

# Baryonic Post-Processing of $N$ -body Simulations, with Application to Fast Radio Bursts

ADNAN KHAN,<sup>1</sup> IAN WILLIAMS,<sup>2</sup> AND MATTHEW MCQUINN<sup>2</sup>

<sup>1</sup>*University of Washington, Department of Physics, 3910 15th Ave NE, Seattle, WA, 98195, USA*

<sup>2</sup>*University of Washington, Department of Astronomy, 3910 15th Ave NE, Seattle, WA, 98195, USA*

Submitted to ApJ

## ABSTRACT

Where the cosmic baryons lie in and around galactic dark matter halos is only weakly constrained. We develop a method to quickly realize models for their distribution. Our approach uses the statistical advantages of  $N$ -body simulations, while painting on the profile of gas around individual halos in ways that can be motivated by detailed models or zoom-in hydrodynamic simulations of galaxies. Possible applications of the algorithm include extragalactic dispersion measures to FRBs, the Sunyaev-Zeldovich effect, baryonic effects on weak lensing, and cosmic metal enrichment. As an initial application, we use this tool to investigate how the baryonic profile of foreground galactic-mass halos affects the statistics of the dispersion measure (DM) towards cosmological FRBs. We show that the distribution of DM is sensitive to the distribution of baryons in galactic halos, with viable models having significantly different probability distributions for DM to a given redshift. We also investigate the requirements to measure statistically the circumgalactic electron profile for FRB analyses that stack DM with impact parameter to foreground galaxies, quantifying the size of the contaminating ‘two halo’ term from correlated systems and the number of FRBs for a high significance detection. Publicly available Python notebooks implement our CGMBrush algorithm.

## 1. INTRODUCTION

In a simpler universe in which the baryons could not cool, they would roughly trace the dark matter. However, our universe deviates far from this simple picture, as the baryons can cool and condense to high densities in dark matter halos. Yet, even adding cooling does not match the real complexity. If there were only cooling, the baryons associated with each dark matter halo would have condensed into stars – the Milky Way galaxy would have been ten times more massive (e.g. [Behroozi et al. 2010](#); [Tumlinson et al. 2017](#)). Instead, galaxies are host for violent phenomena like supernovae and active galactic nuclei that prevent such massive galaxies from forming. These phenomena inject energy and ultimately redistribute the gas, possibly even to distances well outside their dark matter halos. In addition to learning the locations of the baryons at low redshifts, understanding this redistribution is likely key to such big questions as to why some galaxies have stopped accreting gas and are becoming red and dead. It may also be important for realizing precision cosmology with weak lensing.

Most of our understanding of the distribution of baryons around halos is derived from observations of absorption lines against a bright background source, typically a quasar (Bahcall & Spitzer 1969; Thom et al. 2012; McQuinn 2016). This probe has revealed large column densities in atomic hydrogen and in many metal ions to impact parameters that extend 100-200 kpc from their foreground galaxies. However, while photoionization modeling has been used to constrain the amount of  $10^4\text{K}$  gas (e.g. Werk et al. 2014), there is no direct way to use these observations to probe the bulk properties and radial profile of all the gas (most of which is likely at higher temperatures).

Two related probes of the circumgalactic medium (CGM) are the thermal and kinetic Sunyaev Zeldovich (SZ) effects (Sunyaev & Zeldovich 1970, 1972), which measure respectively the inverse Compton heating and Doppler shifting of CMB photons when scattering off of free electrons. When stacking thousands of halos, the thermal SZ has been used to measure the projected pressure profile of the gas in halos to  $\sim 10^{13} M_\odot$  halos (Planck Collaboration et al. 2013; Greco et al. 2015; Schaan et al. 2016). The kinetic SZ is a new frontier that has recently been used to measure the the electron profile in the  $\sim 10^{13.5} M_\odot$  halos of massive luminous red galaxies (Amodeo et al. 2021). One limitation for both SZ effects is that the wide beam of CMB instruments makes it difficult to probe structure below the virial radius of galactic halos. Additionally, the kinetic SZ stacking requires redshifts to reconstruct the velocities, and the largest spectroscopic catalogues are of galaxies considerably more massive than the Milky Way. Both limitations add difficulty to using current and upcoming SZ measurements to probe the gas in halos of the Milky Way’s mass and below.

A new class of sources that provide an avenue for localizing the baryons is fast radio bursts (FRBs), which were rediscovered by Thornton et al. (2013), confirming an initial report (Lorimer et al. 2007). FRBs are a class of bright, extra-galactic radio transients with short, millisecond durations or less, that likely is sourced by magnetars (Petroff et al. 2019; Cordes & Chatterjee 2019). There is a detectable FRB roughly once every minute somewhere in the sky, only a small fraction of which are currently detected. However, many wide-field instruments are now coming online and are expected to increase detections by orders of magnitude (see Petroff et al. 2021, for a review). Unlike the aforementioned probes (aside from the kinetic SZ), FRB observations directly provide the amount of electrons along each sightline as the signal traverses through the intervening medium. The dispersion measure (DM) of this signal is an observed quantity that measures the delay in the arrival time as a function of the frequency caused by the total electron column along the line of sight (McQuinn 2014; Deng & Zhang 2014; Macquart et al. 2020; Simha et al. 2021; Connor & Ravi 2021; Lee et al. 2021). Simple estimates (McQuinn 2014) suggest that FRBs may excel at constraining the gas profiles around galaxies.

This paper develops a new technique to model the distribution of baryons, and is applicable to all these CGM observables. As an initial application, we focus on how the baryonic profile of foreground galactic-mass halos affects the statistics of DM towards cosmological FRBs. Our technique starts with an  $N$ -body simulation, which is the primary numerical method for modeling nonlinear structure formation in hundreds of megaparsec regions (Centrella & Melott 1983; Davis et al. 1985).  $N$ -body simulations do not follow how gas cools and is redistributed by feedback. Hydrodynamic simulations of galaxy formation are the preferred method to capture this feedback physics (e.g. Fielding et al. 2017; Su et al. 2020; Li & Tonnesen 2020; Stern et al. 2019; Lochhaas et al. 2020). Since the energy injection from supernovae and AGN is not resolved, these hydrodynamic simulations adopt varied prescriptions for injecting energy into the gas on resolved scales, around sites of star formation and

AGN activity in the simulation (Somerville & Davé 2015; Faucher-Giguère et al. 2016). The relatively small box of these hydrodynamic simulations means that they still struggle to adequately sample the cosmological density field (Zhu et al. 2018), an issue that is significantly mollified in the largest  $N$ -body simulations. Thus, we opt for a hybrid approach that uses the statistical advantages of large  $N$ -body simulations, while painting on the distribution of gas around individual halos in ways that can be motivated by zoom-in hydrodynamic simulations of galaxies (see also Schneider & Teyssier 2015 for models in a similar spirit).<sup>1</sup> This way, we are able to combine the statistical power of  $N$ -body simulations with the results from higher resolution hydrodynamic simulations (or alternatively semi-analytic models). Our method can further use CGM models based on observational measurements or even phenomenological parameterizations to understand what the data can support.

This paper is organized as follows. We discuss our methodology in § 2. In § 3.1 we describe the simulation, and investigate the application of different profiles and their results. We summarize our conclusions in § 4. Appendix A provides additional justification for our method.

## 2. THE CGMBRUSH METHOD

We present a method that can be applied to any cosmological  $N$ -body simulation (also generalizable to hydrodynamic simulations). The primary requirement for our FRB application is that the  $N$ -body simulation must resolve sub-Milky Way halo masses ( $\sim 10^{11} M_\odot$ ) and be in a  $\gtrsim 100$  comoving Mpc box that is large enough for a cosmologically representative sample of structures. The method realizes where the baryons are located with a significant reduction in the computational cost relative to hydrodynamic simulations of galaxy formation at the expense of some self-consistency.

This method adds to  $N$ -body simulations one thing they lack – a prescription for the locations of the baryons, especially around galaxies where the baryons’ dynamics are distinct from the purely gravitational dynamics of the dark matter owing to cooling and feedback. Algorithmically, we extract the matter associated with dark matter halos from the density field. We then redistribute the 20% baryon fraction of this matter using physically motivated models for the baryonic profiles around halos. For the dark matter that lies outside of dark matter halos, this method assumes that there is a component of the gas that perfectly traces the dark matter with the cosmic ratio. While this is almost certainly true when averaging over megaparsec scales and larger, it may not be on the outskirts of halos; as discussed in Appendix A, the clustering observables that are of most interest are fortunately not very sensitive to how one models the diffuse gas at the outskirts.

Large  $N$ -body simulations follow typically billions of particles, with the largest surpassing trillions. To reduce the size of outputs, the particles are often projected onto an Eulerian grid with some interpolation kernel (such as “nearest grid point” or “cloud-in-cell”; Klypin et al. 2011). Our method has the advantage of working on the combination of these Eulerian grid outputs and a halo field we create from the simulations’ catalogue of collapsed halos. Moreover, the method only requires the 2D projection of these grids, which allows our algorithm to achieve an enormous dynamic range. A fixed grid, while wasteful for many applications, is justified for the projected distribution of baryons as many halos intersect with essentially every sightline.

### 2.1. Algorithm at a single redshift

<sup>1</sup> The method of Schneider & Teyssier (2015) adds forces owing to baryonic pressure and feedback on top of the gravity forces included in traditional  $N$ -body simulations. This contrasts with the quick post-processing method we employ.

We first describe how we generate a baryon field for a single redshift. To start, the algorithm inputs a cosmological density field projected onto an Eulerian grid of  $N^3$  numbers in a box of comoving size  $L$ . Much of our discussion references a particular simulation to show examples. We digress briefly to describe some of its specifications, but please refer to § 3.1 for further details and a discussion of convergence. For the Bolshoi simulation (Klypin et al. 2011) used for the application in this paper,  $L = 250h^{-1}\text{Mpc}$  and  $N = 256$ , where this grid size is what is available publicly. We verify convergence against an  $N = 512$  output that is available for  $z = 0$  in § C. The algorithm projects the 3D density grids to 2D grids with  $N^2$  cells along a selected axis. This is shown for the  $z = 0$  Bolshoi snapshot in the top left panel of Figure 1. The acronym DM on the axes is because the gas column is identical aside from redshift factors to the dispersion measure, an FRB observable considered in detail in later sections.

