Multitracer Cosmological Line Intensity Mapping Mock Light-cone Simulation

Shengqi Yang1, Rachel S. Somerville2,3, Anthony R. Pullen1,2, Gergö Popping4, Patrick C. Breysse1,2, and Abhishek S. Maniyar1

1 Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA; sy1823@nyu.edu
2 Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA
3 Department of Physics and Astronomy, Rutgers University, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA
4 European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748, Garching, Germany

Received 2020 September 24; revised 2021 February 19; accepted 2021 March 6; published 2021 April 27

Abstract

Submillimeter emission lines are important tracers of the cold gas and ionized environments of galaxies and the targets for future line intensity mapping surveys. Physics-based simulations that predict multiple emission lines arising from different phases of the interstellar medium are crucial for constraining the global physical conditions of galaxies with upcoming line intensity mapping observations. In this work, we present a general framework for creating multitracer mock submillimeter line intensity maps based on physically grounded galaxy formation and submillimeter line emission models. We simulate a mock light cone of 2 deg² over a redshift range 0 ≤ z ≤ 10 comprising discrete galaxies and galaxy [C II], CO, and [C I] emission. We present simulated line intensity maps for two fiducial surveys with resolution and observational frequency windows representative of COMAP and EXCLAIM. We show that the star formation rate and line emission scaling relations predicted by our simulation significantly differ at low halo masses from widely used empirical relations, which are often calibrated to observations of luminous galaxies at lower redshifts. We show that these differences lead to significant changes in key summary statistics used in intensity mapping, such as the one-point intensity probability density function and the power spectrum. It will be critical to use more realistic and complex models to forecast the ability of future line intensity mapping surveys to measure observables such as the cosmic star formation rate density.

Unified Astronomy Thesaurus concepts: Intergalactic medium (813); Diffuse radiation (383); Large-scale structure of the universe (902)

1. Introduction

Cosmic microwave background experiments and galaxy surveys have made great progress in advancing our understanding of the origin and evolution of different components of our universe. However, the wealth of data provided by these measurements raises more profound questions. For example, whether the observed accelerating expansion of the universe should be explained by the presence of dark energy or the breakdown of general relativity on a cosmological scale is still under debate (Perlmutter et al. 1999; Dvali et al. 2000; Peebles & Ratra 2003; Carroll et al. 2004). On the astrophysics side, the cause of the cosmic galaxy star formation rate (SFR) density deviating from the continuous growth of dark matter (DM) halos from redshift z ≈ 2 to the present is still unclear (Madau & Dickinson 2014). Moreover, our knowledge of the conditions in the interstellar medium (ISM) and the effects on the global properties of galaxies is limited, which is crucial for understanding the star formation (SF) and galaxy evolution process (Carilli & Walter 2013). Answering these questions requires measurements with higher resolution, larger observational volumes, or new experimental designs.

Line intensity mapping (LIM) is an emerging technique to advance our understanding of both cosmology and extragalactic astronomy in the next decade (Kovetz et al. 2017). Unlike galaxy surveys that resolve individual sources, LIM integrates all of the emission along the line of sight, including the signal contributed by faint sources. The advantages of LIM are threefold. First, LIM can probe vast cosmological volumes at high redshifts, where the emitters are generally too faint to be resolved in galaxy surveys. This feature not only allows tests of gravity on cosmological scales but also provides information about the large-scale structure distributions from the present time all the way back to the epoch of reionization (EoR). Second, since all of the emission along the line of sight is collected in the LIM experiments, LIM surveys sample the entire galaxy population, while traditional galaxy surveys are biased toward the brightest sources. This feature is important for inferring ISM and SF properties at different cosmic eras. Lastly, by not attempting to resolve individual sources, LIM only requires a modest telescope aperture size, which is more economical.

 Originally, LIM was conceived to map the 21 cm line, an important tracer of the matter distribution, emitted by neutral hydrogen during the EoR, as well as the postreionization epoch. The molecular and fine-structure lines emitted from galaxies, such as [C II] and CO lines, have also attracted interest, leading to the design of several LIM experiments. Since different emission lines are unique tracers of the corresponding gas phases, multitracer LIM studies will provide new information about ISM conditions for various cosmic times. In the next decade, numerous LIM surveys will be conducted. Some examples are HIRAX (Newburgh et al. 2016), CHIME (Amiri et al. 2017), and HERA (DeBoer et al. 2017), which probe 21 cm; SPHEREx (Doré et al. 2014), targeting Lyα and Hα lines; and future CO LIM experiments COMAP (Li et al. 2016), TIME (Crites et al. 2014), TIM (Hailey-Dunsheath et al. 2018), CONCERTO (Dumitru et al. 2019), and EXCLAIM (Ade et al. 2020), measuring [C II] emission.

Two unsolved problems for LIM survey data analysis are (1) disentangling the target emission line signal from the Milky Way (MW) foreground, cosmic infrared background (CIB),
interloper lines, and other contamination and (2) connecting the measured LIM signal to the physical properties of line emitters, as well as cosmological quantities of interest. Analytic models and numerical simulations are powerful tools to help answer these questions. Compared to numerical simulations, analytic models of line emission enjoy the advantage of higher computational efficiency but suffer from many limitations. First, empirical models are often calibrated only to local galaxy measurements; thus, they cannot be extended confidently to higher redshifts. Theoretical line emission models that are derived from the statistical balance equation generally make assumptions about the ISM properties and only consider single line emission; thus, estimating the cross-correlations between multiple lines becomes challenging and requires extra assumptions about the correlation index. Numerical hydrodynamic cosmological simulations are based on a more robust underlying physics model, but due to the computational cost, they still must make trade-offs between volume and resolution. For example, the high-resolution FIRE simulations (Hopkins et al. 2014) resolve galaxies with stellar masses down to $10^7 M_\odot$ and can partially resolve the multiphase ISM, but each FIRE simulation only represents one DM halo. Large-volume hydrodynamical simulations such as IllustrisTNG (Nelson et al. 2019), EAGLE (McAlpine et al. 2016), and SMIBA (Davé et al. 2019) can typically represent volumes of 50–300 Mpc on a side, but the mass resolution is limited to $\sim 10^6 – 10^7 M_\odot$, and phenomenological “subgrid” models must be used to treat processes such as SF, stellar feedback, and black hole growth and feedback (see Somerville & Davé 2015 for a review). The quick particle mesh (QPM) mocks used in eBOSS (White et al. 2014) are simulated in a huge box with 2.56 Gpc $h^{-1}$ side$^{-1}$, while the DM halo mass resolution is $10^9 h^{-1} M_\odot$—insufficient to resolve individual galaxies. In all cases, a single cosmological simulation cannot provide information about the large-scale structure ($\sim 100$ Mpc) and the conditions in the ISM that influence line emission ($\sim$ parsecs) simultaneously.

One option to cut down the computational expense is to apply empirical relations between DM halos and galaxy properties, such as SFR, and then use empirically calibrated scaling relations to translate this to line emission. This approach has been used quite extensively in the literature (e.g., Silva et al. 2015; Yue et al. 2015; Li et al. 2016; Serra et al. 2016; Fonseca et al. 2017; Ille et al. 2019; Padmanabhan 2019). The drawbacks to this method, as mentioned before, are that the empirical models are calibrated to observations over a certain luminosity and redshift range but applied over broader ranges in these quantities, over which the models are not well tested. Moreover, empirical line emission models generally focus on a single tracer and do not self-consistently predict multiple lines. They therefore fail to exploit the exciting potential of using multiple tracers for LIM measurements.

An intermediate approach between empirical halo models and numerical simulations is the semianalytic model (SAM) approach. Similar to an $N$-body simulation, an SAM dynamically evolves an ensemble of DM and baryonic components in a cosmological context. To improve the simulation efficiency, the SAM solves the numerically complex, nonlinear physical processes involved in DM halo and galaxy evolution through simplified but physically motivated treatments, which are calibrated to more detailed numerical simulations. The SAMs adopt simplifying assumptions for evolving DM and baryons and, like large-volume hydrodynamic simulations, generally adopt phenomenological recipes to treat processes such as SF and stellar feedback. These recipes contain free parameters that are calibrated to match global observational quantities such as the stellar mass function, galaxy gas fractions, mass–metallicity relation, etc. Therefore, although SAMs have been quite successful at matching a broad range of observations over cosmic time, there are remaining uncertainties about whether inaccuracies in model assumptions might be partly compensated by the freedom to tune free parameters. Studies that compare $N$-body/hydro simulations with SAMs show overall agreement in key global quantities and their evolution over cosmic time (Somerville & Davé 2015), although there are still significant discrepancies among the gas properties and SF efficiencies (Benson et al. 2001; Hirschmann et al. 2012; Mitchell et al. 2018). Popping et al. (2019b) found tension between observed galaxy H$_2$ masses at high redshift and predictions from both an SAM and IllustrisTNG but fairly close agreement between the theoretical predictions from the two methods. Overall, SAMs represent a promising method for providing mock data for upcoming large surveys and LIM experiments.

