An Intensity Mapping Constraint on the CO-galaxy Cross-power Spectrum at Redshift ∼3

Ryan P. Keenan[1,3] , Garrett K. Keating[2] , and Daniel P. Marrone[1]
[1] Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA; rpkeenan@email.arizona.edu
[2] Center for Astrophysics, Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

Abstract

The abundance of cold molecular gas plays a crucial role in models of galaxy evolution. While deep spectroscopic surveys of CO emission lines have been a primary tool for measuring this abundance, the difficulty of these observations has motivated alternative approaches to studying molecular gas content. One technique, line intensity mapping, seeks to constrain the average molecular gas properties of large samples of individually undetectable galaxies through the CO brightness power spectrum. Here we present constraints on the cross-power spectrum between CO intensity maps and optical galaxy catalogs. This cross-measurement allows us to check for systematic problems in CO intensity mapping data, and validate the data analysis used for the auto-power spectrum measurement of the CO Power Spectrum Survey. We place a 2σ upper limit on the band-averaged CO-galaxy cross-power of $P_\times < 540\ \mu\mathrm{K}\,h^{-3}\,\mathrm{Mpc}^3$. Our measurement favors a nonzero $\langle T_{\mathrm{CO}} \rangle$ at around 90% confidence and gives an upper limit on the mean molecular gas density at $z \sim 2.6$ of $7.7 \times 10^8\ M_\odot\ \mathrm{Mpc}^{-3}$. We forecast the expected cross-power spectrum by applying a number of literature prescriptions for the CO luminosity–halo mass relation to a suite of mock light cones. Under the most optimistic forecasts, the cross-spectrum could be detected with only moderate extensions of the data used here, while more conservative models could be detected with a factor of 10 increase in sensitivity. Ongoing CO intensity mapping experiments will target fields allowing for extensive cross-correlation analysis and should reach the sensitivity required to detect the cross-spectrum signal.

1. Introduction

Cold molecular gas represents the raw material for star formation. Understanding its abundance over cosmic time is a necessary ingredient for theories of galaxy formation and evolution (Carilli & Walter 2013; Walter et al. 2020). Blind surveys of emission from CO, the preferred tracer of the total molecular gas content in galaxies, have begun to probe its redshift evolution (Walter et al. 2014; Decarli et al. 2016; Pavesi et al. 2018; Decarli et al. 2019; Lenkić et al. 2020). However, until at least the advent of the next-generation Very Large Array (ngVLA), these types of survey will be limited to sky areas of a few square arcminutes and only tens of secure direct detections (Decarli et al. 2020). As a result, these samples are subject to large statistical uncertainties, which make it difficult to determine trends with redshift (Keenan et al. 2020).

Line intensity mapping (LIM) offers a complementary approach to direct detection efforts. Instead of searching for individual objects at high significance, LIM measures line intensity in large fields. The integrated luminosity of every line-emitting galaxy is extracted from the power spectrum of the three-dimensional line intensity distribution and used to constrain the properties of galaxies too faint to be individually detected (Visbal & Loeb 2010; Gong et al. 2011; Lidz et al. 2011; Breysse et al. 2014; Li et al. 2016). Such data sets can be used to constrain the

luminosity function of CO emission lines and thus the abundance evolution of molecular gas. The first LIM measurements have provided constraints on the total CO abundance at redshifts $1 < z < 6$ (Pullen et al. 2013; Keating et al. 2015, 2016; Uzgil et al. 2019; Keating et al. 2020). With dedicated intensity mapping instruments now taking data (e.g., TIME; Sun et al. 2021; COMAP; Ihle et al. 2019; CONCERTO; Concerto Collaboration et al. 2020), this technique is set to advance our understanding of high redshift molecular gas in the near future.

However, understanding the coevolution of star formation and molecular gas abundance will require synthesis of radio and submillimeter gas measurements with data from optical and IR (OIR) galaxy surveys. Placing the statistical results of intensity mapping in the context of individually detected OIR galaxies represents a potential challenge. One promising path forward is through cross-correlation between catalogs from galaxy surveys and CO maps produced by LIM experiments (Pullen et al. 2013; Wolz et al. 2016, 2017; Chung et al. 2019). By making use of the large-scale clustering present in the maps and the galaxy distribution, cross-correlation can extract information beyond what is derived from stacking spectra extracted at known galaxy positions. Cross-correlations between intensity maps and a broad range of other data sets have been explored as a tool for addressing numerous astrophysical and cosmological problems, including calibration of photometric redshifts (Cunnington et al. 2019), characterizing feedback from active galactic nuclei (Breysse & Alexandroff 2019), determining physical properties of the interstellar medium (ISM; Sun et al. 2019), exploring the Lyα forest (Carucci et al. 2017), measuring baryon acoustic oscillations at high redshift (Cohn et al. 2016), constraining properties of neutrinos (Moradinezhad Dizgah et al. 2022a), mitigating the effects of cosmic variance (Oxholm & Switzer 2021), and

---

exploring models of primordial non-Gaussianity (Moradinezhad Dizgah & Keating 2019), among others.

Cross-correlation also serves as a check on systematics in intensity mapping data, and can be used to validate detections of the autocorrelation power spectrum (Furlanetto & Lidz 2007; Silva et al. 2015). This is true not only for CO but also ongoing hydrogen 21 cm (Pen et al. 2009; Chang et al. 2010; Masui et al. 2013; Switzer et al. 2013, 2015; Anderson et al. 2018; Wolz et al. 2022), [C II] 158 μm (Pullen et al. 2018; Switzer et al. 2019; Yang et al. 2019), and Lyα (Croft et al. 2018) intensity mapping projects.

The CO Power Spectrum Survey (COPSS; Keating et al. 2015, 2016, hereafter K15 and K16) was a first-generation intensity mapping experiment, which obtained thousands of hours of observations targeting emission from the CO(1–0) transition at redshift $z \sim 2.3$–3.3, constraining the total luminosity of CO during the peak of cosmic star formation. The COPSS observations were designed to allow cross-correlation by targeting regions with extensive coverage in optical/infrared spectroscopy. Here we study the cross-correlation between deep 30 GHz observations of the GOODS-N field from COPSS and a large catalog of galaxies with spectroscopic redshifts.

In Section 2 we review the mathematical formalism of the LIM technique and cross-correlation. In Section 3 we describe the 30 GHz and optical data we use. We detail our data analysis procedure for both cross-correlation and stacking in Section 4 and present the results of this analysis in Section 5. In Section 6 we describe tests to verify that our analysis is not unduly affected by systematics. In Section 7 we model the CO-galaxy cross spectrum in the presence of measurement errors, to verify our analysis methodology and quantify measurement uncertainties. We place our results in the context of theoretical expectations and upcoming experiments, and use them to constrain the average CO luminosity and cosmic molecular gas density at $z \sim 3$ in Section 8. We conclude in Section 9. Throughout we assume a flat Λ cold dark matter cosmology with $H = 70$ km s^{-1} Mpc^{-1} and $\Omega_M = 0.27$.

2. Intensity Mapping Formalism

The power spectrum of an intensity field measures the variance in brightness temperature fluctuations as a function of spatial frequency. It encodes information about the total luminosity of line emitting sources as well as information about their clustering. Here we present the formalism in terms of the CO(1–0) line transition, but note that it applies to any emission line. The power spectrum at co-moving scale k can be expressed as the sum of three terms:

$$P_{CO}(k) = P_{2h,CO} + P_{1h,CO} + P_{shot,CO}, \tag{1}$$

where $P_{2h,CO}$ is the "two-halo" term, caused by clustering of dark matter halos, $P_{1h,CO}$ is the "one-halo" term, related to the distribution of CO emitters within individual dark matter halos, and $P_{shot,CO}$ is the "shot noise" due to the discrete nature of individual emitters.

The two-halo term can be expressed as

$$P_{2h,CO} = b_{CO}^2 \left(C_{LT} \int_0^\infty L_{CO} \frac{dn}{dL} dL \right)^2 P_{lin}(k), \tag{2}$$

where b_{CO} is the bias of CO emitters, L_{CO} is CO luminosity, dn/dL is the CO luminosity function, and $P_{CO}(k, z)$ is the power spectrum at co-moving scale k. The factor C_{LT} is a

conversion factor between luminosity and brightness temperature units,

$$C_{LT} = \frac{c^3 (1 + z)^2}{8\pi k_B H(z) \nu_{rest}^3}, \tag{3}$$

c is the speed of light, k_B is the Boltzmann constant, $H(z)$ is the Hubble parameter at redshift z, and ν_{rest} is the rest frequency of CO(1–0) emission, 115.27 GHz. The term in parentheses in Equation (2) is therefore the mean CO brightness temperature of the volume under consideration $\langle T_{CO} \rangle$, and we can write

$$P_{2h,CO} = b_{CO}^2 \langle T_{CO} \rangle^2 P_{lin}(k). \tag{4}$$

This term is the dominant contribution to the power spectrum at small k.

The shot power term can be expressed as

$$P_{shot,CO} = C_{LT}^2 \int_0^\infty L_{CO}^2 \frac{dn}{dL} dL. \tag{5}$$

On scales smaller than a few megaparsecs ($k \gtrsim 1$ Mpc^{-1}), the shot power is the dominant contribution to the power spectrum.

The one-halo term can further enhance the power spectrum at intermediate scales ($k \sim 1$ Mpc^{-1}; Moradinezhad Dizgah et al. 2022b; Schaan & White 2021). This term is, in general, a more complex function of the halo density profile, the distribution of central and satellite galaxies within that profile, and the distribution of CO luminosities over the galaxies in the halo. In the limit where every CO emitter corresponds to a central galaxy, this term is simply the shot power (i.e., there is no one-halo contribution). We refer readers to, e.g., Wolz et al. (2019) and Schaan & White (2021) for parameterizations of the one-halo power.

Analogously to Equation (1), the cross-power spectrum of a galaxy overdensity field and a CO intensity field may be written as

$$P_\times(k) = P_{2h,\times} + P_{1h,\times} + P_{shot,\times}. \tag{6}$$

We can express the two-halo component of the cross-spectrum as

$$P_{2h,\times} = r_{2h} b_{gal} b_{CO} \langle T_{CO} \rangle P_{lin}(k), \tag{7}$$

where r_{2h} is a stochasticity term, reflecting that the CO emission and galaxy catalog may not trace the matter power spectrum in an identical manner, and b_{gal} is the bias of the galaxies in the selected galaxy survey. The r_{2h} term has typically been excluded from formulations of the cross-power spectrum in CO intensity mapping applications (e.g., Breysse & Alexandroff 2019; Schaan & White 2021), but its importance has been explored in the context of 21 cm intensity mapping (Switzer et al. 2013; Wolz et al. 2017; Anderson et al. 2018). In the remainder of this analysis, we assume $r_{2h} = 1$, as the correlation between two tracers is expected to approach unity on scales where the two-halo term dominates. This reduces Equation (7) to

$$P_{2h,\times} = b_{gal} b_{CO} \langle T_{CO} \rangle P_{lin}(k). \tag{8}$$

We revisit and further justify this assumption in Appendix B, where we also show that the main conclusions of our analysis are unchanged when we allow $r_{2h} < 1$.

The cross-shot power is

$$P_{\text{shot},\times} = \frac{1}{n_{\text{gal}}} C_{\text{LT}} \sum_{\text{cat}} \frac{L_{i,\text{CO}}}{V}, \qquad (9)$$

where n_{gal} is the number density of catalog galaxies, $L_{\text{CO},i}$ is the CO luminosity of the ith member of the galaxy catalog, V is the volume sampled, and the sum is taken over the galaxies in the catalog. The cross-shot power can thus be used to measure the average CO luminosity of the optical galaxy catalog, equivalent to measuring the average CO luminosity of catalog galaxies via stacking (Wolz et al. 2017; Breysse & Alexandroff 2019).

The one-halo term is again more complicated than the two-halo or shot terms. The galaxy catalog explored in the remainder of this study is composed of targets from spectroscopic surveys at $z \sim 3$. It can be expected that such surveys are most likely to both target and detect the most massive galaxy in a given halo. In the case that catalog galaxies are all centrals, the one-halo power reaches a maximum of

$$P_{\text{1h},\times}(k \to 0) = \frac{1}{n_{\text{gal}}} C_{\text{LT}} \int_0^\infty L_{\text{CO}}^{\text{sats}}(m) p_{\text{gal}}(m) \frac{dn}{dm} dm \qquad (10)$$

where $L_{\text{CO}}^{\text{sats}}(m)$ is the CO luminosity contributed by satellite galaxies in a halo of mass m, $p_{\text{gal}}(m)$ is the probability for a halo of mass m to contain a galaxy in the galaxy catalog, and dn/dm is the halo mass function (this is a simplification of Equation (21) of Wolz et al. 2019). In the case that the central galaxies also account for most of the CO luminosity of the halo, this implies that the one-halo term will be smaller than the shot power on all scales.

3. Observations and Data

3.1. CO Observations

The CO observations used here were conducted with the Sunyaev–Zel'dovich Array (SZA), an eight element array of 3.5 m antennas, part of the Combined Array for Research in Millimeter-wave Astronomy (CARMA). Observations were conducted from 2013 April to 2015 April as part of COPSS. The fields of the COPSS survey were observed over a frequency range of 27–35 GHz, covered in 16 spectral windows of 500 MHz bandwidth. The field of view of the SZA is $\sim13''$ at the low-frequency end of the band and $\sim10''$ at the high-frequency end. Details of the survey can be found in K16. We use data for the GOODS-N field from both the pilot and primary phases of the COPSS survey.

3.2. Galaxy Catalogs

A number of large spectroscopic surveys have targeted GOODS-N. We have attempted to compile the most complete set of spectroscopic redshifts available by synthesizing these into a single catalog. We draw redshifts from the following sources:

1. The MOSFIRE Deep Evolution Field (MOSDEF) survey of near-IR (NIR) selected galaxies at $z \gtrsim 1.4$ using the Keck MOSFIRE spectrograph (Kriek et al. 2015), giving 124 galaxies.
2. The Team Keck Redshift Survey 2 (TKRS2) with Keck/MOSFIRE (Wirth et al. 2015), with 19 galaxies.
3. A sample of 23 MOIRCS Deep Survey BzK galaxy redshifts from Yoshikawa et al. (2010).