The next step in our algorithm is to remove the mass associated with dark matter halos from the Eulerian density field, which requires that we create a template halo density field. We create the template halo density field from the catalogue of collapsed halos in the simulation, where we require at a minimum their positions and halo masses. First, we divide the halos into  $n_M$  logarithmic mass bins between a minimum and maximum halo mass ( $M_{\min}$  to  $M_{\max}$ ) and assume all halos within a mass bin have the same halo profile. Then we calculate the virial radius of the halos for each mass bin, assuming the density profiles in a given mass bin is described by a single NFW profile (e.g. Cooray & Sheth 2002)

$$\rho(r|M) = \frac{\rho_s}{(r/r_s)[1 + (r/r_s)^2]}, \quad (1)$$

parameterized by  $r_s(M) = r_{\text{vir}}/c$ , where  $r_{\text{vir}}$  and  $c$  are given by,

$$r_{\text{vir}} = \left( \frac{3M}{4\pi[18\pi^2 - 82q(z) - 39q(z)^2]\rho_c} \right)^{1/3}; \quad c(m, z) = \frac{9}{1+z} \left( \frac{M}{M_*} \right)^{-0.13}, \quad (2)$$

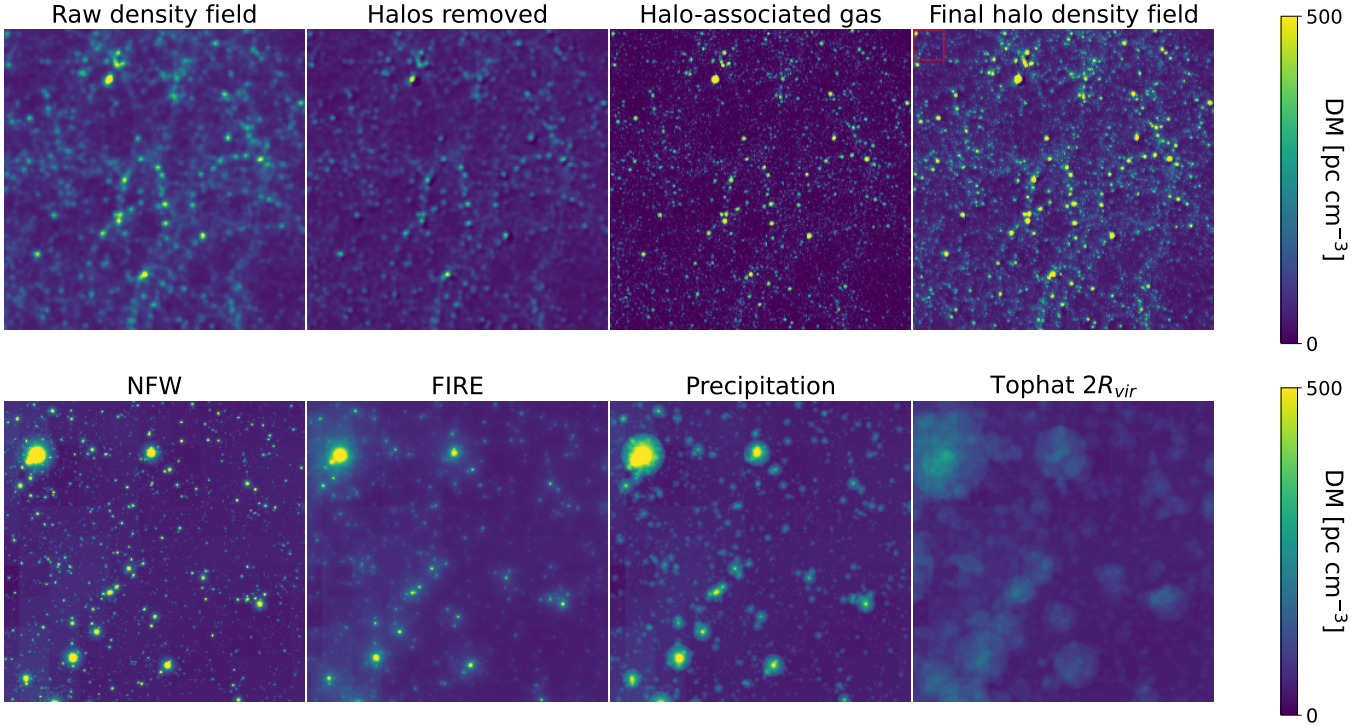
where  $q(z) = \Omega_\Lambda/[\Omega_m(1+z)^3 + \Omega_\Lambda]$ ,  $\rho_c = 3H^2/(8\pi G)$ ,  $M_* = 5 \times 10^{12} M_\odot$ , and this form for the concentration parameter  $c$  is taken from the numerical study of Peacock & Smith (2000). The halo mass determines the characteristic density,  $\rho_s$ , by integrating it out to the virial radius.

We then convolve the halo density profiles with the position of the centers of the halos (the physical location is provided by the simulation), yielding the halo density field template that we subtract off. Our algorithm does this projection using convolutions performed with fast Fourier transforms, which allows it to be very fast. One worry is that we have assumed a single halo profile, when actual relaxed halo profiles show significant scatter in their concentration (Kravtsov 2013). This problem likely only has a minor effect on our results for reasons detailed in the next paragraph.

We note that this step of removing the halo particles could be done prior to creating the Eulerian density grid. This would simplify the method as the user could simply skip the subtraction step.

The original projection of the  $N$ -body particles onto an Eulerian grid is effectively smoothing the density field, and this smoothing must be matched in this subtraction. The simulator has choices in how this smoothing is performed. In our application to the Bolshoi simulation, the field is gridded by cloud-in-cell(CIC) interpolation and Gaussian smoothing with a standard deviation of one cell. To create the template halo field, the algorithm matches the 2D projection of this smoothing on a much finer grid  $N_f$  but with the same physical width for the smoothing kernels, and, then, it





**Figure 1.** Illustration of the method on the Eulerian grid for different profiles. **Top panel:** This panel shows maps of dispersion measure (DM) – which is the total electron column along the line of sight – for the halo subtraction and addition through different stages of our method for our simple  $1\,r_{\text{vir}}$  top-hat model, applied to the  $z = 0$  snapshot of the Bolshoi  $N$ -body simulation for the  $512^2$  grid (results for  $256^2$  grid are shown in Appendix C). Each panel is projected over the simulation boxsize of  $250h^{-1}\text{Mpc}$  and shows a transverse comoving size of 140 Mpc. The left-most panel shows the projection of the raw gridded density field from the simulation. From this density field, we subtract the template halo density field, and the result is shown in the second column. The third column shows the halo-associated gas density field using the  $1r_{\text{vir}}$  top-hat profile that we wish to add back to the subtracted field, where a fine Eulerian grid of  $N_f^2 = 8194^2$  is used. The right most panel shows the projected final halo density field after the finely-gridded halo-associated gas density field are added back to the coarse subtracted field. **Bottom panel:** We zoom in on the region in the red rectangle at the top-right of the final field of transverse comoving size of 14 Mpc, and show this sub-region in the lower panel for different profiles modeled in this work. From left to right in the bottom panels, we show the final field when re-adding baryons with the NFW, FIRE, Precipitation, and  $2\,r_{\text{vir}}$  tophat profiles. These models are described in § 2.2, and we note that the NFW and tophat models are more of a toy character and not meant to be realistic. We can see that the baryonic distribution is the most concentrated in the NFW profile, followed by the FIRE, Precipitation, and then tophat profiles.

downgrids this fine grid to the  $N^2$  resolution of the simulation outputs. This final downgridded halo density field is then used as a template for subtracting the halo contribution to the density field; the second from left panel in the top panel of Figure 1 shows the result of the halo field subtracted from the raw density field (the high DM of a halo when subtracted from the raw density field lead to negative DM regions that can be seen as the dark spots in the figure). Because the size of the coarse Eulerian grid cell (1 Mpc in our applications, which also sets the scale of the additional smoothing) is larger than the virial radius of all but the largest halos, our method is relatively insensitive to our assumption of a single intrinsic halo profile in each mass bin (as given by eqn. 1). The cost

is that structures outside of halos (like sheets and filaments) that are smaller than the Eulerian grid cell are not resolved. Appendix A argues that for our envisioned applications these structures are a subdominant source of fluctuations, and this is further quantified by our convergence test presented in Appendix C that uses different Eulerian grids.

So far we have described how the method creates a halo field to subtract from the density field. The most important step is to create a field that represents the halo-associated gas which we can add back to the subtracted density field. The first step is to choose a profile for halo gas. This can be motivated by simulations or models. We describe some possible choices in § 3.2. To resolve the sub-halo structures, when adding the halos back with our desired profile, we increase the resolution of the 2D grid by a factor of  $\eta$ , so that our halo baryon grid and final total baryon density grid has resolution  $N_f = \eta N$ . The third panel from the right in the top row of Figure 1 shows the halo-associated gas grid for  $\eta = 32$  so that the final output grid is  $N_f = 8194^2$ , assuming a simple 3D top-hat model for the halo-associated baryons that extends  $1 r_{\text{vir}}$  with all the baryons enclosed.

Finally, we add the halo density field back to the raw density field from which the halos were previously removed to create a final grid of size  $N_f^2$  (top right panel in Fig. 1).<sup>2</sup> We use the  $N = 256$  field for the rest of our analysis.

The summary of how the final baryonic halo field is generated for a single redshift is as follows:

1. Start with a simulations density field on a 3D Eulerian grid and project over the “line of sight” dimension to make the  $N^3$  3D grid an  $N^2$  2D grid (top left panel, Fig. 1).
2. Create an  $N^2$  template halo density field by convolving halo centers with a mean NFW halo profile in each halo mass bin and, then, by further applying the same smoothing/interpolation scheme as used to create the Eulerian density grid. (For our application to the Bolshoi simulation, CIC and Gaussian smoothing.)
3. Subtract the template halo density field from the 2D density field to create a halo-subtracted baryon field (top left-middle panel, Fig. 1).
4. Add the baryons associated with halos back to the subtracted density field. First, choose a baryonic profile at each halo mass (see § 2.2 for potential models), and, next, convolve the halo centers with these profiles. As halo gas can have structure tens of kiloparsec scales, do this on a grid  $N_f$  with much higher spatial resolution than the halo subtracted baryon field ( $N_f \gg N$ ) (top middle-right panel, Fig. 1).
5. Add the halo field back to the halo subtracted field to create the full model for the baryonic field on a  $N_f^2$  grid (right panel Fig. 1).

## 2.2. Halo Profiles

The method in the previous section describes a way of modeling the distribution of gas around halos. This allows one to study the effects of different baryonic profiles. Here, we consider several models for the redistribution of gas in and around dark matter halos, some motivated by their simplicity,

<sup>2</sup> In the implementation presented in this paper, we do not subtract the  $\sim 5\%$  of baryons that form stars from the diffuse baryons, even though it does not contribute to the envision statistics such as regarding FRB DM. This omission is justified by the fact that removing this contribution will have a minor effect on any of our results, which are shaped by the large uncertainty in baryonic halo profiles.

others by their analytical nature, and others still by sophisticated zoom-in simulations of the gas around halos. Our code **CGMBrush** can take a user specified profile as well.

In particular, in this paper we consider the following models:

**3D top hat:** This is an unrealistic but instructive density model for the halos. It assumes the closure density in baryons is distributed uniformly within a 3D sphere of radius  $r$  around each halo center. Perhaps the most motivated value is  $r = r_{\text{vir}}(M)$ , but we also consider twice this extent as  $r = 2r_{\text{vir}}(M)$ .

**NFW:** The NFW density profile (eqn. 1) describes the dark matter in halos rather than the gas (Navarro et al. 1996). However, it becomes a better approximation for the gas profile in the most massive galaxy clusters. NFW models overestimate the concentration of the gas in smaller halos because cooling and feedback tend to redistribute the gas to larger radii (Maller & Bullock 2004).

**FIRE:** The FIRE profile is based on cosmological hydrodynamic “zoom-in” simulations with detailed sub-grid physics to account for feedback processes (Hopkins et al. 2018). Their CGM density profiles are well described by an  $r^{-2}$  profile. Our method requires us to redistribute *all* the baryons associated with these halos, whereas Hafen et al. (2019) investigated the profile out to the virial radius, which misses baryons redistributed to further extents. Therefore, we assume an exponential cutoff at  $r_{\text{max}}$  that is determined so that the profile integrates to the total associated gas mass. Thus, our FIRE density profile is given by:

$$\rho = \rho_0 \left( \frac{r}{r_{\text{vir}}} \right)^{-2} \exp[-r/r_{\text{max}}], \quad (3)$$

Hafen et al. (2019) in Fig. 2 considers  $10^{10} M_{\odot}$ ,  $10^{11} M_{\odot}$ ,  $10^{12} M_{\odot}$  and  $z = 0.25$  for CGM, which we extrapolate to higher masses with the help of observations from Davies et al. (2020).<sup>3</sup> We use a normalization determined from assuming this  $r^{-2}$  profile and make sure it integrates to the baryonic mass within  $r_{\text{vir}}$  in Fig. 2 in Hafen et al. (2019) for the respective masses.<sup>4</sup> The fraction of baryons appears constant for different redshifts, as seen in the lower panel of Fig. 2 for  $z = 2$ . Therefore, the only redshift dependence is in  $r_{\text{vir}}$  which scales as in equation (2), and is nearly constant in comoving coordinates except for some contribution from dark energy.