The SAM we choose in this work is the Santa Cruz SAM, developed by Somerville & Primack (1999), Somerville et al. (2008a, 2012), Porter et al. (2014), Popping et al. (2014), and Somerville et al. (2015). The Santa Cruz SAM partitions ISM gas into atomic, molecular, and ionized phases and adopts an H$_2$-based SF recipe motivated by the observed correlation between SFR and molecular gas density (Bigiel et al. 2008, 2011; Schruba et al. 2011). The Santa Cruz SAM successfully reproduces various key UV/optical galaxy observations for redshift $z < 6$ (Somerville et al. 2012, 2015) and has been shown to be in excellent agreement with available observations from $z \sim 6$–10, as well as the reionization history of the universe. The Yung et al. (2019a, 2019b, 2020a, 2020b) paper series also utilizes the Santa Cruz SAM to make high-redshift predictions for the upcoming James Webb Space Telescope.

Popping et al. (2016, 2019a) developed a tool that couples to the Santa Cruz SAM (hereafter referred to as the “submillimeter SAM”) to simulate multiple submillimeter line luminosities for each simulated galaxy. This model has been shown to be successful in reproducing available observations of [C II], CO, and [C I] luminosity versus SFR and stellar mass across various cosmic times back to $z \sim 6$. The combined Santa Cruz SAM and submillimeter SAM pipeline is simultaneously highly computationally efficient yet grounded in physics and able to self-consistently predict a broad suite of observable tracers. It is therefore a particularly powerful tool for generating multiwavelength source catalogs, which can be used to generate mock LIM maps.

In this work, we construct a 2 deg$^2$ light cone over the redshift range $0 \leq z \leq 10$ using DM halos from the Small MultiDark-Planck (SMDPL) $N$-body simulation (Klypin et al. 2016). We then use the Santa Cruz SAM to estimate the merger history for each DM halo and simulate the properties and distribution of galaxies within it. We then apply the submillimeter SAM to estimate the [C II], CO, and [C I] line luminosities for each individual galaxy. By integrating along lines of sight along the light cone, we create synthetic maps over frequency ranges of interest. To the best of our
knowledge, this is the first LIM simulation that self-consistently models multiple far-infrared (FIR) emission lines and links them with UV–optical–near-IR properties of a discrete galaxy source catalog. This mock light-cone catalog is directly relevant to various LIM surveys and will be valuable for future LIM survey design and analysis pipeline development.

The CO emission is an excellent tracer of ISM molecular gas, which is strongly correlated with SF activity in galaxies (Bolatto et al. 2013). Compared with other submillimeter emission lines, the unique advantage of studying CO emission is that the CO molecule simultaneously emits multiple lines from a “ladder” of rotational transitions. Therefore, cross-correlation between the CO emission lines observed in different frequency channels could help remove uncorrelated foreground and interloper contamination and provide rich information about the CO emitters. Several CO $J = 1$–0 empirical models have been proposed in the last decade (Rigby et al. 2008; Visbal & Loeb 2010; Pullen et al. 2013; Li et al. 2016). We will compare the Santa Cruz SAM + submillimeter SAM CO $J = 1$–0 predictions with various empirical CO $J = 1$–0 models and study how CO $J = 1$–0 LIM statistics vary under different models. In this work, we create a mock map with fiducial characteristics similar to the COMAP pathfinder survey, which will probe CO $J = 1$–0 at $z = 2.4$–2.8.

The fine-structure line emitted by ionized carbon, [C II], is another strong tracer of dense gas and SF. The [C II] is the brightest FIR line, contributes 0.1%–1% of the FIR luminosity of the nuclear region of galaxies, and has been modeled analytically by many groups (e.g., Gong et al. 2012; Silva et al. 2015; hereafter Silva15; Pullen et al. 2018; Yang et al. 2019; Sun et al. 2019; Chung et al. 2020). We also create a second fiducial mock map with characteristics representative of the EXCLAIM survey (Ade et al. 2020), which will probe [C II] emitters at $z = 2.5$–3.6.

Two of the most commonly used summary statistics for LIM are the power spectrum and the probability density function (PDF) of intensity values in voxels, also referred to as the one-point intensity PDF or voxel intensity distribution (VID). The VID is a potentially powerful summary statistic based on well-known $P(D)$ analysis methods (Scheuer 1957) for LIM that can provide constraints on the luminosity function of the target line emitters, as well as the SFR density (Bryeysse et al. 2016, 2017). Bryeysse et al. (2019) further proposed that combining a one-point LIM PDF analysis with galaxy surveys, a statistic called the conditional VID (CVID), can not only constrain physical processes but also remove uncorrelated foregrounds, such as the MW continuum emission and the interloper line contamination. Ihle et al. (2019) showed that a joint analysis combining both the power spectrum and VID can yield stronger constraints than either approach independently. In this paper, we present VID and power spectrum predictions computed directly from our fiducial COMAP and EXCLAIM mock maps.

The plan of this paper is as follows. In Section 2, we summarize the method used to create the mock light cone, populate it with sources, and create synthetic maps. Specifically, we briefly introduce the SMDPL $N$-body simulation in Section 2.1. The method to construct a light cone from an $N$-body simulation is explained in Section 2.2. We introduce the Santa Cruz SAM in Section 2.3 and the submillimeter SAM in Section 2.4. The FIR dust emission model is introduced in Section 2.5. We explain how we make mock intensity maps for the submillimeter lines, FIR emission, and MW continuum emission in Section 2.6. Finally, we introduce two fiducial survey designs considered in this work in Section 2.7. In Section 3, we summarize the main results and compare them with observations and other empirical models. We conclude in Section 4. Throughout this work, we assume cosmological parameters consistent with the SMDPL cosmology: $\Omega_m = 0.307$, $\Omega_b = 0.048$, $\Omega_{\Lambda} = 0.693$, $\sigma_8 = 0.829$, $n_s = 0.96$, and $h = 0.678$.

2. Tools and Methods

2.1. SMDPL $N$-body Simulation

The volume of a 2 deg$^2$ light cone along the redshift range $0 \leq z \leq 10$ is about $5.7 \times 10^7$ (Mpc/h)$^3$. To ensure the statistical independence of each region within the mock light cone, we require the $N$-body simulation volume to be no less than the target light-cone volume. Moreover, we expect that below a critical mass, halos will not be able to accrete or retain significant gas reservoirs, so that only halos with masses larger than this critical mass will be relevant for simulating detectable line emission. We therefore require the halo mass resolution in the $N$-body simulation to be at least $10^{10} M_\odot$. A detailed justification of this mass resolution choice is presented in the Appendix. Due to these two aspects of consideration, in this work, we choose the SMDPL cosmological $N$-body simulation for the light-cone construction.

The SMDPL contains $3840^3$ DM particles within a (400 Mpc/h)$^3$ cube and simulates the evolution of DM particles from redshift $z = 19$ to 0. At different redshifts during the simulation, 117 snapshots are taken, with denser sampling in the lower part of the redshift range. The SMDPL assumes a standard $\Lambda$CDM cosmology with cosmological parameters $\Omega_m = 0.307$, $\Omega_b = 0.048$, $\Omega_{\Lambda} = 0.693$, $\sigma_8 = 0.829$, $n_s = 0.96$, and $h = 0.678$. Among all of the Bolshoi/Multidark simulations, SMDPL uses the smallest particle mass, $9.6 \times 10^6 M_\odot$, and has the highest halo mass resolution of $10^{10} M_\odot$. More details can be found in Klypin et al. (2016).

2.2. The DM Halo Light Cone

We adopt the method proposed in Blaizot et al. (2005) to construct a mock light cone of DM halos from the SMDPL $N$-body simulation. As a brief summary, we first apply periodic boundary conditions to the $N$-body data cube. We then randomly select an origin point and a direction for the line of sight to cut out a light cone with a solid angle of 2 deg$^2$ and redshift range $0 \leq z \leq 10$. The redshift of each halo is determined by its comoving distance from the origin and its peculiar velocity along the line of sight. Note that the $N$-body simulation provides snapshots at a set of redshifts $z_1, z_2, \ldots, z_N$, where $N$ is the number of snapshots. For a halo with redshift $z_h \leq z_{\text{halo}} = |z_1 + z_{\text{halo}}|/2$, we read the halo properties from $N$-body snapshot $z_h$. Of the 117 SMDPL snapshots in the redshift range $0 \leq z \leq 10$, 108 are used for the DM halo light-cone construction.

Due to the periodic boundary conditions, DM halos across the mock light cone will generate a repeating pattern and gain extra spatial correlations. To suppress this replication effect, we randomly shift, rotate, and invert the DM halo position in each $N$-body 3D catalog copy while stacking them together. Although this random shuffling breaks the continuity of the
DM overdensity field and adds negative bias to the spatial correlation function, the bias can be accurately estimated for scales smaller than 20% of the box size (Blaiszot et al. 2005).

We adapted the code for the light-cone construction from the one provided in Peter Behroozi’s UNIVERSE_MACHINE package (Behroozi et al. 2019).

