On Estimating the Cosmic Molecular Gas Density from CO Line Intensity Mapping Observations

Patrick C. Breysse[1] , Shengqi Yang[1] , Rachel S. Somerville[2], Anthony R. Pullen[1,2], Gergö Popping[3] , and Abhishek S. Maniyar[1]

[1] Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA; pb2555@nyu.edu
[2] Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA
[3] European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748, Garching, Germany

Abstract

The Millimeter-wave Intensity Mapping Experiment (mmIME) recently reported a detection of excess spatial fluctuations at a wavelength of 3 mm, which can be attributed to unresolved emission of several CO rotational transitions between $z \sim 1$ and 5. We study the implications of these data for the high-redshift interstellar medium using a suite of state-of-the-art semianalytic simulations that have successfully reproduced many other submillimeter line observations across the relevant redshift range. We find that the semianalytic predictions are mildly in tension with the mmIME result, with a predicted CO power $\sim 3.5\sigma$ below what was observed. We explore some simple modifications to the models that could resolve this tension. Increasing the molecular gas abundance at the relevant redshifts to $\sim 10^8 \, M_\odot \, \mathrm{Mpc}^{-3}$, a value well above that obtained from directly imaged sources, would resolve the discrepancy, as would assuming a CO–H$_2$ conversion factor α_{CO} of $\sim 1.5 \, M_\odot \, \mathrm{K}^{-1} \, (\mathrm{km \, s}^{-1})^{-1} \, \mathrm{pc}^2$, a value somewhat lower than is commonly assumed. We go on to demonstrate that these conclusions are quite sensitive to the detailed assumptions of our simulations, highlighting the need for more careful modeling efforts as more intensity mapping data become available.

Unified Astronomy Thesaurus concepts: Galaxy evolution (594); Molecular gas (1073); CO line emission (262); Surveys (1671)

1. Introduction

As the primary fuel for star formation, cold molecular gas plays a critical role in galaxy evolution. From the first detections of molecular gas outside of our own Milky Way (Rickard et al. 1975, 1977; Solomon & de Zafra 1975), there has been an explosion of observations targeting extragalactic molecular gas (see reviews by Carilli & Walter 2013; Combes 2018; Tacconi et al. 2020). Paralleling and informing similar observations of cosmic star formation (Madau & Dickinson 2014), a consensus picture has emerged of the cosmic molecular gas history, with the overall abundance of molecular gas rising to a peak at redshift $z \sim 2$ before declining to the present day (see, e.g., Walter et al. 2020).

However, this wealth of observational data share one key limitation: they can only access sources bright enough to detect individually, usually through their dust continuum or rotational transitions of carbon monoxide. Molecular gas surveys to date have generally applied one of two strategies, either directly targeting a sample of galaxies selected based on other observations (e.g., Tacconi et al. 2013; Freundlich et al. 2019) or blindly scanning a field and reporting candidates above a certain detection threshold (e.g., Walter et al. 2014; Pavesi et al. 2018; González-López et al. 2019). Both of these methods will by their nature detect only a limited number of sources, and those sources will by definition be biased toward the brightest objects, which may or may not be representative of the full population.

Recently, a third observational technique known as line intensity mapping (LIM) has garnered attention as a complement to these approaches. Intensity mapping does not aim to directly image individual objects but instead seeks to make a statistical observation of the aggregate emission from many unresolved objects (see Kovetz et al. 2017, for a review). By targeting spectrally narrow line emission, the target redshift can be finely selected by modifying the observing frequency, enabling tomographic measurements. For molecular gas studies, LIM observations of CO lines are sensitive to fainter galaxies inaccessible to direct imaging, allowing population-wide examination of this crucial interstellar medium (ISM) component. First discussed in Righi et al. (2008), this application of LIM has garnered significant interest (Lidz et al. 2011; Pullen et al. 2013; Breysse et al. 2014, 2016; Breysse & Rahman 2017; Fonseca et al. 2017; Padmanabhan 2018; Breysse & Alexandroff 2019; Chung et al. 2019; Moradinezhad Dizgah & Keating 2020), with a number of experimental efforts planned or in progress (Li et al. 2016; Lagache 2018; Stacey et al. 2018; Sun et al. 2021; Cataldo et al. 2021). Two of these experiments, the CO Power Spectrum Survey (Keating et al. 2016) and the Millimeter Intensity Mapping Experiment (mmIME; Keating et al. 2020, hereafter K20), have reported 2σ and 4σ detections of unresolved CO emission, respectively.

Interpreting the results of CO observations, whether direct or LIM, is a challenging modeling task. The conversion factor α_{CO} between CO luminosity and molecular gas mass is known to depend on metallicity, density, and the level of excitation of the CO-emitting gas (Bolatto et al. 2013; Tacconi et al. 2020). Thus, the molecular abundance inferred from a measurement depends sensitively on modeling assumptions. The LIM measurements add an extra level of complexity to the modeling task. The very thing that gives LIM surveys their sensitivity to

faint populations, that they do not seek to detect individual sources, means that LIM observations are only able to measure quantities integrated over the entire galaxy population. Thus, in addition to modeling the CO-gas conversion for individual sources, a LIM measurement also needs to model how that conversion is distributed across all galaxies, including those too faint to detect individually.

In this work, we study how these modeling choices affect the interpretation of modern CO intensity maps. We will focus on the aforementioned mmIME survey as a test case. A detection of aggregate CO emission at a wavelength of 3 mm using data from the Atacama Large Millimeter/submillimeter Array (ALMA) and the Atacama Compact Array (ACA) was reported by K20. This observed wavelength contains emission from several CO transitions at different redshifts; thus, the claimed detection contains the sum of contributions from each of these lines. Then, K20 made use of several scaling relations from the literature to convert their CO detection into a cosmic molecular gas history, finding an abundance on the high end of but consistent with results from direct-imaging surveys. As we will show, these results are quite sensitive to the details of the assumed scaling relations. The relations used in K20 are all themselves calibrated to relatively small populations of directly detected galaxies, coming from a variety of observed bands and redshifts. They are all empirical in nature; thus, there is no underlying physics behind the various steps in the conversion. Here we will seek to place this interpretation on a more physically grounded footing by employing a semianalytic model (SAM) of galaxy evolution combined with a line spectral synthesis code.

The primary observational quantity reported by the mmIME survey is the total spectral shot power $\tilde{I}_s^2$, which can be directly related to integrals over the CO luminosity function. The observed emission comes from several overlapping CO transitions at different redshifts, so one must model the redshift distribution of the CO transitions to assign a fraction of the total power to each line. Once this is done, each line must be converted to the equivalent CO(1–0) luminosity and, finally, to a molecular gas abundance through an assumed α_{CO} value. Each step must include an assumption about the mass dependence of the relevant conversion. As the mmIME result provides only a single number to constrain this wide model space, the approach taken in K20 is to effectively assume that the shape of all of these distributions is known and fit an overall amplitude to match the observed shot power. We will follow a similar procedure, but we will obtain all of the relevant distributions and scalings from the suite of SAMs described in Yang et al. (2021, hereafter Y20). The SAM procedure seeks to predict CO luminosity in a self-consistent, physically motivated fashion and is calibrated to a wide variety of empirical measurements of multiple lines. This should give them broader applicability and more robustness than any single empirical scaling. We use a two-step process consisting of a galaxy evolution SAM developed by the Santa Cruz group (Somerville et al. 2015) with an additional subgrid model that predicts submillimeter line emission based on galaxy properties (Popping et al. 2019a). We will refer to this suite of models as the "SAM+submillimeter SAM."

Our primary result is a modest tension between the predictions of the SAMs and the overall power detected by mmIME. While the scalings used in K20 for L_{CO} predict a value of $\tilde{I}_s^2$ quite close to the mmIME measurement, the raw SAM+submillimeter SAM prediction is lower by an order of magnitude, corresponding to an $\sim 3.5\sigma$ discrepancy. We find that, given this tension, the same

amplitude-scaling procedure applied in K20 but with an SAM +submillimeter SAM–informed scaling yields a correspondingly high cosmic molecular gas abundance. Taken at face value, this would suggest that the LIM survey sees dramatically more molecular gas in the universe than has been mapped directly.[4] This result is not outside the realm of possibility, as LIM surveys by their nature are sensitive to more emission than traditional surveys, but one must be careful in drawing such a strong conclusion from limited data. To demonstrate this, we go on to examine how varying the model assumptions affects the molecular gas estimates. We find that the tension can be reduced by assuming a conventional molecular gas history and reducing the average α_{CO} value below what is generally assumed for high-redshift galaxies. Similarly, while it does not remove the tension in overall amplitude, changing the assumed density profile of gas clouds in the SAM+submillimeter SAM lowers the amount of molecular gas needed to resolve the tension back down near the level obtained by the original mmIME analysis. Finally, we substitute back in various pieces of the empirical scalings used in K20 and find that the scatter in the resulting molecular gas abundance is significantly larger than the statistical error bars on the measurement, pointing to the need for more data and more careful modeling to resolve this tension.

The organization of this paper is as follows. Section 2 summarizes the mmIME intensity mapping results. Section 3 reviews the details of the SAM+submillimeter SAMs used in this analysis. Section 4 describes how both the K20 scalings and the SAM+submillimeter SAM are used to predict the mmIME signal and illustrates the basic tension described above. Section 5 reviews the K20 procedure for converting the mmIME measurement to a molecular gas abundance and describes the modified version of that method used here, with the comparison of the results of these methods described in Section 6. Section 7 describes the effects of varying our model assumptions. We discuss our results and how they will improve in the future in Section 8 and conclude in Section 9. Throughout this paper, we assume a flat ΛCDM cosmology consistent with the Planck 2018 results (Planck Collaboration et al. 2020).

2. Data

Here we briefly review the relevant details of the mmIME CO measurement.