**Precipitation-regulated:** These are physical models which posit that the density of the CGM is shaped by saturating a threshold ratio of the cooling to dynamical time (Sharma et al. 2012; Voit et al. 2017). Idealized simulations show that exceeding this ratio results in dramatic cooling and fragmentation, that should source star formation and, hence, stellar feedback, and restore balance (McCourt et al. 2012). These models have found success at explaining various observations of gas in the Milky Way CGM (Voit 2019) as well as in other systems (Voit 2019; Voit et al. 2019). Voit (2019) provides a fitting formula to their precipitation-regulated profile

<sup>3</sup> The extrapolated fraction of baryons in CGM for  $10^{13} M_{\odot}$ ,  $10^{14} M_{\odot}$ ,  $10^{15} M_{\odot}$  are 0.5, 0.8, 1, respectively.

<sup>4</sup> From the normalization, we calculate the comoving electron density of  $6.3 \times 10^{-5} \text{ cm}^{-3}$ ,  $1.2 \times 10^{-4} \text{ cm}^{-3}$ ,  $1.9 \times 10^{-4} \text{ cm}^{-3}$ ,  $2.3 \times 10^{-4}$ ,  $2.7 \times 10^{-4}$ , and at  $0.3r_{\text{vir}}$  for  $10^{10} M_{\odot}$ ,  $10^{11} M_{\odot}$ ,  $10^{12} M_{\odot}$ ,  $10^{13} M_{\odot}$ ,  $10^{14} M_{\odot}$ , and  $10^{15} M_{\odot}$ , respectively, and interpolate the trend linearly.

in physical units:

$$n_e(r) = \left\{ \left[ n_1 \left( \frac{r}{1 \text{ kpc}} \right)^{-\zeta_1} \right]^{-2} + \left[ n_2 \left( \frac{r}{100 \text{ kpc}} \right)^{-\zeta_2} \right]^{-2} \right\}^{-1/2} \quad (4)$$

which scales respectively at small and large radii as  $n_e \propto r^{-\zeta_1}$  with  $\zeta_1 \approx 1.2$  and  $n_e \propto r^{-\zeta_2}$  with  $\zeta_2 \approx 2.3$ . The latter scaling is similar to the FIRE profile, although the overall normalization tends to be somewhat smaller for these Voit (2019) profiles. The normalizing factors  $n_1$  and  $n_2$  are interpolated from Table 1 in appendix of Voit (2019), and these factors,  $\zeta_1$  and  $\zeta_2$  are all functions of halo mass. Just like with FIRE, we need to find a way to redistribute all the halo-associated gas. This density profile given by eqn. (4) is integrated out to  $2r_{\text{vir}}$ , and the halo-associated gas mass that is not accounted for is then distributed in a tophat out to  $2r_{\text{vir}}$ . To extrapolate this model to higher redshifts from  $z = 0$  coefficients we use the fact that for an isothermal sphere mass profile, the dark matter density at fixed physical radius scales as  $(1+z)M^{2/3}$ , which means that fixing the dynamical time requires keeping  $(1+z)M^{2/3}$  constant. However, if  $(1+z)M^{2/3}$  is fixed, since  $T_{\text{vir}} \propto (1+z)M^{2/3}$  has exactly the same scaling, this fixes the virial temperature and, hence, the cooling rate. Thus, if we keep the same gas density profile as the  $z = 0$  density profiles of Voit et al. (2018), but identify this profile with a halo with mass such that  $(1+z)M^{2/3}$  is fixed, this essentially fixes the ratio of cooling to dynamical times, which is what we desire to do to maintain the spirit of the model. While we would be hesitant to apply this transformation for high-redshift halos (as of course the transformation gets larger with increasing redshift), all our results are for  $z \leq 1$ .

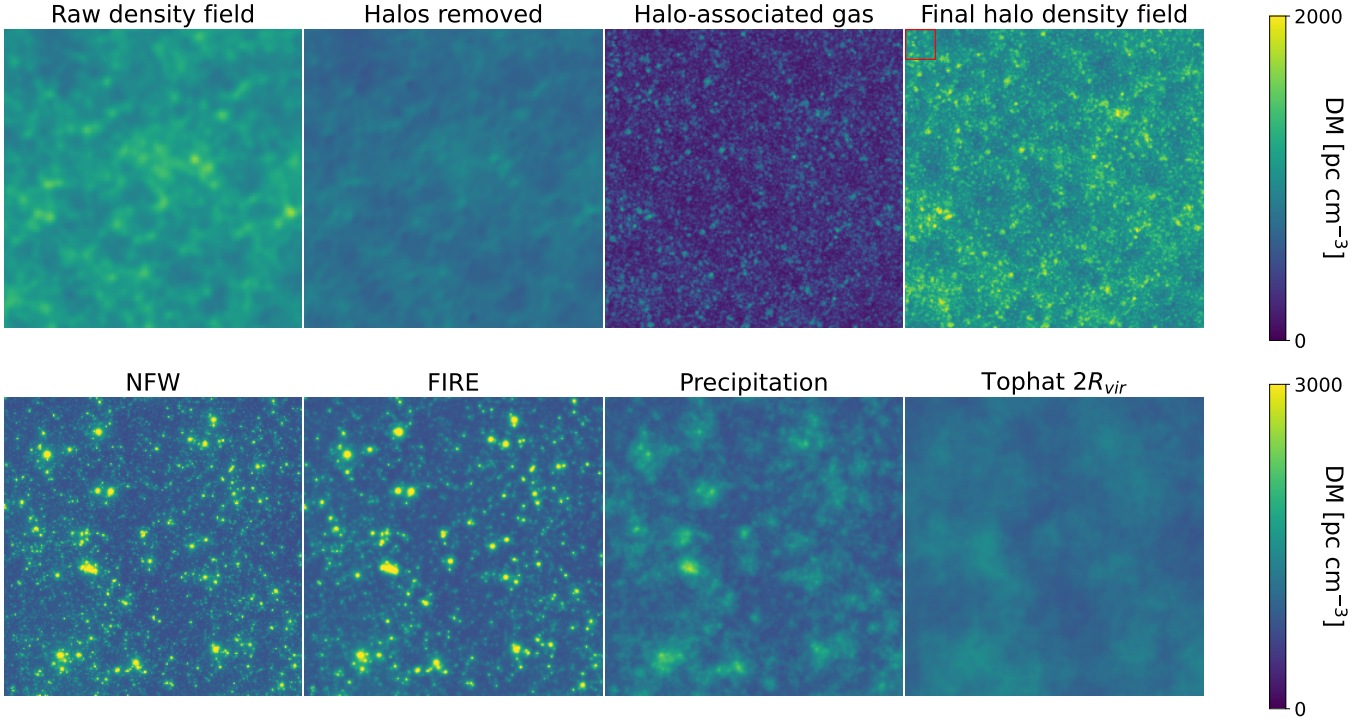
### 2.3. Multiple redshifts and a light-travel image

Most applications of our algorithm require capturing the baryon field across a span in redshift. To achieve this with a periodic simulation box, we follow the standard procedure used in many cosmological analyses of stacking shifted and rotated simulation snapshots that span the desired redshift range (e.g. Saro et al. 2014). The average redshift extent of our example box of size  $250h^{-1}$  Mpc is  $\Delta z \approx 0.1$  at  $z \lesssim 1$ . So to create a map out to higher redshifts than  $z = 0.1$ , we must stack a number of boxes on top of each other. To do so, we use the density and halo field snapshots closest to the mean redshift desired for the next box in the stack.<sup>5</sup> A given snapshot can be selected multiple times if the cadence of simulation outputs is coarse.

Before stacking different simulation outputs, each simulation snapshot is randomly oriented along one of its three major axes and then translated by a randomly oriented vector of random direction and length randomly sampled from  $[0, \sqrt{2}L]$  in the direction transverse to the line of sight. The translation is performed in a manner that preserves the simulation's periodicity. This randomization is done to diminish artifacts from the same structures aligning in the stack because of the periodic boundary conditions. Stacking also assumes that correlations on scales larger than the box size can be neglected.

<sup>5</sup> Since this stacking of a handful of boxes is a rather crude Riemann sum, to guarantee that the average over the box matches the expected  $\overline{DM}(z)$  relation for our FRB application, we choose an effective redshift of each box in the stack to rescale its mean density to  $\bar{n}_{e,0}(1+z_{\text{eff}}^{(n)})^3$  and to assign to the  $n^{\text{th}}$  snapshot, where  $\bar{n}_{e,0}$  is the mean density today, so that the dispersion averages to the mean relation  $\overline{DM}(z)$ . This requires  $z_{\text{eff}}^{(n)} = (\overline{DM}(z_{\text{max}}^{(n)}) - \overline{DM}(z_{\text{min}}^{(n)})) / (L\bar{n}_{e,0}) - 1$ , where  $z_{\text{min}}^{(n)}$  and  $z_{\text{max}}^{(n)}$  are the redshifts for the edges of the  $n^{\text{th}}$  box of comoving size  $L$  in the stack (calculated by converting the comoving size of the  $n$  and  $n+1$  boxes to their corresponding redshifts). This rescaling results in  $z_{\text{eff}}$  being just slightly different than the  $z$  of the snapshot, but is preferable to the 10% errors from  $\overline{DM}(z)$  that result if we do not rescale for the applications presented in this paper. Different weightings should be chosen for other observables.





**Figure 2.** Illustration of the method on the Eulerian grid for a light-rectangle image to a redshift of  $z \approx 1$ , where ‘light-rectangle’ indicates that the projected coordinate follows the light travel delay. **Top panel:** This panel shows maps of dispersion measure (DM) through different stages of halo subtraction and addition for a simple  $1 r_{\text{vir}}$  top-hat model. The top panels show a transverse comoving size of 140 Mpc. The left-most panel shows the projection of the raw gridded density field from the simulation. From this density field, we subtract the template halo density field, and the result is shown in the second column. The third column shows the halo-associated gas density field using the  $1 r_{\text{vir}}$  top-hat profile that we wish to add back to the subtracted field, where a fine Eulerian grid of  $N_f^2 = 8194^2$  is used. The right most panel shows the projected final halo density field after the finely-gridded halo-associated gas density field are added back to the coarse subtracted field. **Bottom panel:** We zoom in on the region in the red rectangle at the top-right of the final lightcone image of transverse comoving size of 14 Mpc, and show this sub-region in the lower panel for different profiles modeled in this work. From left to right in the bottom panels, we show the final field when re-adding baryons with the NFW, FIRE, Precipitation, and  $2 r_{\text{vir}}$  tophat profiles. These models are described in § 2.2, and we note that the NFW and tophat models are more of a toy character and not meant to be realistic. We can see that the baryonic distribution is the most concentrated in the NFW profile, followed by the FIRE, Precipitation, and then tophat profiles, consistent with results from a single box.