4. The catalog of Barger et al. (2008), which compiles most prior redshifts as well as new measurements and includes a total of 2710 redshifts.
5. A survey of optically selected BM/BX galaxies and luminous blue galaxies (LBGs) conducted with the Keck LRIS-B spectrograph (Reddy catalog; Steidel et al. 2004; Reddy et al. 2006).

The same galaxy may appear in multiple catalogs. Therefore, we search for objects within $1''$ (about twice the typical MOSDEF seeing) and $\Delta z/(1+z) < 0.003$ in two or more of the catalogs and remove objects from the older data set, we also remove objects with offsets $<0.''25$ irrespective of redshift.

We then cut the catalog to the redshift range covered by the COPSS frequency band. For CO(1–0) emission, this range is $2.3 < z < 3.2$.

These cuts leave us with a total of 124 MOSDEF galaxies, 17 additional galaxies from TKRS2, 0 galaxies for MOIRCS, 78 galaxies from Barger et al. (2008), and three galaxies from Reddy et al. (2006) for a total of 224 galaxies. Sixteen of these fall in gaps between SZA spectral windows, leaving 208 galaxies.

The positions of all galaxies in our data sample are shown in Figure 1, and their redshift distribution within the spectral windows of our SZA data is shown in Figure 2. Because the primary beam of the telescope tapers from the pointing center, we also count galaxies weighted by the primary beam of the SZA, which gives us an effective total of 145 galaxies.

In addition to spectroscopic catalogs, extensive grism spectroscopy from the 3D-HST program (Momcheva et al. 2016) is available. However, Chung et al. (2019) found that the redshift uncertainty of low-resolution grism spectra results in significant attenuation of the cross-power, and therefore we do not include them. This results in the exclusion of 69 grism redshifts.

Each of the catalogs discussed above is selected on the basis of different properties and for different science goals. Therefore, they do not represent a homogeneous population of galaxies. For the purpose of maximizing the sensitivity of our CO-galaxy cross-power spectrum, we nevertheless use the combined data sets. We also explored cross-correlating the individual catalogs with the COPSS data, but did not detect the cross-power for any subset of galaxies explored. Cross-correlation of subsets of the galaxy population will have to await more sensitive intensity mapping data sets.

4. Data Analysis

The cleaning and analysis pipeline for the SZA data is described in detail in K15 and K16. We provide a brief summary here. The raw interferometric data are recorded as three-dimensional arrays gridded in angular frequency along the sky plane (u and v) and in frequency (ν). During the primary COPSS survey, trailing fields offset by five and ten minutes in R.A. were observed at the same decl. as the GOODS-N field, in order to facilitate removal of contamination from antenna cross-talk and ground emission. Data from all three fields are averaged to produce a model of the contamination, and the model is subtracted from each field. This is not possible for data from the pilot phase of the survey, which were taken without trailing fields, and contamination is a significant problem in short baselines. To reduce this contamination, measurements exceeding 4σ significance are
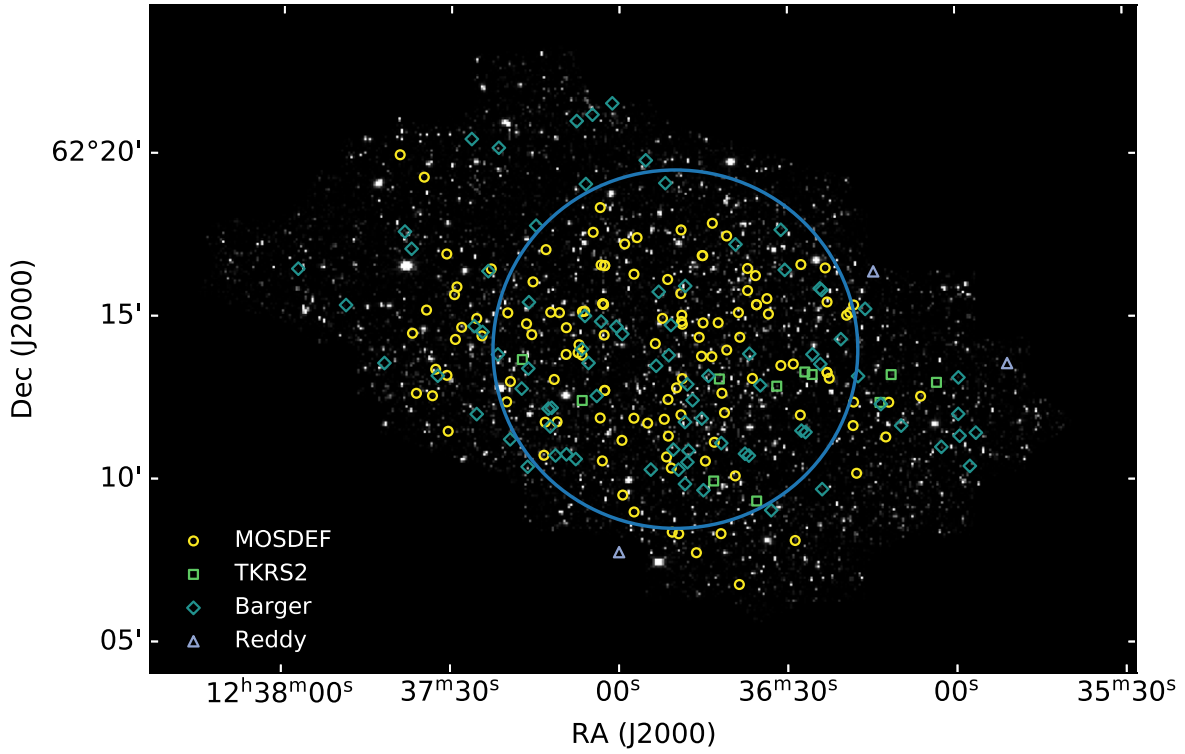
Figure 1. The sample of $2.3 < z < 3.2$ galaxies with spectroscopic redshifts, plotted over the 3D-HST combined F140W+F125W+F160W image of GOODS-N. Galaxies drawn from MOSDEF, TKRS2, Barger et al. (2008), and Reddy et al. (2006) are indicated with yellow circles, green squares, blue diamonds, and lavender triangles, respectively. The large blue circle is the $11'$ primary beam of the SZA at 31 GHz.
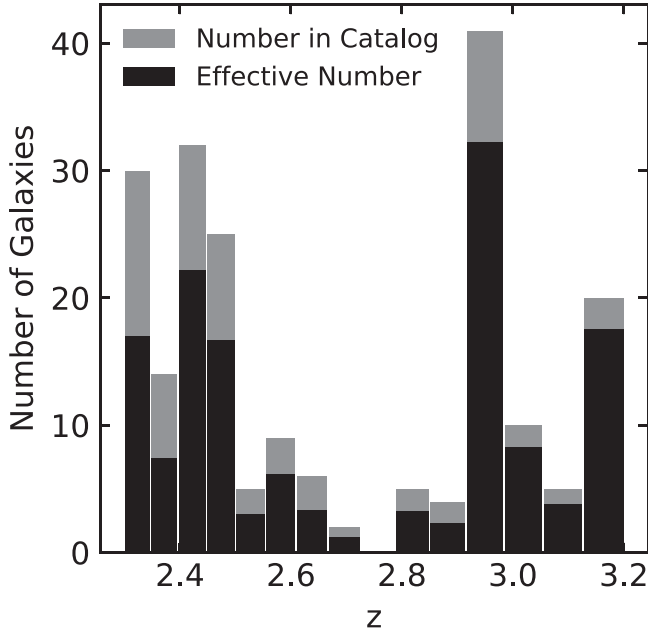

Figure 2. The redshift distribution of our final galaxy sample. The bins correspond to frequency windows of our SZA data. Shown in gray are the total number of galaxies in each window. Due to the tapering of the SZA primary beam, not every galaxy in the field received equal weight in our analysis. We define the effective number of galaxies, with each galaxy weighted by the primary beam response at its position. The black histogram shows this distribution. As the primary beam becomes larger at lower frequency (higher redshift) while the redshift catalog area remains fixed, the effective number of galaxies increases relative to the total in high redshift windows.

removed. The data from the primary survey contain no measurements above this threshold, and thus we do not expect that any pilot measurements are above this threshold due to

astronomical signals. As a result, this data cut should not filter out the signal of interest.

In autocorrelation, ground contamination will be a source of excess power and must be carefully removed from the data to avoid a positive bias. In cross-correlation, the ground contamination does not correlate with galaxy positions, and therefore naturally drops out of the cross-spectrum without introducing any bias. However, random alignments of the galaxies and ground emission introduce additional noise. For our data, we find that cleaning the ground contamination improves the sensitivity of the cross-spectrum despite the noise that is added as a result of field differencing, and we retain this step. We verify in Section 6.2 that neither field differencing nor sigma-clipping alter our final power spectrum.

The grids are next Fourier transformed along the frequency axis within each spectral window, to produce "delay visibilities." For each window, this procedure produces a grid in (u, v, η) where η is the Fourier transform of frequency, also called the delay.

We expect continuum sources to be spectrally smooth over the frequency range of each window. For typical spectral indices seen at 30 GHz, these sources will predominantly show power in the $\eta = 0$ bin. To remove these sources, we drop this η bin from our grids. K16 verify that this procedure produces auto-power spectra without significant contamination from continuum sources.

Next, the grid is converted from observable (u, v, η) coordinates to co-moving megaparsec units at the redshift of the expected CO signal:

$$\boldsymbol{k} = (k_x, k_y, k_z) = \left(\frac{2\pi u}{X(z)}, \frac{2\pi v}{X(z)}, \frac{2\pi \eta}{Y(z, \nu_{\mathrm{rest}})} \right) \quad (11)$$

where $X = D_M(z)$ the co-moving transverse distance at redshift z and $Y = c(1 + z)^2/H(z)\nu_{rest}$, c is the speed of light, $H(z)$ is the redshift-dependent Hubble parameter, and ν_{rest} is the rest frequency of the target line (115.27 GHz for CO(1–0)).

4.1. CO Auto-Spectra

To produce the CO auto-power spectrum, we use the estimator

$$\mathcal{P}(\boldsymbol{k}) = \frac{1}{V_{eff}}\left[\frac{\Sigma_{\boldsymbol{k}'}\sigma_k^{-2}\sigma_{k'}^{-2}C(\boldsymbol{k} - \boldsymbol{k}')(|\tilde{T}^*(\boldsymbol{k})\tilde{T}(\boldsymbol{k}')|)}{\Sigma_{\boldsymbol{k}'}\sigma_k^{-2}\sigma_{k'}^{-2}C(\boldsymbol{k} - \boldsymbol{k}')} - \mathcal{A}_k\right] \quad (12)$$

$$P(k) = \langle\mathcal{P}(\boldsymbol{k})\rangle_{\boldsymbol{k}\cdot\boldsymbol{k}} = k^2 \quad (13)$$

where $\mathcal{P}$ is the three-dimensional power spectrum, $\sigma_{\boldsymbol{k}}$ is the estimated thermal noise, C is the expected normalized covariance matrix, and $\tilde{T}$ is the delay visibility. V_{eff} is the effective volume probed by the measurement, which for a Gaussian beam of solid angle Ω and bandwidth B is given by $V_{eff} = X^2YB\Omega/2$. $\mathcal{A}$ is the sum of the autocorrelations of individual delay visibilities within each grid cell, which is subtracted to remove noise bias. The averaging in Equation (13) is done weighting each cell by the inverse of its estimated thermal noise variance.

4.2. Cross-spectra

To compute the cross-power spectrum, we construct a grid in angular coordinates (θ, ϕ) and frequency (ν) to capture the galaxy distribution. To calculate the latter, we compute the redshifted frequency of the CO(1–0) line given the reported optical redshift of the galaxy, such that $\nu = \nu_{rest}/(1 + z)$. We populate the grid by counting the number N of galaxies falling within each cell. Then we compute the overdensity field $\delta_N(\theta, \phi, \nu) = (N - \bar{N})/\bar{N}$ where $\bar{N}$ is the mean number of galaxies per cell in the region covered by the optical surveys. The resultant grid is then Fourier transformed across all three dimensions and converted into $\boldsymbol{k}$ coordinates via Equation (11).

The finite resolution of the grid results in quantization errors, which will cause some decorrelation between the optical and radio data sets. To mitigate this, we use a grid three times finer in angular resolution, and five times finer in frequency resolution than the SZA data cubes.

To estimate the cross-power spectrum, we use the estimator

$$\mathcal{P}_{\times}(\boldsymbol{k}) = \frac{1}{V_{\times,eff}}\mathcal{R}\left(\frac{\tilde{T}^*(\boldsymbol{k})\tilde{\delta}_N(\boldsymbol{k})}{A(\boldsymbol{k})}\right) \quad (14)$$

$$P_{\times}(k) = \langle\mathcal{P}_{\times}(\boldsymbol{k})\rangle_{\boldsymbol{k}\cdot\boldsymbol{k}=k^2} \quad (15)$$

where $V_{\times,eff}$ is the effective volume of the cross-power measurement, $\mathcal{R}$ denotes the real part of a complex number, $\tilde{T}$ and $\tilde{\delta}_N$ are the Fourier duals of T and δ_N, respectively, and $A(k)$ corrects for attenuation of the observed power spectrum (this correction is detailed Section 4.2.1). The averaging in Equation (15) is done weighting each grid cell by the inverse of its variance, after applying attenuation corrections.