The power spectrum constraints in K20 come from maps of two separate fields. The first data set comes from ALMA as part of the ALMA Spectroscopic Survey in the Hubble Ultra Deep Field (ASPECS; Decarli et al. 2019; González-López et al. 2019). It consists of a 4.6 arcmin^2 area of the Hubble Ultra Deep Field mapped at frequencies between 84 and 115 GHz with a 3.906 MHz spectral resolution. The second observation, made using the ACA, covers a 15 arcmin^2 area within the COSMOS field (Scoville et al. 2007) over the same frequency range with a resolution of 7.813 MHz.

The CO(2–1), (3–2), (4–3), and (5–4) transitions fall into this frequency band at redshift ranges centered at $z = 1.3$, 2.5, 3.6, and 4.8, respectively. Authors K20 reported their baseline power spectrum result in the form of the total spectral shot

---

[4] Most direct CO surveys estimate the molecular gas abundance using only directly detected sources, without extrapolating the luminosity function to fainter emitters (see, e.g., Walter et al. 2014; Decarli et al. 2019; Lenkić et al. 2020).

power $\tilde{I}_s^2$, which can be expressed as the sum of contributions from each of these lines:

$$\tilde{I}_s^2(\nu_{\text{obs}}) = \sum_{\text{all lines}} \frac{P_{\text{shot,line}}(z)}{X^2(z)\,Y(z,\nu_{r,\text{line}})}. \qquad (1)$$

Each line is observed at a redshift $z = \nu_r/\nu_{\text{obs}}$ assuming a rest frequency ν_r and an observed frequency ν_{obs}. The power in each line is rescaled to project the spatial dimensions at its emission redshift to a common observed frame in the angular direction by $X(z) = D_M(z)$, where D_M is the comoving radial distance to redshift z, and the line-of-sight direction by $Y(z,\nu_r) = c(1+z)^2/H_0 E(z)\nu_r$, where c is the speed of light, H_0 is the Hubble constant at $z = 0$, and $E(z)$ is the dimensionless Hubble parameter. On the small ($k \sim 10\,h\,\text{Mpc}^{-1}$ or more) spatial scales relevant to this measurement, the power spectrum of each line is dominated by a scale-independent Poisson shot-noise component P_{shot}, which will be defined in detail in Section 4.

A measured $\tilde{I}_s^2 = 730 \pm 240\ \mu\text{K}^2$ in the ASPECS field and $\tilde{I}_s^2 = 890 \pm 440\ \mu\text{K}^2$ in the ACA field were reported by K20. Keenan et al. (2020) showed that, for fields as small as the ones considered here, the median measured power falls, on average, below the "true" cosmological average due to sample variance effects. Accounting for these effects and combining the two fields gives their final result of $\tilde{I}_s^2 = 1010^{+550}_{-390}\ \mu\text{K}^2$.

3. SAMs and Submillimeter SAMs

The sample of dark matter (DM) halos and galaxies adopted in this paper is provided by the Y20 cosmological mock light cone. Here we briefly summarize the light-cone simulation framework and parameters. We refer readers to Y20 for a more detailed description.

The Y20 workflow can be summarized as a three-step process. In the first step, DM halos are selected from an N-body simulation along a past light cone, as described in Somerville et al. (2021), using the `lightcone` package provided by Behroozi et al. (2019). We choose the Small MultiDark-Planck (SMDPL) N-body simulation (Klypin et al. 2016) to provide the DM halo information for a 2 deg² wide and $0 \leqslant z \leqslant 10$ long cosmological light cone. Since the simulation cube volume of SMDPL is $(400\ \text{Mpc}\ h^{-1})^3$, greater than the volume of the target light cone, $5.7 \times 10^7 (\text{Mpc}\ h^{-1})^3$, the statistical independence among different regions within the light cone is, in principle, guaranteed. Moreover, the $10^{10}\ M_\odot$ halo mass resolution of SMDPL is fine enough to ensure that massive halos that can retain significant gas reservoirs and make a significant contribution to the LIM statistics are not omitted.

In the second step, an SAM developed by the Santa Cruz group (Somerville & Primack 1999; Somerville et al. 2008, 2012, 2015, 2021; Popping et al. 2014; Porter et al. 2014) is used to simulate the DM halo merger history, as well as the formation and evolution of galaxies. The Santa Cruz SAM simulates the halo merger history using a method based on the extended Press–Schechter (EPS) formalism (Somerville & Kolatt 1999) or extracts the DM halo merger trees directly from N-body simulations. We adopt the EPS method in this work, as it allows us to resolve halos down to much lower masses. It also replaces the computationally expensive hydrodynamic simulations with simplified but physically motivated treatments, covering the galaxy merging, gas heating and cooling, ISM gas partitioning, photoionization squelching, and various feedback processes.

The Santa Cruz SAM we use in this work has successfully reproduced various UV/optical galaxy observations up to $z = 8$ (Somerville et al. 2012, 2015, 2021; Yung et al. 2019a, 2019b). However, it is important to keep in mind that, like all current galaxy formation models, there are still many uncertainties in the key physical processes implemented within the Santa Cruz SAMs, and some of their predictions are in tension with some observations (see, e.g., Popping et al. 2019b; Somerville et al. 2021).

In the last step, a subresolution recipe developed by Popping et al. (2019a, hereafter P19) is used to simultaneously model the [C II], CO, and [C I] emission of each of the simulated galaxies, and we refer to this model as the submillimeter SAM in the following. The Santa Cruz SAM divides each simulated galaxy into multiple annuli and uses the predicted gas surface density, metallicity, and star formation rate (SFR) in each annulus to predict the fraction of the ISM that is in the form of molecular, atomic, and ionized gas. The submillimeter SAM then randomly selects molecular cloud (MC) masses in each annulus from a power-law MC mass function until the total mass of the MCs reaches the molecular gas mass predicted by the SAM. The submillimeter SAM then assumes an MC radial density profile and further grids each MC into 25 zones. In each zone, it estimates the emission luminosity of multiple submillimeter lines with a linearly interpolated lookup table given by the line spectral synthesis code DESPOTIC (Krumholz 2014). Finally, the galaxy-wide [C II], CO, and [C I] luminosities are computed by summing over the emission contributed by all of the MC zones in all annuli.

The submillimeter SAM has the advantage of simultaneously and self-consistently estimating multiple submillimeter emission lines and their correlations based on a set of properties predicted by a cosmological galaxy formation model. However, the submillimeter SAM introduces free parameters to describe the MC mass distribution, radial density profile, and external radiation fields. Currently, calibration is roughly performed through manual parameter tuning. Particularly relevant to this work, P19 found that the choice of MC radial density profile has a significant influence on the resulting submillimeter line emission statistics. P19 showed that the submillimeter SAM predictions with an assumed Plummer density profile agree best with the observed line luminosity versus galaxy SFR relations that were available at publication (2018). However, Y20 found that the original submillimeter SAM prediction is about a factor of 3 lower than more recent ALMA [C II] measurements (Zanella et al. 2018; Béthermin et al. 2020). Valentino et al. (2020) also suggested that the submillimeter SAM [C I] luminosity versus IR luminosity predictions at high redshifts are significantly lower than the observations. It was shown by Y20 that using a power-law MC density profile can largely alleviate those tensions. In this work, we therefore adopt the power-law MC radial density profile as our baseline assumption, and we discuss how much the MC density profile assumption influences molecular gas density constraints in Section 7.2.

4. Estimating LIM Shot Power from Line Emission Models

In this section, we outline the formalism used to connect galactic line emission to the spectral shot power in an intensity map. We will compare two methods of modeling the CO signal, one from the original K20 paper and a new method based on the SAM+submillimeter SAM described in the preceding section, and show that the semianalytic prediction is in mild tension with the LIM data.

The shot noise P_{shot} that appears in Equation (1) is a quantity that arises in LIM measurements due to the discrete nature of most line-emitting sources. The CO emission arises virtually entirely from dense gas within the ISM of star-forming galaxies, which, on cosmological scales, effectively form a population of point-source emitters. On large scales, the clustering of DM halos causes the locations of these emitters to be correlated, but on small scales, their power spectra are dominated by random, approximately Poisson density fluctuations. This gives rise to a scale-independent power spectrum on small scales with an amplitude

$$P_{\text{shot}}(z) = C_{LI}^2(z) \int_{M_{\min}}^{\infty} \langle L^2(M) \rangle f_{\text{duty}}(M) \frac{dn}{dM} dM \qquad (2)$$

(Lidz et al. 2011). This differs from the shot noise in a spectroscopic galaxy survey, as each galaxy is weighted by its line luminosity. In Equation (5), $\langle L^2(M) \rangle$ is the mean-square line luminosity of halos with mass M, which is then averaged over the halo mass function dn/dM. The expressions for $\langle L^2(M) \rangle$ and $f_{\text{duty}}(M)$ are the main "moving parts" of the model and may be either based on the physics-based SAM or from empirical considerations, as commonly assumed in the literature. We assume the Tinker et al. (2008) halo mass function throughout this work. The quantity

$$C_{LI}(z) = \frac{c}{4\pi \nu_r H(z)} \qquad (3)$$

converts from luminosity density to intensity units. As is common in the literature, we assume that halos with masses below $M_{\min} = 10^{10} M_{\odot}$ do not emit any appreciable CO, though our results are relatively insensitive to the exact value of this choice.

