The Astrophysical Journal, 933:186 (25pp), 2022 July 10 https://doi.org/10.3847/1538-4357/ac63c7

COMAP Early Science. V. Constraints and Forecasts at $z \sim 3$

Dongwoo T. Chung[1,2], Patrick C. Breysse[3], Kieran A. Cleary[4], Håvard T. Ihle[5], Hamsa Padmanabhan[6], Marta B. Silva[5], J. Richard Bond[1], Jowita Borowska[5], Morgan Catha[7], Sarah E. Church[8], Delaney A. Dunne[4], Hans Kristian Eriksen[5], Marie Kristine Foss[5], Todd Gaier[9], Joshua Ott Gundersen[10], Stuart E. Harper[11], Andrew I. Harris[12], Brandon Hensley[13], Richard Hobbs[7], Laura C. Keating[14], Junhan Kim[4], James W. Lamb[7], Charles R. Lawrence[9], Jonas Gahr Sturtzel Lunde[5], Norman Murray[1], Timothy J. Pearson[4], Liju Philip[9], Maren Rasmussen[5], Anthony C. S. Readhead[4], Thomas J. Rennie[11], Nils-Ole Stutzer[5], Bade D. Uzgil[4,15], Marco P. Viero[4], Duncan J. Watts[5], Risa H. Wechsler[8,16], Ingunn Kathrine Wehus[5], and David P. Woody[7]
(COMAP Collaboration)

[1] Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada; dongwooc@cita.utoronto.ca
[2] Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada
[3] Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY, 10003, USA
[4] California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA
[5] Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, N-0315 Oslo, Norway
[6] Département de Physique Théorique, Université de Genève, 24 Quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland
[7] Owens Valley Radio Observatory, California Institute of Technology, Big Pine, CA 93513, USA
[8] Kavli Institute for Particle Astrophysics and Cosmology & Physics Department, Stanford University, Stanford, CA 94305, USA
[9] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[10] Department of Physics, University of Miami, 1320 Campo Sano Avenue, Coral Gables, FL 33146, USA
[11] Jodrell Bank Centre for Astrophysics, Alan Turing Building, Department of Physics and Astronomy, School of Natural Sciences, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[12] Department of Astronomy, University of Maryland, College Park, MD 20742, USA
[13] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[14] Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany
[15] National Radio Astronomy Observatory, Pete V. Domenici Array Science Center, P.O. Box 0, Socorro, NM 87801, USA
[16] SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
Received 2021 November 19; revised 2022 February 24; accepted 2022 March 4; published 2022 July 13

Abstract

We present the current state of models for the $z \sim 3$ carbon monoxide (CO) line intensity signal targeted by the CO Mapping Array Project (COMAP) Pathfinder in the context of its early science results. Our fiducial model, relating dark matter halo properties to CO luminosities, informs parameter priors with empirical models of the galaxy–halo connection and previous CO (1–0) observations. The Pathfinder early science data spanning wavenumbers $k = 0.051$–$0.62\ \mathrm{Mpc}^{-1}$ represent the first direct 3D constraint on the clustering component of the CO (1–0) power spectrum. Our 95% upper limit on the redshift-space clustering amplitude $A_{\mathrm{clust}} \lesssim 70\ \mu\mathrm{K}^2$ greatly improves on the indirect upper limit of $420\ \mu\mathrm{K}^2$ reported from the CO Power Spectrum Survey (COPSS) measurement at $k \sim 1\ \mathrm{Mpc}^{-1}$. The COMAP limit excludes a subset of models from previous literature and constrains interpretation of the COPSS results, demonstrating the complementary nature of COMAP and interferometric CO surveys. Using line bias expectations from our priors, we also constrain the squared mean line intensity–bias product, $\langle Tb \rangle_2 \lesssim 50\ \mu\mathrm{K}^2$, and the cosmic molecular gas density, $\rho_{\mathrm{H2}} < 2.5 \times 10^8\ M_\odot\ \mathrm{Mpc}^{-3}$ (95% upper limits). Based on early instrument performance and our current CO signal estimates, we forecast that the 5 yr Pathfinder campaign will detect the CO power spectrum with overall signal-to-noise ratio of 9–17. Between then and now, we also expect to detect the CO–galaxy cross-spectrum using overlapping galaxy survey data, enabling enhanced inferences of cosmic star formation and galaxy evolution history.

Unified Astronomy Thesaurus concepts: CO line emission (262); Cosmological evolution (336); High-redshift galaxies (734); Molecular gas (1073); Radio astronomy (1338)

1. Introduction

Line intensity mapping (LIM) surveys propose to map 3D fluctuations in integrated redshifted spectral line emission across large cosmological volumes (see reviews by Kovetz et al. 2017, 2019). These survey designs generally focus on statistical measurements of the line emitters as a whole, including faint populations of galaxies that cannot be detected in isolation but may be inferred in aggregate. Investigating what measurements and inferences this perspective enables, previous literature has studied the potential of high-redshift LIM with carbon monoxide (CO) lines for over a decade (see, e.g., Righi et al. 2008; Visbal & Loeb 2010; Lidz et al. 2011; Visbal et al. 2011; Pullen et al. 2013; Li et al. 2016; Padmanabhan 2018; Moradinezhad Dizgah & Keating 2019; Sun et al. 2019; Yang et al. 2021, 2022).