2.3. Santa Cruz SAM

In this work, we use the Santa Cruz SAM described in Somerville & Primack (1999), Somerville et al. (2008a, 2012), Porter et al. (2014), Popping et al. (2014), and Somerville et al. (2015) to simulate physical and observable properties of galaxies and how they evolve self-consistently over cosmic time. The Santa Cruz SAM is a comprehensive galaxy formation model that uses a simplified but physical treatment of the key processes that shape galaxy evolution. It divides cold gas into ionized, atomic, and molecular phases and applies an H2-based SF recipe. It also contains a model accounting for DM halo merging history, evolution of subhalos and galaxy mergers, shock heating and radiative cooling of hot gas within virialized DM halos, supernova feedback, black hole growth and active galactic nucleus (AGN) feedback, photoionization squelching, and other processes involved in galaxy evolution. The evolution of galaxies is tightly related to the merger history of the DM halo. The halo merger history is commonly represented by “merger trees,” which can be either extracted directly from N-body simulations or estimated through other semianalytic formalism. In this work, the Santa Cruz SAM estimates the merger history through a multibranch tree algorithm based on the extended Press–Schechter (EPS) formalism (Somerville & Kolatt 1999). The advantage of using the EPS formalism over extracting the DM halo merger history from N-body simulations is that merger trees provided by N-body simulations have limited mass resolution, while the EPS formalism can extend merger trees to progenitors with arbitrarily small mass. In this work, we record “root halos” down to M_{root,min} = 10^{10} M_{\odot}, and follow merger histories down to a hundredth of the root mass, or 10^{9} M_{\odot}, whichever is smaller. It was shown in Porter et al. (2014) that the predictions of the SAM when using N-body–based merger trees from the Bolshoi simulation are very similar to the EPS-based predictions.

Before the EoR, the SAM assumes that each DM halo contains hot gas with a mass equal to the baryonic mass. After the universe is fully ionized by z = 11, a fraction of the baryonic mass is allowed to accrete into the halo, based on the filtering mass obtained from hydrodynamic simulations by Okamoto et al. (2008). The hot gas then experiences radiative cooling and collapses into the central galaxy, where it is assumed to form a rotationally supported disk. The disk size is estimated following Somerville et al. (2008b), and the radial distribution of the cold gas disk is described by an exponential profile.

In the most up-to-date version of the Santa Cruz SAM (Popping et al. 2014; Somerville et al. 2015), cold gas in the galaxy disk is divided into ionized (H II), atomic (H I), and molecular (H2) phases. The SFR surface density is modeled by molecular hydrogen-based recipes that are calibrated to observations. Popping et al. (2014) and Somerville et al. (2015) explored a variety of different recipes for gas partitioning and SF efficiency. They found that the metallicity-based, UV background–dependent recipe based on Gnedin & Kravtsov (2011; the GK model) combined with an H2-based SF relation with a density-dependent slope (the Bigiel2 model) gives results that best match observations from the local universe to z = 4. Yung et al. (2019a, 2019b, 2020a, 2020b) then confirmed that this model produces the best agreement with observations up to z ~ 10. We therefore adopt the GK + Bigiel2 model in this work.

Other ingredients, such as stellar feedback, heavy-element generation, and black hole growth, are also included. We refer readers to Somerville et al. (2008a), Popping et al. (2014), and Somerville et al. (2015) for more details. Other galaxy evolution parameters are identical to the values presented in Somerville et al. (2015) and Popping et al. (2019a).

2.4. Submillimeter Emission Line Modeling

The major source of the emission lines we consider in this LIM simulation, i.e., the [C II], CO, and [C I] lines, is dense molecular clouds (MCs) in the ISM. The Santa Cruz SAMs predict the scale length of the cold ISM gas in the disk and the fraction of gas in a dense molecular phase but do not provide predictions on the properties of MCs. In this work, we adopt the subresolution recipe developed in the submillimeter SAM proposed by Popping et al. (2019a, hereafter GP19) to simultaneously model multiple submillimeter lines. Specifically, each simulated galaxy is divided into radial annuli, and for each annulus, the masses of ionized, atomic, and molecular gas are computed following the GK model. The submillimeter SAM then randomly generates MCs with masses in the range 10^{4} M_{\odot} < M_{MC} < 10^{7} M_{\odot} following a power-law mass function:

$$\frac{dN}{dM} \propto M^{-\beta}. \tag{1}$$

It was shown by GP19 that the specific value of \( \beta \) does not influence the line emission predictions much. In this work, we assume \( \beta = 1.8 \) based on local observations of cloud distribution functions. The random MC generating process stops when the total mass of H2 in the simulated MCs reaches the mass of H2 in the corresponding galaxy annulus. The submillimeter SAM then divides each MC into multiple zones and uses DESPOTIC (Krumholz 2014), a code that solves the energetics of optically thick interstellar clouds, to compute the line emission spectrum. DESPOTIC treats each MC zone as a spherical shell with uniform physical and chemical properties. Given the compositions and physical conditions of an MC zone, as well as the external radiation field, DESPOTIC then solves the heating and cooling processes, chemical processes, and profile of the spectral lines. The dominant heating processes are the external radiation field and grain photoelectric heating. The main cooling processes are line cooling and dust thermal radiation. Chemical reactions and the corresponding rate coefficients are provided by a reduced carbon–oxygen chemical network (Nelson & Langer 1999) and a nonequilibrium hydrogen chemical network (Glover & Clark 2007; Glover & Clark 2012). The external ultraviolet (UV) radiation field \( G_{UV} \) and the ionization rate by cosmic rays (CRs) \( \zeta_{CR} \) are scaled according to the local SFR surface density

---

5 https://bitbucket.org/pbehroozi/universe_machine.git
\[ \Sigma_{\text{SFR}} \text{ predicted by the SAM:} \]

\[ G_{UV} = G_{UV,MW} \times \frac{\Sigma_{\text{SFR}}}{\Sigma_{\text{SFR,MW}}} \]

\[ \xi_{\text{CR}} = 0.1 \xi_{\text{CR,MW}} \times \frac{\Sigma_{\text{SFR}}}{\Sigma_{\text{SFR,MW}}} . \]  

Here the MW SFR surface density \( \Sigma_{\text{SFR,MW}} = 790 M_{\odot} \text{ Myr}^{-1}\text{kpc}^{-2} \) (Bonatto & Bica 2011), UV radiation field \( G_{UV,MW} = 9.6 \times 10^{-4} \text{erg cm}^{-2}\text{s}^{-1} \) (Seon & Witt 2012), and CR ionization rate in the diffuse ISM \( \xi_{\text{CR,MW}} = 10^{-16} \text{s}^{-1} \) (Narayanan & Krumholz 2017). Following Narayanan & Krumholz (2017), a factor of 0.1 is introduced to \( \xi_{\text{CR,MW}} \) to account for the CR shielding in the interiors of MCs. We refer readers to Krumholz (2014) for more details about DESPOTIC and GP19 for the parameters we use to compute the line luminosities.

In this work, we grid each MC into 25 zones, which is sufficient for producing convergent [C II], CO, and [C I] luminosities. The density profile assumed for the MCs is another crucial parameter in the line emission simulation. It is assumed by GP19 that all of the MCs have a Plummer density profile, which was shown to produce a range of line luminosity versus SFR relations in the best agreement with the observations available at the time of publication (2018). However, we find that the [C II] luminosities predicted by the original version of a submillimeter SAM are lower than more recent Atacama Large Millimeter/submillimeter Array observations by a factor of \( \sim 3 \) at high redshifts. Valentino et al. (2020) also suggested that at high redshifts, the [C I] luminosity versus infrared luminosity relations \( L_{\text{C I}} - L_{\text{IR}} \) predicted by GP19 are significantly lower than observations. Motivated by these discrepancies, we replace the Plummer density profile of MCs with the power-law density profile introduced in GP19. Variation of the cloud radial density profile does not significantly influence the CO luminosity predictions, but it effectively boosts the [C II] and [C I] luminosities so that all [C II], CO, and [C I] scaling relations predicted by a submillimeter SAM are in better agreement with current observations. A submillimeter SAM also accounts for the atomic diffuse ISM emission by modeling this ISM phase as one-zone clouds. The hydrogen and column densities of the diffuse atomic gas are fixed as \( n_{H} = 10 \text{ cm}^{-3} \) and \( N_{H} = 10^{21} \text{ cm}^{-2} \), respectively. Finally, the luminosity of the [C II], CO, and [C I] lines emitted from all MC zones and galaxy annuli are summed over to provide the total line luminosity of each galaxy. It was shown by GP19 that this fiducial model produces [C II], CO, and [C I] luminosity versus SFR relations in good agreement with the available observations over a broad range of cosmic time.

2.5. Dust Continuum Emission Modeling

The CIB signal is the dominant correlated contamination in LIM experiments. Models that account for absorption and emission by dust in the ISM of galaxies are implemented in the Santa Cruz SAM in a manner similar to that described in Somerville et al. (2012). The Santa Cruz SAM then assumes that all of the absorbed energy is reradiated in the IR and computes the total IR luminosity \( L_{\text{IR}} \) of each galaxy. Based on the hypothesis that the dust spectral energy distribution is well correlated with \( L_{\text{IR}} \), we use standard dust SED templates to compute the SED spectrum of each galaxy given \( L_{\text{IR}} \).