The large box sizes our algorithm envisions (i.e.  $250h^{-1}$  Mpc in the application we present), and the  $\lesssim 1$  Mpc halo scale correlations of interest, suggest that this is a good approximation. (See also Appendix A, which quantifies the size of fluctuations at  $k \lesssim 2\pi/L$ .) For each stacked box, we repeat the above procedure of removing and adding halos to attain the final density field. By doing so, we simulate the matter fields to high redshifts as shown in Figure 2.<sup>6</sup> The right most image of the final

<sup>6</sup> Currently our maps are in position and not angular space. Position space is convenient for the applications in this paper. However, that we observe in angular space means that a given angle samples larger physical scales at low redshifts. To be used for angular statistics, like the angular power spectrum of the Sunyaev Zeldovich anisotropies or weak lensing, the algorithm needs to be generalized. Angular maps are also interesting for our FRB application to understand the variance from local structures and to potentially constrain gas profiles around very low redshift galaxies (Connor & Ravi 2021).

halo density field is a 140 Mpc region, and shows that our method is able to resolve structures and distribute baryons quite effectively without introducing any artifacts from redshift stacking. The zoomed in region of 14 Mpc in the lower panel allows us to see the concentration of baryons for different profiles, with NFW and tophat as toy models for the redistribution of baryons, and FIRE and Precipitation models capturing some of the CGM physics. As in Figure 1, we see that the NFW profile is the most concentrated, followed by the FIRE, Precipitation, and then tophat profiles. We discuss the PDFs from these stacked fields in 3.2.

### 3. APPLICATION OF CGMBRUSH TO FRBS

#### 3.1. *N-body simulation specifications*

To demonstrate our method, we use here the publicly-available Bolshoi simulation at *CosmoSim*, which meets our basic requirements: it resolves sub-Milky Way halos ( $\sim 10^{11} M_\odot$ ), which we require for convergence on our statistics (Appendix B), and it has greater than a  $\gtrsim 100$  Mpc box, which is large enough for a cosmologically representative sample of structures. This simulation is a  $2048^3$  particle  $N$ -body simulation in the  $\Lambda$ CDM cosmology (Klypin et al. 2011) with a box of size  $250 \text{ Mpc } h^{-1}$  and a particle mass  $1.35 \times 10^8 M_\odot h^{-1}$ . It is run with a mass and force resolution to follow halos down to the completeness limit of  $V_{\text{circ}} = 50 \text{ km s}^{-1}$ . It adopts the  $\Lambda$ CDM parameters  $h = 0.7$ ,  $\Omega_m = 0.27$ ,  $\Omega_b = 0.0469$ ,  $n = 0.95$  and  $\sigma_8 = 0.82$ . The outputs are available for a large number of snapshots between a redshift of 0 and 17.

*CosmoSim* provides the Bolshoi simulation density field in two resolutions,  $256^3$  and  $512^3$ . Since the higher resolution of  $512^3$  is only provided at  $z = 0$ , we use the  $256^3$  grid for our primary calculations; we compare the results of using the two resolutions in Appendix C, and show that there are relatively minor differences in our final results using the different resolutions. Thus, we project the  $256^3$  3D density field to a 2D field of  $256^2$ , and use this field for the rest of the procedure. In the left most panel in Figure 1, we project over one box in a field of view of 70 Mpc, where ‘DM’ is a measure of the projected baryonic density. It is important to note how the density field in the simulation is smoothed, as we must smooth the halo density field in the same manner: it is calculated using cloud-in-cell interpolation, and then Gaussian smoothing of one-cell-sigma is applied. These choices are relatively standard, but the method may work somewhat for different interpolation kernels.

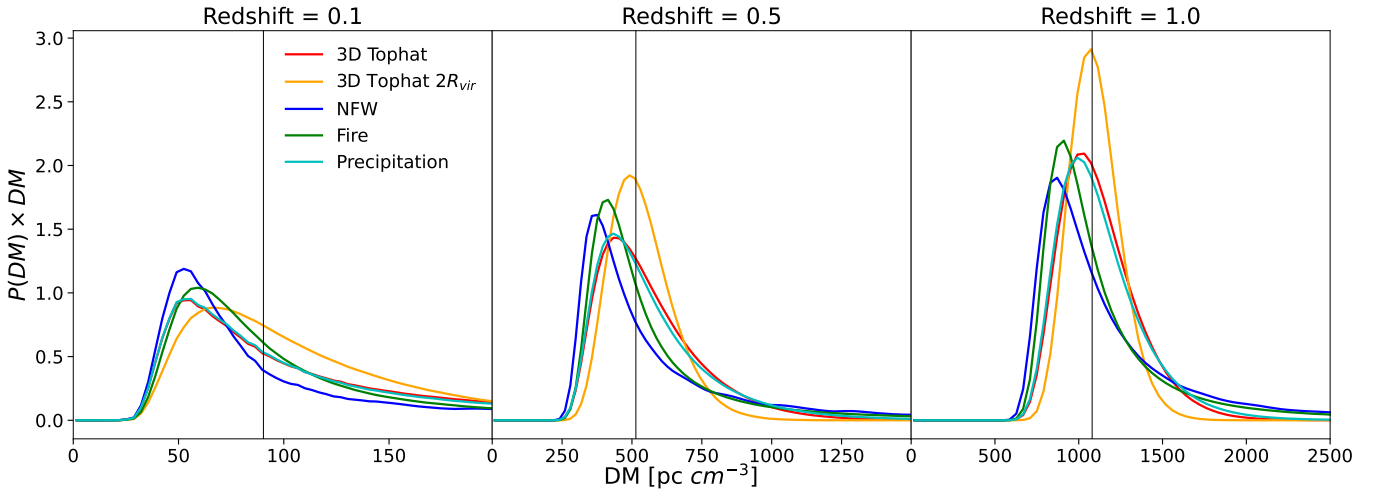
The *CosmoSim* halo positions are specified on a  $1024^3$  grid, rather than the likely more common choice of providing the exact positions. Since this is four times the resolution of our density grid, this coarse-graining of positions does not affect our results. We bin the halos into  $n_M = 30$  logarithmic mass bins between a minimum and maximum halo mass of  $10^{10} - 10^{14} M_\odot$ , resulting in  $\Delta M \approx 0.3M$ . For each mass bin we calculate an NFW profile using the mean mass of the bin, assuming that all halos within a mass bin have the same profile, and convolve the profiles with the positions of the halos to create a halo density field. That our algorithm does this projection using convolutions allows it to be very fast. To match the size and smoothing scheme of the provided *CosmoSim* density field outputs, we apply Gaussian smoothing to the halo density field. This smoothing is first done on a  $1024^2$  grid. When subtracting the halo field from the density field, we coarsen the halo density field from a finer grid to a  $256^2$  grid by averaging over neighboring cells. The top panel of Figure 1 demonstrates these steps in the two leftmost panels.

We can place the gas distribution of halos in the simulation on a much finer grid than the Bolshoi simulation density field – an approach that is motivated by the fact that the gas distribution in halos

is the only field that clumps on much smaller scales. To resolve sub-halo structure, we increase the resolution to  $(\eta \times 256)^2$ , where  $\eta = 4, 8, 16, 32, 64, 128$ . Using the same method for calculating profiles as described in the methods section, we create a halo density field on the higher resolution grid, and add the halos back to the subtracted field using the profiles we describe in section 2.2.

The central regions have of course the highest DM values, but this decreases with increasing distance. How well the central regions are resolved depends on our final grid resolution. A grid of size  $1024^2$  has a cell-size of 340 kpc, which is too low resolution: for halos lighter than  $10^{12} M_\odot$  with virial radii less than 250 kpc, we cannot resolve the internal structure at all. Increasing the resolution to 8, 16, and 32 times the 1024 grid, with  $\eta = 32$ ,  $\eta = 64$ , and 124 respectively, decreases the size of the cells to 44, 22 and 11 kpc, capturing the rapid fall in density of the central regions of halos down to  $10^{10} M_\odot$  with virial radii of 40 kpc, enabling the sightlines to better sample different regions of halos in the entire range of masses in consideration as evident in the right most panels of Figure 9 in Appendix C. Refer to Appendix C to examine how our results converge as we increase the resolution of the final grid.

### 3.2. The probability distribution of $DM(z)$



**Figure 3.** Probability distribution functions (PDF) of the cosmic contribution to DM assuming different baryonic profiles around halos with  $M_{\min} > 10^{10} M_\odot$ . The different curves illustrate the dependence on different profiles including spherical tophat, FIRE, NFW, and precipitation (the description of these profiles is given in § 3.2). The redshift dependence is shown across the three panels for  $z = 0.1, 0.5$  and 1. The vertical line shows the mean DM to the given redshift. The fact that the PDFs are significantly different shows that our method is sensitive to the different baryonic profiles around halos. Calculated using the 8192 resolution.

One of the main goals of our method is to analyze the effect of different profiles on the distribution of dispersion measure (DM) as probed by FRBs. In this vain, we are able to use any well-defined profile to create a halo density field and see its effect on the PDF of DM. The DM of this signal is an observed quantity that measures the delay in the arrival time as a function of the frequency caused by the total electron column along the line of sight. For FRBs at a redshift  $z_s$  (McQuinn 2014),

$$DM(z_s, \hat{\mathbf{n}}) = \int_0^{\chi(z_s)} \frac{\rho_e(z, \hat{\mathbf{n}})}{(1+z)^2} d\chi \quad (5)$$

Standard deviation of the dispersion measure (pc cm <sup>-3</sup> )			
CGM Profile	$z = 0.1$	$z = 0.5$	$z = 1.0$
3D Tophat	72	159	214
3D Tophat $2r_{\text{vir}}$	49	108	148
NFW	178	377	480
FIRE	130	337	513
Precipitation	79	174	233

**Table 1.** Standard deviation of dispersion measure (DM) in units of pc cm<sup>-3</sup> for the halo gas density profiles studied in this work. DM is a redshift-weighted electron column.

where  $d\chi = c dz/H(z)$  is the differential of the conformal distance ( $\chi$ ),  $\rho_e(z, \hat{n})$  is the electron number density at redshift  $z$  in direction  $\hat{n}$ , and  $z_s$  is the source redshift.

Our algorithm is capable of being used over an arbitrary redshift interval, limited only by what redshifts snapshots the simulations provide. However, in the application of our algorithm, we only consider the universe in the interval  $z = 0 - 1$ . This interval likely spans the range of most detected FRBs; the highest redshift FRB that has so far been localized to a galaxy has  $z = 0.66$  (Ravi et al. 2019).<sup>7</sup> Since the length of a single box is  $250h^{-1}\text{Mpc}$ , we need approximately nine boxes to simulate the effect of going to the desired  $z = 1$ .

Observations of DM from FRB’s probe the fluctuations in the DM from sightline to sightline. From Figure 2 we can create PDF out to a redshift of 1. Figure 3 shows the PDF of the cosmic distribution of DM to redshifts of 0.1, 0.5, and 1, respectively, for the profiles described above. All models we have implemented predict a high-DM tail to the PDF, which turns out to be dominated by the most massive systems with  $m_h \gtrsim 10^{13}M_\odot$  (see Appendix B where we transition from a  $2r_{\text{vir}}$  spherical tophat profile to an NFW profile above varying mass thresholds to show that the transition has the effect of essentially eliminating the ‘1-halo’ contribution of the small mass halos to the width of the PDF). The more diffuse the gas around halos or the rarer the halos that hold onto their gas, the more concentrated is the PDF. This trend arises because each sightline intersects a more statistically representative set of structures for the models where the baryonic profiles of the halos is more diffuse (McQuinn 2014). That the profiles lead to significantly different PDFs of DM shows that we are able to probe the baryonic structure. The standard deviation in DM for the profiles described above are provided in Table 1.

### 3.3. DM versus impact parameter to galaxies

Ultimately, this baryon science from FRBs will not use the PDF of DM, but instead will stack foreground galaxies with different impact parameters to background FRBs to measure statistically the gas profile around halos. Unlike the PDF of DM, stacking is unbiased by any DM intrinsic to the host system. Such a stacking measurement was considered in McQuinn (2014), but it has not been investigated using simulations. A concern is the importance of correlating material in the stack, as stacking does not just measure the distribution of gas that is associated with the fraction of mass in baryons associated with each halo of mass  $M$ . Our method is useful for understanding the size of this effect.