The history of direct power spectrum measurements of CO intensity is somewhat shorter, as surveys like the CO Power Spectrum Survey (COPSS; Keating et al. 2016) and the Millimeter-wave Intensity Mapping Experiment (mmIME; Keating et al. 2020a) have only begun to publish results relatively

recently. Both of these surveys leverage community instruments to make interferometric measurements of the CO line intensity field over fields of ∼10–100 arcmin^2 (∼10^{-3}–10^{-2} deg^2), with both claiming measurements of CO power slightly beyond a 2σ level of significance. In addition to the CO auto-spectrum measurements from COPSS and mmIME, Keenan et al. (2022) demonstrated the feasibility of cross-correlation between COPSS data and galaxy surveys, placing an upper limit on the CO–galaxy cross-spectrum. However, neither COPSS nor mmIME probes sufficiently large scales to constrain CO fluctuations shaped by clustering, instead measuring the shot noise chiefly expected to arise from the stochastic bright end of the luminosity function (LF).

Results from the first observing season (Y1) of the CO Mapping Array Project (COMAP) Pathfinder (Cleary et al. 2022), the first instrument specifically designed for single-dish CO LIM, provide the first direct constraints on the clustering component of the high-redshift CO line intensity power spectrum. COMAP Pathfinder observations at 26–34 GHz measure CO (1–0) (rest frequency 115.27 GHz) at $z = 2.4$–3.4 in three fields of 4 deg^2, allowing characterization of larger transverse scales than with previous interferometric LIM surveys. Other papers associated with these results[17] describe the instrument (Lamb et al. 2022), data processing and mapmaking procedures (Foss et al. 2022), and power spectrum methodology and results (Ihle et al. 2022). This paper aims to convert these measurements into astrophysical inferences and consider forecasts for the remainder of the initial 5 yr Pathfinder campaign, with a separate paper by Breysse et al. (2022a) considering potential realizations of COMAP beyond the Pathfinder.

LIM surveys like COMAP tend to assume approximate isotropy in the spatial structure of line emission across the survey volume and so report the spherically averaged power spectrum across comoving wavenumber k. The clustering component of this power spectrum, proportional to the first moment of the LF, traces the large-scale structure underlying the CO emitters and dominates at low k. The (independent) shot-noise component, proportional to the second moment of the LF, describes scale-independent fluctuations arising from the Poisson statistics of line emitters and dominates at high k. Due to how each component relates to the CO LF, shot-noise measurements probe CO stochasticity and bright emitters, but clustering measurements are less skewed toward the bright end of the LF, meaning they will be more sensitive to the properties of faint emitters. We expect COMAP Pathfinder results to be significant in this context because our high-sensitivity instrumentation is purpose-built for CO LIM and enables access to large angular scales unavailable to COPSS or mmIME, thus allowing measurements of the power spectrum at clustering-dominated ranges of k previously inaccessible to CO LIM surveys.

While current COMAP Pathfinder measurements are consistent with white noise and thus provide an upper limit for the spherically averaged CO power spectrum $P(k)$ at $k \sim 10^{-1}$ Mpc^{-1}, several years remain in the observing campaign, during which we anticipate a detection based on previous models in the literature. Furthermore, members of the COMAP collaboration have worked on updating our own fiducial CO models and expectations for LF and molecular gas density constraints at the conclusion of the COMAP Pathfinder survey. In this context this paper aims to answer questions about the COMAP Pathfinder campaign following Y1 and early science verification:

1. What inferences do our early science verification data enable about the $z \sim 3$ CO (1–0) power spectrum and molecular gas abundance?
2. Given early science sensitivities and updated $z \sim 3$ models, what are our present expectations for constraints on these same quantities, and others like the CO LF, at the end of 5 yr of COMAP Pathfinder observations?

We organize the paper as follows. In Section 2 we outline our fiducial model for CO emission at $z \sim 3$, chiefly in comparison to the model of Li et al. (2016) and to observational results. Then, in Section 3 we consider implications of the current COMAP Pathfinder $P(k)$ limit in relation to other models and observational results in the literature. Finally, in Section 4 we outline simulated constraints of our new CO model based on expected 5 yr results from the COMAP Pathfinder, before outlining overall conclusions in Section 5.

Unless otherwise stated, we assume base-10 logarithms and a ΛCDM cosmology with parameters $\Omega_m = 0.286$, $\Omega_\Lambda = 0.714$, $\Omega_b = 0.047$, $H_0 = 100\ h$ km s^{-1} Mpc^{-1} with $h = 0.7$, $\sigma_8 = 0.82$, and $n_s = 0.96$, to maintain consistency with previous COMAP simulations (Li et al. 2016; Ihle et al. 2019). The cosmology is also broadly consistent with 9 yr WMAP results (Hinshaw et al. 2013). Distances carry an implicit h^{-1} dependence throughout, which propagates through masses (all based on virial halo masses, proportional to h^{-1}) and volume densities ($\propto h^3$).

2. Devising a Model for CO at Redshift 3

As we discussed in Section 1, substantial astrophysical literature over the past decade has formulated models for the line emission fluctuations probed by CO LIM, some of which we will consider later in this paper in relation to COMAP Pathfinder observations. While a detailed examination of these prior models is not within the scope of the present work, we distinguish two general approaches to dark-matter-halo-based models of CO fluctuations, as follows:

1. *Indirect models of the halo–CO connection via intermediate properties:* These models connect the mass distribution of dark matter (DM) halos to a different property like star formation rate (SFR) or infrared luminosity (usually as a proxy for SFR), often but not necessarily through simulations or abundance matching. This intermediate property then relates to CO luminosity via simulations or fits to data. This group includes the models of Pullen et al. (2013), which make use of an empirical relation from Wang et al. (2010) between CO luminosity and bolometric far-infrared luminosity, and the model of Li et al. (2016), which instead uses a CO–SFR relation derived in part from empirical correlations described by the review of Carilli & Walter (2013). (This group therefore also includes the modification of this last model by Keating et al. 2020a, which replaces the Carilli & Walter 2013 relation with results from Kamenetzky et al. 2016).
2. *Direct models of the halo–CO connection, motivated by observed CO emitter abundances:* The recent emergence of direct constraints on the high-redshift CO LF now

[17] Beyond CO LIM, see also early results from continuum observations composing the COMAP Galactic Plane Survey (Rennie et al. 2022).

enables direct abundance matching of such measurements to the halo mass distribution to obtain a halo model for CO emission. Padmanabhan (2018) first undertook formulation of a model of this kind applicable at $z \sim 3$ (using data compiled across a wide redshift range), and the present work now aims to follow this approach also (albeit only for CO (1–0) at $z \sim 3$).