Specifically, we first integrate over all of the SED templates to compute their total IR luminosity \( L_{\text{IR}}(i) \) \( i = 1, \ldots, N \), where \( N \) is the number of SED templates. We then compare \( \log L_{\text{IR}} \) with \( \log L_{\text{temp}} \) and estimate the dust emission SED of each galaxy through linear interpolation. We next integrate the interpolated dust emission SED for each galaxy to compute the IR luminosity in the fiducial survey observable frequency window, which will be specified in Section 2.7. In this work, we use dust SED templates provided by Chary & Elbaz (2001).

We present a comparison between the integrated extragalactic background light (EBL) spectrum of the 2 deg\(^2\) mock light cone predicted by the Santa Cruz SAM and the observational results in Figure 1. The observational estimates of the EBL are provided by Berta et al. (2011), Béthermin et al. (2012), Aravena et al. (2016), Wang et al. (2017), and Zavala et al. (2017), as summarized in Maniyar et al. (2018). As shown previously by Somerville et al. (2012), this approach produces a reasonable agreement with observational EBL constraints. Somerville et al. (2012) confirmed that halos with masses less than \( 10^{10} M_{\odot} \) make a negligible contribution to the dust emission.

2.6. Mapmaking

With the submillimeter line modeling and dust continuum emission modeling introduced above, we can construct realistic intensity maps for various target frequencies, including interloper and CIB dust continuum contamination under arbitrary angular and spectral resolution. Below, we describe how we create the integrated maps for the emission line and dust continuum emission, include bulk velocities associated with rotation within individual galaxies, and model the finite angular and frequency resolution of an observational map. We also describe how we compute the MW foreground and the configuration of two fiducial surveys that we will use to construct our maps.

\footnote{In this work, \( \log \) denotes a base-10 logarithm.}
2.6.1. Emission Line and FIR Intensity Map

In this section, we provide the details of how we make a mock intensity map. Consider a galaxy at redshift $z$ with disk rotation velocity $v_{\text{disk}}$, emitting one line of interest at rest-frame frequency $\nu_0$. Suppose the galaxy inclination angle is $\beta$ (a face-on galaxy corresponds to $\sin \beta = 0$, while an edge-on galaxy has $\sin \beta = 1$); due to the Doppler effect, the emission line profile width increases by

$$\frac{d\nu}{\nu_0} = \frac{v_{\text{disk}} \cos \beta}{c},$$

where $c$ is the speed of light. We ignore the line broadening caused by thermal motion because the bulk motion is the dominant source of the galaxy velocity dispersion. In this work, we use a simple normalized top-hat function with width $d\nu/(1+z)$ and mean $\nu_0/(1+z)$ to describe the redshifted line profile $\Phi(\nu)$. The intensity contributed by this single galaxy to a map in the frequency range $[\nu_\text{min}, \nu_\text{max}]$ and angular resolution $\theta_{\text{pix}}$ is

$$I = \frac{L}{4\pi \chi^2(z)(1+z)^2(\nu_\text{max} - \nu_\text{min}) \Omega_{\text{pix}}} \int_{\nu_\text{min}}^{\nu_\text{max}} \Phi(\nu) d\nu,$$

where $L$ is the target line luminosity of the emitter, $\chi(z)(1+z)$ is the luminosity distance, and $\Omega_{\text{pix}} = \theta_{\text{pix}}^2$ is the solid angle of each map pixel. The intensity of the FIR emission contributed by a single galaxy at redshift $z$ is computed using a similar method,

$$I_{\text{CIB}} = \frac{\int_{\nu_\text{min}}^{\nu_\text{max}} F(\nu) d\nu}{4\pi \chi^2(z)(1+z)^2(\nu_\text{max} - \nu_\text{min}) \Omega_{\text{pix}}},$$

where the dust SED $F(\nu)$ is estimated following Section 2.5. We repeat this procedure for all galaxies within the light cone and sum up all contributions from different galaxies along the line of sight.

We grid the 2 deg$^2$ field in the R.A. and decl. dimensions with a bin width of $\theta_{\text{pix}}$, which is 10 times smaller than the LIM survey beamwidth $\theta_{\text{FWHM}}$ (the angular resolution of the fiducial LIM surveys we consider in this work is specified in Section 2.7, Tables 1 and 2). We treat the spatial distribution of the galaxy light as a delta function for each galaxy, which is a good approximation, as our pixels are much larger than the expected extent of individual emitters. We then convolve the maps with a Gaussian with FWHM $\theta_{\text{FWHM}}$ to represent beam smearing. We additionally grid the 3D intensity maps in the frequency direction with the bin width of the LIM survey frequency resolution. Pixels near the borders of the Gaussian smeared images are sensitive to the choice of boundary condition. In this work, we apply periodic boundary conditions because this choice preserves the LIM total intensity.

2.6.2. MW Foreground Intensity Map

Cleaning the MW foreground will be one of the great challenges for intensity mapping experiments. Therefore, we provide the option of including a realistic MW foreground in our mock maps. We use the Python Sky Model PYSM package (Thorne et al. 2017) to simulate the continuum foreground mock map. The simulated foregrounds are combinations of synchrotron, free–free, anomalous microwave, and thermal dust emission components. Each component is simulated by the simplest model 1 of PYSM. We construct MW foreground maps with $n_{\text{side}} = 2048$ to ensure that the angular resolution is higher than the current LIM surveys. Since pixel locations in the full sky maps are different from the other 2 deg$^2$ intensity maps we constructed in Section 2.6.1, we interpolate the MW foreground intensity on the [R.A., decl.] grids using bilinear interpolation and then smooth the interpolated foregrounds to the target angular resolution $\theta_{\text{FWHM}}$.

2.7. Fiducial Surveys

In order to provide a demonstration of our mapmaking tool and compare our simulation with predictions from other models, in this work, we make intensity maps for two fiducial surveys.

The first fiducial survey is designed to align with the COMAP pathfinder (Li et al. 2016), which probes CO $J = 1-0$ emitters at redshift $z = 2.4-2.8$. This fiducial survey has an observed frequency window of 30–34 GHz, a spectral resolution of $\delta\nu = 40$ MHz, and an angular resolution of $\theta_{\text{FWHM}} = 6'$.

The second fiducial survey is designed to align with EXCLAIM (Ade et al. 2020), an upcoming balloon mission, which will observe the frequency range 420–540 GHz at spectral resolution $\delta\nu = 937.5$ MHz and angular resolution $\theta_{\text{FWHM}} = 4'$.

We assume a 2 deg$^2$ sky area centered at R.A. = 10° and decl. = −35° as the mapping region for both fiducial surveys, located within the 408 deg$^2$ EXCLAIM survey area. We model the instrumental noise for the COMAP survey as a Gaussian PDF with zero mean and standard deviation $\sigma_N$. We estimate $\sigma_N$ as the final map sensitivity multiplies $10/\sqrt{8\ln(2)}\delta\nu$ (Li et al. 2016). Here the $10/\sqrt{8\ln(2)}$ factor is caused by the fact that the map grid in this work is 1 order smaller than the actual LIM survey beam, and we compute the pixel size as $\theta_{\text{pix}}$, instead of $\theta_{\text{FWHM}}/(8\ln(2))$, which is slightly different from Li et al. (2016). In Figure 4 of this article we show the simulated signal, interloper lines, and continuum contamination of the EXCLAIM survey. If we further include the instrumental noise in Figure 4, the presented intensity maps will be dominated by the noise and it will be very difficult to see the

---

**Table 1**

Summary of Parameters for the COMAP Pathfinder Fiducial Survey, Adopted from Li et al. (2016)

| Parameter | COMAP Fiducial Survey |
|-----------|-----------------------|
| Frequency band ($\Delta\nu$) | 30–34 GHz |
| Frequency channels ($\Delta\nu$) | 40 MHz |
| Beamwidth ($\theta_{\text{FWHM}}$) | 6' |
| Final map sensitivity | 41.5 $\mu$K MHz$^{\text{s}2}$/2 |

**Table 2**

Summary of Parameters for the EXCLAIM Fiducial Survey (Ade et al. 2020)

| Parameter | EXCLAIM Fiducial Survey |
|-----------|-------------------------|
| Frequency band ($\Delta\nu$) | 420–540 GHz |
| Frequency channels ($\Delta\nu$) | 937.5 MHz |
| Beamwidth ($\theta_{\text{FWHM}}$) | 4' |
large scale structure (LSS). EXCLAIM is allowed to have such high noise because its goal is to detect statistical anisotropy rather than directly imaging line LSS, as stated in this sentence. The instrumental noise will be removed during the signal measurement through cross correlating EXCLAIM to BOSS. Since the LIM-galaxy survey cross-correlation technique is beyond the scope of this work, we leave a detailed EXCLAIM noise model to future works.