Many simpler LIM models assume that all halos of mass M emit exactly the same luminosity. Here we relax this assumption in two important ways. First, we allow the halo luminosity at a given mass to scatter about its mean value, to account for a dependence on properties other than halo mass. We assume that the actual luminosity of a halo takes the form $L = xL_0(M)$, where $L_0(M)$ is the mean luminosity for all halos of a given mass. Following Li et al. (2016), the scatter value x is assumed to be a lognormally distributed random variable with probability distribution[5]

$$P(x) = \frac{1}{x\sigma \ln 10 \sqrt{2\pi}} \exp\left[-\frac{(\ln x + \sigma^2 \ln 10/2)^2}{2\sigma^2} \right]. \qquad (4)$$

The amplitude of the scatter is set by σ, which for the K20 model takes a mass-independent value of 0.37 dex. Averaging over this distribution gives a final shot power of

$$P_{\text{shot}}(z) = C_{LI}^2(z) e^{[\sigma \ln(10)]^2}$$
$$\times \int_{M_{\min}}^{\infty} L_0^2(M) f_{\text{duty}}(M) \frac{dn}{dM} dM. \qquad (5)$$

Our second change is to assume that only a fraction $f_{\text{duty}}(M)$ of all halos of a given mass are emitting at any given time (Lidz et al. 2011; Pullen et al. 2013). This is particularly important at lower redshifts, where many galaxies may be quenched and emit effectively zero CO. We have slightly generalized the f_{duty} factor here both by allowing it to explicitly depend on halo mass and by including both it and a scatter at the same time.

Keating et al. (2016) discussed accounting for all of the variation around $L_0(M)$ with a single choice of σ, but we find that this does not provide an accurate description of the galaxy population produced by the SAM+submillimeter SAM.

In the simulated galaxies, the CO luminosity depends on the density of molecular gas; therefore, the "duty cycle" is largely sensitive to the gas content of the galaxy (Somerville et al. 2008). Cold gas can be accreted by cooling from the hot halo or mergers of gas-rich satellites. Feedback from the radiatively inefficient "jet mode" of an active galactic nucleus (AGN) can reduce or stop cooling and therefore quench the central galaxy. Cold gas can be ejected via supernova-driven winds and the radiatively efficient "bright mode" of an AGN. The f_{duty} factor accounts for the fraction of galaxies that are quenched by these processes at any given time.

We predict P_{shot} for the SAM+submillimeter SAM by binning all of the halos in our simulated catalog by halo mass. We then account for quenching by removing all halos with a specific SFR (sSFR) below $1/3t_H(z)$, where sSFR is the ratio between SFR and halo mass, and $t_H(z)$ is the Hubble time at redshift z. The SAM +submillimeter SAM equivalent of f_{duty} is then the fraction of galaxies that remain in a mass bin after this cut. Each mass bin then inherently includes a scatter about the mean halo mass, which arises from the variations in halo merger history within the ensemble of halos. We could use this to estimate a lognormal scatter width $\sigma(M)$ for comparison with other models, but when we predict the shot power, we can obtain a more accurate result by directly estimating $\langle L^2(M) \rangle$ in each mass bin and applying Equation (2) (i.e., using the full predicted distribution of luminosities within each mass bin, instead of assuming that it has a lognormal form). We could further improve the accuracy by removing the explicit f_{duty} and dn/dM dependence from Equation (2) and simply integrating over an interpolated luminosity function, but leaving the expression in this form lets us use the same mass function for all of our calculations, bringing us closer to an apples-to-apples comparison between the two approaches.

Figure 1 shows the mean mass–luminosity relationships for the dominant CO transitions considered here. As we are concerned with an observation at a single frequency, in this figure and for the rest of this paper, we will assume that each line is sourced only at the redshift corresponding to ν_{obs}, with CO(2–1) at an average redshift of 1.3, CO(3–2) at 2.5, CO(4–3) at 3.6, and CO(5–4) at 4.8. Shaded regions around each curve in Figure 1 denote the 68% confidence interval of the scatter around the mean relation. The higher-redshift predictions for the K20 model cut off above some halo mass due to the mass limits of the Behroozi et al. (2013) SFR catalog. Figure 2 shows the $f_{\text{duty}}(M)$ values inferred from the SAM at the central redshifts where mmIME observes the four CO lines. This value is quite close to unity at high redshift but drops for lower-redshift lines as quenching effects become more important.

With these relations in hand, we can predict the total spectral shot power present in these two models and compare to the value reported in K20, as shown in Figure 3. As noted in K20, the modified Li et al. (2016) scaling model predicts a total shot power comparable to the mmIME measurement. The physically motivated SAMs, on the other hand, produce a prediction a factor of ~20 smaller, corresponding to an ~3.5σ discrepancy under the quoted mmIME sensitivity.[6] This difference can be

---

[5] Note that our scatter PDF differs slightly from the similar calculation in Sun et al. (2019), as we define $P(x)$ such that the mean luminosity is preserved, i.e., $\langle L(M) \rangle = L_0(M)$ for all values of σ.

[6] The error quoted in K20 would correspond to an ~2.5σ discrepancy under the assumption of Gaussian uncertainty. The full non-Gaussian likelihood leads to a stronger tension (G. K. Keating 2022, private communication).
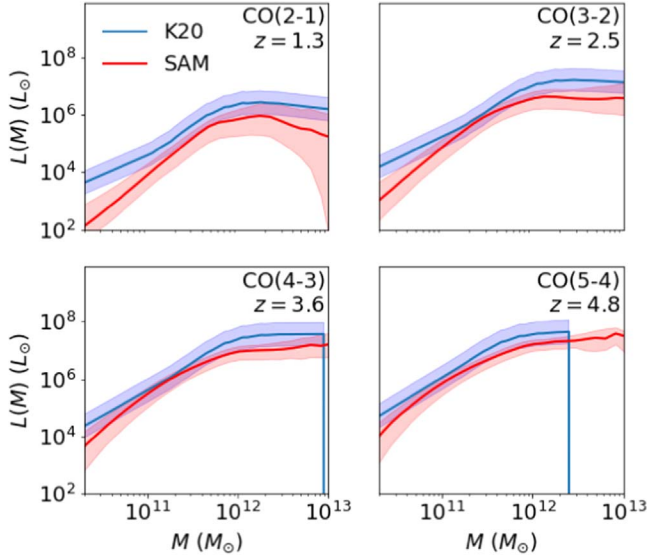
Figure 1. Mass–luminosity relations for the four CO transitions we consider here assuming the modified Li model (blue) and the SAM prediction (red). Shaded bands show the width of the 1σ scatter. The K20 models cut off at the maximum mass available in the Behroozi et al. (2013) SFR catalogs.
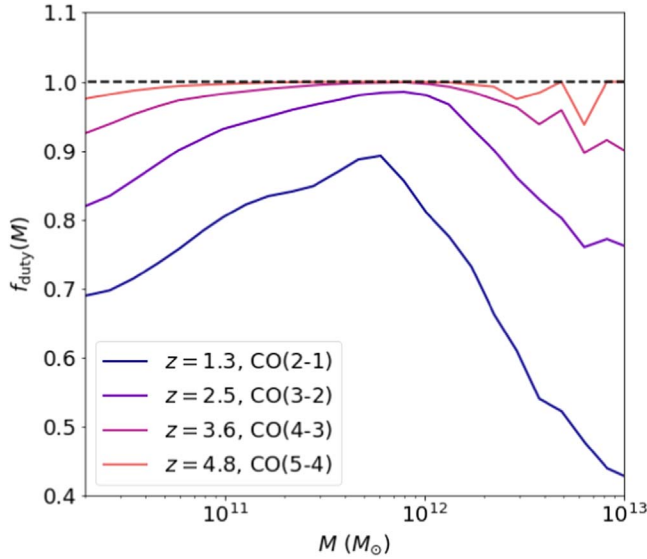

Figure 2. Fraction $f_{\rm duty}$ of halos with nonnegligible star formation as a function of halo mass at the redshifts of each of the four CO transitions we consider here inferred from the SAM+submillimeter SAM. Though $f_{\rm duty}$ is a property of halos rather than specific CO transitions, we have labeled the relevant transition in the legend to track which $f_{\rm duty}$ curve is used for each line.

largely attributed to the nearly universally lower amplitude of the $L_0(M)$ relationships seen in Figure 1. The lowest-redshift CO lines make up a bit of power in the SAM+submillimeter SAM due to extra scatter about the mean relation, but this increase is offset by the inclusion of the subunity $f_{\rm duty}$ from Figure 2. The separation between the semianalytic prediction and the LIM observation seen in Figure 3 is the tension we seek to explore in the rest of this work.

5. Connection to Molecular Gas

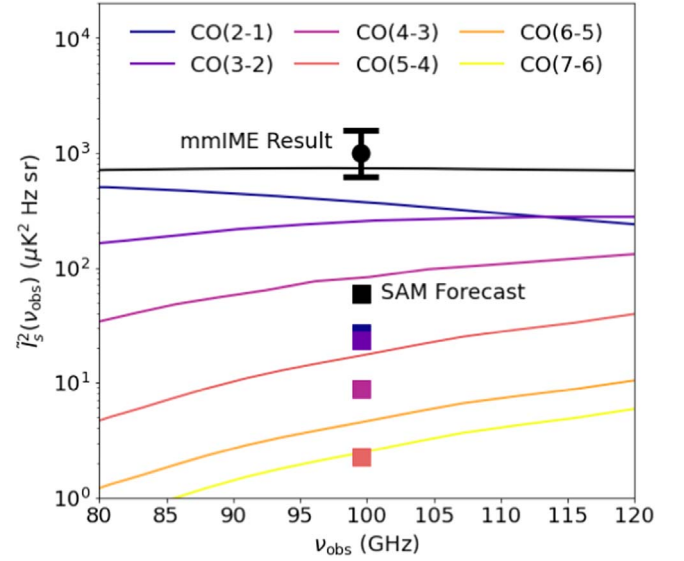If we are to understand the consequences of this possible tension, we need to connect the integrated CO observations

Figure 3. Spectral shot power $I_s^2(\nu)$ for CO rotational transitions (colored curves/points) and the total power observed at a given frequency (black curves/points). The circular point and error bars show the best-fit value and 60% confidence interval from K20. Lines show the predictions as a function of observed frequency for the modified Li et al. (2016) model used in K20 (compare to their Figure 7). Squares show the values predicted by the SAM +submillimeter SAM. Note that the semianalytic prediction falls significantly below the mmIME measurement.

discussed above to the physical properties of the emitting galaxies. For maps of CO emission, the most obvious quantity to explore is the molecular gas abundance, as we expect MCs to source virtually all of the CO emission we detect. A prescription for constraining the cosmic molecular gas history from the mmIME $I_s^2(\nu)$ measurement was laid out by K20. In this section, we will briefly review this procedure, then describe a more streamlined method enabled by the internal consistency of the SAM+submillimeter SAM catalog.