Previous forecasting efforts for COMAP have used the fiducial model of Li et al. (2016). Since then, we have gained new insight into CO (1–0) emitters at high redshift through two important surveys: the CO Luminosity Density at High Redshift survey (COLDz; Riechers et al. 2019), which provides the strongest constraints on the CO (1–0) LF at $z = 2$–3 to date; and COPSS (Keating et al. 2016), which made a tentative detection of shot-noise power from small-scale CO fluctuations. Other surveys such as the aforementioned mmIME, the ALMA SPECtroscopic Survey (ASPECS) in the Hubble Ultra Deep Field, and the Plateau de Bure High-z Blue-Sequence Survey 2 (PHIBBS2) lend insight into emission in higher-J CO lines at these redshifts (Decarli et al. 2020; Keating et al. 2020a; Lenkić et al. 2020).

Here we present a new fiducial model that takes into account the COPSS and COLDz measurements—as well as priors from empirical models of the halo mass–SFR relation and the SFR–CO luminosity relation already used in Li et al. (2016)—and uses a double power-law parameterization modified from Padmanabhan (2018) (removing redshift dependence[18]). The new parameterization models the halo mass–CO luminosity relation with greater flexibility and directness compared to Li et al. (2016).

We first provide an overview of the parameterization in Section 2.1 and then present priors on the model parameters in Section 2.2. An additional aspect of our model is a basic treatment of line broadening, as described in Section 2.3, which is highly approximate but acceptable when considering the sensitivity expected especially from our early data. Only after laying this groundwork can we discuss our procedure for inferring parameter constraints from COMAP Pathfinder measurements in Section 4, through Markov Chain Monte Carlo (MCMC) runs using our fiducial parameterization and priors to inform forward models of one- and two-point statistics.

2.1. Fiducial Parameterization of the Halo–CO Connection

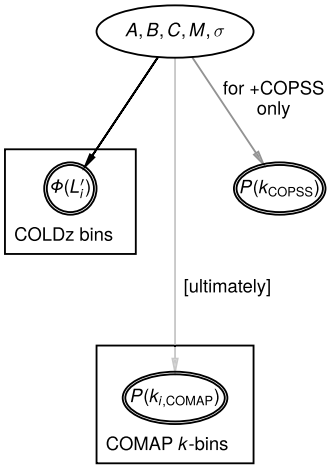
The double power-law parameterization of the halo mass–CO luminosity relation $L(M_h)$ approximates the composition of a series of scaling relations connecting halo mass M_h to CO luminosity L or $L_{\rm CO}$ similar to the series considered by Li et al. (2016).

1. As in Li et al. (2016), we consider a single power law relating IR luminosity and CO luminosity:
$$\log L_{\rm IR} = \alpha \log L'_{\rm CO} + \beta, \tag{1}$$
where for CO (1–0)
$$\frac{L_{\rm CO}}{L_\odot} = 4.9 \times 10^{-5} \frac{L'_{\rm CO}}{\rm K\ km\ s^{-1} pc^2}. \tag{2}$$
This form of scaling relation is commonly fitted to observational data in the literature (see, e.g., the reviews of Solomon & Vanden Bout 2005; Carilli & Walter 2013). Li et al. (2016) found values of α anywhere between 1.00 and 1.37 from the high-redshift CO studies available at the time of that work.
2. We also relate IR luminosity to SFR as in Li et al. (2016):
$$\frac{\rm SFR}{M_\odot\ {\rm yr}^{-1}} = \delta_{\rm MF} \times 10^{-10} \left(\frac{L_{\rm IR}}{L_\odot} \right), \tag{3}$$
for some coefficient $\delta_{\rm MF}$ whose value depends on the initial mass function (IMF). This is set to 1 in Li et al. (2016), but we refer the reader to reviews by Calzetti (2013), Casey et al. (2014), and Madau & Dickinson (2014) for more information about the IMF dependence of this SFR calibration in the style of Kennicutt (1998).
3. The UniverseMachine (UM) framework of Behroozi et al. (2019) models the average SFR for a star-forming galaxy hosted in a halo with maximum circular velocity at peak halo mass $v_{M\rm peak}$ as
$$\frac{\langle {\rm SFR_{SF}} \rangle (v_{M_{\rm peak}})}{M_\odot\ {\rm yr}^{-1}} = \epsilon \left[\frac{1}{v^{\alpha_{\rm UM}} + v^{\beta_{\rm UM}}} + \gamma \exp\left(-\frac{\log^2 v}{2\delta^2} \right) \right], \tag{4}$$
where
$$v = \frac{v_{M_{\rm peak}}}{V\,[{\rm km\ s^{-1}}]}. \tag{5}$$
(Note that we have added "UM" subscripts to α and β from Behroozi et al. 2019, denoting that these are UM parameters, to avoid confusion with α and β from Li et al. 2016.) This is a double power law with a Gaussian component added to it. However, here we assume that the effect of the Gaussian component is negligible (i.e., $\gamma \approx 0$) and consider only the double power-law component. The above equations are for the star-forming galaxy population rather than the quenched population, but according to the model of Behroozi et al. (2019), the latter is a small enough portion of galaxies at the redshifts we probe that we do not consider it for this exercise.
4. Behroozi et al. (2019) also provide a relation (although approximate) between peak halo mass (which for these redshifts is essentially the same as halo mass) and $v_{M\rm peak}$:
$$v_{M_{\rm peak}}(M_h) = (200\ {\rm km\ s^{-1}}) \left[\frac{M_h}{M_{200\ {\rm km\ s^{-1}}}(a)} \right]^{0.3}, \tag{6}$$
where $a = 1/(1+z)$ is the scale factor at redshift z and
$$M_{200\ {\rm km\ s^{-1}}} = \frac{1.64 \times 10^{12}\ M_\odot}{(a/0.378)^{-0.142} + (a/0.378)^{-1.79}}. \tag{7}$$