Parameters of the COMAP and EXCLAIM fiducial surveys are summarized in Tables 1 and 2, respectively. The line emission and redshift range of emitters observed in these two fiducial surveys are summarized in Table 3.

### Table 3

| Line   | \( \nu_0 \) [GHz] | Redshift Range for COMAP Fiducial Survey | Redshift Range for EXCLAIM Fiducial Survey |
|--------|------------------|------------------------------------------|-------------------------------------------|
| C II   | 1901.0           | ...                                      | 2.5–3.6                                   |
| CO \( J = 1–0 \) | 115.3            | 2.4–2.8                                  | ...                                       |
| CO \( J = 2–1 \) | 230.5            | 5.8–6.7                                  | ...                                       |
| CO \( J = 3–2 \) | 345.8            | 9.2–10.0                                 | ...                                       |
| CO \( J = 4–3 \) | 461.0            | ...                                     | 0.0–0.1                                   |
| CO \( J = 5–4 \) | 576.3            | ...                                     | 0.0–0.4                                   |
| C \( 1 J = 1–0 \) | 492.2            | ...                                     | 0.0–0.2                                   |
| C \( 1 J = 2–1 \) | 809.4            | ...                                     | 0.4–1.0                                   |

Note. The second column shows the rest-frame frequencies of emission lines simulated by a submillimeter SAM. The third column shows line emitter redshift ranges for the COMAP fiducial survey, for which the observed frequency window is 30–34 GHz. The fourth column shows line emitter redshift ranges for the EXCLAIM 420–540 GHz fiducial survey. Ellipses indicate that the corresponding emission line will not be observable in the frequency window of the relevant fiducial survey.

### 3. Results

#### 3.1. Geometry and Intensity Maps

The mock light-cone geometry is shown in Figure 2. The longest axis represents redshift, while the other two axes show the R.A. and decl. of each DM halo. The color of each voxel shows the number counts of DM halos with \( M_{\text{halo}} > 10^{10} M_\odot \) in the corresponding spatial grid. Since the physical volume of the light cone increases along the redshift direction while the DM halo number density is decreasing, the number counts of DM halos reach a maximum at redshift \( 2 < z < 3 \), while there are very few halos at \( z < 1 \) and \( z > 8 \).

We present intensity map slices of the CO \( J = 1–0 \) ([C II]) signal, interloper lines, and dust continuum background simulated by the Santa Cruz SAM + submillimeter SAM, together with the MW foreground given by PYSM for the COMAP (EXCLAIM) fiducial survey in Figure 3 (Figure 4). The CO \( J = 1–0 \) and [C II] signals trace the underlying DM density distribution.

#### 3.2. LIM Statistics

In this work, we only include halos with masses larger than the \( N \)-body resolution \( M_{\text{halo}} > 10^{10} M_\odot \) in all statistics.

We present Santa Cruz and submillimeter SAM predictions of [C II] and CO \( J = 3–2 \) luminosity versus SFR relations in different redshift ranges in Figure 5. In Figure 5, we only include central galaxies of the mock light cone that satisfy \( sSFR > 1/(3t_H(z)) \). Here \( sSFR \) is the galaxy-specific SFR defined as the ratio between SFR and stellar mass, and \( t_H(z) \) is the Hubble time at the galaxy redshift. This selection criterion picks out star-forming galaxies that are comparable to the individual sources targeted in the observed samples. We present the 14th, 50th, and 86th percentiles of line luminosity in each plot. The [C II] observational data are provided by

---

Figure 2. Spatial distribution of DM halos in the mock light cone. The longest axis shows redshift, while the other two axes show the R.A. and decl. of DM halos in the mock. Voxel color is determined by the number counts of DM halos in the corresponding spatial cell. We also highlight the redshift range where the EXCLAIM survey will observe [C II] emitters \((2.5 < z < 3.6)\).
Zanella et al. (2018) for $1.7 < z_{\text{obs}} < 2.0$ and Capak et al. (2015), Knudsen et al. (2016), Willott et al. (2015), Decarli et al. (2017), González-López et al. (2014), Kanekar et al.
The age of the universe at redshift $z_{\text{obs}}$ is $1.7 < z_{\text{obs}} < 2.0$.

The second row shows the CO $J = 3 - 2$ luminosity of galaxies as a function of their SFR at different redshifts. The CO $J = 3 - 2$ observations in panels (3) and (4) are from Tacconi et al. (2010) and Tacconi et al. (2013), respectively. Observational data are shown as blue dots. We only select central star-forming galaxies that satisfy sSFR $> 1/(3H(z))$. The upper, middle, and lower red curves show the 14th, 50th, and 86th percentiles of line luminosity, respectively, predicted by the SAM $+$ submillimeter SAM. The joint distribution of luminosity vs. SFR of the mock is shown by the gray 2D histogram. Since the observations are mostly distributed among the high-SFR ranges, a zoomed inset is added to each panel to better present the SAM-observation comparisons. Submillimeter line luminosity vs. SFR relations predicted by the SAM $+$ submillimeter SAM are in good agreement with observations over a broad redshift range.

We compare our CO $J = 1 - 0$ predictions with models from the literature proposed by Righi et al. (2008), Visbal & Loeb (2010), Pullen et al. (2013), and Li et al. (2016) in the redshift range $2.4 < z_{\text{obs}} < 2.8$. This redshift range covers the CO $J = 1 - 0$ emitters for the COMAP fiducial survey introduced in Section 3. The comparisons of the SFR versus halo mass $M_{\text{halo}}$ relations and the luminosity of CO $J = 1 - 0$ $L_{\text{COJ}=1-0}$ versus $M_{\text{halo}}$ relations between our simulation and the CO models considered in this paper are presented in Figure 6. For the SFR--$M_{\text{halo}}$ comparison, we also consider models proposed in Silva et al. (2015), Righi et al. (2008), Visbal & Loeb (2010), and Pullen et al. (2013) all assumed simple linear relations between logSFR and log$M_{\text{halo}}$ with slopes and other free parameters calibrated to various observations, while Silva et al. (2015) and Li et al. (2016) modeled the SFR--$M_{\text{halo}}$ relation as double power laws or more complex functionals to capture the SFR flatness at high halo masses caused by the quiescent galaxy population. We multiply the CO $J = 1 - 0$ luminosity from Visbal & Loeb (2010) and Pullen et al. (2013) by a duty cycle factor $f_{\text{duty}} = 10^8$ yr $/f_{\text{age}}(z)$ to compute the time-averaged CO $J = 1 - 0$ intensities for a consistent model comparison. Here $f_{\text{age}}(z)$ is the age of the universe at redshift $z$. The SFR--$M_{\text{halo}}$ relation predicted by our simulation is in better agreement with the double power-law behavior. Similarly, the trend of $L_{\text{COJ}=1-0}$--$M_{\text{halo}}$ given by our simulation is closer to the most updated CO model introduced in Li et al. (2016; hereafter Li16). However, our simulation predicts lower $L_{\text{COJ}=1-0}$ at $M_{\text{halo}} < 10^{11} M_\odot$. 

Figure 5. Fine-structure and molecular line luminosity vs. galaxy SFR. The first row shows the [C II] luminosity of galaxies as a function of their SFR at different redshifts. The [C II] observations in panel (1) are from Zanella et al. (2018). Observations in panel (2) are provided by Capak et al. (2015) and Bethermin et al. (2020). The second row shows the CO $J = 3 - 2$ luminosity of galaxies as a function of their SFR at different redshifts. The CO $J = 3 - 2$ observations in panels (3) and (4) are from Tacconi et al. (2010) and Tacconi et al. (2013), respectively. Observational data are shown as blue dots. We only select central star-forming galaxies that satisfy sSFR $> 1/(3H(z))$. The upper, middle, and lower red curves show the 14th, 50th, and 86th percentiles of line luminosity, respectively, predicted by the SAM $+$ submillimeter SAM. The joint distribution of luminosity vs. SFR of the mock is shown by the gray 2D histogram. Since the observations are mostly distributed among the high-SFR ranges, a zoomed inset is added to each panel to better present the SAM-observation comparisons. Submillimeter line luminosity vs. SFR relations predicted by the SAM $+$ submillimeter SAM are in good agreement with observations over a broad redshift range.
We compare the [C II] luminosity–$M_{\text{halo}}$ relation predicted by the Santa Cruz SAM + submillimeter SAM in the redshift range $2.5 < z < 3.5$, which covers the [C II] emitter redshift range of the EXCLAIM fiducial survey, with the [C II] models proposed by Silva et al. (2015), Pullen et al. (2018), and Padmanabhan (2019) at $z = 3$ in Figure 7. Silva et al. (2015) introduced two [C II] models. One (Silva15M) is based on Gong et al. (2012; hereafter Gong12) but with an improved galaxy hot-gas metallicity model calibrated to SMGs (De Lucia & Blaizot 2007; Guo et al. 2011), while the other (Silva15L) is a combination of empirical $L_{\text{CII}}$–SFR relations calibrated to local as well as high-redshift observations (Malhotra et al. 2001; Kanekar et al. 2013; González-López et al. 2014; Ota et al. 2014; De Looze et al. 2014) and an SFR model constructed with the previously mentioned SAMs (De Lucia & Blaizot 2007; Guo et al. 2011). The [C II] model of Pullen et al. (2018; hereafter Pullen18) is also a modified version of Gong12 but with a different [C II] number density model. Both Silva15M and Pullen18 require an ISM gas temperature $T_k$ and electron number density $n_e$ to determine the level abundance of [C II] $^3P_{3/2}$ and further predict [C II] luminosity. In Figure 7, the upper bounds of Silva15M and Pullen18 correspond to $T_k \to \infty$ and $n_e \to \infty$, while the lower bounds are predicted assuming $T_k = 100$ K and $n_e = 1$ cm$^{-3}$. The upper and lower bounds of Silva15L correspond to the $m_1$ and $m_3$ $L_{\text{CII}}$–SFR model specified in Table 1 of Silva et al. (2015). Padmanabhan (2019) adopted a power-law form for the $L_{\text{CII}} - M_{\text{halo}}$ relation with an exponential cutoff, which is significantly different from the Santa Cruz SAM + submillimeter SAM simulation predictions over most of the host halo mass range studied in this work.