5.1. Redshift Evolution

The first hurdle we need to clear is to separate the single-frequency mmIME measurement into its contributions from the individual CO lines. Many methods have been proposed in the literature for separating the signal from a target line from interloper emission (Gong et al. 2014; Breysse et al. 2015; Cheng et al. 2016, 2020; Lidz & Taylor 2016; Sun et al. 2018). Most of these tools, however, require either an additional cosmological tracer for masking or cross-correlation or a larger, deeper map than those produced by mmIME. Thus, K20 adopted a simpler estimate in which each line is assigned a fraction

$$f_{\rm tot} = \frac{\tilde{I}_{s,{\rm line}}^2}{\tilde{I}_s^2} \qquad (6)$$

of the total observed shot power. They then assigned the $f_{\rm tot}$ that would be obtained under the modified Li et al. (2016) model. In other words, they assumed that the scaling model had the correct redshift evolution, and only the overall amplitude needed to be shifted to match the measured shot power. We can carry out the same prescription using our semianalytically derived relations and calculate a predicted $f_{\rm tot}$ from the mass–luminosity relations plotted in Figure 1. The results of this

Table 1
Predicted Contributions to the Total Shot Power $\tilde{I}_s^2$ from Each Line Considered Here under the Modified Li et al. (2016) Model Used and the SAM +Submillimeter SAM

Transition	$\langle z \rangle$	K20 $\tilde{I}_s^2$	SAM $\tilde{I}_s^2$	K20 f_{tot}	SAM f_{tot}
CO(2–1)	1.3	340	27.5	0.43	0.44
CO(3–2)	2.5	310	23.5	0.4	0.38
CO(4–3)	3.6	120	8.8	0.14	0.14
CO(5–4)	4.8	30	2.3	0.03	0.04
Total	⋯	790	62.1	1.00	1.00

Note. All $\tilde{I}_s^2$ values have the dimension μK^2 Hz sr.

calculation are given in Table 1. One can immediately see that the two models assign similar fractional contributions to each line. This may be due to both models having been calibrated on similar multiline source data. Both models assign <1% of the total signal to lines with higher frequency than CO(5–4), so we will adopt the same assumption as K20 and neglect any contribution from higher-order CO transitions.

This method of separating interloper lines is obviously highly simplified, and despite the similarity between these two models, the data would certainly permit a substantially different redshift evolution. Furthermore, we see in Table 1 that an ∼1% change in the total intensity corresponds to an ∼30% increase in the CO(5–4) shot power. This is a problem intrinsic to any measurement of a faint signal overlapping with brighter emission. We clearly should not trust these models to be accurate to the percent level, so especially the highest-redshift estimates made under these assumptions should be taken with a grain of salt. However, our purpose here is a rough exploration of this pair of models, not a precision measurement. Between this and the fact that it is difficult to derive a better redshift distribution without further measurements, we leave a more detailed exploration of the redshift evolution of the LIM signal to future work.

5.2. Scaling Models

Now that we have an estimate for at least one moment of the CO luminosity function in our four redshift bins, we can convert that value into a measurement of the cosmic molecular gas abundance ρ_{H2}. First, we will review the method used in K20 for this conversion. Typically, the molecular gas content M_{H2} of a galaxy is related to its CO luminosity by the parameter

$$\alpha_{\text{CO}} = \frac{M_{\text{H2}}}{L'_{\text{CO(1-0)}}}, \tag{7}$$

where the observer-unit CO luminosity L' is related to the physical version by

$$\frac{L_{\text{CO}(J)}}{L_\odot} = 4.9 \times 10^{-5} \left(\frac{\nu_J}{\nu_{(1-0)}} \right)^3 \left(\frac{L'_{\text{CO}(J)}}{\text{K km s}^{-1} \text{ pc}^2} \right). \tag{8}$$

Authors K20 assumed a typical extragalactic value of $\alpha_{\text{CO}} = 3.6\, M_\odot\, (\text{K km s}^{-1}\, \text{pc}^2)^{-1}$, with wider estimates ranging from a Milky Way–like $\alpha_{\text{CO}} = 4.3\, M_\odot\, (\text{K km s}^{-1}\, \text{pc}^2)^{-1}$ down to a ULIRG-like $\alpha_{\text{CO}} = 0.8\, M_\odot\, (\text{K km s}^{-1}\, \text{pc}^2)^{-1}$.

Since α_{CO} is typically reported relative to CO(1–0) luminosity, it is necessary to convert the higher-J CO intensities predicted here to their CO(1–0) equivalents. Authors K20 adopted the mean

line ratios $r_{J,1} = L'_J / L'_{(1-0)}$ from a sample of $z = 1.5$ optically selected galaxies (Daddi et al. 2015) and averaged them over the galaxy population by assuming

$$P_{\text{shot},J} = r_{J,1}^2 P_{\text{shot,CO(1-0)}}, \tag{9}$$

where $r_{J,1} = 0.76 \pm 0.09$, 0.42 ± 0.07, 0.31 ± 0.06, and 0.23 ± 0.04 for the CO(2–1), (3–2), (4–3), and (5–4) transitions, respectively.

Finally, we need a prescription for fitting a model for $L_{\text{CO},J}$ to the measured $\tilde{I}_s^2$. As we have only a single number to base this on, we can handle only a single independent parameter in our fitting. Authors K20 assumed a broken power-law mass–luminosity relation for this step parameterized by

$$L_{\text{CO}}(M) = \begin{cases} A_{\text{CO}} \frac{M^2}{M_0} & M \leqslant M_0 \\ A_{\text{CO}} M_0 & M \geqslant M_0, \end{cases} \tag{10}$$

where $M_0 = 10^{12}\, M_\odot\, h^{-1}$ is the location of the turnover, and the overall amplitude A_{CO} is the free parameter used for the fit. When applying Equation (10), the K20 model assumes the default mass-independent scatter of $\sigma = 0.37$ dex from Li et al. (2016). As with all of the K20 scaling relations, we continue to assume $f_{\text{duty}} = 1$.

Thus, the overall procedure to connect the measured $\tilde{I}_s^2$ to ρ_{H2} goes as follows.

1. Find the value of A_{CO} in Equation (10) that best fits the measured shot power $\tilde{I}_s^2$.
2. Convert the resulting CO luminosity to the CO(1–0) equivalent using Equation (9).
3. Convert the CO(1–0) luminosity to M_{H2} assuming a constant α_{CO}.
4. Integrate the resulting $M_{\text{H2}}(M)$ over a halo mass function to get a total estimated ρ_{H2}.

5.3. Alternate Procedure Using SAMs

The full range of detailed galaxy properties available in the SAM+submillimeter SAMs enables a more nuanced procedure for constraining ρ_{H2} from $\tilde{I}_s^2$. The above scaling approach necessarily assumes that most of its relevant quantities are both mass- and redshift-independent. Due to the sparsity of empirical data, many of the scalings are calibrated at different redshifts than the mmIME observation. Furthermore, all of the scalings should have some galaxy-to-galaxy scatter about the given mean relation, just like that of the mass–luminosity relation discussed previously. Scatter in one relation can easily be correlated with scatter in another, further complicating attempts to model these relationships analytically.

With our simulated halos in hand, we can calculate the full mass- and redshift-dependent forms of Equations (7)–(10) based on the SAM. However, while we will make use of these later, we do not actually need them to get a ρ_{H2} measurement. Following the broad strokes of the K20 model, we will assume that we know the shape of the $L_0(M)$ relation from the SAM results and fit for an overall amplitude offset $A_{\text{CO}}^{\text{SAM}}$ such that

$$L_{0,\text{fit}}(M) = A_{\text{CO}}^{\text{SAM}} L_0(M). \tag{11}$$

In other words, we want to find the value of $A_{\text{CO}}^{\text{SAM}}$ that brings the SAM+submillimeter SAM forecast for $\tilde{I}_s^2$ into agreement with the mmIME result plotted in Figure 3,

assuming the semianalytic predictions for the mean and scatter on $L_0(M)$, as well as $f_{\rm duty}$.

Since we know the molecular content of each SAM galaxy, we know the value of the total cosmic molecular gas density $\rho_{\rm H2,0}$ that corresponds to the default SAM+submillimeter SAM results. By definition, then, $A_{\rm CO}^{\rm SAM} = 1$ corresponds to a universe with $\rho_{\rm H2} = \rho_{\rm H2,0}$, and the total shot power $\tilde{I}_s^2$ scales as $(A_{\rm CO}^{\rm SAM})^2$. Since we are rescaling the CO luminosities linearly, we can get a K20-equivalent $\rho_{\rm H2}$ measurement simply from

$$\rho_{\rm H2} = A_{\rm CO}^{\rm SAM} \rho_{\rm H2,0}. \qquad (12)$$

Note that, in the SAM, we of course know the predicted value of the conversion from any excitation-state CO line to the CO(1–0) line luminosity, and we know the value of $\alpha_{\rm CO}$ for each galaxy. Indeed, these conversions vary galaxy by galaxy and also indirectly depend on redshift due to changes in the ISM properties and the cosmic microwave background radiation field strength (see, e.g., Popping et al. 2016). With the procedure described above, we are assuming a population-averaged effective conversion between observed line emission and CO(1–0) and, similarly, a population-averaged value of $\alpha_{\rm CO}$. This implicitly assumes that these conversion factors are correctly predicted by the SAM, and that the SAM correctly predicts the galaxy population properties that are relevant for determining them.