Across all of these relations, we can in principle list the independent parameters $\{\alpha, \beta, \delta_{\rm MF}, \epsilon, \alpha_{\rm UM}, \beta_{\rm UM}, V\}$. However, many of these are degenerate in the context of CO LIM data, and for our analyses it makes more sense to deal with combinations of these parameters, in a simplified reparameterization.

If we make the assumption that α is close to unity, or at least $\alpha \gg 1$ and $\alpha \ll 1$—which seems a justifiable one, given that the prior on this parameter in Li et al. (2016) was $\alpha = 1.17 \pm 0.37$—then we can collapse all of the above scaling relations into a

[18] Whereas Padmanabhan (2018) sought to model CO over a broad redshift range of $z \sim 0$–3, we concentrate on a narrower range where redshift evolution is expected to be much less significant. Therefore, a redshift-dependent parameterization would complicate the model for little benefit.



Step 1: Initial priors devised (Section A.1)

- "flat" (conservative, uninformative)
- "UM" (based on empirical fits and models)
- "P18" (high-information, strong assumptions about CO redshift evolution)

Step 2: Condition priors on current observations

- Likelihood functions based on COLDz LF alone ("+COLDz") or COPSS $P(k)$ constraint also ("+COLDz+COPSS")—cf. Section A.2
- Infer updated priors via MCMC (Section A.3)

Step 3: Use resulting posteriors as new data-driven priors

- {flat,UM,P18}+{COLDz,COLDz+COPSS} priors can now be meaningfully conditioned on COMAP data once they reach sufficient sensitivity
- Can also be used to generate best estimate models for COMAP forecasting

Figure 1. Simplified, annotated graphical representation of the derivation of our model priors for the $z \sim 3$ CO (1–0) $L(M_h)$ relation, which is considered in much greater detail throughout Appendix A.

single $L'_{\rm CO}(M_h)$ relation (which then exactly corresponds to the intrinsic $L_{\rm CO}$ via Equation (2)) with four free parameters:

$$\frac{L'_{\rm CO}(M_h)}{\rm K\ km\ s^{-1} pc^2} = \frac{C}{(M_h/M)^A + (M_h/M)^B} \equiv \frac{C}{m^A + m^B}. \quad (8)$$

Additionally, we assume that there is some lognormal scatter σ (in units of dex) about this relation, which is taken to be the (linear) mean at fixed halo mass.

Note also that α does not need to be exactly 1, i.e., the CO–IR or CO–SFR relation does not need to be exactly linear, for Equation (8) to be a reasonable approximation to (even if not an exact description of) the true composition of Equations (1)–(7). With $\alpha = 1$ the true composition will clearly deviate somewhat from our simplification, but principally around $M_h = M$ and not so much for $M_h \ll M$ or $M_h \gg M$. In any case, our simplification is flexible enough to reflect highly nonlinear trends of CO emission against other properties. This is important because simulations and observational analyses continue to explore the complex environmental factors and physical processes that drive CO emissivity (or lack thereof) relative to H_2 content in molecular clouds and galaxies and thus determine the physical interpretation of CO LIM observations (Li et al. 2018; Gong et al. 2020; Inoguchi et al. 2020; Keating et al. 2020b; Madden et al. 2020; Seifried et al. 2020; Breysse et al. 2022b).

2.2. Priors from Previous Models and Observations

We want to formulate a set of priors for our model parameters for two reasons. The first is that they serve as a range of fiducial expectations for forecasting the CO signal at this relatively speculative time. The other is that they will serve as the ground level for Bayesian inferences from COMAP data.

The details of these priors are somewhat ancillary to the primary results of this work, and so they are discussed largely in Appendix A. However, we present a broad overview in Figure 1. In short, we begin with one of three possible sets of initial priors on our CO model parameters ("flat," "UM," "P18") and then condition these priors on either the COLDz LF constraints alone or both the COLDz constraints and the COPSS $P(k)$ measurement. These posterior distributions, obtained via MCMC, then act as data-driven priors for COMAP and can be conditioned on COMAP data at some later date to yield updated posterior distributions.

The data-driven priors can also be used to generate analytic estimates for the real-space $P(k)$ at $z \sim 2.4$, the central redshift for the COLDz observations. The `lim` package[19] can generate $P(k)$ estimates for every sample of each MCMC. We use a minimum halo mass of $10^9\,M_\odot$ for CO emission in these calculations, but as our models strongly favor a steep superlinear faint-end power law for $L(M_h)$ (i.e., $A < -1$), shifting the minimum halo mass up to $10^{10}\,M_\odot$ has minimal effect on our predictions, including for $P(k)$. We therefore use the higher minimum mass for the remainder of this work, as it matches the value used in our previous fiducial model devised by Li et al. (2016) and as the cosmological simulations we use for simulated COMAP inferences in Section 4.2 will only resolve halos with mass $\sim 10^{10}\,M_\odot$ (reproducing correct statistics for halos with $M_h > 2.5 \times 10^{10}\,M_\odot$).