<sup>7</sup> It is both harder to find FRBs at higher redshift owing to the added computational expense to search for these in FRB backends as well as to localize them to a galaxy.

Unlike for  $P(\text{DM})$ , where we needed to trace through multiple redshifts, for the mean stacked profile of DM in a mass bin, we only need to consider a single box, as material at distances greater than the  $\approx 100$  Mpc simulation box – to which our method is applied – is weakly correlated with the halo used in the stack. Material along each FRB sightline that is in the foreground or background will add to the variance of the stack, whose effect we can calculate from the  $P(\text{DM})$  reported in the previous section.

We calculate the radial profile of halos in our final density field for different impact parameters as follows: for all halos within a mass bin, we trim out a grid around the halo from the final density field; then, we stack the trimmed grid on top of each other for each mass bin; from this stack of halos, we calculate a density profile for the given mass bin. Figure 4 shows the radial profiles of DM for  $10^{11} M_\odot$ ,  $10^{12} M_\odot$ , and  $10^{13} M_\odot$  halos. The dashed lines show the DM of the baryonic profile applied to the halos, and the solid lines show the DM of the final density field. The vertical dashed line shows the virial radius of the chosen halo profile. To the extent that there is no contribution of correlated material that lies outside of the halo, these two profiles should agree. The differences between the solid and dashed curves illustrates the contribution of all the matter correlated with the halo, the “two-halo” term. The two-halo term is much larger than the theoretical calculations by [McQuinn \(2014\)](#). In addition to the diffuse matter correlated with the halos, higher mass halos can host substantial populations of satellite galaxies around them as well.

In each of the three panels, we can see that NFW profile is the most concentrated; this fact is most prominently seen in the bottom panel for  $10^{13} M_\odot$ , where it diverges most significantly from other profiles in the central regions. Consistent with the images of the final density field in Figure 1, we can see that after NFW, FIRE profile is the most concentrated above  $10^{12} M_\odot$ , followed by precipitation, and tophat. While our method is sensitive to the profiles of halos an order of magnitude below a Milky Way mass halo, the sensitivity increases for higher mass halos as evident in the bottom panel.

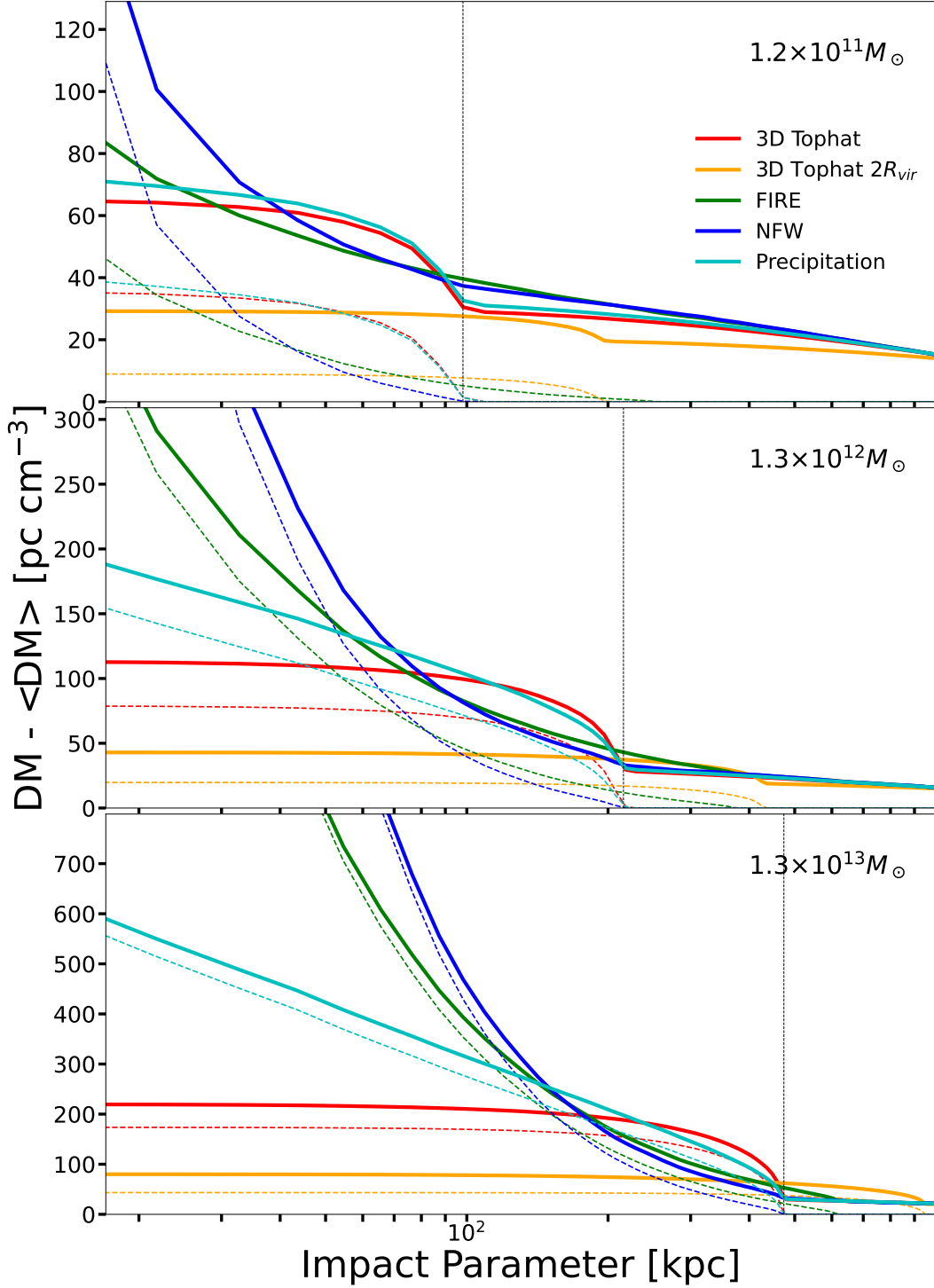
We next consider the sensitivity of stacking analyses if we assume that a sample of  $N_{\text{FRB}}$  localized FRBs have been localized. The  $1\sigma$  error bar in the bin described by impact parameter of  $r$  is

$$\sigma(r) = \frac{\sqrt{\sigma_{\text{Cosmic}}^2 + \sigma_{\text{Host}}^2}}{\sqrt{N_{\text{inter}}(r)}}; \quad N_{\text{inter}} = N_{\text{FRBs}} \times N_{\text{halos}} \times P(r) \quad (6)$$

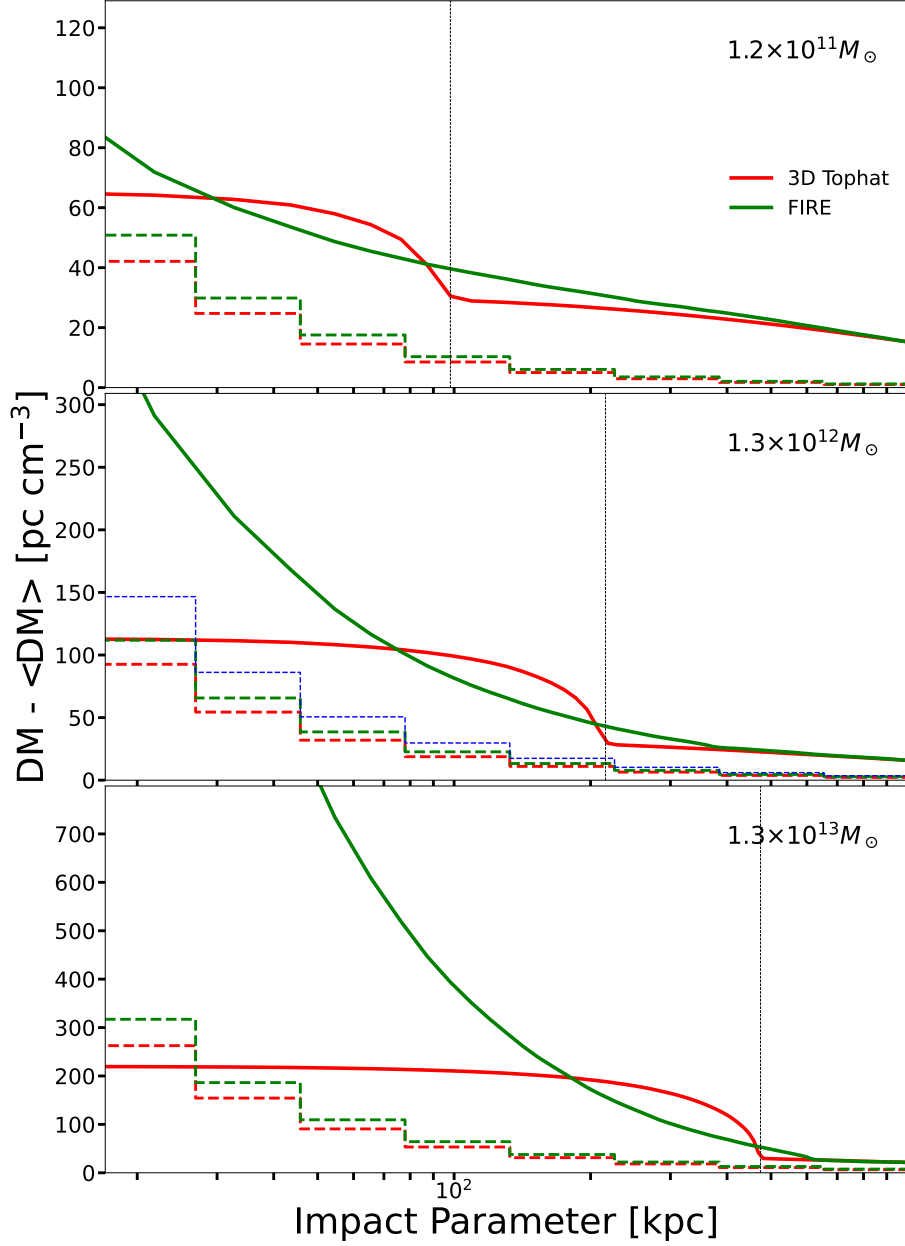
where  $N_{\text{inter}}$  is the number of FRB sightlines that fall within  $[r, r + \delta r]$  for a halo of mass bin of  $[M - \delta M/2, M + \delta M/2]$  where  $M$  is the central halo mass in the bin that we use as our label, and  $N_{\text{halos}}$  is the average number of halos intersected by each sightline to a redshift of  $z \approx 0.5$  – characteristic of a typical redshift of current FRB samples.  $N_{\text{halos}}$  as read off from Figure 1 in [McQuinn \(2014\)](#) are 4, 2, and 0.6 for halos with mass above  $10^{11} M_\odot$ ,  $10^{12} M_\odot$ , and  $10^{13} M_\odot$ , respectively. However, in our calculation we assume more conservative estimates for halos intersected by each sightline to a redshift. First, [McQuinn \(2014\)](#) estimates are for intersecting halos above a mass threshold, whereas we are calculating the error for a mass bin; second, small halos are harder to detect. Therefore, we reduce 4 and 2 to 1 and 1 for  $10^{11} M_\odot$  and  $10^{12} M_\odot$  halos. The  $P(r)$  factor in  $N_{\text{inter}}$ , gives the probability of the FRB signal passing through the gas for a given impact parameter bin at distance  $r$  from the center of the halo and is given by  $P(r) = (\pi(r + \delta r)^2 - \pi r^2)/(\pi r_{\text{vir}}^2)$ .