6. Results

Here we present the molecular gas history inferred by using the procedure based on the semianalytic mocks, as outlined in Section 5.3, in place of the K20 scaling relations, as described in Section 5.2. Figure 4 shows the resulting $\rho_{\rm H2}(z)$ estimate compared to that obtained in the original mmIME analysis. We further compare those measurements to a compilation of direct-imaging results from Walter et al. (2020, hereafter W20), which includes CO-based measurements from the xCOLD GASS (Fletcher et al. 2021), ASPECS (Decarli et al. 2020), COLDz (Riechers et al. 2019), and VLASPECS surveys (Riechers et al. 2020), along with several dust continuum observations (Berta et al. 2013; Scoville et al. 2017; Magnelli et al. 2020). The black line shows a fit to these direct observations. The original mmIME results quoted in K20 are broadly consistent with the direct-imaged molecular gas surveys, if slightly on the high side. The values of $\rho_{\rm H2}$ obtained by directly integrating the intrinsic molecular gas masses in all galaxies in the SAM are also consistent with the values estimated from direct-imaging surveys.

Due to the discrepancy between the predicted and measured $\tilde{I}_s^2$, however, the SAM+submillimeter SAM requires an $A_{\rm CO}^{\rm SAM}$ of ~ 4 to match the mmIME $\tilde{I}_s^2$ measurement. Thus, we find that using the scalings derived from the SAM and the mmIME measurement of $\tilde{I}_s^2$, instead of the K20 scalings, increases the inferred value of $\rho_{\rm H2}$ from a value consistent with the direct measurements to one substantially above them.

It should be noted that, again because we are only fitting a single data point, the four redshift bins plotted in Figure 4 are all exactly correlated. Thus, what we are seeing is the same $\sim 3.5\sigma$ tension from Figure 3 being carried over into the estimates of $\rho_{\rm H2}$. What we have effectively done is assumed that the relationship between H_2 mass and CO luminosity is exactly known from the
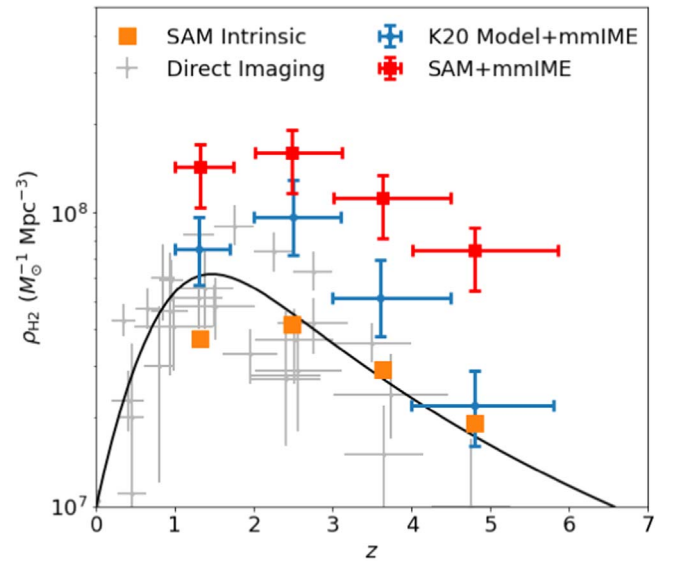

Figure 4. Measurements of the cosmic molecular gas history. Gray points show direct measurements collected in W20; the black line shows a fit to these points. Blue points are the values quoted in K20 assuming empirical scalings. Orange squares show the intrinsic values of $\rho_{\rm H2}$ in the SAM, and red points show the values inferred from the mmIME measurement using conversions based on the SAM. Note that, because all of the LIM measurements are derived from a single measured data point, all of the errors on the red and blue points are perfectly correlated.

SAM+submillimeter SAM and linearly increased the amount of molecular gas needed to match the mmIME $\tilde{I}_s^2$ value.

If all of the above assumptions are taken at face value, this would seem to indicate that mmIME has seen a substantial reservoir of high-redshift molecular gas that has not been seen by traditional surveys. This is not implausible on the surface, as one of the primary motivators for intensity mapping surveys like mmIME is to map the properties of sources below the detection thresholds of individual images.

It is clearly premature to make such a strong statement, though, at least without substantial caveats. A single data point in modest tension with a single model is not enough justification to claim a detection of a factor of several increase in the molecular gas content of the universe, particularly when modeling a system as complicated and poorly understood as the high-redshift ISM. We will devote the next section to an exploration of some of the ways in which the results of this analysis change depending on the model assumptions we adopt.

7. Varying the Model

As noted in Section 3, both the SAM and submillimeter SAM contain free parameters that characterize the physical processes in the galaxy formation model (e.g., star formation efficiency, stellar feedback efficiency, AGN feedback efficiency, etc.) and the subgrid properties of the MCs that populate the ISM and give rise to the observed CO emission (cloud mass function, radial profile, etc.). These parameters have, up to now, been manually tuned to match the properties of individual observed sources like the ones used to derive the W20 direct estimate of $\rho_{\rm H2}$. In principle, we could try varying these parameters, perhaps using something like a Markov Chain Monte Carlo sampler, to explore whether there are parts of the complex and high-dimensional parameter space that can satisfy the direct-imaging and LIM constraints

simultaneously. This would be a very interesting exercise but is well beyond the scope of the current analysis.

Thus, we will instead examine a few simple modifications to our SAM-based calculation and examine their implications. We do not claim that either the results presented above or any in this section are definitive statements about the nature of the high-redshift universe, but they do offer a picture of the kind and scale of model-dependent effects that exist in this and future LIM analyses.

7.1. Fixed ρ_{H2}, Varying α_{CO}

As stated previously, our primary assumption when deriving Figure 4 was that we know precisely the relationship between CO and molecular gas, so that an increase in CO intensity must be accompanied by an increase in ρ_{H2}. However, the same result could be accomplished by instead altering the relationship between CO and H_2. The most obvious place to start would be the choice of α_{CO}. We can hold the total ρ_{H2} constant and simply increase the CO luminosity per unit of molecular gas.

In the SAM+submillimeter SAM, we know both the CO luminosity and H_2 mass of each simulated galaxy. Thus, we can access the full distribution of $\alpha_{CO}(M)$ for halos of different mass. To compare to the values quoted above, we need the population-averaged value. Averaging only over the star-forming population, this takes the form

$$\overline{\alpha}_{CO} = \frac{1}{\overline{n}} \int f_{duty}(M) \alpha_{CO}(M) \frac{dn}{dM} dM, \quad (13)$$

where $\overline{n} = \int f_{duty}(M)(dn/dM)dM$ is the mean number density of star-forming galaxies.

To fit the mmIME result by varying α_{CO}, we start with the end result of the previous section, a fitted value of A_{CO}^{SAM} that brings the semianalytic prediction of $\tilde{I}_s^2$ into agreement with the mmIME measurement. As we saw, this alone yields a ρ_{H2} estimate well in excess of that obtained through direct observations. We then make a new assumption, that the W20 ρ_{H2} fit in fact provides the true value of the molecular gas density. Finally, we linearly scale the SAM α_{CO} parameter to a value that brings the red SAM+mmIME points from Figure 4 into agreement with the W20 fit. Symbolically, this is given by

$$\alpha_{CO}^{fit} = \alpha_{CO}^{SAM} \frac{\rho_{H2}^{W20}}{A_{CO}^{SAM} \rho_{H2,0}}. \quad (14)$$

The results of this computation are plotted in Figure 5, with a comparison to both the redshift-independent range of α_{CO} assumed in K20 and the mean values obtained from the unmodified SAM+submillimeter SAM. Note that, because the intrinsic SAM+submillimeter SAM predictions for ρ_{H2} are fairly close to the W20 fit, this is similar to the result we would obtain simply by scaling α_{CO}^{SAM} by $1/A_{CO}^{SAM}$ to directly cancel out the effect of increasing the CO luminosity.

We find that, under these assumptions, we would need an $\overline{\alpha}_{CO}$ value of ~ 1.5–$2 \, M_{\odot} \, (K \, km \, s^{-1} \, pc^2)^{-1}$, roughly a factor of 2 lower than that assumed in K20. The $\sim 3.5\sigma$ offset between the red points and black line in Figure 5 is again the same initial tension in $\tilde{I}_s^2$ but this time entirely attributed to a change in CO emission per unit of molecular gas.

Note that throughout this section, we have continued to hold all aspects of the SAM+submillimeter SAM predictions to their intrinsic values with only two exceptions: the linear A_{CO}^{SAM}

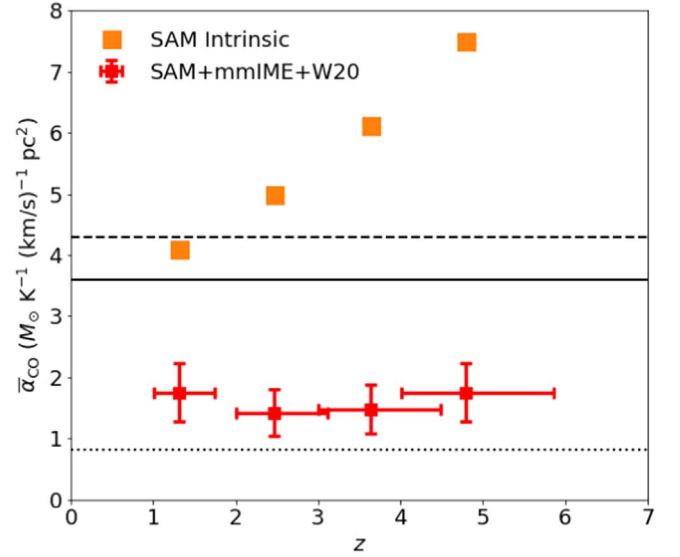**Figure 5.** Population-averaged conversion constants α_{CO} between CO(1–0) luminosity and molecular gas mass for the four mmIME redshift bins. The black lines show the values assumed in K20 for all galaxies (solid; the default value used in their analysis), the Milky Way (dashed), and ULIRGS (dotted). Orange squares show the values from the SAM alone, and red squares show the value needed to bring the SAM+submillimeter SAM into agreement with both the mmIME $\tilde{I}_s^2$ and W20 ρ_{H2} measurements.

scaling on $L_0(M)$ needed to fit the mmIME measurement and a linear scaling on α_{CO} to bring the resulting ρ_{H2} fit into agreement with W20. The shape of $L_0(M)$, the scatter around that average, and the conversions between CO(1–0) and each of the higher transitions are all unchanged.