We plot the 68% credibility intervals in Figure 2 alongside both the model of Li et al. (2016), which previously acted as the fiducial model for COMAP simulations, and observational LIM results from COPSS and mmIME. The COPSS result is in some tension ($\approx 2\sigma$–3σ) with our other priors, as are the mmIME estimates ($\approx 1\sigma$–2σ). One proposition by Keating et al. (2020a) was that clustering could contribute significantly to the COPSS measurement and thus the best estimate for the shot-noise power spectrum should be adjusted down to $2.0^{+1.1}_{-1.2} \times 10^3\,h^{-3}\ \mu{\rm K}^2\,{\rm Mpc}^3$. We discuss the balance of clustering versus shot noise for the CO power spectrum further in Section 3.2, in the context of current COMAP constraints.

There is one caveat related to this tension that we should consider about our data-driven priors. At present, surveys like COLDz principally constrain CO emitter abundances around or above the knee of the LF and do not meaningfully constrain the faint-end slope of the LF. The COLDz data prefer neither a positive faint-end slope that would suggest fewer faint CO emitters nor a negative one that would suggest more faint emitters. Splitting the difference necessarily results in a highly tempered estimate of the total abundance of CO emitters and thus a highly tempered estimate of the total CO power spectrum.

[19] https://github.com/pcbreysse/lim/tree/pcbreysse

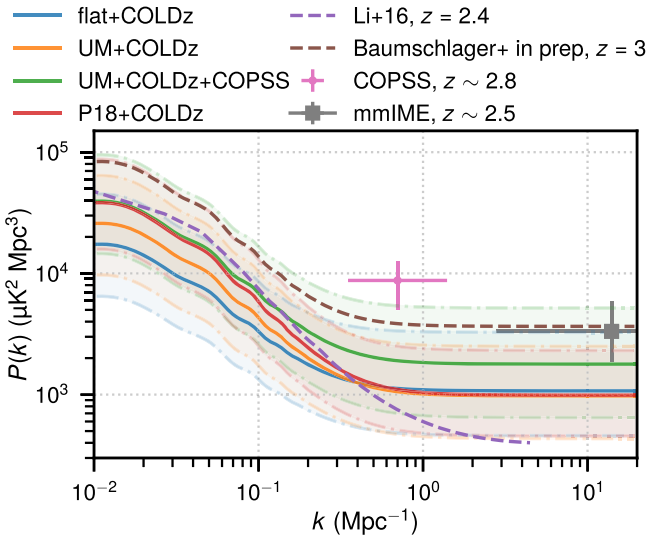


Figure 2. Predictions for the real-space $P(k)$ of the CO (1–0) line intensity field at $z \sim 2$–3. Alongside our data-driven priors and their 68% credibility intervals (solid lines and shaded areas bounded by dashed–dotted lines), we also show predictions from Li et al. (2016) and the `TNG300_2` model of B. Baumschlager et al. (2022, in preparation; used by Silva et al. (2021) for forecasting COMAP–HETDEX synergies), as well as results from COPSS (Keating et al. 2016) and mmIME (Keating et al. 2020a).

This tempered nature affects not only comparisons of our data-driven priors with observational results but also comparisons with previous models informing some of our priors. The best-fit model of Padmanabhan (2018), for instance, also uses observational data to drive an abundance-matched $L(M_h)$ model. At $z \sim 3$ the principal driver is the COPSS data from Keating et al. (2016), but in the form of constraints on the Schechter parameterization of the CO LF. The prior on the faint-end slope that Keating et al. (2016) used is loose but asymmetric and does prefer negative values, and their overall estimate of the LF knee lies higher in both abundance and luminosity than the COLDz constraints. Thus, these data drive the original model $P(k)$ of Padmanabhan (2018) orders of magnitude above our P18+COLDz model $P(k)$, which the COLDz data temper significantly.

One tempting resolution of the tension between our COLDz-driven priors and the COPSS and mmIME results, then, is in interpretation of results from CO line searches like COLDz as a kind of lower bound when considering quantities that involve the faint end of the LF, including the CO power spectrum. We find this idea mirrored in the interpretation of ASPECS data by Uzgil et al. (2019), who quoted a lower limit on the mean total CO line temperature based on the individual line detections from that survey. For ASPECS CO (2–1) detections, Uzgil et al. (2019) were able to use CO–galaxy cross statistics with external optically selected spectroscopic redshifts to constrain the faint-end slope of the $z \sim 1$ CO LF. However, they elected not to claim similar constraints for CO (3–2) at $z \sim 2.5$ owing to potential unreliability of such constraints given the percentage of ASPECS detections without matching optical counterparts. Therefore, any clustering amplitude constraint from direct detections depends strongly on the selection characteristics. Since CO LIM surveys trade this dependence away for the price of potential systematics and contamination, the discrepancy between COLDz and COPSS could be considered a natural result of these caveats.

Table 1
CO Model Point Estimates Based on Data-driven Priors

Data-driven Prior	Point Estimates for: A	B	$\log C$	$\log \frac{M}{M_\odot}$	σ
"flat+COLDz"	−3.7	7.0	11.1	12.5	0.36
"UM+COLDz"	−2.75	0.05	10.61	12.3	0.42
"UM+COLDz+COPSS"	−2.85	−0.42	10.63	12.3	0.42
"P18+COLDz"	−2.4	−0.5	10.45	12.21	0.36

Note. Values are determined at $z \sim 2.8$ to match the median $P(k)$ and LF values from each data-driven prior. We indicate our fiducial choice in boldface.