To determine the effective volume, we define a survey footprint function $F(\theta, \phi)$ to be 1 at positions (θ, ϕ) within the footprint of the optical galaxy catalogs of GOODS-N and 0 elsewhere. This footprint is determined primarily by HST imaging coverage of the field (Skelton et al. 2014), with slight
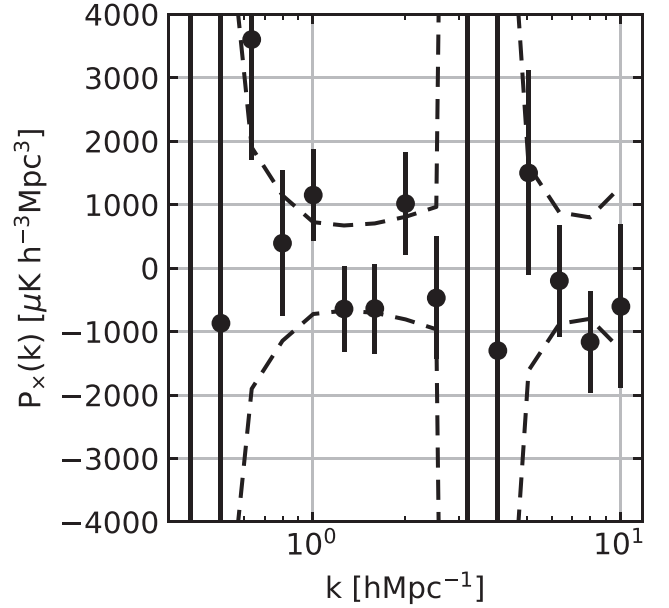

Figure 3. The results of our cross-spectrum analysis between the COPSS 30 GHz observations of GOODS-N and our spectroscopic catalog of OIR galaxies. The dashed curves show the 1σ sensitivity level of our measurement.

expansions to cover a small number of galaxies from Reddy et al. (2006) that lie beyond the edges of the HST coverage. We then compute

$$\Omega_{\times,eff} = \int F(\theta, \phi) \times W(\theta, \phi)d\Omega, \quad (16)$$

where $W(\theta, \phi)$ is the primary beam response of the SZA. This allows us to compute the effective volume as $V_{\times,eff} = X^2YB\Omega_{\times,eff}$.

To estimate the noise in the cross-power spectrum of each window, we randomized the phases of the SZA delay visibilities and recomputed the cross-power using the same galaxy grid. The phase randomization removes all position information from the radio data sets, so this is equivalent to computing the cross-spectrum of the galaxy field with random noise in the radio data set. We perform 10^4 trials to generate the expected noise distribution for each frequency window. We then determine the 1σ noise level from the 15.9 and 84.1 percentiles of the distribution.

We construct final power spectra by combining each spectral window and determining the average power by further averaging across all k-modes. Our cross-power spectrum results are reported in Section 5 and Figure 3.

4.2.1. Attenuation Correction

In practice, a number of observational effects are known to attenuate the observed auto- and cross-spectra. Errors in spectroscopic redshifts can shift galaxies between cells of our overdensity field, causing position mismatches with the CO temperature field that cause decorrelation along the line of sight for small scale modes of the cross-spectrum (Chung et al. 2019). Moreover, galaxy rotation and bulk motions of the ISM broaden CO emission lines, resulting in typical line profiles hundreds of kilometers per second wide. As these line widths are comparable to the size of the spectral channels in our SZA data, we expect the signal to spread into adjacent channels.

This will result in attenuation of both the auto- and cross-spectra along the line of sight (Chung et al. 2021).

The attenuation of the COPSS auto-power spectrum has been explored in Keating et al. (2020). This effect is found to be fairly small, and as we make no effort to interpret our auto-spectrum here, we do not consider attenuation corrections to the auto-power estimator from Section 4.1. However, the additional effect of redshift errors can significantly attenuate the cross-spectrum signal, particularly at high k. Therefore we attempt to correct this attenuation in our three-dimensional cross-power spectrum.

The attenuation of the cross-spectrum can be expressed as

$$A(\boldsymbol{k}) = \exp\left[-\frac{1}{2}k_z^2 \times \left(\frac{\sigma_{\rm lw}^2}{c^2} + \sigma_z^2\right)\left(Y(z, \nu_{\rm rest})\frac{\nu_{\rm rest}}{1+z}\right)^2\right] \quad (17)$$

where k_z is the line-of-sight component of $\boldsymbol{k}$, $\sigma_{\rm lw}$ describes the width of the CO emission lines and for a Gaussian line profile is equal to the FWHM divided by 2.355, and σ_z is the redshift uncertainty. The final term puts these widths of these uncertainties in units of co-moving length.

Kriek et al. (2015) compared the spectroscopic redshifts from the MOSDEF survey and a compilation of prior spectroscopic measurements, finding a typical $\sigma_z/(1+z)$ of 0.001. Most other catalogs we draw from do not provide detailed redshift uncertainties; however, the TKRS2 redshifts are measured with the same telescope and instrument as MOSDEF, and the remaining catalogs are included in the comparison catalog used by Kriek et al. (2015) to determine the MOSDEF uncertainties. Therefore we take the MOSDEF value as our estimate of the redshift error for the whole sample.

Chung et al. (2021) found that using a single, characteristic line profile is adequate for accounting for the effect of line width in many applications. We therefore estimate the typical FWHM of CO emitters in our data cube to be ~300 km s^{-1}. This value is broadly consistent with the median line width of 360 km s^{-1} from a compilation of blindly detected CO lines from the ALMA Spectroscopic Survey in the Hubble Ultra Deep Field (ASPECS; González-López et al. 2019) and CO Luminosity Density at High-z Survey (COLDz; Pavesi et al. 2018) deep fields in the redshift range $1.0 \lesssim z \lesssim 3.6$.

In Section 7 we revisit our choices in modeling line width and redshift uncertainty, finding that across the range of likely values, the uncertainty introduced to the final power spectrum is negligible in comparison to the instrumental noise, and smaller than the effects of other uncertainties such as cosmic variance.

4.3. Stacking Analysis

The shot power in the CO-galaxy cross-spectrum is proportional to the mean CO luminosity of the galaxies in the catalog. This mean luminosity can also be measured from our CO data cube via stacking. To facilitate comparison between these two approaches, we construct a stack on the positions of our catalog galaxies.

To perform this analysis, we image the GOODS-N data using natural weighting. The resulting image has a typical 1σ sensitivity of 0.27 Jy km s^{-1} at the center of the field and a synthesized beamwidth of $1\rlap{.}''6$. We then extract a cutout from the primary beam-corrected image at the position and expected frequency of each source. We use only galaxies with expected

CO(1–0) line frequencies at least 31.25 MHz (one channel) from the edge of a spectral window and within $15''$ of the map center. We convert each extracted snapshot to units of luminosity per spectral channel using

$$L'_{\rm chan} = 3.25 \times 10^7 S \Delta v_{\rm chan} \frac{D_L(z)^2}{(1+z)^3 \nu_{\rm obs}^2} \text{ K km s}^{-1} \text{ pc}^2 \quad (18)$$

where S is the channel flux density in janskys, $\Delta v_{\rm chan}$ is the velocity width of the channel in kilometers per second, D_L is the luminosity distance in megaparsecs, z is the redshift of the source, and $\nu_{\rm chan}$ is the frequency of the channel. The channel width of the COPSS data is 31.25 MHz, which corresponds to a velocity widths between 270 and 340 km s^{-1}. We then compute the inverse variance weighted sum of the extracted spectra for all of our cutouts.

A redshift uncertainty of $\Delta z/(1+z) \sim 0.001$, appropriate for the spectroscopic redshifts in our catalog, corresponds to a shift in expected CO line frequency of approximately one channel. CO line widths correlate with the line luminosity, and typically do not exceed 1000 km s^{-1} even for the most luminous high redshift galaxies (Carilli & Walter 2013). Most galaxies within our stack will have line widths well below this limit. Therefore we expect that most of the flux will fall within the three central channels of our stack. We therefore sum over these three channels to compute the luminosity of the stack.

5. Results

We find an auto-power of $P_{\rm CO} = 1700 \pm 3500$ $\mu K^2 h^{-3} Mpc^3$, corresponding to a 2σ upper limit of $P_{\rm CO} < 8700$ $\mu K^2 h^{-3} Mpc^3$ for the GOODS-N field. This result is consistent with the measurement of K16 using the full COPSS data set of $P_{\rm CO} = 3000 \pm 1300$ $\mu K^2 h^{-3} Mpc^3$. We note that the "GOODS-N" result reported in K16 also includes data from the trailing fields, which are not included in our analysis because they have no corresponding optical data. As our auto-power results are not constraining, we do not consider them further in this paper. When constructing a joint constraint from the auto- and cross-power spectra, we instead use the final COPSS auto-power spectrum from K16.

Figure 3 shows our computed cross-power spectrum for the GOODS-N field. Taking the inverse variance weighted average over all k, we find a cross-power of $P_\times = -4 \pm 270$ $\mu K h^{-3} Mpc^3$. This corresponds to a 2σ upper limit of $P_\times < 540$ $\mu K h^{-3} Mpc^3$.

Figure 4 shows the stacked spectrum of our galaxy catalog. We find no evidence of a detection of CO emission in the stack, finding $L' = (-0.5 \pm 2.5) \times 10^{10}$ K km s^{-1} pc^2 corresponding to an upper limit of $L' < 5.0 \times 10^{10}$ K km s^{-1} pc^2.

6. Data Validation

Optical and radio observations are subject to very different systematics. Radio intensity mapping measurements can be corrupted by terrestrial and astronomical foregrounds, but these are unrelated to the optical/IR galaxy catalogs used in the cross-correlation measurement. We therefore expect that contamination problems seen in the auto-power spectrum will be greatly diminished in the cross-spectrum. Here we search for evidence of contamination in our cross-spectra and compare these results to known contamination issues in the auto-spectra.
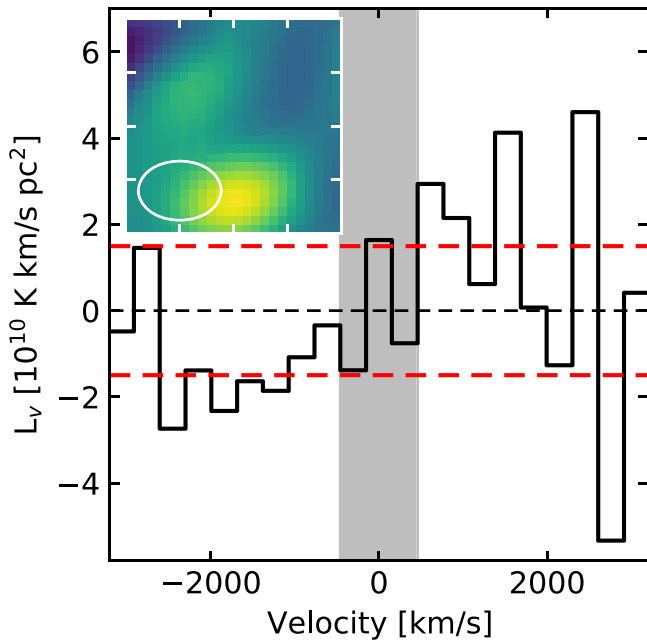
Figure 4. The average spectrum of our GOODS-N galaxy catalog, produced by stacking at the expected frequency of the CO(1–0) emission line. The red dashed lines indicate the average 1σ noise level in a channel (which matches the channel-by-channel uncertainty within a few percent), while the gray shaded region shows the three channels integrated to obtain our constraint on the average CO luminosity luminosity. Inset: $4' \times 4'$ image of the central three channels of our stack. The spectrum is extracted from the central pixel of the map. The synthesized beam is shown in the lower-left corner.

6.1. Continuum Source Contamination

We perform a number of additional randomizations as a check on noise and systematics. First, we randomize the phase of the galaxy grid and compute the cross-power with the non-randomized radio data. We verify that this produces results consistent with the radio phase randomizations, finding that 1σ noise curves from the galaxy phase randomizations match within a few percent.

Next we randomize the redshifts (or equivalently frequencies) of galaxies within each window. This serves as a check of our continuum removal as galaxies matched in angular position to a continuum source will remain matched when shifted in frequency, producing excess low-k power (Switzer et al. 2019). For comparison, we produce a set of randomizations where each galaxy is assigned a new angular position within the footprint of the galaxy survey, in addition to a new redshift. Figure 5 shows how the 16th and 84th percentiles of each set of randomized power spectra compare with the three-dimensional position randomization. The curves generally match. Continuum contamination would appear as a positive bias in the redshift-only distribution, but the redshift-only randomization shows no excess relative to the other two tests. The median redshift-randomized spectrum (not shown) is consistent with zero at all k and in all windows, while the 16th and 84th percentiles are symmetric around zero. We therefore conclude that the contribution from continuum sources to our measurement is negligible.

6.2. Mitigating Ground Contamination

In our reduction of the SZA data, we used observations from trailing fields to model and remove ground-correlated emission, and when this was not possible, we flagged baselines

containing any measurement exceeding 4σ significance. To test the effect of the contaminated modes on the cross-signal, we re-ran our analysis without the trailing field subtraction, using only flagging at a range of significance cuts, including an analysis where no data are excluded. Figure 6 shows the cross-power spectrum of the GOODS-N data for each cut.

Interestingly, the sensitivity curves for the data with no trailing field subtraction fall slightly above those with the subtraction at low k, even though differencing fields increases the thermal noise in the CO data. This suggests that, while not the dominant source, systematics can contribute a non-negligible level of noise relative to the thermal noise.

Even without any corrections, there is no significant excess power in the cross spectrum. The results are relatively insensitive to the significance cut applied, and the power spectrum with no corrections is consistent within errors of that constructed with both trailing field subtraction and a significance cut. By contrast, performing auto-power spectrum calculations without correcting for ground contamination produces power at many times the expected noise level, which swamps any astrophysical signal (see Appendix A and K15). This demonstrates the utility of cross-correlation for cleaning contaminated data; we could use the data in cross-correlation analyses, without introducing biases and with only a moderate increase in the noise level of the resulting power spectra, even if no correction for ground contamination was possible.