According to Figures 6 and 7, the scaling relations predicted by the SAM are in best agreement with models proposed by other groups at $10^{11.5}M_\odot \lesssim M_{\text{halo}} \lesssim 10^{13}M_\odot$, where high-redshift galaxy observations are available. Galaxies in less massive DM halos are generally too faint to be detected, while halos more massive than this range are rare. As a result, currently, the scaling relations beyond this halo mass range are not well constrained, and we are unable to tell which set of model predictions is more reliable. However, unlike empirical models that are not physics-grounded, the shape of the scaling relations predicted by the SAM is determined by the underlying physical processes included in the galaxy formation model. For example, the slope of the luminosity versus halo mass relation at low halo masses is largely determined by the stellar feedback strength, while the $L$–SFR slope at high halo masses is influenced by the galaxy-quenching treatments connected to AGN feedback in this model. Since the upcoming LIM surveys will produce constraints on various scaling relations over a broader halo mass range, the SAM will assist in interpreting these constraints in the context of a development of a better understanding of various feedback mechanisms involved in galaxy formation.

In Figure 8, we present the [C II] intensity of the 128 EXCLAIM fiducial survey frequency channels in our EXCLAIM-like simulated map, together with predictions given by Gong12, Silva15, and Pullen18. Similar to Figure 7, the upper bounds of the Gong12, Silva15M, and Pullen18 predictions assume $T_k \to \infty$ and $n_e \to \infty$, while the lower bounds correspond to $T_k = 100$ K and $n_e = 1$ cm$^{-3}$. The upper and lower bounds of the Silva15L predictions correspond to the $m_1$ and $m_4$ models, respectively. We find a significant spread in predictions by different models (nearby 2 orders of magnitude). Among the [C II] models presented in Figure 8, the Silva15L empirical models assume a redshift-independent, linear $L$–SFR relation and calibrate model parameters to different galaxy samples. Models $m_1$–$m_4$ are provided to characterize the scatter in the $L$–SFR relation caused by the variation of galaxy redshift and ISM properties. In practice, the $L$–SFR relation is redshift-dependent. This is the main reason that the SAM [C II] prediction at redshift $2.5 < z < 3.5$ is much higher than the Silva15L $m_4$ model, which is calibrated to local galaxies, while closer to $m_1$, which is calibrated to high-redshift galaxy samples. Other models estimate [C II] intensity through assuming a uniform photon-dominated region (PDR) environment and solving the statistical balance equation between the two [C II] fine-structure energy levels (hereafter, we refer to this type of model as a collisional excitation model). Although in
Figure 7. The [C II] luminosity vs. halo mass scaling relation at redshift 2.5 < z < 3.5. The empirical [C II] relations are from Silva et al. (2015), Pullen et al. (2018), and Padmanabhan (2019). The gray band shows the 68% confidence level of the scaling relations predicted by this work. The upper and lower bounds of bands corresponding to the Pullen18 and Silva15M models are predicted assuming ISM parameters \( T_k = \infty, n_e = \infty \) and \( T_k = 100 \text{ K}, n_e = 1 \text{ cm}^{-3} \), respectively. The upper and lower bounds of the green band correspond to the Silva15L \( m_1 \) and \( m_4 \) models, respectively. The differences between the luminosity vs. halo mass relations predicted by this work and empirical models in the literature are most significant for \( M_{\text{halo}} < 10^{11.5} \) and \( > 10^{12} M_\odot \).

Figure 8. The [C II] intensity in the EXCLAIM fiducial survey observed frequency window predicted by the SAM, compared with models and observational constraints from Gong et al. (2012), Silva et al. (2015), Pullen et al. (2018), and Yang et al. (2019). The upper and lower bounds of the Gong12, Silva15M, and Pullen18 models are predicted assuming ISM parameters \( T_k = \infty, n_e = \infty \) and \( T_k = 100 \text{ K}, n_e = 1 \text{ cm}^{-3} \), respectively. The upper and lower bounds of the Silva15L models correspond to the Silva et al. (2015) \( m_1 \) and \( m_4 \) empirical [C II] models, respectively. The SAM prediction is in best agreement with the lower bounds of collisional excitation models.

Practice, the PDR properties are complex and nonuniform, the good agreement between SAM predictions and the lower bounds of collisional excitation models indicates that, on average, the MCs generated by the submillimeter SAM have a typical gas density \( n_H \approx 1 \text{ cm}^{-3} \) and temperature \( T_k \approx 100 \text{ K} \). Additionally, our predicted intensity signal is more than a factor of 10 lower than the measurement of Yang et al. (2019). One possible explanation for this disagreement is that the signal excess measured by Yang19 could be a combination of [C II] emission and continuum CIB emission that is not well captured by the simple CIB model used for parameter constraints. The large fluctuations of the SAM predictions seen in Figure 8 are caused by variations in large-scale structure probed by different EXCLAIM observed frequency channels.

We present predictions of the normalized VID or one-point PDF for our COMAP- and EXCLAIM-like fiducial mock maps in the left and right panels of Figure 9. Before computing the VID, we downgrid the map with resolution \( \theta_{\text{pix}} \) to the LIM survey spatial resolution \( \theta_{\text{FWHM}} \) in order to minimize correlations between histogram bins (Vernstrom et al. 2014). This is done to ensure reliable VID noise estimation, used in the later reduced \( \chi^2 \) test. In the left panel of Figure 9, we separately show the contribution to the normalized VID from CO \( J = 1-0 \), as well as from the primary interlopers, CO \( J = 2-1 \) and CO \( J = 3-2 \), for the COMAP-like map. We also show the shot noise, MW, and dust continuum (CIB) components. We show a comparison with the predictions of the Li16 model, which only includes the primary CO \( J = 1-0 \) signal. As we showed in Figure 6, the Li16 model input \( L_{\text{COJ } 1-0} = 0.4 M_\odot \) relation is actually quite similar to the one that emerges from our simulation, but it has a shallower slope and higher amplitude at halo masses below \( \sim 10^{11.5} M_\odot \). Due to the low luminosity of less massive emitters, the higher amplitude of the \( L_{\text{COJ } 1-0} = 0.4 M_\odot \) relation given by Li16 does not lead to an overall higher signal. As a result, the CO \( J = 1-0 \) VID of the SAM is higher than that of Li16 for \( J \gtrsim 30 \text{ Jy sr}^{-1} \), mainly because the SAM CO \( J = 1-0 \) is slightly brighter than the Li16 predictions in the halo mass range \( 10^{11.2} M_\odot \lesssim M_{\text{halo}} \lesssim 10^{12} M_\odot \). The CO \( J = 2-1 \) line is the primary interloper line contaminant but still about 1 order less luminous than the CO \( J = 1-0 \) signal. The MW and dust continuum emission are about 2 orders of magnitude higher than the signal.