It is, however, possible that the resolution of any tension specifically involving the shot-noise-dominated measurement of COPSS actually lies in a lower-abundance faint end of the CO LF. If reweighted based on the COPSS measurement, the COLDz LF Schechter parameter posterior would actually weakly prefer larger positive values of the faint-end log-slope of the LF. The Schechter function as used by Riechers et al. (2019) models the CO emitter number density per log-luminosity bin as proportional to a power law $L_{\rm LF}^{\alpha}$ times an exponential cutoff $\propto \exp(-L/L_*)$. Then, the shot noise is proportional to the average integrated squared luminosity of the emitters, which is roughly proportional to $\Gamma(\alpha_{\rm LF}+2)$. This function reaches a local minimum at $\alpha_{\rm LF} \approx -0.54$ but will be greater for lower or higher values of $\alpha_{\rm LF}$. We can make sense of this physically: a CO LF with fewer faint emitters and more emitters near or above the knee—e.g., from low-mass halos hosting low-metallicity systems with high CO dissociation rates—leads to enhanced contrast of CO line intensity fluctuations at small scales and thus a greater shot-noise amplitude of the CO power spectrum. Without a clustering measurement like COMAP, independent of both direct-detection surveys like COLDz and shot-noise LIM surveys like COPSS, we have limited ability to bound the faint end of the CO LF from either below or above.

Ultimately, we drive our fiducial UM+COLDz+COPSS estimate with the best and most relevant observational data available, but this model is conservative by nature of the COLDz data (which hold far higher total statistical weight than the COPSS data). In future forecasting and forward models, we should be entirely open to the possibility that faint CO emitters are far more abundant—and thus that the integrated cosmic average CO intensity is considerably higher—than direct CO line searches suggest at the time of writing.

For now, reverting to the COMAP central redshift of $z \sim 2.8$, we can identify specific parameter values to approximately match the median $P(k)$ and $\phi(L')$ values for each set of priors (as shown in Figure 2 and Appendix A). This assumes that the CO signal is relatively insensitive to cosmology and redshift (within the COMAP survey range), which is true when compared to our model uncertainties. We show the parameter point estimates corresponding to each data-driven prior in Table 1.

2.3. Incorporating Line Broadening

Not only are CO emitters not point sources, but their extent in a data cube does not correspond to their extent in physical or comoving space. Some of this is due to instrumental resolution, but some of this is due to observations being in redshift space rather than in real space. One key effect to consider are the

peculiar velocities of the gas within each galaxy—due both to overall galactic rotation and to turbulent gas motion separate from this rotation—which results in Doppler broadening of the CO line emission.

Chung et al. (2021) provide some methods to account for line broadening, providing an empirical line width model for CO (1–0) under the assumption that CO emitters are rotation-dominated, mostly disk-like sources. The inclination angle i of each emitter's axis of rotation relative to the observer line of sight is assumed to be random and independent, with a uniform distribution of $\cos i \in (0,1)$. Using this model, we set the FWHM of the CO line profile for a host halo of virial mass M_h to the circular velocity of the halo at the median inclination angle of $i = \pi/3$. In this work, we use either numerical calculations based on an analytic model or approximate N-body simulations using the peak patch method (Stein et al. 2019) that we consider further in Section 4.2. In both cases, the halo maximum circular velocity is unavailable and we use the virial velocity $v_{\rm vir}$ instead. Chung et al. (2021) preferred the former but compared using one versus the other and found the choice to not affect results significantly.

The CO line FWHM estimated from the host halo's virial velocity and randomized inclination is

$$\begin{aligned} v(M_h, z, i) &= v_{\rm vir}(M_h, z)\left[\frac{\sin i}{\sin(\pi/3)}\right] \\ &= \left(\frac{\Delta_c}{2}\right)^{1/6} [GM_h H(z)]^{1/3}\left(\frac{\sin i}{\sqrt{3}/2}\right) \\ &\approx 35\ {\rm km\ s^{-1}}\left(\frac{\sin i}{0.866}\right)\left(\frac{\Delta_c}{200}\right)^{1/6} \\ &\quad \times \left(\frac{M_h}{10^{10}\,M_\odot}\frac{H(z)}{100\ {\rm km\ s^{-1}\ Mpc^{-1}}}\right)^{1/3}. \end{aligned} \tag{9}$$

Here Δ_c is the spherical overdensity within the virial radius of the halo, relative to the critical density of our cosmology. The value used by Chung et al. (2021) is 180, whereas 200 is also common (being historically considered canonical for a cosmology with critical matter density—see White 2001, 2002). This difference in Δ_c is of minimal concern, as the resulting difference in $v(M_h)$ is only a few percent.

When a forecast of *only* the spherically averaged $P(k)$ is required, a single Gaussian filter with an effective velocity scale $v_{\rm eff}$ is sufficient to describe the smearing of the total CO line intensity cube. This comes at the cost of some accuracy but will bring significantly improved computational speed in any contexts where the approximation is applicable. Including adjustments for random inclinations, the appropriate effective velocity given by Equation (46) of Chung et al. (2021) is

$$v_{\rm eff} = \frac{1}{2}\left(\frac{\langle L^2 v_{\rm vir}\rangle}{\langle L^2\rangle} + \frac{4}{\pi\sqrt{3}}\frac{\langle L^2\rangle}{\langle L^2 v_{\rm vir}^{-1}\rangle}\right), \tag{10}$$

where $\langle x\rangle \equiv \int dM_h\,(dn/dM_h)\,x$.