6.3. Search for Interloper Line Contamination

Emission lines from other molecules in galaxies at lower redshift can appear in intensity mapping data cubes and be confused for emission from the target line. This emission contributes to the auto-power spectrum in a manner that cannot easily be disentangled from the power of the target line (Cheng et al. 2016, 2020). However, the sources of these lines will not appear in the galaxy density cubes used in cross-correlation, and therefore these interlopers will not contribute to the cross-power spectrum. In addition, if galaxy catalogs are available at the redshift of the confusing sources, they can be used to measure the contributions of those sources and assess the level at which they contribute to the auto-power.

At the redshifts studied by COPSS, no significant interlopers are expected for the CO(1–0) transition (Chung et al. 2017), although dense gas tracers such as HCN $J = 1 - 0$ and CS $J = 1 - 0$ have been proposed as a potential contaminant (Breysse et al. 2015). HCN emission in our SZA data corresponds to sources at $1.5 < z < 2.3$. Our galaxy catalogs contain 227 galaxies in this redshift range. We cross-correlate these galaxies with the SZA data in the same manner as Section 4.2 and find results consistent with zero. This allows us to place a 2σ upper limit on the total HCN cross-power of $P_{\times,\mathrm{HCN}} < 566\ \mu\mathrm{K}\,h^{-3}\,\mathrm{Mpc}^3$.

CS $J = 1 - 0$ emission would appear in the SZA bandpass at $0.4 < z < 0.8$. In this redshift range, our galaxy catalogs are considerably more complete. We cross-correlate the SZA data with a sample of 835 galaxies from our GOODS-N catalog. The measured power is again consistent with zero, and we place a 2σ upper limit on the shot noise in the CS cross-power spectrum of $P_{\times,CS} < 44\ \mu\mathrm{K}\,h^{-3}\,\mathrm{Mpc}^3$.

In practice, contamination of the auto-power spectrum by low redshift interlopers is expected to be primarily driven by a handful of the brightest objects, while our cross-power limits reflect the average luminosity of all objects in our catalog. We
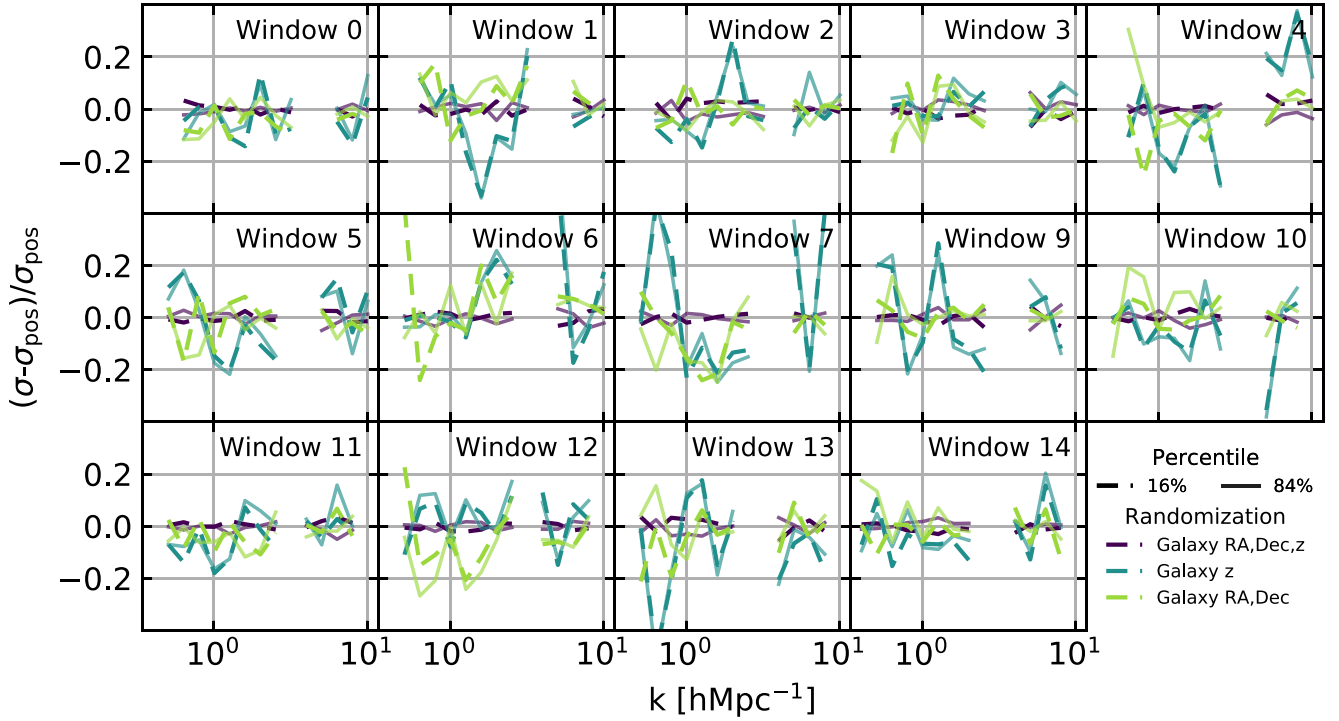
Figure 5. The 16th (dashed) and 84th (solid) percentiles of 2500 power spectra generated with randomized galaxy position (dark blue), redshift (blue–green), and position angle (green). Each curve has been normalized by the 1σ curve of the position randomization. Results of the three randomizations match, indicating no excess of power due to continuum sources that would appear in the randomizations for frequency but not position or position angle.
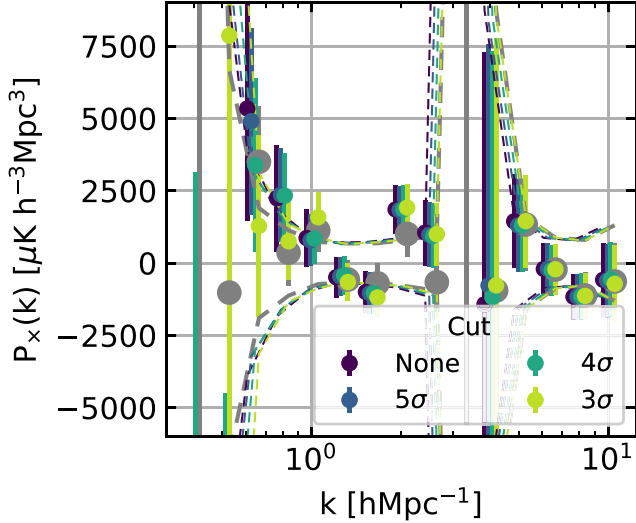

Figure 6. The cross-power spectrum for GOODS-N, reproduced using a range of different significance cuts to clean ground-correlated emission from the COPSS data. Values for each cut are offset slightly in k for clarity. The gray points show our final power spectrum, produced with a 4σ cut along with trailing field subtraction. The cross-power levels are consistent for all spectra shown.

therefore cannot rule out the possibility that a few extremely luminous objects are present in our field. However, models where contamination is expected to be significant require objects with HCN luminosities $\gtrsim 10^{10}$ K km s^{-1} pc^2 (Chung et al. 2017). Even if such galaxies do exist, they would be so rare as to be unlikely to appear in our survey volume; therefore, it seems unlikely that they contaminate the auto-power results of K16 significantly. This is further confirmed by analysis of the COPSS image cubes, which show no evidence for bright emission lines.

7. Modeling

Here, we use mock observations to better understand any systematics in our data analysis procedure. In Section 7.1 we describe a fiducial model of the galaxy-CO cross-power spectrum. Then in Section 7.2 we assess the degree of attenuation caused by redshift uncertainties, spectral line emission being spread over multiple frequency channels, and the SZA's sampling of the uv-plane. We use these results to validate the attenuation corrections applied to our observed three-dimensional power spectrum. In Section 8.1 we extend this analysis to literature models of the CO power spectrum in order to contextualize the sensitivity of our data.

7.1. Model Description

We follow the approach of Li et al. (2016, hereafter L16; see also Silva et al. 2015; Chung et al. 2017), assigning CO luminosities to halo catalogs from dark matter simulations using a series of scaling relations. This approach is particularly advantageous for modeling cross-correlation, as each CO emitter can be matched to other halo properties in order to construct mock galaxy catalogs for cross-correlation. We use the scaling relation between halo mass and CO luminosity fit to the COPSS auto-power spectrum by Keating et al. (2020) as our fiducial model:

$$L_{\mathrm{CO}} = \begin{cases} A_{\mathrm{CO}} \dfrac{M^2}{M_0} & M \leqslant M_0 \\ A_{\mathrm{CO}} M_0 & M > M_0 \end{cases} \qquad (19)$$

where M is the halo mass, A_{CO} is the mass to CO luminosity ratio, and M_0 is a cutoff mass, above which the CO luminosity of halos remains approximately constant. M_0 is set to be 10^{12} h$^{-1}M_\odot$, and the fit to the COPSS auto-power spectrum

yields $A_{CO} = 1.5 \times 10^{-6}$ $L_\odot$ $M_\odot^{-1}$. To account for galaxy-to-galaxy variations, a log-normal scatter with $\sigma_{CO} = 0.37$ dex is added to the M_{halo}-L_{CO} scaling for each halo. We implement this model using galaxy catalogs from the publicly available IllustrisTNG-300 simulation (Naiman et al. 2018; Nelson et al. 2018; Marinacci et al. 2018; Pillepich et al. 2018; Springel et al. 2018).

To simulate observations of our models, we draw random lines of sight through the simulation box, selecting galaxies within $40''$ of the field center to include in light cones, and allowing redshift evolution by stepping through simulation snapshots. The procedure for generating light cones is described in detail in Keenan et al. (2020). The light cones are subdivided into frequency and angular resolution elements matched to the resolution of our data, and for each element, a brightness temperature is calculated. The light cones are then apodized according to the primary beam response of the SZA antennas.

Next, we select a catalog of galaxies to utilize for cross-correlation. To approximately reproduce the galaxy distribution in our catalog, we group spectral windows into three broad ranges: two high-density bins from $2.3 \lesssim z \lesssim 2.6$ and $2.8 \lesssim z \lesssim 3.2$ and a low-density bin in the $2.6 \lesssim z \lesssim 2.8$ gap. The high-density bins correspond to the redshift ranges in which [O III] emission falls in an NIR window and can be used to make secure redshift determinations in IR spectroscopic surveys such as MOSDEF (Kriek et al. 2015). The redshift ranges covered in Reddy et al. (2006) also match approximately to this selection. Within each bin, we calculate the total number of galaxies in our real redshift catalog and require that our mock catalogs have this number of galaxies in the matching redshift range. We then populate our galaxy catalogs by drawing halos with $M > 10^{11}$ $M_\odot$, which roughly corresponds to the 10^9 $M_\odot$ stellar mass limit of the MOSDEF survey. MOSDEF prioritizes high-mass targets, and so we weight our selection similarly.

We use these catalogs to construct galaxy density fields using the R.A., decl., and observed redshift in a manner analogous to Section 4.

7.2. Evaluation of Signal Attenuation and Systematics

As discussed in Section 4.2.1, a number of factors are known to distort observed intensity mapping power spectra from their intrinsic shape and amplitude. Here we explore this in greater depth by implementing the following effects in our model:

1. Redshift uncertainty: we account for the decorrelation caused by redshift uncertainties by assigning each simulated galaxy an observed redshift drawn from a Gaussian distribution with a width $\sigma_z/(1 + z)$ and centered at the true redshift.
2. CO line width: we simulate the spreading of the CO signal across multiple spectral channels by assigning all galaxies a Gaussian line profile characterized by a single FWHM.
3. Imperfect attenuation corrections: our attenuation corrections assume a redshift error and line width. If these assumptions are incorrect, they will result in over- or under-correction of the attenuation. To account for this effect, we run versions of our model where we increase or decrease the redshift errors and line widths, but still

correct them assuming the values of σ_z and σ_{lw} given in Section 4.2.1.

Our inclusion of these effects allows us to assess how well our attenuation corrections work and estimate the uncertainty introduced by this step in our analysis.

In the upper panel of Figure 7 we show the cross-spectra of a series of models run with different combinations of these effects. Table 1 summarizes the parameters of each model.

The solid black line shows the "true" power spectrum, while the gray dashed line shows the power spectrum attenuated by 300 km s^{-1} line widths and redshift errors of $\sigma_z/(1+z) = 0.001$. Comparing these models illustrates the expected signal attenuation in the absence of corrections. The attenuation increases with k, as our survey accesses low k modes primarily perpendicular to the line of sight where the k_z component is small and attenuation is minimal. However at large k the power spectrum is attenuated by roughly 70%. The dashed black line shows the power spectrum with our attenuation correction applied, demonstrating that we successfully recover most of the attenuated power.

The filled regions in the top and middle panels of Figure 7 show the effects of choosing the wrong values of σ_z or σ_{lw} when applying the attenuation correction. These simulations are run with line widths of 100 and 500 km s^{-1} (yellow regions) or redshift uncertainties of $\sigma_z/(1 + z) = 0.0007$ or 0.0013 (blue regions), but are attenuation corrected assuming the fiducial values from Section 4.2.1 (FWHM = 300 km s^{-1}, $\sigma_z/(1 + z) = 0.001$). The middle panel shows that both of these effects result in factional errors of less than 20%.

Finally, the bottom panel of Figure 7 shows the ratio of the corrected power spectrum with the true power spectrum for 10 individual light cones, along with the distribution from all 1000 simulated fields. Averaged over all light cones, the correction works within a few percent, while the attenuation-corrected cross-spectra for 68% of individual light cones match the true power spectra to within $\pm25\%$ at a given k.

7.2.1. Error Budget for SZA Cross-correlation

To this point, we have assumed uniform coverage of the uv-plane by our observations. In practice however, our data can only be used to measure the cross-spectrum at points in uv-space sampled by the SZA. We therefore recomputed our simulated power spectra using the same uv-sampling and cell-weighting as our data.

Figure 8 shows the fractional uncertainty due to the attenuation correction, imperfect knowledge of σ_z and σ_{lw}, thermal noise, and sample variance for our simulations including the SZA uv-coverage. The σ_z uncertainties correspond to a $\pm30\%$ change in σ_z, while the σ_{lw} uncertainties correspond to ±200 km s^{-1} in FWHM. The sample variance is computed from our ensemble of simulated light cones.