We perform a reduced \( \chi^2 \) test to estimate whether the sensitivity of the upcoming COMAP “pathfinder” survey is high enough to distinguish different CO \( J = 1-0 \) models. We clean the bright continuum foreground from the mock intensity map by subtracting the mean intensity averaged over all voxels (Breysse et al. 2017). This simple foreground removal method is not optimal, since it only estimates the continuum amplitude at the first order, and it also shifts the VID by a constant after cleaning. More accurate VID continuum treatment requires further work. However, Breysse et al. (2017) showed that this method can effectively clean the continuum. It does not significantly bias or weaken the VID constraints on the Schechter CO luminosity function parameters. Assuming that the intensities of all voxels are almost independent and follow a binomial distribution, the VID variance can be estimated as \( \sigma^2(I) = B(I)(1 - B(I)/N_{\text{vox}}) \approx B(I) \), where \( B(I) \) is the number of voxels with intensity falls in the bin centered at \( I \), and \( N_{\text{vox}} \) is the total voxel number of the mock LIM data. These assumptions break down at high signal-to-noise ratios but should suffice for the case considered here (Ihle et al. 2019). We bin the cleaned intensity maps with CO \( J = 1-0 \) signals predicted by the SAM and Li16 into \( N_0 = 30 \) logarithmically spaced bins within \( 4 \lesssim I_{[\text{Jy sr}^{-1}]} \lesssim 200 \). Voxels with intensities beyond this range are ignored. The normalized VID predicted by the SAM and Li16 after the continuum cleaning are presented in the middle panel of Figure 9. The reduced \( \chi^2 \)
between the SAM and Li16 VID is

$$\chi^2_{\nu} = \frac{1}{N_b} \sum (B^{Li16}(I) - B^{SAM}(I))^2 / \sigma_{SAM}(I)/4 = 6.8,$$

(6)

where $B(I)$ consists of CO $J = 1$–0, interloper lines, and instrumental noise. We reduce the VID variance by a factor of 4 because the COMAP survey measures four sky patches. This shows that the COMAP pathfinder fiducial survey can distinguish the SAM + submillimeter SAM and the empirical model Li16 through the auto CO $J = 1$–0 power spectrum.

Finally, we show the spherically averaged power spectrum of the fiducial COMAP and EXCLAIM mock surveys in the left and right panels of Figure 10. Dotted lines show the power spectrum of emission lines for the mock LIM with an angular resolution $1$ order of magnitude smaller than the beamwidth of the corresponding LIM survey, while solid curves account for the power spectrum attenuation caused by smoothing. The black dashed line shows the $1\sigma$ power spectrum error contributed by the COMAP instrumental noise,

$$\sigma_n(k) = \frac{P_n(k)}{\sqrt{N_{\text{modes}}}},$$

(7)

where $P_n$ is the power spectrum of the instrumental noise, and $N_{\text{modes}}$ is the number of $k$ modes in each $k$ bin. We reduce the instrumental noise for the COMAP pathfinder fiducial survey by a factor of 2, since the power spectrum will be averaged over four independent sky patches in practice.

Similar to the VID foreground cleaning process, the strong continuum emissions presented in the power spectrum can be removed through taking advantage of their smooth frequency spectra. In this work, we follow Wang et al. (2006) and estimate the continuum intensity dependence of frequency by a second-order log–log polynomial for the foreground removal. We confirm that the continuum components in the power spectrum are effectively removed at $k > 0.1$ Mpc$^{-1}$. In the middle panel of Figure 10, we compare the cleaned power spectrum with the CO $J = 1$–0 signal simulated by the SAM and Li16. We estimate the variance of the power spectrum as $\sigma^2(k) = P^2(k)/N_{\text{modes}}$, where $P(k)$ is the spherically averaged power spectrum consisting of smoothed signal, smoothed interloper emission, and instrumental noise, and compute the reduced $\chi^2$ between the Li16 model prediction and the SAM + submillimeter SAM over the $N_b = 5k$ bins within

Figure 9. Normalized one-point PDF of the LIM of the COMAP and EXCLAIM fiducial surveys. Left: line signal, continuum emission, interloper lines, and shot noise components of COMAP VID. Middle: comparison of the COMAP VID with the CO $J = 1$–0 signal predicted by the SAM and Li16. Continuum emission is removed by subtracting the mean intensity averaged over all voxels. Right: line signal, continuum emission, and interloper lines of EXCLAIM VID. For the COMAP fiducial survey, the CO $J = 2$–1 line is the primary interloper contamination source, while for the EXCLAIM fiducial survey, the dominant interloper contamination comes from CO $J = 4$–3 and CO $J = 5$–4. The MW and dust continuum emission are much higher than the signal for both surveys. A reduced $\chi^2$ test shows that the COMAP pathfinder survey will be able to distinguish between the SAM and Li16 models with VID statistics.

Figure 10. Fiducial COMAP and EXCLAIM power spectra. Solid (dotted) curves show the emission line power spectra with (without) the map smoothing representing the beam. The black dashed curve shows the $1\sigma$ error contributed by the COMAP instrumental noise. Left: line signal, continuum emission, interloper lines, and shot noise components of the COMAP power spectrum. Middle: comparison of COMAP power spectrum with the CO $J = 1$–0 signal predicted by the SAM and Li16. The continuum is estimated and subtracted by fitting the continuum intensity dependence of frequency by a second-order log–log polynomial. Right: line signal, continuum emission, and interloper lines of the EXCLAIM power spectrum. We confirm that with the presence of interloper and instrumental noise contamination, the COMAP pathfinder fiducial survey is able to distinguish the SAM + submillimeter SAM and the empirical model Li16 through the auto CO $J = 1$–0 power spectrum.
\[ 0.1 \leq k/[\text{Mpc}^{-1}] \leq 0.3; \]
\[
X^2 = \frac{1}{N_0} \sum_{k} \left( \frac{\rho^{Li16}(k) - \rho^{SAM}(k)}{\sigma^{SAM}_s} \right)^2 = 9.1. \tag{8}
\]

Here again, we reduce the power spectrum variance by a factor of 4 due to the four independent sky patches measured by the COMAP survey. We ignore the power spectrum at \( k > 0.3 \) Mpc\(^{-1}\), since the shot noise dominates at a small scale. The \( X^2 \) test result shows that the COMAP pathfinder fiducial survey can distinguish the Li16 model and the SAM + submillimeter SAM. Ultimately, we want to use the LIM summary statistics to constrain physical processes in galaxy formation. The ability of the LIM statistics to discriminate between two models that both fit observations of bright submillimeter line emission sources but differ in the mapping between DM halo mass and line luminosity for lower-mass halos demonstrates the promise of this approach.

4. Conclusion

In this work, we present a framework for constructing synthetic multitracer LIM maps based on a mock light cone extracted from an N-body simulation and populated with a physics-based semianalytical galaxy formation model. The workflow is as follows. (1) Construct a DM halo light cone from the N-body simulation catalog. (2) Use the Santa Cruz SAM to simulate the DM halo merger history and generate galaxy formation histories. (3) Use the submillimeter SAM to estimate the luminosity of the [C\(II\)], CO, and [C\(II\)] lines for each simulated galaxy. (4) Grid the discrete galaxy and emission line catalog in the [R.A., decl., \( \nu_{\text{obs}} \)] space and generate 3D intensity maps. Following this procedure, we have constructed a mock light cone that covers a 2 deg\(^2\) sky area and extends over the redshift range \( 0 \leq z \leq 10 \). We check that the emission line luminosity versus SFR relations for [C\(II\)] and CO \( J = 3-2 \) predicted by the mock light cone are in good agreement with observations at various cosmic times. The integrated CIB spectrum predicted by the SAM is also consistent with observational constraints over a wide frequency range.

We show that the widely used scaling relations such as SFR–\(M_{\text{halo}}\) and \(L–M_{\text{halo}}\) predicted by our simulation are in good agreement with empirical models in the literature in the halo mass range \( 10^{10.5}M_{\odot} \leq M_{\text{halo}} \leq 10^{12}M_{\odot} \). However, the differences become significant beyond this halo mass range, where no observations are currently available and the scaling relations are not well constrained. Due to the physics-based nature of the SAM + submillimeter SAM approach, future LIM observations will constrain scaling relations over wider halo mass ranges and provide more information about important mechanisms in the galaxy formation process, such as stellar and AGN feedback.

Based on this mock light cone, we simulate intensity maps for two fiducial LIM experiments with instrumental parameters aligned with the upcoming COMAP and EXCLAIM LIM surveys. Our simulation shows that the MW and CIB continuum emission are 2–3 orders of magnitude brighter than the signal. The CO lines are the dominant interloper contamination sources for both the COMAP and EXCLAIM surveys. We also show that with the presence of instrumental noise and interloper contamination, the CO \( J = 1-0 \) line auto power spectrum predicted by the SAM + submillimeter SAM is significantly different from the prediction of the Li16 empirical model. The ability of LIM summary statistics to discriminate between models that are calibrated/validated on bright sources but differ in the predicted properties of fainter objects demonstrates the promise of this approach for constraining physical processes in galaxy evolution. There are well-studied frameworks on using the one-point LIM PDF, CVID, or power spectrum to constrain high-redshift galaxy properties such as the line luminosity function and SFR density. We will use the 2 deg\(^2\) mock light-cone and intensity maps simulated in this work to test different methods to extract these physical quantities, as well as foreground removal techniques, in future works. Although we only constructed mock LIM data for the COMAP and EXCLAIM surveys as two examples in this work, this simulation framework can be easily applied to other upcoming LIM surveys.

We thank Trevor M. Oxoholm for beneficial discussions about the EXCLAIM instrumental noise estimation. S.Y. thanks L. Y. Aaron Yung for SAM usage guidance. We gratefully acknowledge support from the Simons Foundation. This work made use of the Flatiron Institute Computing Cluster. A.R.P. was supported by NASA under award Nos. 80NSSC18K1014 and NNH17ZDA001N.