As Chung et al. (2021) make clear, stark shortcomings in approximating the effect of line broadening with only $v_{\rm eff}$ exist in the context of projections made in the present work for future analyses, which will deal not only with $P(k)$ but also with the voxel intensity distribution (VID). Therefore, in mocks of the CO line intensity field using approximate N-body simulations, we bin halos by virial velocity and broaden the CO emission from each bin by its median velocity. We use the two-tier approach outlined in Chung et al. (2021), which ignores line broadening for halos below a certain mass whose line profiles are not possible to resolve with the COMAP Pathfinder science channelization of 32 MHz (equivalent to $\approx$320 km s^{-1} in velocity space for 30 GHz observations). To recap the procedure in full:

1. Divide the halos into a low-mass subset with $M_h < 10^{11}\,M_\odot$ and a high-mass subset with $M_h > 10^{11}\,M_\odot$. The cut point is equivalent to $v_{\rm vir} \approx 107$ km s^{-1}, so the low-mass subset includes all halos whose CO line widths should span less than one-third of a COMAP science voxel.
2. Generate a CO cube from the low-mass subset without applying any Gaussian filters.
3. Divide the high-mass subset into 16 equally spaced linear bins in virial velocity.
4. For each bin, generate a CO cube with a Gaussian filter applied to approximate line broadening. The median virial velocity across all halos within the bin sets the Gaussian width. This results in 16 CO cubes, one for each velocity bin.
5. Sum all 17 CO cubes, including the low-mass CO cube, for the final simulated product.

Simulations by Chung et al. (2021) show that this approach keeps $P(k)$ within 10% of the reference simulation (using 64 bins in halo circular velocity) and the VID approximately within Poisson error of the reference simulation. The increase in time for the CO cube computation is around a factor of 30, but the computation is still sufficiently fast when considering the other steps involved in simulations such as power spectrum evaluation. Thus, this will be our approach to simulating line broadening for anything more complicated than simple $P(k)$ forecasts.

Note that we did not apply this correction above when constraining our priors with observational results. First, the COLDz data set used is of distributions of discrete emitters, and our COLDz-based likelihood does not need models of any line profiles. Second, the UM+COLDz+COPSS calculation needs in principle to correct for the effect of line broadening on the CO (1–0) power spectrum, especially as it will attenuate the apparent power spectrum less for wavenumbers where COMAP measures $P(k)$ compared to COPSS. However, even for COPSS the effect at $k \sim 1\ h$ Mpc^{-1} is typically $\sim$30% and thus is subdominant to the overall uncertainty in the $\sim 2\sigma$ COPSS $P(k)$ result. Therefore, we err on the conservative side and do not correct for line broadening in devising our priors.

3. Implications of COMAP Early Science Power Spectrum Measurements

The present state of COMAP observations does not yet allow for the kinds of analyses that we forecast in Section 4. However, the $P(k)$ result[20] obtained by Ihle et al. (2022)

[20] Strictly speaking, as Ihle et al. (2022) note in their Section 3.1, the result is based on a pseudo-power-spectrum measurement and may have some residual mode-mixing bias. However, their Figure 1 also shows that this mode-mixing bias likely is a small effect (5%–30%) that enhances the pseudo-spectrum relative to the true signal. The measurement obtained from this pseudo-spectrum result should thus still be a valid, if possibly conservative, upper limit on the true CO power spectrum.

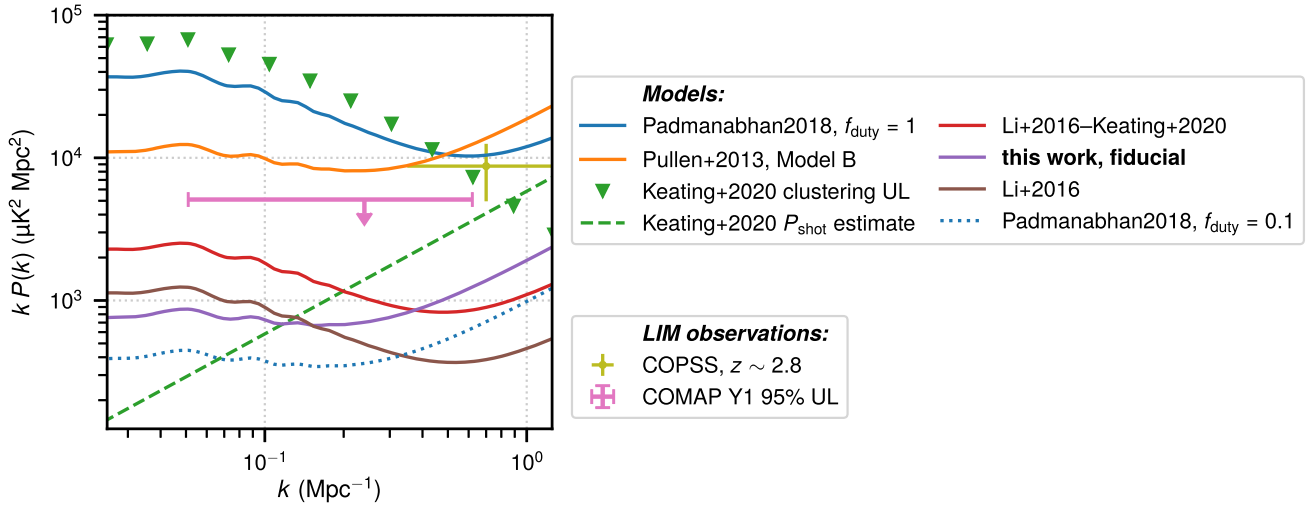


Figure 3. COMAP Pathfinder early science constraint (pink) on the redshift-space CO (1–0) power spectrum at $z \sim 3$, alongside model predictions from this work (UM+COLDz+COPSS) and those recalculated based on $L(M_h)$ relations from Padmanabhan (2018), Pullen et al. (2013), and Li et al. (2016), as well as a variation on the latter from Keating et al. (2020a). We also show interpretations of the COPSS result both as a direct $P(k)$ measurement (yellow error bars; Keating et al. 2016) and as a constraint on clustering (triangles) and shot-noise amplitudes (dashed line; Keating et al. 2020a).

already has constraining power that strongly complements the COPSS result. We do note that after the original analysis of Keating et al. (2016), which essentially assumed the measurement to be entirely shot noise dominated, Keating et al. (2020a) reanalyzed the COPSS power spectrum result as a combination of clustering and shot-noise components, setting an upper limit on clustering and revising down the best estimate for shot noise. Nonetheless, the COPSS measurement still reflects a k-range where the majority of the signal is shot noise, whereas the majority of any potential power spectrum measurement in the COMAP k-range would be from clustering.