At all k, the thermal noise is dominant, and equals or exceeds the expected signal. Sample variance and discrepancies between the true and corrected power spectrum are the largest of the remaining sources of uncertainty at around 20%–30% for most k, while the remaining effects are smaller. The sample variance and attenuation correction-related uncertainties and SZA residuals all also scale with the expected signal, and as our fiducial model is at the bright end of the range of possible models (see Section 8.1), these errors would produce an even smaller contribution to the total error budget for fainter signals. We therefore consider only thermal noise in the analysis of our
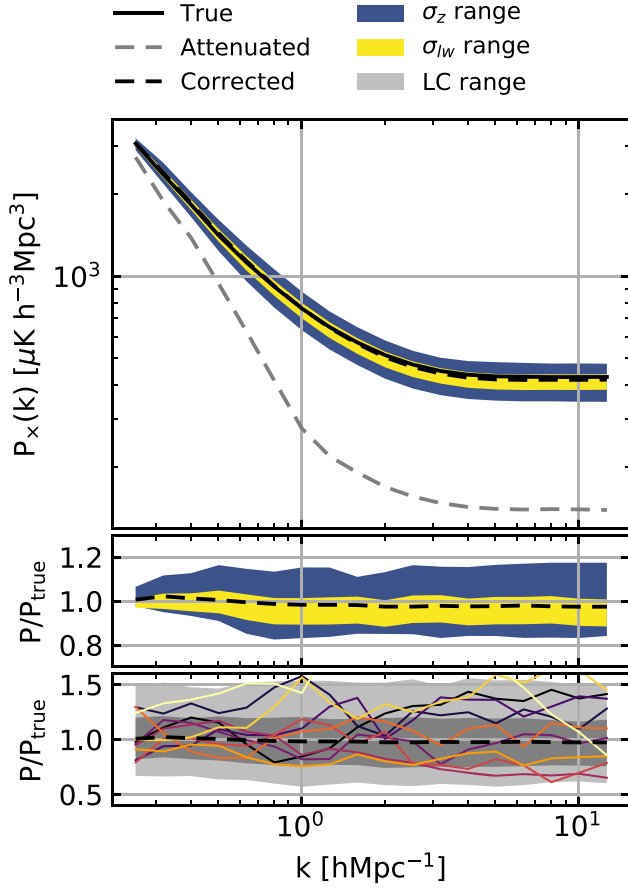
Figure 7. Top: the mean cross-power spectra of our simulated light cones. The solid black line shows the true spectrum with no attenuation. The gray dashed lines shows the power spectrum attenuated by redshift errors of $\sigma_z/(1 + z) = 0.001$ and Gaussian line profiles with 300 km s^{-1} FWHM, and the black dashed line shows this same power spectrum after applying the attenuation corrections described in Section 4.2.1. The yellow region shows the effect of assuming a 300 km s^{-1} line width when the true width ranges from 100–500 km s^{-1}, while the blue region shows the effect of assuming $\sigma_z/(1 + z) = 0.001$ when the true value ranges from 0.0007 to 0.0013. Middle: the ratios of attenuation-corrected spectra to the true power spectrum, showing the degree of uncertainty introduced by incorrect assumptions in the attenuation correction. Colors are the same as the top panel. Bottom: the ratios of attenuation-corrected spectra to their corresponding true power spectra. The black dashed line is for the mean of our 1000 light cones, while the darker and lighter gray regions show the range containing 68% and 95% of the light cones, illustrating the uncertainty introduced by attenuation corrections to an individual light cone. The ratios for 10 individual light cones are shown by colored lines.

Table 1
Summary of Simulation Parameters Used in Section 7.2

Name	FWHM	$\frac{\sigma_z}{(1+z)}$	Corrected
True	0	0	No
Attenuated	300	0.001	No
Corrected	300	0.001	Yes
σ_{lw} range	100	0.001	Yes
	500		
σ_z range	300	0.0007	Yes
		0.0013	

Note. Attenuation corrections, when applied, assume an FWHM of 300 km s^{-1} and $\sigma_z/(1 + z)$ of 0.001, even when the input values of these parameters differ.
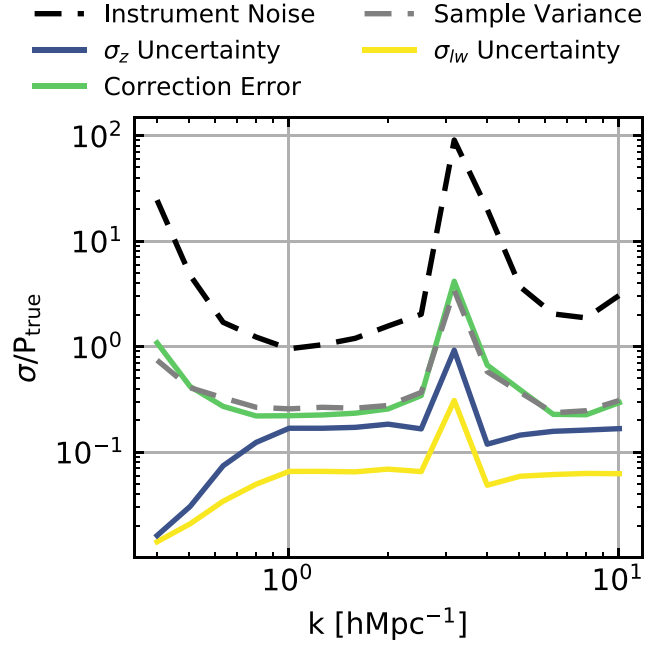

Figure 8. The fractional uncertainty in our power spectrum due to thermal noise (black dashed line), sample variance (gray), attenuation correction (green), and imperfect knowledge of galaxy redshift errors (blue) and line widths (yellow).

observed power spectrum, but note that these other errors will be important for more sensitive future experiments.

8. Discussion

Our cross-power spectrum includes contributions from shot, one-halo, and two-halo components. The difference in the k-space shape of these contributions makes them potentially separable. We fit our observed power spectrum with a two-component model consisting of a constant shot power term and a two-halo term consisting of the linear matter power spectrum multiplied by a scale factor.[4]

We compute the linear matter power spectrum appropriate for each window in our data using the HALOMOD package (Murray et al. 2021) and combine these into a single spectrum using the same weights as our data. We then fit $P_{\mathrm{shot},\times}$ and the scale factor multiplying P_{lin}: $b_{\mathrm{CO}}b_{\mathrm{gal}}\langle T_{\mathrm{CO}}\rangle$ to our observed cross-power spectrum.

The gray curves in Figure 9 show the parameter distribution for this fitting procedure. We find $b_{\mathrm{gal}}b_{\mathrm{CO}}T_{\mathrm{CO}} = 122 \pm 60\ \mu$K and $P_\times = -482 \pm 357\ \mu$K h^{-3} Mpc^3 (medians and limits containing the central 68% of the marginalized probability distribution for each parameter). The fit is degenerate, and the data marginally favor a negative shot power, pushing the two-halo term toward more positive values. Negative shot power

---

[4] Here we do not attempt to model the one-halo term, as its shape and amplitude depend on the details of the tracer galaxy populations in a more complex manner than the other two terms. We expect that most of our catalog galaxies are the brightest and most massive optical galaxy in their halo, and CO luminosity has been found to correlate with galaxy mass (Inami et al. 2020). Based on these expectations and the discussion at the end of Section 2, it is likely that the amplitude of the one-halo term is lower than the shot term at all scales. Even if the one-halo term contributes substantially to the power spectrum, the first-order effect would be to make our upper limits on the other two terms more stringent; therefore, we do not expect the omission of the one-halo term to alter our conclusions.
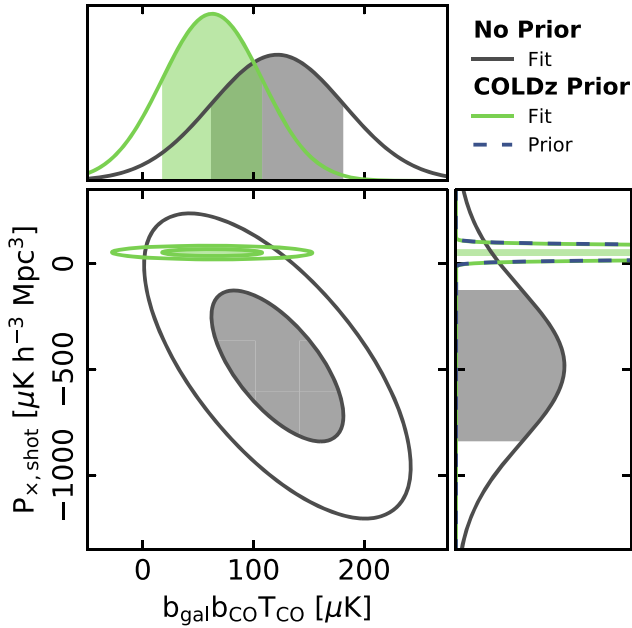
Figure 9. Lower left: the joint distribution of $\langle T_{CO} \rangle b_{CO} b_{gal}$ and $P_{shot,\times}$ for our fits with no priors (gray) and with a prior on $P_{shot,\times}$ based on the Pavesi et al. (2018) stack (green). Contours show 1σ and 2σ levels of the distribution. Upper left: the likelihood distribution of $b_{gal} b_{CO} \langle T_{CO} \rangle$. The filled regions show the central 68% of the probability distribution. Lower right: the distribution of $P_{shot,\times}$. The blue dashed line shows the prior based on the COLDz stack.

would correspond to negative average CO luminosity, and is therefore nonphysical.

We can improve our fit by using an external constraint on the average line luminosity of galaxies in our catalog to construct a prior on the $P_{shot,\times}$. Pavesi et al. (2018) performed a stacking analysis of CO(1–0) emission from GOODS-N galaxies at $2.0 < z < 2.8$ using a catalog similar to ours. Their stack was constructed using only objects undetected in their CO data, so we correct the average luminosity to include the two detected galaxies in their sample and use this to place a prior of $P_{shot,\times} = 54 \pm 16~\mu\text{K}\,\text{h}^{-3}\,\text{Mpc}^3$. Rerunning the fit then gives $b_{gal} b_{CO} T_{CO} = 63 \pm 45~\mu\text{K}$. The full parameter distribution is shown in green in Figure 9. We find good agreement between our results with the COLDz prior and a prior that simply requires a positive shot power term.

8.1. Comparison to Literature Models

Models of the CO power spectrum disagree by as much as an order of magnitude in signal amplitude. We therefore re-run the mock observations described in Section 7 with three additional models proposed in recent theoretical literature.

Our first literature model comes from Padmanabhan (2018) who define a double power law relating M_{halo} to L_{CO}. They tune this power law to reproduce observational constraints on the CO luminosity function at low redshift (Keres et al. 2003) as well as the COPSS auto-power constraint. The model also includes an f_{duty} term, describing the fraction of galaxies that produce CO emission. This term is poorly constrained observationally, but the fraction of galaxies with active star formation at $z \sim 2.5$ is high (e.g., Behroozi et al. 2019). Therefore we set $f_{duty} = 1$.

Our next model comes from Chung (2019) who also use a double power-law model, this time calibrated to reproduce the M_{halo}–L_{CO} relation of Li et al. (2016), which in turn is

calibrated using the halo mass to SFR relation of Behroozi et al. (2013) and the SFR–L_{CO} correlation summarized in Kennicutt (1998) and Carilli & Walter (2013). This model includes a log-normal scatter of $\sigma_{CO} = 0.40$ similar to that in our fiducial model. As the results of the Li et al. (2016) and Chung (2019) models are by design very similar, we only consider the more recent Chung (2019) model here.

Our third model is that of Yang et al. (2021a) who provide parametric formulas for L_{CO} as a function of halo mass and redshift calibrated to reproduce the results of the semi-analytic model (SAM) of far-infrared emission presented in Yang et al. (2021b).

Additionally, in order to explore how our galaxy catalog affects the signal, we consider the results of modifying our optical galaxy catalog to contain fewer of the highest-mass galaxies. We compute the cross-power spectra of all three CO models with this modified galaxy catalog along with the fiducial catalog described in Section 7.1.

The simulated cross-spectra are shown in Figure 10. We use our 2σ sensitivity levels for the shot and two-halo terms to set upper limits on the signal amplitude. We also use the lower limit on ρ_{mol} at $z \sim 2.5$ reported by Decarli et al. (2020) and our estimates of b_{gal} and b_{CO} from Section 8.3.4 to estimate a lower limit on the two-halo power. The regions excluded by these limits are shown in gray in Figure 10.

The models based on Chung (2019) and Yang et al. (2021a) are well below the sensitivity of the current data set. The SAM underlying the Yang et al. (2021a) model was compared to recent CO intensity mapping observations in Breysse et al. (2021) and found to be in tension with current observational results. The Chung (2019) model is also in moderate tension with the tentative detection of the CO auto-power in COPSS, producing an auto-power signal approximately an order of magnitude fainter. Our simulations show that in cross-correlation, both models also run up against the lower limit on the two-halo power from direct detection. These discrepancies suggest that the power spectra of these two models represent lower limits on the range of possible signals.

On the other hand, the K20 and Padmanabhan (2018) models lie near the upper edge of the allowed parameter space. Averaging across the full range in k, the expected signal-to-noise ratio (S/N) on $P_{tot,\times}$ for K20 and our fiducial galaxy catalog ranges from 1.6 to 2.6 across the central 68% of simulated light cones. For the lower-mass catalog, this range is 1.0–1.7. The expected S/N for the Padmanabhan (2018) model ranges from 2.4 to 3.6 for the fiducial catalog and 1.5–2.5 for the lower-mass catalog. Our non-detection of the band-averaged CO-galaxy cross-power is therefore in moderate tension with this model (although consistent with versions using a lower f_{duty} or lower-mass catalogs).