7.2. Varying the Submillimeter SAM Subgrid Ingredients

The submillimeter SAM and the underlying Santa Cruz SAM are both highly sophisticated frameworks with many tunable parameters necessary to model the complexity of the ISM. For the same reason that we do not attempt any fully automated parameter search, an exhaustive study of all of the possible SAM and submillimeter SAM variations that could affect $\tilde{I}_s^2$ is beyond the scope of this paper. We will instead highlight a single change to show how sensitive the interpretation is to the choice of model. Specifically, we will alter the density profile assumed for individual MCs from the power-law shape we have assumed thus far to the Plummer density profile originally suggested by Popping et al. (2019a).

Using a new galaxy catalog produced under this changed assumption, we obtain a predicted $\tilde{I}_s^2$ of 198.7 μK^2 Hz sr, a factor of 3 or so higher than our original result. The tension with the mmIME value of $1010^{+550}_{-390} \, \mu K^2$ Hz sr is lessened but still lingers at $\sim 2\sigma$. However, when we repeat the calculation of the cosmic molecular gas density from Figure 4, we see that the inferred ρ_{H2} values are now much closer to those obtained through the K20 scaling model, as shown in Figure 6. We still see a modest increase for the highest-redshift CO(5–4) line, but, as mentioned above, the highest-redshift points are likely the least trustworthy under our limited ability to separate interlopers.

As we have discussed in Section 3, both the power-law and Plummer profiles produce results for the L_{CO}–SFR relation that are in reasonable agreement with the observational constraints within the rather large scatter. However, the $L_{[C \, II]}$–SFR and
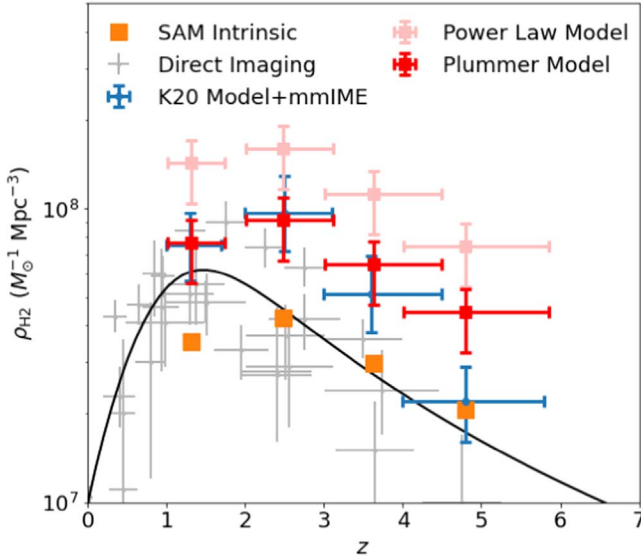
Figure 6. Comparison of molecular gas mass density inferred from the mmIME measurement using a Plummer MC profile (red) to results from Figure 4 (pink) that assumes a power-law profile for the submillimeter SAM subgrid model.

$L_{[C\ I]}$–L_{IR} relations predicted by the SAM+submillimeter SAM assuming the Plummer model are significantly lower than the current observations. The purpose of showing the results with the Plummer radial profile assumption is not to suggest that it is the correct resolution to the tension we have pointed out. Instead, we show this variant to illustrate how sensitive the inferred ρ_{H_2} can be to details about the subgrid properties of the ISM.

7.3. Varying the K20 Scalings

For the final step in our exploration, we will examine how the SAM+submillimeter SAM predictions for the individual scalings used in Section 5.2 differ from those assumed in K20 and estimate how these differences impact the estimate of ρ_{H2}. Though we did not explicitly use either the α_{CO} or $r_{J,1}$ scalings to produce Figure 4 from the SAMs, we can of course read these quantities directly from the simulated catalog. The most obvious difference between the SAM-inferred relations and those in K20 is that both the CO–H_2 and line luminosity ratios are mass- and redshift-dependent. Figure 7 shows the two primary ratios used to convert the CO measurement to ρ_{H2}, comparing the mass-dependent semianalytic values to the constant empirical values assumed in K20.

We can immediately see clear and significant differences between the SAM predictions and the K20 assumptions. For α_{CO}, the SAMs predict higher values at all masses and redshifts than the K20 assumption. The $r_{J,1}$ values are generally higher as well, the redshift dependence is a bit different, and there is a significant halo mass dependence. As the entire point of the SAM+submillimeter SAM is that all of these quantities are computed self-consistently based on a physical model, we cannot simply manually change these relations to see how they affect the measured $\tilde{I}_s^2$; any changes would have to be obtained by varying parameters like the cloud profile choice discussed in the previous section. In order to get a sense of the role of the different pieces of the model for obtaining an estimate of ρ_{H2} from $\tilde{I}_s^2$, we need to artificially break the internal consistency of the semianalytic predictions. This will lose the key advantage
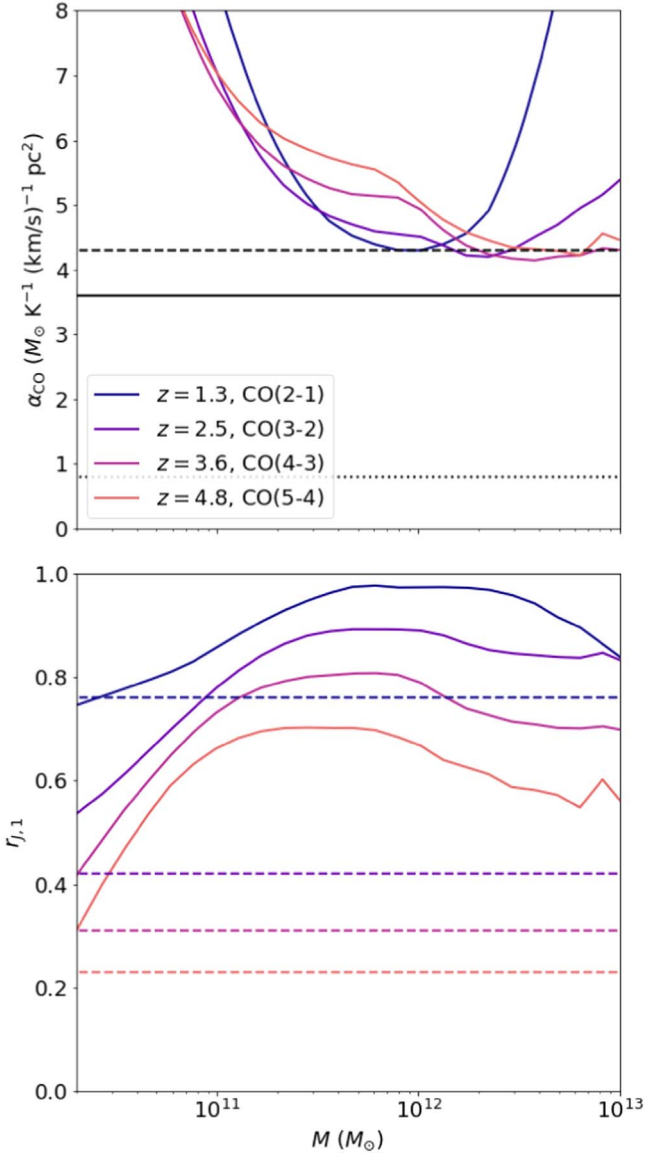

Figure 7. Comparing the mass dependence of SAM scaling relations to those assumed in K20. (Top) CO(1–0) luminosity–H_2 mass ratio $\alpha_{CO}(M)$ at the central redshifts of our four CO transitions (colored lines) compared to the typical, Milky Way, and ULIRG values used in K20 (black solid, dashed, and dotted lines, respectively). All α_{CO} averages are computed using only the star-forming galaxies from the SAM+submillimeter SAM. (Bottom) CO line luminosity ratios $r_{J,1}(M)$ for the same lines (solid lines) compared to the constant Daddi et al. (2015) values from K20 (dotted lines). For both relations, note that the halo mass function tends to weight 10^{11}–$10^{12}\ M_\odot$ halos the strongest when computing average quantities.

of the SAM+submillimeter SAM procedure, but it will give us an idea of where the differences come from and how big they might be.

Combining all of the scalings from Section 5.2 allows us to explicitly write the dependence of ρ_{H2} on the fitted A_{CO}^{SAM}, yielding

$$\rho_{H2} = A_{CO}^{SAM} \frac{1}{x'} J^3$$
$$\times \int \frac{\alpha_{CO}(M)}{r_{J,1}(M)} f_{duty}(M) L_J(M) \frac{dn}{dM} dM, \quad (15)$$

where $x' \equiv 4.9 \times 10^{-5} \, L_\odot \, (\mathrm{K \, km \, s^{-1} \, pc^2})^{-1}$ is the conversion factor between L and L' from Equation (8). The factor of J^3 comes from the frequency dependence of the same conversion. Since this final computation is even more approximate than the previous ones, we will neglect the scatter in these quantities for now. Equation (15) without scatter gives a ρ_{H2} result within 5% of the original calculation using the full SAM+submillimeter SAM results with scatter. We can see that the ρ_{H2} inferred from a given line scales as $\alpha_{\mathrm{CO}}/r_{J,1}$ up to an integral over mass. We can modify this ratio to change ρ_{H2} without changing the assumed mass–luminosity relation $L_J(M)$.