The error is plotted in Figure 5. The dashed red and green piece-wise lines show estimates for the sensitivity to the mean DM for a tophat and FIRE profile for a sample of  $N_{\text{FRB}} = 100$  and  $z \approx 0.5$  localized FRBs for halos of masses  $10^{11} M_\odot$ ,  $10^{12} M_\odot$ , and  $10^{13} M_\odot$  in the top, middle, and bottom





**Figure 4. DM vs Impact Parameter:** The angular profile of the dispersion measure (DM) for different gas profiles is shown in each panel, which can be measured by stacking DM measurements by their impact parameter to foreground galaxies. The middle panel is for a Milky Way-like halo of  $10^{12} M_{\odot}$ , whereas the other panels consider galactic halos that are an order of magnitude larger and smaller. The dashed lines show the DM for the baryonic profile we applied to the mass bin, and the solid lines show the DM extracted from the final density field. The difference between the solid and dashed curves owes to the contribution of the gas correlated with halos that the mask does not capture (i.e. the ‘two halo’ term). The calculations are made with 32,768 resolution.



**Figure 5.** The angular profile of the dispersion measure for the  $1r_{\text{vir}}$  3D tophat and FIRE gas profile (solid curves) alongside our  $1\sigma$  error bar estimates for a mock stacking survey of localized FRBs (dashed peicewise line). Namely, this survey assumes 100 FRB localizations with a typical redshift of  $z \approx 0.5$  plus models for the foreground incidence rate of different halo masses (see § 3.2). The blue dashed segmented line in the middle panel shows how the error for FIRE profile increases assuming a host DM contribution for a host galaxy contribution of  $\sigma_{\text{Host}} = 300 \text{ pc cm}^{-3}$  whereas the other curves are in the limit of the host galaxy being a subdominant contribution to the variance. The calculations are made at 32,768 resolution.

panel, respectively. These estimates use the quoted numbers for  $z = 0.5$  bursts as well as the  $\sigma_{\text{Cosmic}}$  in Table 1. The significance that each profile can be detected in a radial bin is assessed by the ratio of the profile to the amplitude of the per piece-wise lines. The intrinsic contribution from the host galaxy is set to zero ( $\sigma_{\text{Host}} = 0$ ) in the red and green dashed perce-wise lines. The vertical line again

indicates the virial radius. For 100 FRBs, we are sensitive to all models from a third of a virial radius out to a virial radius. In the top panel, we can see that the profile is well constrained to just over a tenth of the virial radius for halos of mass  $10^{11} M_{\odot}$ . In the middle panel, for halos of mass Milky Way, the profile is well constrained down to a tenth of a virial radius also. Upon adding the  $\sigma_{\text{Host}} = 300 \text{ pc cm}^{-3}$  as the intrinsic contribution from the host galaxy, the error increases significantly as shown by the dashed blue line, but the profile is still well constrained down to a quarter of the virial radius. This value for the host galaxy contribution is below the estimates of [Macquart et al. \(2020\)](#). In the bottom panel, for systems of mass  $10^{13} M_{\odot}$ , the profile is constrained with 100 bursts to half a virial radius.

Thus, with as few as 100 FRBs and optical follow-up to find foreground galaxies, we predict that the DM profile can be well constrained over an impact parameter of  $0.3 - 1 r_{\text{vir}}$  for halo masses of  $10^{10} - 10^{13} M_{\odot}$ , with the sensitivity in detail depending on the true gas profile. While, our method can measure the gas profile outside of the virial radius with such a sample, the two-halo term from correlating gas is generally dominant even over our gas profiles that distribute the most gas at  $\gtrsim 1 r_{\text{vir}}$  radii (and for sub-Milky Way halos this can also be the case even at  $< 1 r_{\text{vir}}$ ). The limiting factor for measuring the halo gas profile towards even dwarf galaxies is having a deep enough survey to identify these galaxies; if they can be identified, we conclude that their gas profile can be measured statistically with FRBs. Because of this, the assumptions that went into our sensitivity estimates for sub-Milky Way mass halos are likely optimistic. It may also be, especially towards lower masses, that the mapping from galaxy properties to halo mass is less one-to-one. In this case, the interpretation of such a stacking experiment may be more complex.

#### 4. CONCLUSIONS

Constraining the locations of the cosmic baryons is important for both understanding galaxy formation and achieving precision cosmology from large-scale structure surveys. We have developed a method to redistribute baryons around halos in  $N$ -body simulations and presented its application to FRBs. The method is fast, only requiring convolutions in 2D, and it generates large Eulerian 2D grids, which are easier to work with than adaptive outputs. A fixed grid, while wasteful for many applications, is more justified for the projected distribution of baryons as many halos intersect with essentially every sightline. The method enables modeling the baryons in much larger volumes and around smaller galaxies than can be achieved with modern cosmological hydrodynamic simulations, and it can be used to quickly survey a suite of potential models.

The distribution of halo-associated baryons that are used can be informed by the results of semi-analytic models, zoom-in hydrodynamic simulations, or previous observations. We implemented several such distributions in this study. In addition to toy analytic models, we considered models for the baryonic profile motivated by the precipitation-limited models of [Voit \(2019\)](#) and the FIRE galaxy formation simulations [Hafen et al. \(2019\)](#). Our publicly available Python package **CGMBrush** allows the user to create custom models.

By applying our method to the statistics of DM towards cosmological FRBs, we have shown that the DM is quite sensitive to the distribution of baryons in galactic halos, with different plausible models producing significantly different probability distributions of DM to a given redshift. We have also investigated the projected profile that could be measured in a stacking analysis using foreground galaxies. Our investigation included the first assessment of the effect of the contaminating matter from correlating systems, which we find complicates inferring the gas profile towards smaller mass

systems. With as few as 100 FRBs and optical follow-up to find foreground galaxies, we predict that the DM profile can be well constrained over an impact parameter of  $0.3 - 1 r_{\text{vir}}$  for halo masses of  $10^{10} - 10^{13} M_{\odot}$ , with the sensitivity in detail depending on the true gas profile. The gas profile outside of the virial radius would also be measured with such a sample, although the two-halo term from correlating gas is generally dominant even over our gas profiles that distribute the most gas at  $\gtrsim 1 r_{\text{vir}}$  radii (and for sub-Milky Way halos this can also be the case even at  $< 1 r_{\text{vir}}$ ). Unlike other CGM observables which are most sensitive to large halo masses, the limiting factor for measuring the halo gas profile towards even dwarf galaxies is having a deep enough survey to identify these galaxies; if they can be identified, we conclude that their gas profile can be measured statistically with FRBs. With the likely “exponential” increase of FRB detections with localizations, with CHIME and HIRAX predicting thousands in the next few years (Petroff et al. 2021), our estimates suggest these measurements have the potential to finally constrain the total gas profile around galaxies of different types.

Our algorithm is not just relevant to FRB DM science that this paper focuses on. With extensions that allow for magnetic fields and small-scale gas inhomogeneities, it could also be used to predict the rotation and scattering measures to FRBs (Prochaska et al. 2019). Furthermore, our technique is relevant for modeling the thermal and kinetic Sunyaev-Zeldovich effects, which also probe the gas profiles around halos. Our tool can straightforwardly be extended to generate fast, flexible models for the Sunyaev-Zeldovich angular power spectra. Additionally, our method could be applied to understand the distribution of metals in the universe owing to galactic feedback. Indeed, a related post-processing of tophat enrichment profiles around simulated galaxies was used in Booth et al. (2012) to quantify the extent of enrichment – a technique they devised because their suite of hydrodynamic simulations could not reproduce the scope of metal enrichment. We think additional investigations in this spirit can be used to better understand enrichment. Finally, our models can be used to model the weak lensing power spectrum, where the uncertain gas distribution around halos is a major systematic for accurately estimating cosmological parameters. Current lensing analyses mainly use the predictions of hydrodynamic simulations to account for this uncertainty (Chisari et al. 2018; Foreman et al. 2020; Osato et al. 2021), whereas our models have more versatility. A next step needed for many of these observables is to make the CGMBrush lightcone images in angular space.

We thank Zachary Hafen for helping with the prescription for the FIRE-inspired profile, G. Mark Voit for help in implementing the precipitation model, and the CosmoSim Database for making the results of simulations publicly available. We acknowledge support from NSF award AST-2007012.

## REFERENCES

- |  |  |
|--|--|
| <p>Amodeo, S., Battaglia, N., Schaan, E., et al. 2021, PhRvD, 103, 063514</p> <p>Bahcall, J. N., &amp; Spitzer, Lyman, J. 1969, ApJL, 156, L63</p> <p>Behroozi, P. S., Conroy, C., &amp; Wechsler, R. H. 2010, ApJ, 717, 379</p> <p>Booth, C. M., Schaye, J., Delgado, J. D., &amp; Dalla Vecchia, C. 2012, MNRAS, 420, 1053</p> | <p>Centrella, J., &amp; Melott, A. L. 1983, Nature, 305, 196</p> <p>Chisari, N. E., Richardson, M. L. A., Devriendt, J., et al. 2018, MNRAS, 480, 3962</p> <p>Connor, L., &amp; Ravi, V. 2021, arXiv e-prints, arXiv:2107.13692</p> <p>Cooray, A., &amp; Sheth, R. 2002, PhR, 372, 1</p> <p>Cordes, J. M., &amp; Chatterjee, S. 2019, ARA&amp;A, 57, 417</p> |
|--|--|

- Davies, J. J., Crain, R. A., Oppenheimer, B. D., & Schaye, J. 2020, *MNRAS*, 491, 4462
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *ApJ*, 292, 371
- Deng, W., & Zhang, B. 2014, *ApJL*, 783, L35
- Faucher-Giguère, C.-A., Feldmann, R., Quataert, E., et al. 2016, *MNRAS*, 461, L32
- Fielding, D., Quataert, E., McCourt, M., & Thompson, T. A. 2017, *MNRAS*, 466, 3810
- Foreman, S., Coulton, W., Villaescusa-Navarro, F., & Barreira, A. 2020, *MNRAS*, 498, 2887
- Greco, J. P., Hill, J. C., Spergel, D. N., & Battaglia, N. 2015, *ApJ*, 808, 151
- Hafen, Z., Faucher-Giguère, C.-A., Anglés-Alcázar, D., et al. 2019, *MNRAS*, 488, 1248
- Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, *MNRAS*, 480, 800
- Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, *ApJ*, 740, 102
- Kravtsov, A. V. 2013, *ApJL*, 764, L31
- Lee, K.-G., Ata, M., Khrykin, I. S., et al. 2021, *arXiv e-prints*, arXiv:2109.00386
- Li, M., & Tonnesen, S. 2020, *ApJ*, 898, 148
- Lochhaas, C., Bryan, G. L., Li, Y., Li, M., & Fielding, D. 2020, *MNRAS*, 493, 1461
- Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, *Science*, 318, 777
- Macquart, J. P., Prochaska, J. X., McQuinn, M., et al. 2020, *Nature*, 581, 391
- Maller, A. H., & Bullock, J. S. 2004, *MNRAS*, 355, 694
- McCourt, M., Sharma, P., Quataert, E., & Parrish, I. J. 2012, *MNRAS*, 419, 3319
- McQuinn, M. 2014, *ApJL*, 780, L33
- . 2016, *ARA&A*, 54, 313
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, *ApJ*, 462, 563
- Osato, K., Liu, J., & Haiman, Z. 2021, *MNRAS*, 502, 5593
- Peacock, J. A., & Smith, R. E. 2000, *MNRAS*, 318, 1144
- Petroff, E., Hessels, J. W. T., & Lorimer, D. R. 2019, *A&A Rv*, 27, 4
- . 2021, *arXiv e-prints*, arXiv:2107.10113
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013, *A&A*, 557, A52
- Prochaska, J. X., Macquart, J.-P., McQuinn, M., et al. 2019, *Science*, 366, 231
- Ravi, V., Catha, M., D’Addario, L., et al. 2019, *Nature*, 572, 352
- Saro, A., Liu, J., Mohr, J. J., et al. 2014, *MNRAS*, 440, 2610
- Schaan, E., Ferraro, S., Vargas-Magaña, M., et al. 2016, *PhRvD*, 93, 082002
- Schneider, A., & Teyssier, R. 2015, *JCAP*, 2015, 049
- Sharma, P., McCourt, M., Quataert, E., & Parrish, I. J. 2012, *MNRAS*, 420, 3174
- Simha, S., Tejos, N., Prochaska, J. X., et al. 2021, *arXiv e-prints*, arXiv:2108.09881
- Somerville, R. S., & Davé, R. 2015, *ARA&A*, 53, 51
- Stern, J., Fielding, D., Faucher-Giguère, C.-A., & Quataert, E. 2019, *MNRAS*, 488, 2549
- Su, K.-Y., Hopkins, P. F., Hayward, C. C., et al. 2020, *MNRAS*, 491, 1190
- Sunyaev, R. A., & Zeldovich, Y. B. 1970, *Ap&SS*, 7, 3
- . 1972, *Comments on Astrophysics and Space Physics*, 4, 173
- Thom, C., Tumlinson, J., Werk, J. K., et al. 2012, *ApJL*, 758, L41
- Thornton, D., Stappers, B., Bailes, M., et al. 2013, *Science*, 341, 53
- Tumlinson, J., Peebles, M. S., & Werk, J. K. 2017, *ARA&A*, 55, 389
- Voit, G. M. 2019, *ApJ*, 880, 139
- Voit, G. M., Donahue, M., Zahedy, F., et al. 2019, *ApJL*, 879, L1
- Voit, G. M., Ma, C. P., Greene, J., et al. 2018, *The Astrophysical Journal*, 853, 78
- Voit, G. M., Meece, G., Li, Y., et al. 2017, *ApJ*, 845, 80
- Werk, J. K., Prochaska, J. X., Tumlinson, J., et al. 2014, *ApJ*, 792, 8
- Zhu, W., Feng, L.-L., & Zhang, F. 2018, *ApJ*, 865, 147



