Shull & van Steenberg (1985).

3. RESULTS

By \(z = 6\), we find that the 2LPT-initialized volume produces 25% more Pop III stars than the identical Zel’dovich approximation volume. This is synonymous with saying that high-mass DMHs collapsed earlier and more often. Indeed, at \(z = 24\), there are three times the number of DMHs above
Figure 4. Response of neutral hydrogen in the universe to the radiation produced by seed BHs in Zel’dovich approximation (left) and 2LPT (right) initialized runs at three times in one realization. Both simulations used the same BH seeding and growth mechanism, and sampled the same initial phases. We used ray-tracing to propagate the BH radiation throughout the box, calculating the electron temperature of the gas and ionization state.

(A color version of this figure is available in the online journal.)

10⁷ M⊙, roughly a 5σ peak, and by z = 20, the difference in the halo mass function above ∼10⁷ M⊙ (∼4−σ) is ∼40%. Note that at z = 6, the difference in the number of newly formed Pop III-star-hosting halos has virtually vanished—this suggests that the simulation has undergone enough e-folding times to damp away transients.

3.1. The Changes in the Seed Black Hole Population

An overall 25% increase in the number of Pop III-hosting halos may not sound huge, but the consequences of this early and generous sowing of Pop III stars on protosupermassive BH growth are more profound. When we grow seed BHs with
the major-merger-driven prescription outlined above, the minor difference in the halo mass function magnifies. Figure 1 shows the BH mass density, $\rho_{\text{BH}}$, as a function of redshift. We calculate a factor of $\sim 2$ difference in $\rho_{\text{BH}}$ by $z = 6$. If we allow a very vigorous, and perhaps unrealistic, BH growth scheme that fuels the BH continuously at the Eddington rate, then $\rho_{\text{BH,2LPT}}$ would be more than 10 times larger than the Zel’dovich approximation. Future space-based gravitational wave observatories may detect these BH inspirals; the gravitational wave signal in this pre-reionization era could feature more numerous and louder sources than predicted, depending on the detector configuration.

3.2. Implications for the Reionization Epoch

Perhaps the most interesting consequence of the 2LPT DMH mass function is the effect of the more numerous and more massive BHs on reionization. Figure 2 shows the probability density of the BH mass function at $z = 6$ over all 2LPT and Zel’dovich approximation initialized volumes. We find that massive BHs ($\log M_{\text{BH}} \gtrsim 5.5$) are $\sim 1.4$ times more common in 2LPT volumes.

Although our simulation is purely collisionless, we can estimate the contribution from the first BHs on reionization. By $z = 6$, we find that 2LPT BHs emit 75% more ionizing photons per hydrogen atom than Zel’dovich approximation BHs. This alone implies that the ionization history will be dramatically different. We next present the results from three-dimensional radiative transfer calculations on these six volumes.

Figure 3 shows the evolution of the total BH luminosity and mass-weighted electron fraction in the Zel’dovich approximation and 2LPT simulations in one realization. At $z > 10$, the total luminosity fluctuates but is roughly equivalent. Afterward, the total luminosity in the 2LPT simulation is consistently $\sim 3$ times greater than Zel’dovich approximation. This results in a larger ionized fraction at $z < 9$, and yields a total ionized fraction of 0.16 and 0.25 for Zel’dovich approximation and 2LPT at $z = 6$, respectively. However, this enhanced contribution to reionization only translates to a change in the Thomson optical depth of $\Delta \tau_e = 1.7 \times 10^{-3}$, if the universe is completely ionized at $z < 6$. In the other two realizations, the ionized fraction increases by 0.12 and 0.09 at $z = 6$.

The density-weighted projections of electron fraction in Figure 4 depict the spatial differences between Zel’dovich approximation and 2LPT at three redshifts. The H II regions around the BHs are generally larger and more abundant in the 2LPT case. Overall, the 2LPT-ionized volume is $\sim 2.5$ times larger, where we define ionized as $x_e > 0.5$. Furthermore, this additional luminosity globally warms the IGM outside these H II to 17,000 K in underdense regions ($\delta < 1$), compared to 14,000 K in Zel’dovich approximation. This IGM global warming by pre-reionization BHs may be important in regulating the formation and growth of later, lower mass BHs (Tanaka et al. 2012).