Appendix

Minimum Halo Mass

Here we show that halos with masses less than \( 10^{10}M_{\odot} \) make a negligible contribution to the average line intensity (and similarly, to the other LIM statistics) for all ISM submillimeter emission models with a steep low-mass slope, including the submillimeter SAM used in this work.

Most of the current line luminosity versus halo mass relations, including the SAM simulation results, can be approximated by double power-law relations:

\[
L(M) = N \left[ \left( \frac{M}{M_i} \right)^{-b} + \left( \frac{M}{M_i} \right)^a \right]^{-1}. \tag{A1}
\]

In Figure 11 (left panel), we present the mean SAM \( L(M) \) relations at redshift \( z = 3 \), together with a double power-law relation \( L(M)/[L_{\odot}] = 10^8\left(M/M_{\odot}\right)^{-1.6} \). Here we ignore the scatter in the \( L(M) \) relation because it will not influence the average emission line intensity. Further study shows that the \( a \) and \( b \) parameters, which characterize the SAM \( L(M) \) slopes for high and low halo mass ranges, do not vary significantly with redshift.

For cases where the low-mass halos in the system are negligible, the assumed \( L(M) \) models have a relatively large value for the \( b \) parameter, in which case, the halo luminosity decreases very rapidly as the halo mass decreases. In Figure 11 (right panel), we show the fractional loss of mean intensity as a function of the halo mass integration lower bound:

\[
\frac{\langle I \rangle (M_{\text{min}})}{\langle I \rangle (10^{10}M_{\odot})} = \frac{\int_{M_{\text{min}}}^{\infty} dM \langle dn/dM \rangle L(M)}{\int_{10^{10}M_{\odot}}^{\infty} dM \langle dn/dM \rangle L(M)}. \tag{A2}
\]
Here the $L(M)$ models are the double power law defined in Equation (A1) with fixed parameters $N = 10^6 \, [L_\odot]$, $M_1 = 10^{11.6} \, [M_\odot]$, and $a = -0.5$. We adopt the halo mass function $dn/dM$ given by the model of Sheth & Tormen (2002). This toy model test shows that when the $L(M)$ relation low-mass slope $b > 2$, halos with masses less than $10^{10} \, M_\odot$ make a negligible contribution to the average intensity. Notice that $L(M)$ models such as Visbal & Loeb (2010) and Lidz et al. (2011) assume linear relations, i.e., $b = 1$. For those linear $L(M)$ models, the low-mass halos contribute significantly to the average line intensity. However, SAM predictions prefer $b \geq 3.5$ for all emission lines probed in this paper, and we therefore confirm that setting the mass resolution of our simulations as $10^{10} \, M_\odot$ will not significantly influence the LIM statistics.

**References**

Ade, P. A. R., Anderson, C. J., Barretinge, E. M., et al. 2020, JLTP, 199, 1027
Amiri, M., Bandura, K., Berger, P., et al. 2017, ApJ, 844, 161
Aravena, M., Decarli, R., Walter, F., et al. 2016, ApJ, 833, 71
Berta, S., Magnelli, B., Nordon, R., et al. 2011, A&A, 532, A49
Bethermin, M., Fudamoto, Y., Ginolfi, M., et al. 2020, A&A, 643, A2
Béthermin, M., Le Floch, E., Ilbert, O., et al. 2012, A&A, 542, A58
Bigiel, F., Leroy, A., Walter, F., et al. 2008, AJ, 136, 2846
Bigiel, F., Leroy, A. K., Walter, F., et al. 2011, ApJL, 730, L13
Blazek, J., Wadadekar, Y., Guiderdoni, B., et al. 2005, MNRAS, 360, 159
Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207
Bonatto, C., & Bica, E. 2011, MNRAS, 415, 2827
Bradač, M., Garcia-Appadoo, D., Huang, K.-H., et al. 2017, ApJL, 836, L2
Brysye, P. C., Anderson, C. J., & Berger, P. 2019, PrRvL, 123, 231105
Bryssy, P. C., Kovetz, E. D., Behroozi, P. S., Dai, L., & Kamionkowski, M. 2017, MNRAS, 467, 2996
Bryssy, P. C., Kovetz, E. D., & Kamionkowski, M. 2016, MNRAS, 457, L127
Capak, P. L., Carilli, C., Jones, G., et al. 2015, Nat, 522, 455
Carilli, C. L., & Walter, F. 2013, ARA&A, 51, 105
Carniani, S., Maiolino, R., Smit, R., & Amorin, R. 2018, ApJL, 854, L7
Carroll, S. M., Duvvuri, V., Trodden, M., & Turner, M. S. 2004, PhRvD, 70, 043528
Chary, R., & Elbaz, D. 2001, ApJ, 555, 161
Chung, D. T., Viero, M. P., Church, S. E., & Wechsler, R. H. 2020, ApJ, 892, 51
Crites, A. T., Bock, J. J., Bradford, C. M., et al. 2014, Proc. SPIE, 9153, 91531W
Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827
De Looze, I., Cormier, D., Lebouteiller, V., et al. 2014, A&A, 568, A62
De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2
DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, PASP, 129, 045001
Decarli, R., Walter, F., Venemans, B. P., et al. 2017, Natur, 545, 457
Doré, O., Bock, J., Ashby, M., et al. 2014, arXiv:1412.4872
Dumitraș, C., Kulkarni, G., Lagache, G., & Haehnelt, M. G. 2019, MNRAS, 485, 3486
Dvăi, G., Gabadadze, G., & Portaì, M. 2000, PhLB, 485, 208
Fonseca, J., Silva, M. B., Santos, M. G., & Cooray, A. 2017, MNRAS, 464, 1948
Glover, S. C. O., & Clark, P. C. 2012, MNRAS, 421, 116
Glover, S. C. O., & Mac Low, M. M. 2007, ApJS, 169, 239
Gnedin, N. Y., & Kravtsov, A. V. 2011, ApJ, 728, 88
Gong, Y., Cooray, A., Silva, M., et al. 2012, ApJ, 745, 49
Gonzalez-Lopez, J., Riechers, D. A., Decarli, R., et al. 2014, ApJ, 784, 99
Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, MNRAS, 413, 101
Hailey-Dunsheath, S., Barlis, A. C. M., Aguirre, J. E., et al. 2018, JLTP, 193, 968
Hirschmann, M., Naab, T., Somerville, R. S., Burkert, A., & Oser, L. 2012, MNRAS, 419, 3200
Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581
Inoue, A. K., Tamura, Y., Matsu, H., et al. 2016, Sci, 352, 1559
Kanekar, N., Wagg, J., Chary, R. R., & Carilli, C. L. 2013, ApJL, 771, L20
Klypin, A., Yepes, G., Gottlöber, S., Prada, F., & Heß, S. 2016, MNRAS, 457, 4340
Knudsen, K. K., Richard, J., Kneib, J.-P., et al. 2016, MNRAS, 462, L6
Knudsen, K. K., Watson, D., Frayer, D., et al. 2017, MNRAS, 466, 138
Kovetz, E. D., Viero, M. P., Lidz, A., et al. 2017, arXiv:1709.09066
Krumholz, M. R. 2014, MNRAS, 437, 1662
Li, T. Y., Wechsler, R. H., Devaraj, K., & Church, S. E. 2016, ApJ, 817, 169
Lidz, A., E. F., Prada, F., et al. 2018, MNRAS, 474, 492
Mitchell, P. D., Lacey, C. G., Lagos, C. D. P., et al. 2018, MNRAS, 474, 492
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Maiolino, R., Carilli, C., Fontana, A., et al. 2015, MNRAS, 452, 54
Maiolino, R., Carniani, S., Fontana, A., et al. 2015, MNRAS, 452, 54
Malhotra, S., Kaufman, M. J., Hollenbach, D., et al. 2001, ApJ, 561, 766
Maniyar, A. S., Béthermin, M., & Lagache, G. 2018, A&A, 614, A39
McAlpine, S., Helly, J. C., Schaller, M., et al. 2016, A&AR, 15, 72
Mitchell, P. D., Lacey, C. G., Lagos, C. D. P., et al. 2018, MNRAS, 474, 492
Maiolino, R., & Krumholz, M. R. 2017, MNRAS, 467, 50
Nelson, D., Springel, V., Pillepich, A., et al. 2019, CoM&AC, 6, 2
Nelson, R. P., & Langer, W. D. 1999, ApJ, 524, 923

![Figure 11. Left: line luminosity vs. halo mass relations $L(M)$. Mean $L(M)$ relations simulated by the SAM are shown as solid curves. The black dashed line shows a double power-law $L(M)$ relation defined in Equation (A1), with parameters $N = 10^6 \, [L_\odot]$, $M_1 = 10^{11.6} \, [M_\odot]$, and $a = -0.5$. Right: fractional loss of averaged line intensity as a function of the lower bound of halo mass integration. The low halo mass slope of the $L(M)$ relation is varied. Other double power-law parameters are fixed as $N = 10^6 \, [L_\odot]$, $M_1 = 10^{11.6} \, [M_\odot]$, and $a = -0.5$.](image-url)