## APPENDIX

## A. THE SCALES THAT CONTRIBUTE TO THE VARIANCE OF THE BARYONIC FIELD

In this appendix, we discuss what structures contribute to the variance of the cosmological baryon field. We show that the variance of the dispersion measure mainly comes from gas in galactic dark matter halos and, to a lesser extent, gas that traces  $\approx 100$  Mpc cosmic structures. Structures that are somewhat less diffuse than halos are less important to model, which is fortunate because the method detailed in this paper is the least consistent in modeling them. A similar conclusion holds for the 3D electron power spectrum at all wavenumbers.

The left panel in Fig. 6 shows  $k^2 P_e / (2\pi)^2$  at  $z = 0.5$  for models where different halos retain their gas with the same NFW halo profile of the dark matter above the specified halo mass and that halos are largely evacuated below this mass. These calculations use the halo model (e.g. Cooray & Sheth 2002).<sup>8</sup> Here  $k$  is the wavenumber and  $P_e$  the electron overdensity power spectrum with the standard Fourier convention in cosmology such that the area under the curves is proportional to the variance of DM (McQuinn 2014):

$$\sigma_{\text{DM}}^2 \approx \int_0^z \frac{cdz}{H(z)} (1+z)^2 \bar{n}_{e,0}^2 \int d \log k \left[ \frac{k^2 P_e(k, z)}{(2\pi)^2} \right], \quad (\text{A1})$$

where  $\bar{n}_{e,0}$  is the cosmic mean electron number density today. The left panel shows that much of the variance potentially owes to the profiles in galactic mass halos. A lesser contribution to the variance is from large-scale structure shown by the thick black curve. (This thick black curve is the halo model's ‘two halo term’.) Much of the area under this curve is at  $2\pi/100\text{Mpc}^{-1} < k < 2\pi/10\text{Mpc}^{-1}$ .

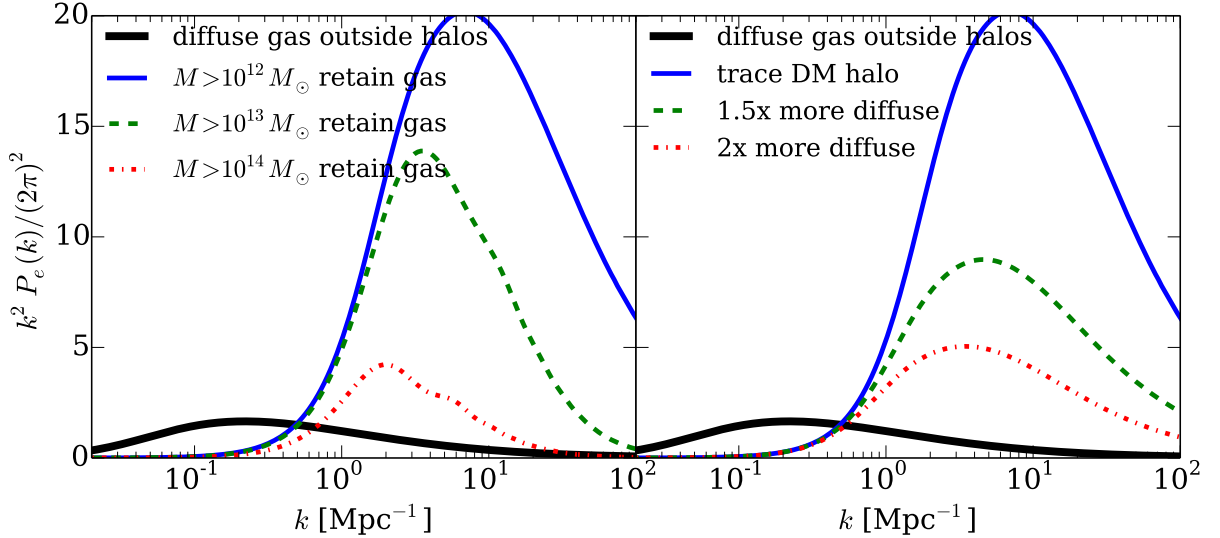
Large-scale structure becomes more important as the gas around halos becomes more diffuse. The right panel in Fig. 6 considers the previous model where halos with mass satisfying  $M > 10^{12} M_\odot$  retain their gas, but where an illustrative model for the distribution of the gas is assumed where it has an NFW profile that is dilated by a factor of 1, 1.5 and 2 compared to the dark matter. This illustrates how the more diffuse the gas, the smaller contribution to the variance from sightlines intersecting individual halos, and the relatively larger contribution of the large-scale structure term.

In both panels, the large-scale structure term is calculated assuming the linear theory density power spectrum, as is traditional in the halo model. The linear power spectrum is inaccurate of course on nonlinear wavenumbers of  $k \sim 2\pi/(1-10)\text{Mpc}^{-1}$ , but fortunately these scales are not a significant contribution to the variance. That ‘two halo correlations’ are a somewhat small contribution to the power at  $k \sim 2\pi/(1-10)\text{Mpc}^{-1}$  also holds for other baryonic clustering statistics and not just those most relevant for FRBs. Indeed, the unimportance of few Mpc scale structures is the reason why the halo model successfully describes the 3D nonlinear matter power spectrum even at mildly nonlinear wavenumbers.

Since the method presented in this paper uses cosmological simulations, it does a much better job than the halo model at capturing nonlinear matter clustering. However, on  $\sim 1$  Mpc scales the baryons can be affected by feedback, which we are not modeled except for gas that is associated with

<sup>8</sup> Our calculation uses the same code as McQuinn (2014), except here the Press-Schechter rather than Sheth-Tormen mass function is used. Press-Schechter somewhat overpredicts the abundance of rare massive halos and underpredicts less rare ones.

a dark matter halo. We are certainly not modeling the interplay between more pristine cosmological gas and outflows from galactic halos. Our justification is that regions somewhat outside of halos are less important for capturing many statistical measures of the cosmic baryon field.

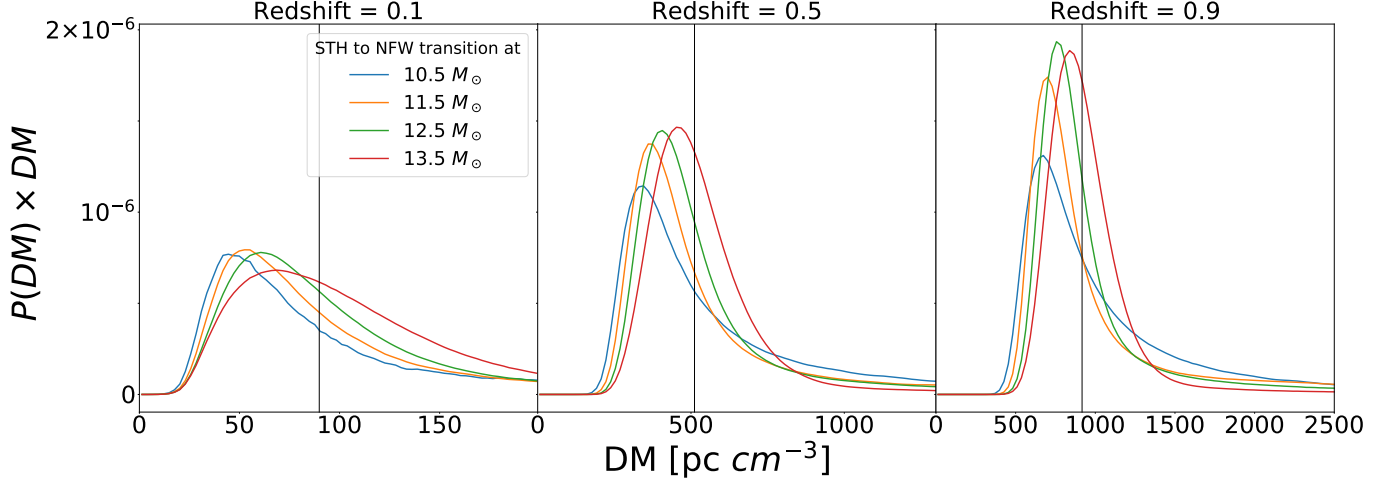


**Figure 6. Left panel:** Halo model calculation for the  $z = 0.5$  contribution of the nonlinear power spectrum of baryons assuming a toy model where they trace the NFW halo profile above the specified halo mass and where halos are largely evacuated below this mass. The colored curves show the one halo term that depends on the profile, and the black shows the two halo term (which traces the large-scale matter field). The area under these curves is proportional to the variance of DM (eqn. A1). **Right panel:** Considering only the model where  $> 10^{12} M_{\odot}$  halos retain their gas, but where the distribution of the gas is assumed to have an NFW profile that extends 1.5 or 2 times more in radius than the dark matter halo. This illustrates how the more diffuse the gas, the smaller the contribution of this term.