The COPSS auto-power measurement consisted of ten times more CO data than the subset used here and achieved an S/N of 2.3 for the auto-power. It is noteworthy that, for models derived from the COPSS result, the expected cross-power S/N is as high as 2.9, even with our much smaller data set. If we could use the full COPSS survey area in cross-correlation, or if all survey time had been dedicated to the GOODS-N field, we would expect to detect the cross-power for the K20 model at S/N > 6 or the Padmanabhan (2018) model at S/N > 9. While no re-imagining of COPSS would detect the Chung (2019) and Yang et al. (2021a) models, a number of ongoing CO intensity mapping experiments studying multiple transitions of the CO line will achieve much greater depths than the data presented
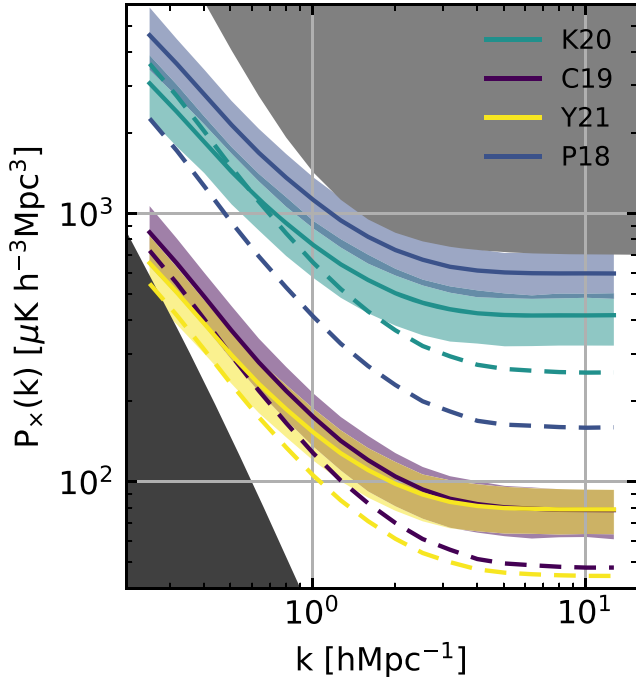
Figure 10. Cross-power spectra simulated using light cones from the TNG300 simulation and the K20 (blue–green), Padmanabhan (2018, blue), Chung (2019, purple), and Yang et al. (2021a, yellow) prescriptions for CO luminosity. Solid lines show the mean power spectrum of 1000 light cones, and bands show the area containing 68% of the realizations, while dashed lines show the result of using a galaxy catalog consisting of lower-mass galaxies. The darker gray region at the lower left is ruled out by lower limits on $\langle T_{CO} \rangle$ from Decarli et al. (2020). The lighter gray filled region at the upper right shows the parameter space excluded by our non-detection.

here. Many of these studies are targeting fields with significant spectroscopic redshift information (Chung et al. 2019; Sun et al. 2021). These experiments can therefore be expected to detect CO-galaxy cross-power and, for optimistic forecasts, constrain its shape in order to extract valuable physical information about the high redshift galaxy population. It is possible that such results could be rendered well before these experiments reach the sensitivity required to measure the auto-power spectrum.

8.2. Limits on Average CO Luminosities

Using Equation (9), we can convert the shot component of the cross-power into an average CO luminosity of the galaxy catalog used for cross-correlation. Assuming the shot power dominates the power spectrum across the full range of k probed here, our upper limit on the total power corresponds to a 2σ upper limit on the mean CO luminosity of our galaxies of $L' < 3.3 \times 10^{10}$ K km s^{-1} pc^2. Our best-fit power spectrum (with no prior on $P_{shot,\times}$) allows us to separate out the shot term and results in an upper limit of $L' < 4.4 \times 10^{10}$ K km s^{-1} pc^2. These are comparable to the limit from our stack of $L' < 5.0 \times 10^{10}$ K km s^{-1} pc^2.

In Figure 11 we present these upper limits on the mean CO luminosity of our catalog, along with the CO luminosities of confirmed galaxies detected in blind searches for CO emission at comparable redshifts. Pavesi et al. (2018) reported CO(1–0) emission from five galaxies at $2.0 < z < 2.8$ spread over regions GOODS-N and COSMOS surveyed with the Very Large Array. González-López et al. (2019) reported an additional five galaxies in CO(3–2) at $2.0 < z < 3.1$ detected in the Hubble Ultra Deep
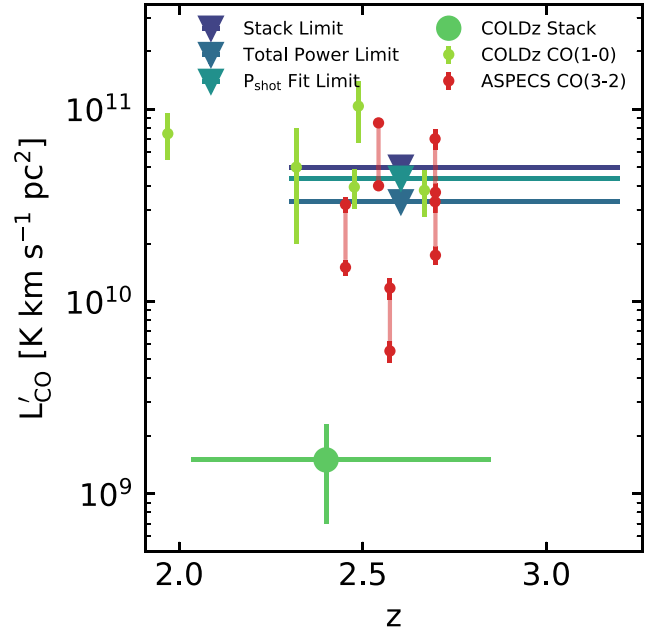

Figure 11. Upper limits on the CO luminosity of galaxies in our catalog derived from both our power spectrum analysis and our stacked spectrum (blue), compared to literature luminosity measurements at comparable redshifts. The stack of 78 GOODS-N galaxies from COLDz is shown as the large green point, while individually detected galaxies from the COLDz survey are shown in light green. Galaxies detected in CO(3–2) by the ASPECS survey are shown in red. As a range of values for the CO(3–2)/CO(1–0) line ratio have been reported, we convert each ASPECS luminosity to CO(1–0) assuming line ratios of 0.85 and 0.40 and show these two points connected by a vertical line for each source.

Field with the Atacama Large Millimeter/submillimeter Array (ALMA). We convert CO(3–2) luminosities to CO(1–0) assuming a CO(3–2) to CO(1–0) luminosity ratio of 0.85 or 0.40 (Daddi et al. 2015; Boogaard et al. 2020). Our upper limits sit around the median luminosity of these direct detections. These results are in agreement with Pavesi et al. (2018), who performed a stacking analysis in GOODS-N using a similar galaxy catalog (Skelton et al. 2014), and measured a mean luminosity about an order of magnitude below our limit.

8.3. Interpreting the Two-halo Constraint

Measuring the average abundance of molecular gas in galaxies too faint for direct study at high redshift is a major objective of LIM. Assuming a constant molecular gas mass to CO luminosity ratio α_{CO}, the molecular gas density can be written in terms of the mean CO brightness temperature as

$$\rho_{mol} = \alpha_{CO} \frac{H(z)}{(1+z)^2} \langle T_{CO} \rangle \qquad (20)$$

where $H(z)$ is the Hubble parameter at redshift z.

Intensity mapping measurements constrain $\langle T_{CO} \rangle$ through the two-halo term of the power spectrum, which is proportional to $b_{gal} b_{CO} \langle T_{CO} \rangle$ for the cross-spectrum. If b_{gal} and b_{CO} are known or can be measured, then a constraint on the amplitude of the two-halo power becomes a constraint on $\langle T_{CO} \rangle$ alone.

In the remainder of this section, we discuss our procedure for determining b_{gal} and b_{CO} in order to constrain $\langle T_{CO} \rangle$ from our cross-spectrum results alone and in combination with the COPSS auto-power spectrum.

8.3.1. Galaxy Bias

We can measure $b_{\rm gal}$ directly from our catalog by measuring the galaxy-galaxy auto-power spectrum

$$P_{\rm gal}(k, z) = b_{\rm gal}(z)^2 P_{\rm lin}(k, z) + P_{\rm shot,gal}, \quad (21)$$

where $P_{\rm shot,gal}$ is the galaxy shot power and, for individual spectral windows, will be equal to $\bar{n}^{-2}$, the inverse of the galaxy number density squared. We compute the auto-power spectrum of our galaxy density grid using the estimator

$$\mathcal{P}_{\rm gal}(\boldsymbol{k}) = \frac{1}{V_{\rm cat}}(\tilde{\delta}_N^*(\boldsymbol{k})\tilde{\delta}_N(\boldsymbol{k})) \quad (22)$$

$$P_{\rm gal}(k) = \langle \mathcal{P}_{\rm gal}(\boldsymbol{k}) \rangle_{\boldsymbol{k}\cdot\boldsymbol{k}=k^2} \quad (23)$$

and fit the one-dimensional galaxy power spectrum with two-halo and shot components.[5] The resulting estimate of $b_{\rm gal}$ is $3.5^{+0.6}_{-0.8}$.

The bias of OIR selected, star-forming galaxies has been studied extensively, allowing us to cross-check our fit with literature results. Durkalec et al. (2015) measured the galactic bias of a large sample of rest-frame ultraviolet selected galaxies with spectroscopic redshift from the VIMOS Ultra Deep Survey (VUDS), and find $b_{\rm gal} \sim 2.7$–2.8 (depending on fitting method) at a mean redshift of $z = 2.95$. Herrero Alonso et al. (2021) measured $b_{\rm gal} \sim 2.8$–3.0 for Lyα emitting galaxies at a median redshift $z = 3.8$. Geach et al. (2012) found $b_{\rm gal} \sim 2.4$ for Hα emitting galaxies at $z = 2.2$. Adelberger et al. (2005) found $b_{\rm gal} \sim 2.1$, 2.4, and 2.6 at $z = 1.7$, 2.2, and 2.9 for galaxies identified using $U_n G \mathcal{R}$ color selection. Kashikawa et al. (2006) found $b_{\rm gal} \sim 5$ for $z = 4.1$ LBGs.

These results are generally consistent with our own fit. The overall convergence of these results around $b_{\rm gal} \sim 3$ masks a number of difficulties in determining $b_{\rm gal}$ for a given galaxy sample. The galaxy bias is known to be a function of galaxy properties, with brighter, more massive, and redder galaxy samples all tending to be more clustered (Kashikawa et al. 2006; Khostovan et al. 2019). The bias also appears to increase from redshifts ~ 2–6 (Durkalec et al. 2015; Khostovan et al. 2019). In particular, because of the inhomogeneous nature of the galaxy catalog we assemble, the appropriate galaxy bias for our analysis may differ from these results. In addition, galaxy clustering studies typically use correlation function (i.e., real space) analyses to determine $b_{\rm gal}$, which may introduce different systematics compared to our power spectrum derived constraint. Therefore we use the value derived directly from our catalog in the following analysis.

8.3.2. CO Tracer Bias

The CO tracer bias is less studied. In theory it can be extracted from anisotropies in the three-dimensional CO power spectrum; however, such a measurement requires much more sensitive data than is currently available.

Instead we can compute $b_{\rm CO}$ as

$$b_{\rm CO} = \frac{\int L_{\rm CO}(M) b(M) \frac{dn}{dM} dM}{\int L_{\rm CO}(M) \frac{dn}{dM} dM} \quad (24)$$

where $b(M)$ is the mass-dependent halo bias, dn/dM is the halo mass function, and $L_{\rm CO}(M)$ is the halo mass to CO luminosity relation. To estimate $b_{\rm CO}$ we use the halo mass function of Tinker et al. (2008, 2010), the halo bias prescription of Tinker et al. (2010), and our halo mass to CO luminosity model outlined in Section 7 and Equation (19). This gives $b_{\rm CO} = 2.3$. Using a similar procedure, but a different halo mass to CO luminosity scaling, Chung et al. (2019) found $b_{\rm CO} = 2.7$. We take the mean of these two values as our estimate.

8.3.3. Joint Fit of Galaxy, CO, and Cross-power Spectra

The CO auto-spectrum and galaxy-CO cross spectrum both depend on $\langle T_{\rm CO} \rangle$, but with different exponents and different combinations of bias terms: $P_{\rm 2h,CO} \propto b_{\rm CO}^2 \langle T_{\rm CO} \rangle^2$ and $P_{\rm 2h,\times} \propto b_{\rm CO} b_{\rm gal} \langle T_{\rm CO} \rangle$. Meanwhile the galaxy auto-spectrum contains additional information about the galaxy bias term with $P_{\rm 2h,gal} \propto b_{\rm gal}^2$. Fitting all three spectra jointly can extract the maximum information from an intensity mapping data set (see Switzer et al. 2013 for a previous exploration of jointly fitting auto- and cross-spectra in the context of H I intensity mapping).

We therefore perform a joint fit to the final COPSS CO auto-spectrum (K20), our cross-spectrum, and our galaxy auto-spectrum using a model with five parameters: $b_{\rm CO}\langle T_{\rm CO} \rangle$, $b_{\rm gal}$, $P_{\rm shot,CO}$, $P_{\rm shot,\times}$, and $P_{\rm shot,gal}$. To account for sample variance between the COPSS spectrum measured over many fields, we add a nuisance term in our model, which allows $b_{\rm CO}\langle T_{\rm CO} \rangle$ to vary between the CO auto-spectrum and the cross-spectrum. For this term, we set a Gaussian prior with a mean difference of zero and a standard deviation of 10% of $b_{\rm CO}\langle T_{\rm CO} \rangle$, chosen based on the sample variance model of Keenan et al. (2020). We fit this model using a Markov Chain Monte Carlo procedure implemented with the EMCEE package (Foreman-Mackey et al. 2013).

The resulting five-parameter fit is shown and compared to fits to the cross-spectrum or CO auto-spectrum alone for parameters that can be constrained by individual spectra in Figure 12. The results are largely consistent with the individual spectrum fits presented at the beginning of this section and in Keating et al. (2020). However, the joint fit results in significantly tighter constraints on $b_{\rm CO}\langle T_{\rm CO} \rangle$ and rules out extremely negative values of $P_{\rm shot,\times}$ even without the use of a prior on this term. We summarize the fit results in Table 2.