4. DISCUSSION

We found that pre-reionization simulations are particularly sensitive to the phase-space initialization method, given that the astrophysically interesting DMHs—ones that host the first stars, BHs, and protogalaxies—are all expected to collapse at very high redshifts. This means that the transient errors in the initial positions and velocities have only a few $e$-folding times to damp away before we need to make reliable measurements of the collapsed DMHs. We found that the Zel’dovich approximation underestimated the number and mass of the high-mass DMHs during the dark ages, and this $\sim 25\%$ difference in the extreme end of the halo mass function magnified, increasing the BH mass density $\sim 1.4$ times for $\log M_{\text{BH}} \gtrsim 5.5$ at $z = 6$ (see Figure 2). This difference depends both on the fact that seed BHs form earlier, which allows them to grow for a longer time through mergers and accretion, and also that they are sown more often.

One of the largest effects of this more accurate 2LPT technique is apparent in the reionization history of the early universe. The volume of the universe reionized at $z = 6$ more than doubles, only due to the initialization method. We caution that our main goal was to study differences in our two testbed simulations using simple prescriptions for BH feedback and mass accretion. We neglect cold gas inflows, mechanical feedback, and indeed all gas and reionization physics is treated in post-processing. We are also neglecting a potentially critical effect of DM streaming motions on baryons at $z = 1000$, which will also delay galaxy formation compared to a standard approach (Tseliakhovich & Hirata 2010). While the actual reionized volume of the universe is not robust here, the critical point is that both Zel’dovich approximation and Press–Schechter approaches underestimate the volume merely because they both assume a linear evolution of DM overdensities.

We focused on Pop III-star-hosting halos here, simply because the predicted DMH mass thresholds are easily defined. However, a Zel’dovich approximation-initialized simulation will always underestimate the very high end of the DMH mass function during the pre-reionization era, and this will have an effect on any prediction that relies on accurate high-mass DMH number counts—an obvious example is seed BH formation from direct collapse (Bromm & Loeb 2003; Lodato & Natarajan 2006; Belgeleman et al. 2006).

Note, too, that our simulation is $\sim 10^6$ times too small to model billion solar mass BHs, whose number density is 1 per Gpc$^3$ (Jiang et al. 2009). Although WMAP-5 can constrain the number of $\gtrsim 10^{13} M_\odot$ halos, our volume is too small to sample these halos. We do predict that these even rarer peaks will be more numerous and will be, on average, more massive in the pre-reionization epoch than predicted by $N$-body simulations initialized with Zel’dovich approximation. In principle, observations will be able to constrain the true halo mass function in the dark ages. While the relatively small difference in $\tau_e$ we observed may not be distinguishable with PLANCK and Herschel (Zahn et al. 2011), the global IGM temperature difference may be detectable (Theuns et al. 2002), and Square Kilometer Array (SKA) could easily probe this difference in the 21 cm power spectrum (Barkana & Loeb 2005).

This work was conducted at the Advanced Computing Center for Research and Education at Vanderbilt University, Nashville, TN. We also acknowledge support from NSF CAREER award AST-0847696. We would like to thank the referee for helpful comments.

REFERENCES

Abel, T., Bryan, G. L., & Norman, M. L. 2002, Science, 295, 93
Alvarez, M. A., Wise, J. H., & Abel, T. 2009, ApJ, 701, L133
Anninos, P., Zhang, Y., Abel, T., & Norman, M. L. 1999, New Astron., 2, 209
Barkana, R., & Loeb, A. 2005, ApJ, 626, 1
Begelman, M. C., Volonteri, M., & Rees, M. J. 2006, MNRAS, 370, 289
Bromm, V., Coppi, P. S., & Larson, R. B. 1999, ApJ, 527, L5
Bromm, V., & Loeb, A. 2003, ApJ, 596, 34
Clark, P. C., Glover, S. C. O., Smith, R. J., et al. 2011, Science, 331, 1040
Couchman, H. M. P., & Rees, M. J. 1986, MNRAS, 221, 53