## B. EFFECT OF HALO MASS SCALES ON PDF

Here we try to understand what mass scales matter to the variance of the PDF of DM, using a similar toy model for baryonic gas as the previous appendix. In fig. 7, we plot a series of PDFs in which we assume a  $2r_{\text{vir}}$  spherical tophat profile below a mass threshold and dark-matter-tracing NFW profile above. This transition has the effect of essentially eliminating the ‘1-halo’ contribution of the small mass halos to the width of the PDF. The shift of the PDFs as a result of increasing the mass threshold shows that all mass scales matter from  $10^{13.5} M_{\odot}$  down to  $10^{10.5} M_{\odot}$  halos. There is a small number of halos above the highest mass bin, their contribution to the DM is significant as observed by the shift in the curves from  $10^{12.5} M_{\odot}$  to  $10^{13.5} M_{\odot}$ . While no mass scales are negligible, we observe a gradual convergence towards the smallest halos, which are host to dwarf galaxies that are unable to hold onto the gas flung out but supernovae and AGNs. These PDFs are evaluated using the baryonic redistribution method as described in the paper and applied to the publicly-available Bolshoi simulation at *CosmoSim*<sup>9</sup>

<sup>9</sup> <https://www.cosmosim.org/cms/simulations/bolshoi/>

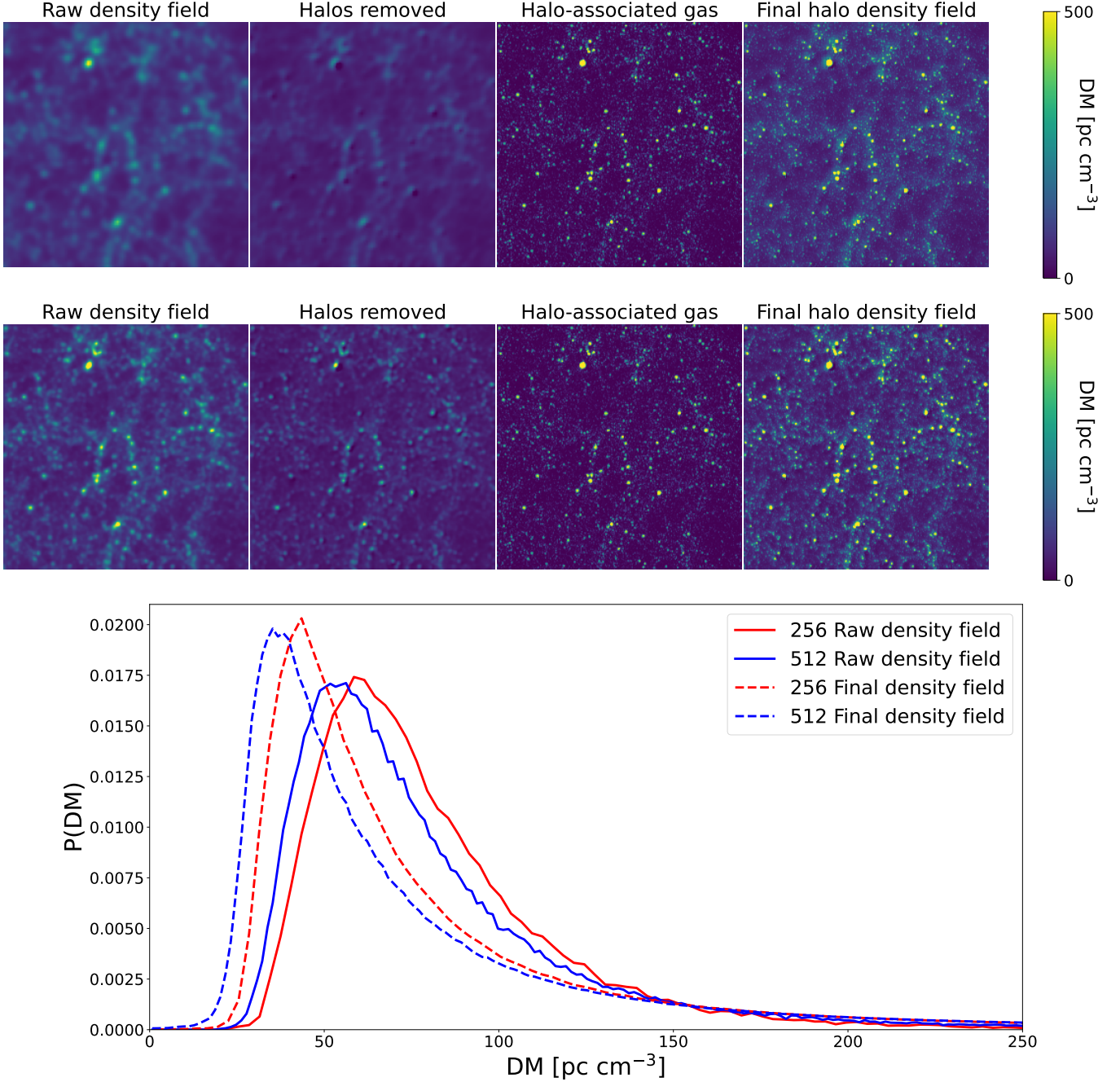


**Figure 7.** The PDFs are plotted for a prescription where we transition from a  $2r_{\text{vir}}$  spherical tophat profile to an NFW profile above the mass thresholds listed in the legend. The former profile take as the maximally evacuated case and the later NFW as the maximally concentrated. This transition has the effect of essentially eliminating the halo profile contribution to the PDM (the ‘1-halo’ contribution). This toy model illustrates where different mass halos likely contribute to the PDF and what mass scales need to be captured to reach a converged result.

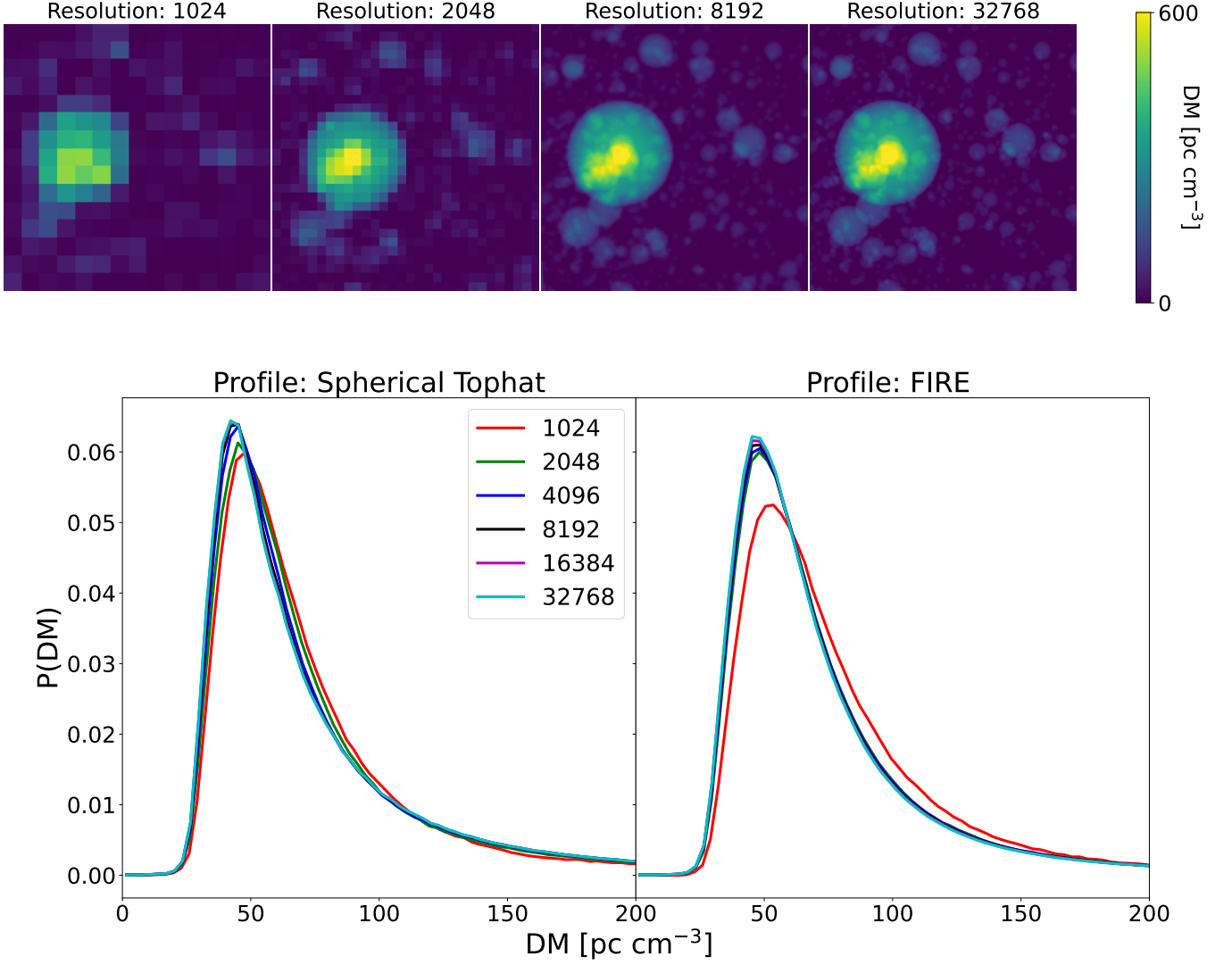
### C. RESOLUTION TESTS

We first consider how sensitive our results are to the resolution that the  $N$ -body simulation is gridded. Figure 8 shows the steps in our calculations for two different Eulerian grids for the  $N$ -body outputs. The top panel is for  $N = 256$  and the bottom panel for  $N = 512$ . Visually, we can see that the differences are modest in the final algorithmic output in the right most panels. More quantitatively, we test convergence for the two base grid resolutions by creating a probability distribution function of the gas column density ‘DM’ across the simulation snapshots shown in the top panel in figure 1. We have done this for both the  $1r_{\text{vir}}$  tophat profile and the FIRE simulation. Both panels show that the two fields result in similar PDFs with the same standard deviation after the halos are re-added of  $\approx 38$  for  $N = 256$  and  $\approx 51$  for  $N = 512$ . We take this to mean that our algorithm is only somewhat sensitive to the base gridding. The calculations in the main text use  $N = 256$ .

We also consider how sensitive our results are to the resolution of our up-sampled fine grid that we create to add halos back to the grid. In the top panel of Figure 9, where we have added the halos back by using a tophat profile (described in more detail in 2.2), we can see that as we increase the resolution to  $\eta = 32$  in the second image, which corresponds to resolution of  $\eta N = 8192$ , not only do we resolve the internal structure of larger halos of size  $r_{\text{vir}} \approx 1$  Mpc significantly, but also of the much smaller halos that were previously not well resolved as seen the leftmost image in the panel with  $\eta = 16$ . The leftward shift of the PDFs in the bottom panel of section 9 shows the effect of increasing resolution. The differences are quite small, with convergence by  $\eta = 32$  corresponding to a resolution of 8,192. All calculations in the main text use  $\eta = 32$  or higher resolution.



**Figure 8.** Illustration of method on two Eulerian base grid resolutions: the fiducial  $N^2 = 256^2$  grid (top) and the higher resolution  $N^2 = 512^2$  grid (middle). The images show the halo subtraction and addition through different stages of our method, applied to the  $z = 0$  snapshot of the Bolshoi  $N$ -body simulation. Each panel is projected over the simulation boxsize of  $250h^{-1}\text{Mpc}$  and zooms in on a subregion with transverse comoving size of 140 Mpc. The left panels shows the raw gridded density field from the simulation. From this density field, we subtract the halo density field, and the result is shown in the second column. The third column shows the halo density field that we add back to the subtracted field, where a fine Eulerian grid of  $N_f^2 = 8194^2$  is used. The last panel shows the net field after the finely-gridded halos are added back to the coarse subtracted field. In this case, we use our simple  $1 r_{\text{vir}}$  top-hat model, as the differences are smaller for more concentrated models. The bottom panel is a probability distribution of the gas column density, ‘DM’, across the  $250h^{-1}\text{Mpc}$  box.



**Figure 9.** The images at the top show how the PDF converges when increasing the resolution of halo addition grid. Each image is computed by projecting the  $z = 0$  box over  $250h^{-1}$  Mpc and has a transverse comoving size of  $5h^{-1}$  Mpc. The bottom panel shows how PDFs for different profiles tend towards convergence as we increase the resolution.