In Appendix B we explore the effect of including the stochasticity term $r_{\rm 2h}$ in our fit. The upper limits on $b_{\rm CO}\langle T_{\rm CO} \rangle$ and by extension $\langle T_{\rm CO} \rangle$ and $\rho_{\rm mol}$ are largely unaffected.

8.3.4. Limits on the Mean Molecular Gas Abundance

Based on our fit to the cross-spectrum alone, we derive a constraint on $\langle T_{\rm CO} \rangle$ of 8.4 ± 6.0 μK or an upper limit of $\langle T_{\rm CO} \rangle < 20.4$ μK (assuming $r_{\rm 2h} = 1$, the values of $b_{\rm gal}$ and $b_{\rm CO}$ from Sections 8.3.1 and 8.3.2, and using the Pavesi et al. 2018 prior on the cross-shot power). Our joint fit to the cross- and auto-spectra allows us to derive a tighter constraint of $\langle T_{\rm CO} \rangle = 4.7^{+3.5}_{-4.8}$ μK and an upper limit of $\langle T_{\rm CO} \rangle < 10.9$ μK (again using the Pavesi et al. 2018 prior and assuming $r_{\rm 2h} = 1$ and the $b_{\rm CO}$ from Section 8.3.2, but now deriving $b_{\rm gal}$ directly from the fit). The inclusion of the auto-power spectrum results in a factor of two reduction in our upper limit.

Pullen et al. (2013) cross-correlated quasars with the cosmic microwave background maps of Wilkinson Microwave Anisotropy Probe (WMAP) to constrain the CO-quasar cross-power
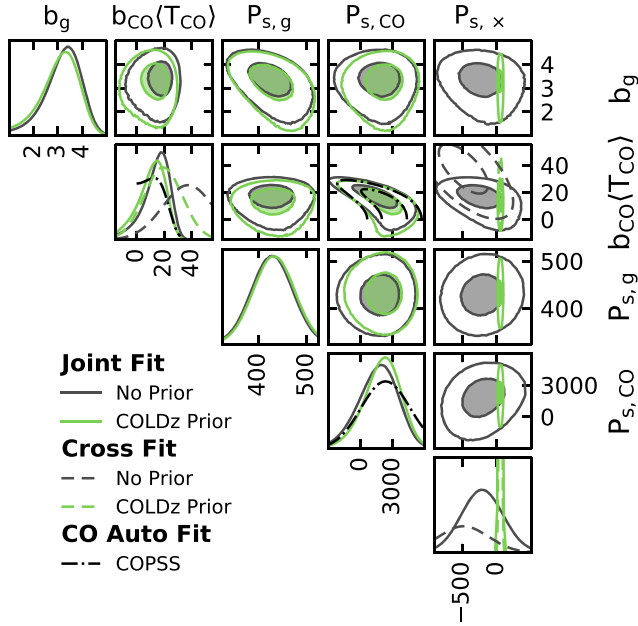
---

[5] We again exclude the one-halo term from our fitting. Because the mean number of galaxies per halo is less than one for our small galaxy catalog, the one-halo term should be small compared to the shot term (Schaan & White 2021).

Figure 12. The first panels of each row show the marginalized probability distributions for b_{gal}, $b_{CO}\langle T_{CO}\rangle$, $P_{shot,gal}$, $P_{shot,CO}$, and $P_{shot,\times}$ (from top to bottom) based on our joint fitting of the CO, galaxy, and cross-power spectra. The remaining panels show the joint distributions of pairs of parameters, with contours at the 1σ and 2σ levels of the distribution. Gray shows the fit performed with no prior on the cross-shot power, while green shows fits using the Pavesi et al. (2018) stack as a prior. Dashed lines in the $b_{CO}\langle T_{CO}\rangle$ and $P_{shot,\times}$ show the results from fitting only the cross-power spectrum, highlighting the improved constraining power of the joint fit. Black dotted–dashed lines in the $b_{CO}\langle T_{CO}\rangle$ and $P_{shot,CO}$ show the results of fitting only the COPSS CO auto-power spectrum. The auto-spectrum alone only constrains $b_{CO}^2\langle T_{CO}\rangle^2$, and therefore the contours do not extend below $b_{CO}\langle T_{CO}\rangle = 0$.

Table 2
Estimates for Power Spectrum Terms Based on Fits to P_{CO}, $P_\times$, and/or P_{gal}

Parameter	Units	Value (w/ prior)	Value (no prior)
		Joint fit of P_{CO}, $P_\times$, P_{gal}	
b_{gal}		$3.1^{+0.7}_{-0.9}$	$3.3^{+0.7}_{-0.9}$
$b_{CO}\langle T_{CO}\rangle$	μK	12^{+9}_{-12}	15^{+8}_{-11}
$P_{shot,gal}$	h^{-3}Mpc^3	430 ± 40	430 ± 40
$P_{shot,CO}$	μK^2 h^{-3} Mpc^3	2100^{+1600}_{-1800}	1700^{+1700}_{-1900}
$P_{shot,\times}$	μK h^{-3} Mpc^3	53 ± 16	-210 ± 300
		Fit of $P_\times$	
$b_{gal}b_{CO}\langle T_{CO}\rangle$	μK	63 ± 45	122 ± 60
$P_{shot,\times}$	μK h^{-3} Mpc^3	53 ± 16	-482 ± 357
		Fit of P_{CO}	
$b_{CO}\langle T_{CO}\rangle$	μK	⋯	<31
$P_{shot,CO}$	μK^2 h^{-3} Mpc^3	⋯	2000^{+1100}_{-1200}
		Fit of P_{gal}	
b_{gal}		⋯	$3.5^{+0.6}_{-0.8}$
$P_{shot,gal}$	h^{-3}Mpc^3	⋯	430 ± 40

Note. Values are shown both with and without use of a prior on $P_{shot,\times}$ of 54 ± 16 μK h^{-3} Mpc^3 derived from the results of Pavesi et al. (2018).

spectrum. In the upper panel of Figure 13, we show their reported limits on $b_{CO}\langle T_{CO}\rangle$, compared with our measurement. Our uncertainties are roughly a factor of two smaller than those reported by Pullen et al. (2013), largely due to our access to a three-dimensional CO data set optimized for intensity mapping analyses, compared to the two-dimensional maps available from WMAP.

To translate our results into mean molecular gas densities, we must assume a value of α_{CO}. Star-forming galaxies at $z \gtrsim 1$ have been found to have Milky Way–like α_{CO} (Daddi et al. 2010; Carleton et al. 2017; Cassata et al. 2020). This class of galaxy is expected to account for a significant fraction of the total CO luminosity (Uzgil et al. 2019; Inami et al. 2020), and therefore we follow other recent works on the mean molecular gas density and adopt $\alpha_{CO} \sim 3.6$ $M_\odot$ (K km s^{-1} pc^2)$^{-1}$ (Daddi et al. 2010). For our cross-spectrum-only constraint, this results in an upper limit of $\rho_{H_2} < 1.5 \times 10^9 M_\odot$ Mpc^{-3} at $z \sim 2.6$. The corresponding limit for our joint fit is $\rho_{mol} < 7.7 \times 10^8 M_\odot$ Mpc^{-3}.

In the lower panel of Figure 13 we compare our ρ_{mol} constraints to a number of literature results.

Multiple recent studies have produced deep spectroscopic (sub)millimeter maps of large regions of sky in order to search individual CO emission lines (Pavesi et al. 2018; González-López et al. 2019). The sources detected by these studies probe the contribution to $\langle T_{CO}\rangle$ from the portion of the CO luminosity function above the survey detection threshold, and therefore represent lower limits on $\langle T_{CO}\rangle$ and ρ_{mol} (Decarli et al. 2019; Riechers et al. 2019; Decarli et al. 2020; Lenkić et al. 2020). In Figure 13 we plot these lower limits as gray boxes. Stacking on

optically selected galaxies has also been used to explore the fraction of ρ_{mol} accounted for by direct detections in CO deep fields, suggesting that they account for around 50% (Inami et al. 2020; Walter et al. 2020). However, these measurements are only sensitive to the galaxies included in the stacks, and galaxy catalogs are incomplete at $z \sim 2$–3 where determining spectroscopic redshifts is difficult.

K20 fit models of the halo mass to CO luminosity relation to the shot power measured in the mmIME and COPSS CO autospectra, and combined these with the halo mass function to produce estimates of ρ_{mol}. These results, presented as open circles in Figure 13, show a molecular density higher than reported by direct detection studies at $z \sim 2.5$. The auto-power spectra of mmIME and COPSS contain information about all emitting galaxies; however, the shot power is most sensitive to the brightest objects (Keenan et al. 2020), and the constraints are sensitive to modeling assumptions (Breysse et al. 2021).

In addition to CO emission line studies, searches for CO absorption in spectra of background quasars can be used to constrain molecular gas column densities. The broad distribution of quasars on the sky eliminates the effects of cosmic variance, but these measurements are difficult due to the small cross section of molecular gas reservoirs and small number of bright background sources (Balashev & Noterdaeme 2018). Klitsch et al. (2019) searched for CO absorption toward quasars in the ALMA calibrator database, and used their non-detection to place upper limits on the molecular gas density at $z < 2$. At $z \sim 1.4$, these limits suggest that direct detection studies recover $\gtrsim 25\%$ of the total molecular gas density; however, the calibrator database does not contain enough $z > 2$ sources to extend these results to higher redshifts.

On the other hand, our clustering power measurements are sensitive to all CO emission without any bias toward bright objects, and are also largely independent of modeling assumptions. This makes them the most reliable tracer of the contribution
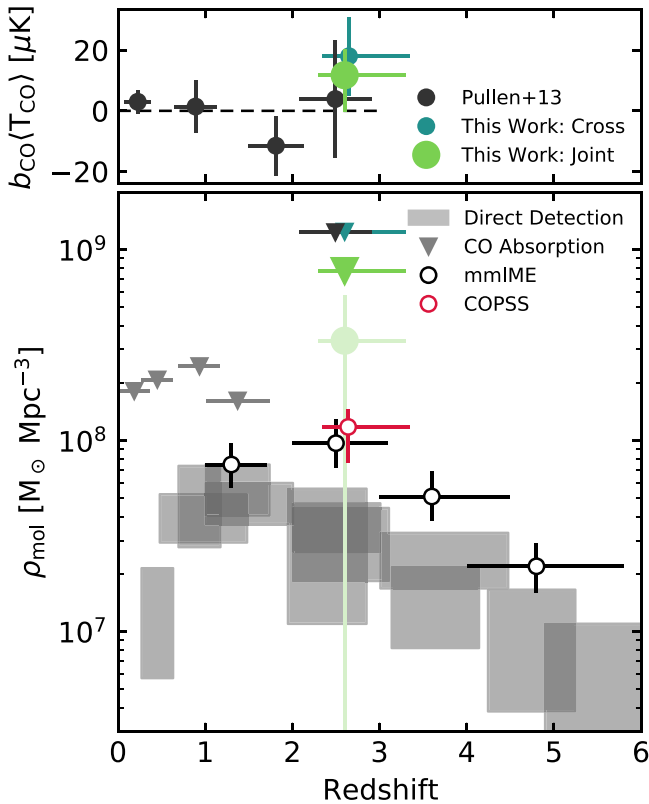
Figure 13. Measurements of the CO mean brightness temperature and molecular gas density presented as a function of redshift based on various CO survey methodologies. Top: $b_{CO}\langle T_{CO}\rangle$ constraints from CO-quasar cross-correlation measurements by Pullen et al. (2013, black) and the results reported here based on CO-galaxy cross-correlation (blue–green point) and based on a joint analysis of the CO and galaxy auto- and cross-power spectra (large green point). Bottom: our upper limits on ρ_{mol} (colored triangles) along with the upper limit of Pullen et al. (2013, black triangle), the upper limits from CO absorption studies (Klitsch et al. 2019, gray triangles), and lower limits from direct detection in CO deep fields (gray boxes; Riechers et al. 2019; Decarli et al. 2020; Lenkić et al. 2020). Open circles correspond to CO auto-power spectrum experiments COPSS (black) and mmIME (red), which are used to fit a halo mass to CO luminosity relation and then integrated over the halo mass function to extract $\langle T_{CO}\rangle$. We also show the central value and 1σ uncertainties of our measurement as the light green point but note that this point is not a significant detection.

of faint galaxies at this redshift, and allows us to set a robust upper limit on the mean molecular gas density. While our survey is not particularly constraining at the current sensitivity, it confirms that the total molecular gas density cannot be much more than an order of magnitude above the direct detection limits. On the other hand, we cannot rule out the possibility that galaxies below the direct detection threshold represent a substantial contribution to the total molecular gas abundance, as suggested by the Keating et al. (2020) auto-power results. Upcoming LIM studies will have much greater sensitivity than the results presented here, and should help to fully contextualize direct detection results.

9. Conclusion

We have presented upper bounds on the CO-galaxy cross-power spectrum at $z \sim 3$. We find a 2σ upper limit on the band-averaged cross-power of $P_\times < 540 \, \mu K \, h^{-3} \, Mpc^3$. This limit lies near models for the cross-power based on the tentative CO auto-power measurement of K16. This suggests that the cross-power may be detectable in current intensity mapping data sets with modest increases in the coverage of spectroscopic redshift

catalogs, and should easily be within the reach of the next generation of LIM experiments, which are currently collecting data. Our measurement is sensitive to the CO emission of *all* galaxies via the clustering component of the cross-power spectrum, allowing us to set a robust upper limit on the mean CO brightness temperature and constrain the mean molecular gas abundance at $z \sim 2.6$ to be $\rho_{mol} < 7.7 \times 10^8 \, M_\odot \, Mpc^{-3}$. This is a factor of two deeper than previous cross-correlation-based constraints.

Cross-correlation can be used to check for systematics and foregrounds in intensity mapping auto-power measurements. We demonstrate this by verifying the data reduction and cleaning of the CO Power Spectrum Survey auto-power measurement. In particular, we search for and find no evidence of bright CS or HCN foregrounds. We also demonstrate the effectiveness of using cross-correlation to remove instrumental effects, finding that strong ground-correlated emission in the COPSS data is cleaned to below the noise level by cross-correlation even with no other cleaning applied.