To get a rough idea of how much this choice matters, in Figure 8, we alternately replace the $\alpha_{\mathrm{CO}}(M)$ and $r_{J,1}(M)$ relations in Equation (15) with the constant values used in the analysis of K20. If we use the commonly assumed $\alpha_{\mathrm{CO}} = 3.6 \, M_\odot \, (\mathrm{K \, km \, s^{-1} \, pc^2})^{-1}$, our ρ_{H2} values fall by 1σ–1.5σ and end up much closer to the K20 values. This effectively restates the point from Figure 5 that a smaller α_{CO} will alleviate the tension in our ρ_{H2} estimates. On the other hand, if we consider ourselves free to modify the model in this fashion, keeping the built-in mass-dependent α_{CO} and instead using the Daddi et al. (2015) constant $r_{J,1}$ ratios increases the inferred molecular gas abundance even further beyond the estimates from direct imaging. All in all, the range between the different estimates in Figure 8 can be seen as an extremely approximate estimate of the "error bar" on the modeling uncertainty in this problem, which is substantially larger than the statistical uncertainties.

We should note that this last type of uncertainty is far from unique to LIM measurements. Direct-imaging measurements of CO transitions must also make assumptions about the values of $r_{J,1}$ and α_{CO}, and these choices are a source of significant uncertainty (see Boogaard et al. 2021, for an example). One could thus envision a similar amount of model-based scatter on the gray literature points in Figure 8 as on the red LIM points. As discussed above, however, the need to choose mass-dependent values for these quantities presents and will continue to present a unique challenge for integrated LIM measurements.

8. Discussion

The core result presented here is both simple and fairly robust, at least at the level of detail we consider here. The mmIME survey produced a measurement of the total shot power from high-redshift CO lines that is roughly 3.5σ, in tension with predictions from the Santa Cruz SAM+submillimeter SAM framework, taken as face value. The tension is present in both versions of the SAM+submillimeter SAM discussed above, which adopt different MC density profiles. There are, of course, many other modeling choices baked into the semianalytic prediction that we have assumed here to be essentially fixed. Most of these parameters, however, are set by comparison to a wide variety of other galaxy survey measurements, and changing any of them to fit the mmIME data could easily create discrepancies elsewhere. A full statistical treatment would require fitting the mmIME $\tilde{I}_s^2$ value and all of the other galaxy measurements simultaneously while varying all of the underlying SAM parameters. Such an effort would be highly enlightening but is beyond the scope of this work, so for this initial exploration, we confine ourselves to reproduction and modest extension of the original K20 mmIME analysis using our semianalytic framework.
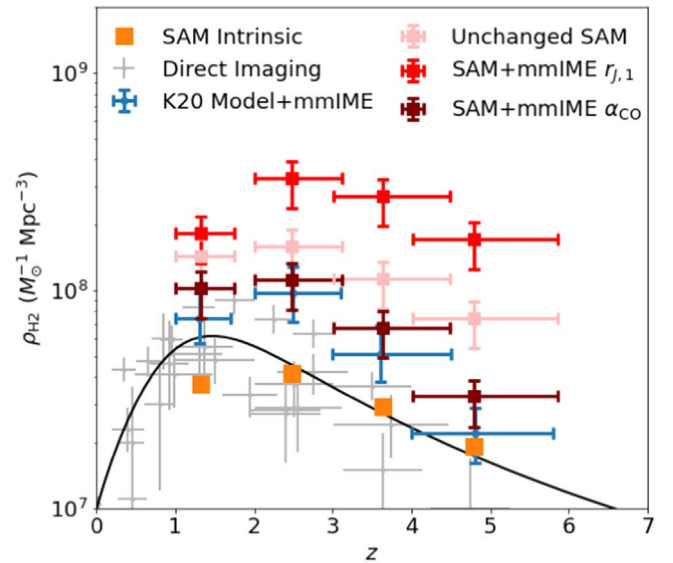

Figure 8. Estimates of ρ_{H2} obtained by combining the semianalytic and empirical scaling relationships in Equation (15), compared to the same K20 mmIME and direct-imaging constraints plotted in Figures 4 and 6. Pink points show our original constraints, which use the unmodified SAM+submillimeter SAM to go from $A_{\mathrm{CO}}^{\mathrm{SAM}}$ to ρ_{H2}. Red and dark red points respectively substitute in the constant line ratio $r_{J,1}$ and CO–H_2 ratio α_{CO} values from K20 while retaining the other model elements from the SAM+submillimeter SAM predictions.

When we repeat an analysis similar to that of K20 but replace their empirical scalings with predictions from the SAM for the relationship between halo mass and CO luminosity and assume that any difference between the mmIME measurement and our model prediction of $\tilde{I}_s^2$ can be explained by a linear change in H_2 content across all halo masses, we infer the presence of a large amount of excess molecular gas at high redshift compared to direct measurements from individual CO line detections or dust continuum measurements. If true (which is far from certain given the limitations of these results), the distribution of H_2 within galaxies would almost certainly not look like a simple linear increase in every galaxy above the original SAM+submillimeter SAM predictions. In order to have been missed in traditional surveys, the excess molecular gas would have to be predominantly located in individually faint lower-mass galaxies, otherwise surveys like COLDz and ASPECS would see it (though there may be sample variance limitations as well; see Keenan et al. 2020). If tensions like this continue to exist and are confirmed in the future, a simple extension to this analysis would be to apply something like the empirical models presented in Padmanabhan (2018), where we are free to add more CO emission to galaxies of different masses rather than uniformly increasing it across the board, as was done here.

The mmIME measurement is not the only source of tension between intensity mapping measurements and semianalytic forecasts. Yang et al. (2019) published an estimate of the 158 μm C II integrated line intensity by cross-correlating data from the Planck satellite with spectroscopic galaxy measurements. As shown in Figure 8 of Y20, the result is also somewhat brighter than the prediction from the SAM+submillimeter SAM. While this tension is also quite modest, it is further evidence that some of the physical processes in the current Santa Cruz SAM+submillimeter SAM may need to be modified.

As emphasized previously, there are many uncertainties in the physical processes and parameter values for both the SAM and submillimeter SAM. There are also other signs of observational tensions with the Santa Cruz SAMs and other models; for example, Popping et al. (2019b) showed that both the Santa Cruz SAM and the IllustrisTNG hydrodynamic simulations show a deficit of large H_2 reservoirs at intermediate redshift $z \sim 1$–2. This could reflect inadequacies in the modeling of processes such as star formation and stellar feedback in the galaxy formation models or the fairly simple assumptions about conversion from the observed CO lines to H_2 mass adopted in that work. As already discussed, addressing a disagreement with a particular set of observations by varying the physics in the models is tricky, because the SAM has been developed over many years to produce agreement with a broad set of observations over a wide range in redshift. Ultimately, a more automated calibration procedure needs to be developed that can account for many observations simultaneously. Similarly, there are many assumptions that go into the subgrid models that need to be adopted to describe the ISM properties and carry out the line spectral synthesis. The properties of the ISM, such as the cloud mass function and cloud radial profiles, may vary in a complex way with environment or other galaxy properties. Again, as already noted, changing one of the subgrid ingredients to fit a particular line may result in worse agreement for a different line arising from a different part of the ISM, as we found when changing the cloud radial profile from a power law to a Plummer profile.

Key to resolving all of these challenges will be, of course, acquiring more data. Fortunately, there will be a dramatic increase in the volume and quality of LIM data in the next few years. The Carbon Monoxide Mapping Array Project (COMAP; Li et al. 2016) targets the CO(1–0) emission line at $z = 2.4$–3.4, right in the middle of the mmIME redshift range and near the excess C II emission seen in the Planck data, and will provide a highly complementary molecular gas measurement over a much larger intensity mapping volume. The Experiment for Cryogenic Large-Aperture Intensity Mapping (EXCLAIM; Cataldo et al. 2021) will probe the C II line in a similar redshift range and offer an improved measurement of that line to improve upon the Planck measurement. Several other reionization-focused projects have an mmIME-like ladder of CO lines as a "foreground" to their higher-redshift science (Lagache 2018; Stacey et al. 2018; Sun et al. 2021).

These new experiments will offer opportunities for more varied analysis of the intensity mapping data beyond the shot-noise power spectra used in mmIME. Some surveys, including COMAP and EXCLAIM, are designed with the possibility of cross-correlating their intensity maps with spectroscopic galaxy surveys in mind. This has the primary utility of improving confidence in the detection of a cosmological signal, but cross-correlations can also provide windows into unique aspects of galaxy evolution beyond those accessible to a single line (Serra et al. 2016; Breysse & Rahman 2017; Breysse & Alexandroff 2019; Chung et al. 2019). Cross-correlating pairs of tracers can also isolate a single line from interlopers that fall into the same frequency, obviating the need to model the sum of several lines at once, as has been done in this work.

With or without cross-correlating, future surveys covering larger areas can make use of larger scales in the power spectrum, which is sensitive to the DM-induced clustering of the galaxy population. This clustering component gives more weight to fainter galaxies, providing independent information about line luminosity functions (see derivation in the Appendix of Breysse & Alexandroff 2019). More detailed analyses like the anisotropic power spectrum (Cheng et al. 2016; Lidz & Taylor 2016; Bernal et al. 2019; Chung 2019) and the one-point intensity distribution (Breysse et al. 2017, 2019; Ihle et al. 2019) permit more detailed probes of the distribution of line intensity than the single shot power measurement available in mmIME. In turn, these measurements will constrain the relationship between halo mass and line luminosity for various tracers, thus placing constraints on the physical processes that are included in physics-based models. For example, the efficiency of stellar feedback will have a strong impact on the slope of the $L(M)$ relation at halo masses below around 10^{12} $M_\odot$, while the details of AGN feedback will impact the $L(M)$ relation and duty cycle at higher halo masses. In future work, we plan to systematically explore how variations in the uncertain components of these physical processes translate into LIM observables.

9. Conclusion

We have demonstrated here that a modest tension exists between new LIM observations of high-redshift galaxies from the mmIME survey and state-of-the-art SAMs of the line emission from the dense ISM in galaxies. At $\sim 3.5\sigma$, this tension is relatively minor and may weaken with time and new data. However, these are some of the first intensity mapping results at these redshifts, and the SAM+submillimeter SAM routines have been calibrated to date primarily on direct galaxy survey measurements, so even this modest tension is interesting. We therefore go on to explore the ways in which this tension could be caused by physical processes that impact the LIM measurements.