COPSS was a pathfinder intensity mapping experiment, conducted without the benefit of optimized instrumentation. The tentative detection of the CO auto-power spectrum in Keating et al. (2016), and the constraining limits on the CO-galaxy cross-power spectrum reported here demonstrate the power of the intensity mapping technique for studying galaxy evolution at high redshift, and in regimes not accessible via traditional observing modes. Since the completion of COPSS, a number of purpose-built intensity mapping instruments have been fielded and are now collecting early science data. These instruments will greatly improve the sensitivity of intensity mapping analyses, resulting in competitive constraints on the CO luminosity of average galaxies, the evolution of the mean molecular gas density, and a wide array of astrophysics and cosmology.

Appendix A
Additional Details on Ground Contamination

In Section 6 we demonstrated that the cross-spectrum is largely unaffected by the ground-correlated emission contaminating the SZA data. For comparison, we show auto-power spectrum of the GOODS-N field without corrections for this contamination in Figure 14. These are produced using the same σ-cuts as Figure 6.

Two things are noteworthy, first, at all cuts, the power from ground contamination in low-k modes is significant, and second, removing modes significantly alters the level of contamination and causes large changes in the power spectrum.
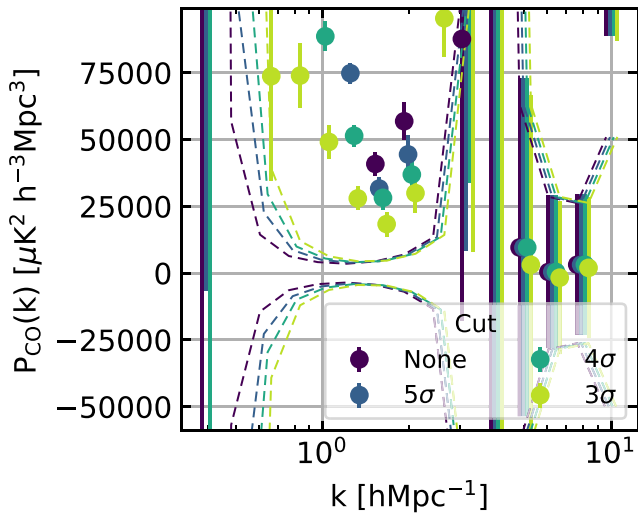
Figure 14. The auto-power spectrum for the GOODS-N field reproduced using a range of different significance cuts to clean the COPSS 30 GHz data. Values for each cut are offset slightly in k for clarity. Contamination is significant for all cuts and changes considerably depending on the level of cleaning performed.

These changes are not observed in the cross-power spectrum, suggesting that the level of contamination does not affect the cross-power spectrum results, and confirming that cross-correlation is statistically immune to biases from uncorrelated contamination.

Appendix B
The Effect of Stochasticity on the Two-halo Term

Our two tracers (galaxies and CO intensity) may not be perfectly correlated with one another, which will lead to a reduction in the cross-power spectrum amplitude relative to the geometric mean of the two auto-spectra (Wolz et al. 2017; Anderson et al. 2018; Schaan & White 2021). This can be quantified in terms of a k-dependent cross-correlation coefficient

$$r(k) = \frac{P_\times(k)}{\sqrt{P_{CO}(k)P_{gal}(k)}}. \tag{B1}$$

For perfectly correlated tracers $r(k) = 1$, and $|r(k)| \leqslant 1$ for any pair of tracers. Since both galaxies and CO trace peaks in the matter density, we also expect $r(k) > 0$.

In the shot power regime, the reduced correlation is a result of scatter in the CO luminosity of the optically selected galaxies, and the effect is implicitly accounted for in our formula for the cross-shot power, Equation (9). In the two-halo regime, $r(k)$ is expected to approach unity, but may be somewhat smaller due to randomness of optically selected and CO-emitting galaxy locations relative to the matter density field. Our cross-power spectrum in Equation (7) accounts for this possibility explicitly via the inclusion of r_{2h} in the clustering term.

We have assumed $r_{2h} = 1$ when interpreting our results. Investigating stochasticity in the galaxy–dark matter cross-spectrum, Bonoli & Pen (2009) found that after subtracting the shot power component of their spectra, $r(k) \sim 1$ for all scales considered (up to $k \sim 2$). Our simulated power spectra (Section 7) also suggest $r_{2h} > 0.7$ for all k where the clustering power is larger than the shot power. However, these

simulations do not include sophisticated treatment of all astrophysical processes that might alter $r(k)$. Chung et al. (2017) explicitly calculate $r(k)$ for simulated cross-power spectra, also finding $r \sim 1$ in the clustering-dominated regime.

If $r_{2h} < 1$, then constraints on $b_{CO}\langle T_{CO}\rangle$ from fitting the cross-spectrum are actually constraints on $r_{2h}b_{CO}\langle T_{CO}\rangle$. The "fit of $P_\times$" values we report in Section 8 and Table 2 should be scaled by $1/r_{2h}$.

In our joint fit of the cross- and auto-spectra, $r_{2h} < 1$ would stretch the range of $b_{CO}\langle T_{CO}\rangle$ favored by the cross-power spectrum to higher values. However, as can be seen from the $b_{CO}\langle T_{CO}\rangle$ panel of Figure 12, the CO auto-spectrum already sets an upper limit on $b_{CO}\langle T_{CO}\rangle$ that rules out much of the parameter space from the fit to the cross-spectrum alone. Thus we would expect $r_{2h} < 1$ to not significantly alter the final limits on $\langle T_{CO}\rangle$ and ρ_{mol} that we give in Section 8.3.

We can verify this by adding r_{2h} to the model we use in the joint fit of the auto- and cross-spectra (i.e., using Equation (7) in place of Equation (8) to model the two-halo term of the cross-power spectrum). We therefore re-fit the spectra with this six-parameter model and the COLDz prior on the cross-shot power. We adopt an additional prior of $0.25 \leqslant r_{2h} \leqslant 1.0$, with the lower limit corresponding to the correlation coefficient in the shot-dominated regime of our simulated cross-power spectra.

We find $b_{CO}\langle T_{CO}\rangle = 10^{+10}_{-16}$. The median and 84th percentile of the distribution are in agreement with the fit without the r_{2h} parameter while the 16th percentile is somewhat lower. The value of r_{2h} itself is not usefully constrained beyond the prior distribution. As expected, upper limits on the molecular gas abundance are largely unchanged by the inclusion of the additional term: $\langle T_{CO}\rangle < 11.1\ \mu K$ and $\rho_{mol} < 7.8 \times 10^8\ M_\odot\ Mpc^{-3}$.

ORCID iDs

Ryan P. Keenan ⓘ https://orcid.org/0000-0003-1859-9640
Garrett K. Keating ⓘ https://orcid.org/0000-0002-3490-146X
Daniel P. Marrone ⓘ https://orcid.org/0000-0002-2367-1080

References

Adelberger, K. L., Steidel, C. C., Pettini, M., et al. 2005, ApJ, 619, 697
Anderson, C. J., Luciw, N. J., Li, Y. C., et al. 2018, MNRAS, 476, 3382
Balashev, S. A., & Noterdaeme, P. 2018, MNRAS, 478, L7
Barger, A. J., Cowie, L. L., & Wang, W. H. 2008, ApJ, 689, 687
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
Bonoli, S., & Pen, U. L. 2009, MNRAS, 396, 1610
Boogaard, L. A., van der Werf, P., Weiss, A., et al. 2020, ApJ, 902, 109
Breysse, P. C., & Alexandroff, R. M. 2019, MNRAS, 490, 260
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., Yang, S., Somerville, R. S., et al. 2021, arXiv:2106.14904
Carilli, C. L., & Walter, F. 2013, ARA&A, 51, 105
Carleton, T., Cooper, M. C., Bolatto, A. D., et al. 2017, MNRAS, 467, 4886
Carucci, I. P., Villaescusa-Navarro, F., & Viel, M. 2017, JCAP, 2017, 001
Cassata, P., Liu, D., Groves, B., et al. 2020, ApJ, 891, 83
Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Natur, 466, 463
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., Breysse, P. C., Ihle, H. T., et al. 2021, ApJ, 923, 188
Chung, D. T., Li, T. Y., Viero, M. P., Church, S. E., & Wechsler, R. H. 2017, ApJ, 846, 60
Chung, D. T., Viero, M. P., Church, S. E., et al. 2019, ApJ, 872, 186

Cohn, J. D., White, M., Chang, T.-C., et al. 2016, MNRAS, 457, 2068

Concerto Collaboration, Ade, P., Aravena, M., et al. 2020, A&A, 642, A60

Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., Blomqvist, M., & Pieri, M. 2018, MNRAS, 481, 1320

Cunnington, S., Harrison, I., Pourtsidou, A., & Bacon, D. 2019, MNRAS, 482, 3341

Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686

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., Aravena, M., et al. 2016, ApJ, 833, 69

Decarli, R., Walter, F., Gónzalez-López, J., et al. 2019, ApJ, 882, 138

Durkalec, A., Le Fèvre, O., Pollo, A., et al. 2015, A&A, 583, A128

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Furlanetto, S. R., & Lidz, A. 2007, ApJ, 660, 1030

Geach, J. E., Sobral, D., Hickox, R. C., et al. 2012, MNRAS, 426, 679

Gong, Y., Cooray, A., Silva, M. B., Santos, M. G., & Lubin, P. 2011, ApJL, 728, L46

González-López, J., Decarli, R., Pavesi, R., et al. 2019, ApJ, 882, 139

Herrero Alonso, Y., Krumpe, M., Wisotzki, L., et al. 2021, A&A, 653, A136

Ihle, H. T., Chung, D., Stein, G., et al. 2019, ApJ, 871, 75

Inami, H., Decarli, R., Walter, F., et al. 2020, ApJ, 902, 113

Kashikawa, N., Yoshida, M., Shimasaku, K., et al. 2006, ApJ, 637, 631

Keating, G. K., Bower, G. C., Marrone, D. P., et al. 2015, ApJ, 814, 140

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

Kennicutt, Robert C. 1998, ApJ, 498, 541

Keres, D., Yun, M. S., & Young, J. S. 2003, ApJ, 582, 659

Khostovan, A. A., Sobral, D., Mobasher, B., et al. 2019, MNRAS, 489, 555

Klitsch, A., Péroux, C., Zwaan, M. A., et al. 2019, MNRAS, 490, 1220

Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, ApJS, 218, 15

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

Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, MNRAS, 480, 5113

Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJL, 763, L20

Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, ApJS, 225, 27

Moradinezhad Dizgah, A., & Keating, G. K. 2019, ApJ, 872, 126

Moradinezhad Dizgah, A., Keating, G. K., Karkare, K. S., Crites, A., & Choudhury, S. R. 2022a, ApJ, 926, 137

Moradinezhad Dizgah, A., Nikakhtar, F., Keating, G. K., & Castorina, E. 2022b, JCAP, 2022, 026

Murray, S. G., Diemer, B., Chen, Z., et al. 2021, A&C, 36, 100487

Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, MNRAS, 477, 1206

Nelson, D., Pillepich, A., Springel, V., et al. 2018, MNRAS, 475, 624

Oxholm, T. M., & Switzer, E. R. 2021, PhRvD, 104, 083501

Padmanabhan, H. 2018, MNRAS, 475, 1477

Pavesi, R., Sharon, C. E., Riechers, D. A., et al. 2018, ApJ, 864, 49

Pen, U.-L., Staveley-Smith, L., Peterson, J. B., & Chang, T.-C. 2009, MNRAS, 394, L6

Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, MNRAS, 475, 648

Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15

Pullen, A. R., Serra, P., Chang, T.-C., Doré, O., & Ho, S. 2018, MNRAS, 478, 1911

Reddy, N. A., Steidel, C. C., Erb, D. K., Shapley, A. E., & Pettini, M. 2006, ApJ, 653, 1004

Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, ApJ, 872, 7

Schaan, E., & White, M. 2021, JCAP, 2021, 068

Silva, M., Santos, M. G., Cooray, A., & Gong, Y. 2015, ApJ, 806, 209

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24

Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676

Steidel, C. C., Shapley, A. E., Pettini, M., et al. 2004, ApJ, 604, 534

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

Switzer, E. R., Anderson, C. J., Pullen, A. R., & Yang, S. 2019, ApJ, 872, 82

Switzer, E. R., Chang, T. C., Masui, K. W., Pen, U. L., & Voytek, T. C. 2015, ApJ, 815, 51

Switzer, E. R., Masui, K. W., Bandura, K., et al. 2013, MNRAS, 434, L46

Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878

Uzgil, B. D., Carilli, C., Lidz, A., et al. 2019, ApJ, 887, 37

Visbal, E., & Loeb, A. 2010, JCAP, 2010, 016

Walter, F., Carilli, C., Neeleman, M., et al. 2020, ApJ, 902, 111

Walter, F., Decarli, R., Sargent, M., et al. 2014, ApJ, 782, 79

Wirth, G. D., Trump, J. R., Barro, G., et al. 2015, AJ, 150, 153

Wolz, L., Blake, C., & Wyithe, J. S. B. 2017, MNRAS, 470, 3220

Wolz, L., Murray, S. G., Blake, C., & Wyithe, J. S. 2019, MNRAS, 484, 1007

Wolz, L., Pourtsidou, A., Masui, K. W., et al. 2022, MNRAS, 510, 3495

Wolz, L., Tonini, C., Blake, C., & Wyithe, J. S. B. 2016, MNRAS, 458, 3399

Yang, S., Popping, G., Somerville, R. S., et al. 2021a, arXiv:2108.07716

Yang, S., Pullen, A. R., & Switzer, E. R. 2019, MNRAS, 489, L53

Yang, S., Somerville, R. S., Pullen, A. R., et al. 2021b, ApJ, 911, 132

Yoshikawa, T., Akiyama, M., Kajisawa, M., et al. 2010, ApJ, 718, 112