We carry out a modified version of the K20 analysis of the mmIME-based measurement of the cosmic molecular gas abundance, using our internally consistent, physically motivated SAM+submillimeter SAM scaling relations at all steps. In doing so, we find implied ρ_{H2} values roughly twice as high as those obtained in K20, corresponding to an $\sim 2\sigma$ increase. As the K20 ρ_{H2} fit is already a bit higher than the measurements obtained from direct-imaging galaxy surveys, our result is quite a bit higher than any previous measurement. In other words, if we assume, as in K20, that any tension between mmIME and the model predictions is due to a change in the molecular gas abundance, the modest tension in the intensity map becomes a similar tension in ρ_{H2}. If this interpretation is valid, it implies that there is a substantial population of CO-emitting galaxies that are statistically detectable to LIM but have not been seen in galaxy surveys.

However, we go on to show that the tension could similarly be resolved by a number of other changes in the modeling assumptions. Most simply, we show that we could equivalently account for the tension by reducing the value of α_{CO} in high-redshift galaxies without changing the molecular gas abundance at all. We also show that small changes to the SAM subgrid ingredients can solve the tension in ρ_{H2} even if the overall tension with mmIME remains. Finally, we relax our reliance on the SAM+submillimeter SAM slightly and show that substituting some of the not-unreasonable scaling relations originally used in K20 can shift our ρ_{H2} result over a range quite a bit larger than the statistical error bars on the measurement, with some choices weakening the tension and others leading to even higher ρ_{H2} values.

It is thus clear that, while current data and model predictions have perhaps begun to differ in interesting ways, neither has

progressed to the point of being able to confidently understand the source of those differences. This state of affairs is rapidly changing, however, particularly on the experimental side. Intensity mapping data in particular will dramatically expand in the next few years, as many dedicated LIM instruments begin to produce results. Perhaps the most important conclusion of this work is that substantial effort will be required on the modeling and simulation fronts if we are to fully understand the new regimes probed by these measurements.

ORCID iDs

Patrick C. Breysse https://orcid.org/0000-0001-8382-5275
Shengqi Yang https://orcid.org/0000-0002-0782-9116
Gergö Popping https://orcid.org/0000-0003-1151-4659
Abhishek S. Maniyar https://orcid.org/0000-0002-4617-9320

References

Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57
Bernal, J. L., Breysse, P. C., Gil-Marín, H., & Kovetz, E. D. 2019, PhRvD, 100, 123522
Berta, S., Lutz, D., Nordon, R., et al. 2013, A&A, 555, L8
Béthermin, M., Fudamoto, Y., Ginolfi, M., et al. 2020, A&A, 643, A2
Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207
Boogaard, L. A., Bouwens, R. J., Riechers, D., et al. 2021, ApJ, 916, 12
Breysse, P. C., & Alexandroff, R. M. 2019, MNRAS, 490, 260
Breysse, P. C., Anderson, C. J., & Berger, P. 2019, PhRvL, 123, 231105
Breysse, P. C., Kovetz, E. D., Behroozi, P. S., Dai, L., & Kamionkowski, M. 2017, MNRAS, 467, 2996
Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2014, MNRAS, 443, 3506
Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2015, MNRAS, 452, 3408
Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2016, MNRAS, 457, L127
Breysse, P. C., & Rahman, M. 2017, MNRAS, 468, 741
Carilli, C. L., & Walter, F. 2013, ARA&A, 51, 105
Cataldo, G., Ade, P., Anderson, C., et al. 2021, arXiv:2101.11734
Cheng, Y.-T., Chang, T.-C., Bock, J., Bradford, C. M., & Cooray, A. 2016, ApJ, 832, 165
Cheng, Y.-T., Chang, T.-C., & Bock, J. J. 2020, ApJ, 901, 142
Chung, D. T. 2019, ApJ, 881, 149
Chung, D. T., Viero, M. P., Church, S. E., et al. 2019, ApJ, 872, 186
Combes, F. 2018, A&ARv, 26, 5
Daddi, E., Dannerbauer, H., Liu, D., et al. 2015, A&A, 577, A46
Decarli, R., Aravena, M., Boogaard, L., et al. 2020, ApJ, 902, 110
Decarli, R., Walter, F., Gónzalez-López, J., et al. 2019, ApJ, 882, 138
Fletcher, T. J., Saintonge, A., Soares, P. S., & Pontzen, A. 2021, MNRAS, 501, 411
Fonseca, J., Silva, M. B., Santos, M. G., & Cooray, A. 2017, MNRAS, 464, 1948
Freundlich, J., Combes, F., Tacconi, L. J., et al. 2019, A&A, 622, A105

Gong, Y., Silva, M., Cooray, A., & Santos, M. G. 2014, ApJ, 785, 72
González-López, J., Decarli, R., Pavesi, R., et al. 2019, ApJ, 882, 139
Ihle, H. T., Chung, D., Stein, G., et al. 2019, ApJ, 871, 75
Keating, G. K., Marrone, D. P., Bower, G. C., et al. 2016, ApJ, 830, 34
Keating, G. K., Marrone, D. P., Bower, G. C., & Keenan, R. P. 2020, ApJ, 901, 141
Keenan, R. P., Marrone, D. P., & Keating, G. K. 2020, ApJ, 904, 127
Klypin, A., Yepes, G., Gottlöber, S., Prada, F., & Heß, S. 2016, MNRAS, 457, 4340
Kovetz, E. D., Viero, M. P., Lidz, A., et al. 2017, arXiv:1709.09066
Krumholz, M. R. 2014, MNRAS, 437, 1662
Lagache, G. 2018, in Peering towards Cosmic Dawn, ed. V. Jelić & T. van der Hulst, Vol. 333 (Cambridge: Cambridge Univ. Press), 228
Lenkić, L., Bolatto, A. D., Förster Schreiber, N. M., et al. 2020, AJ, 159, 190
Li, T. Y., Wechsler, R. H., Devaraj, K., & Church, S. E. 2016, ApJ, 817, 169
Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, ApJ, 741, 70
Lidz, A., & Taylor, J. 2016, ApJ, 825, 143
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Magnelli, B., Boogaard, L., Decarli, R., et al. 2020, ApJ, 892, 66
Moradinezhad Dizgah, A., & Keating, G. K. 2020, ApJ, 872, 126
Padmanabhan, H. 2018, MNRAS, 475, 1477
Pavesi, R., Sharon, C. E., Riechers, D. A., et al. 2018, ApJ, 864, 49
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A1
Popping, G., Narayanan, D., Somerville, R. S., Faisst, A. L., & Krumholz, M. R. 2019a, MNRAS, 482, 4906
Popping, G., Pillepich, A., Somerville, R. S., et al. 2019b, ApJ, 882, 137
Popping, G., Somerville, R. S., & Trager, S. C. 2014, MNRAS, 442, 2398
Popping, G., van Kampen, E., Decarli, R., et al. 2016, MNRAS, 461, 93
Porter, L. A., Somerville, R. S., Primack, J. R., & Johansson, P. H. 2014, MNRAS, 444, 942
Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15
Rickard, L. J., Palmer, P., Morris, M., Turner, B. E., & Zuckerman, B. 1977, ApJ, 213, 673
Rickard, L. J., Palmer, P., Morris, M., Zuckerman, B., & Turner, B. E. 1975, ApJL, 199, L75
Riechers, D. A., Boogaard, L. A., Decarli, R., et al. 2020, ApJL, 896, L21
Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, ApJ, 872, 7
Righi, M., Hernández-Monteagudo, C., & Sunyaev, R. A. 2008, A&A, 489, 489
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Scoville, N., Lee, N., Vanden Bout, P., et al. 2017, ApJ, 837, 150
Serra, P., Doré, O., & Lagache, G. 2016, ApJ, 833, 153
Solomon, P. M., & de Zafra, R. 1975, ApJL, 199, L79
Somerville, R. S., Gilmore, R. C., Primack, J. R., & Domínguez, A. 2012, MNRAS, 423, 1992
Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, MNRAS, 391, 481
Somerville, R. S., & Kolatt, T. S. 1999, MNRAS, 305, 1
Somerville, R. S., Olsen, C., Yung, L. Y. A., et al. 2021, MNRAS, 502, 4858
Somerville, R. S., Popping, G., & Trager, S. C. 2015, MNRAS, 453, 4337
Somerville, R. S., & Primack, J. R. 1999, MNRAS, 310, 1087
Stacey, G. J., Aravena, M., Basu, K., et al. 2018, Proc. SPIE, 10700, 107001M
Sun, G., Chang, T.-C., Uzgil, B. D., et al. 2021, ApJ, 915, 33
Sun, G., Hensley, B. S., Chang, T.-C., Doré, O., & Serra, P. 2019, ApJ, 887, 142
Sun, G., Moncelsi, L., Viero, M. P., et al. 2018, ApJ, 856, 107
Tacconi, L. J., Genzel, R., & Sternberg, A. 2020, ARA&A, 58, 157
Tacconi, L. J., Neri, R., Genzel, R., et al. 2013, ApJ, 768, 74
Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709
Valentino, F., Magdis, G. E., Daddi, E., et al. 2020, ApJ, 890, 24
Walter, F., Carilli, C., Neeleman, M., et al. 2020, ApJ, 902, 111
Walter, F., Decarli, R., Sargent, M., et al. 2014, ApJ, 782, 79
Yang, S., Pullen, A. R., & Switzer, E. R. 2019, MNRAS, 489, L53
Yang, S., Somerville, R. S., Pullen, A. R., et al. 2021, ApJ, 911, 132
Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., Popping, G., & Davé, R. 2019a, MNRAS, 483, 2983
Yung, L. Y. A., Somerville, R. S., Popping, G., et al. 2019b, MNRAS, 490, 2855
Zanella, A., Daddi, E., Magdis, G., et al. 2018, MNRAS, 481, 1976