Coadding constant-elevation scan (CES) data across all fields, the all-scale measurement is $P(k = 0.051\text{–}0.62\,\mathrm{Mpc}^{-1}) = (-2.7 \pm 1.7) \times 10^4\,\mu\mathrm{K}^2\,\mathrm{Mpc}^3$. Asserting $P(k) > 0$ on top of this measurement, we obtain a 95% upper limit of $k\,P(k) < 5.1 \times 10^3\,\mu\mathrm{K}^2\,\mathrm{Mpc}^2$ at $k = 0.24\ \mathrm{Mpc}^{-1}$, shown in Figure 3.

Note that the bulk of the present sensitivity derives from Field 1 CES data, which alone yield $P(k = 0.051\text{–}0.62\,\mathrm{Mpc}^{-1}) = (-4.6 \pm 2.2) \times 10^4\,\mu\mathrm{K}^2\,\mathrm{Mpc}^3$, or a 95% upper limit of $k\,P(k) < 5.4 \times 10^3\ \mu\mathrm{K}^2\,\mathrm{Mpc}^2$ at $k = 0.24\ \mathrm{Mpc}^{-1}$ when requiring $P(k) > 0$.

The current COMAP constraint already excludes the predictions of Padmanabhan (2018) (assuming a CO emission duty cycle—i.e., the fraction of time that any given galaxy is CO luminous—of 1) and Model B of Pullen et al. (2013) at 95% confidence,[21] and overall it constrains the clustering component of the power spectrum better than the COPSS reanalysis of Keating et al. (2020a) by roughly an order of magnitude. We first consider the model exclusions in Section 3.1 before considering clustering constraints in more detail in Section 3.2 and translating these into molecular gas constraints in Section 3.3.

[21] We also exclude but do not consider other models in the literature that are based on outdated assumptions, which subsequent works often supersede. For instance, COPSS data also excluded predictions from the model of Lidz et al. (2011) (even in the pilot analysis done by Keating et al. 2015). However, the Lidz et al. (2011) model had already been reformed at $z \sim 3$ into Model A of Pullen et al. (2013) with a revised halo mass–SFR scaling that was more applicable at these redshifts.

3.1. Excluded Models

Model B of Pullen et al. (2013) was in principle one of the models already excluded by the COPSS measurement of Keating et al. (2016), but we exclude it in the clustering regime, whereas the COPSS results excluded it in the shot-noise regime ($k = 0.5$–$10\ h\ \mathrm{Mpc}^{-1}$). This is a meaningful distinction particularly for this model, as Pullen et al. (2013) implement a duty cycle f_{duty} for CO-bright activity, with which the shot noise scales inversely. Therefore, the constraint of Keating et al. (2016) can only be on some combination of the halo mass–CO luminosity scaling and f_{duty}, encapsulating the shot-noise amplitude. The correction required for Model B of Pullen et al. (2013) to be made consistent with the COPSS measurement could thus be either a different CO–SFR scaling from what Pullen et al. (2013) used—which was a fit to local and high-redshift galaxies by Wang et al. (2010)—or a different value of f_{duty}.

In a typical halo model, the clustering amplitude scales directly with f_{duty}. However, Model B of Pullen et al. (2013) derives the cosmic average CO temperature $\langle T \rangle$ from using the Wang et al. (2010) CO–SFR relation to directly scale the integrated SFR density obtained via Schechter fits to the SFR function tabulated by Smit et al. (2012). As Pullen et al. (2013) assume that the duty cycle for CO-bright activity matches the duty cycle for star formation activity, f_{duty} does not modify $\langle T \rangle$ for Model B of Pullen et al. (2013) and thus should not modify the lower-k power spectrum values that we constrain.

Our results thus suggest that the Wang et al. (2010) CO–SFR relation is not globally applicable to galaxies at $z \sim 3$, in the sense that it cannot be used to connect the SFR functions of Smit et al. (2012) to CO luminosity at this redshift range. Indeed, while the Wang et al. (2010) relation suggests $\mathrm{SFR} \sim L_{\mathrm{CO}}^{1.67}$, this is much steeper than the general correlation at high redshift inferred from data reviewed by Carilli & Walter (2013), which includes some data not available at the time of Wang et al. (2010).

Also of interest is our exclusion of the model of Padmanabhan (2018) with $f_{\mathrm{duty}} = 1$, which explicitly folded the Keating et al. (2016) result into its derivation. In comparison to other models,

this model predicts a higher clustering amplitude relative to the shot-noise amplitude. Without other significant data available to drive the abundance matching carried out at $z \sim 3$ by Padmanabhan (2018), it was perfectly reasonable for the resulting model to account for the COPSS result through a very high overall power spectrum prediction—including a high clustering amplitude—as opposed to additional parameterization of stochasticity to further decouple the shot-noise and clustering amplitudes. This once again highlights the value of having COMAP data to separately constrain the power spectrum at lower k.

Note that Pullen et al. (2013) and Padmanabhan (2018) each present an alternate model that we do not exclude. Model A of Pullen et al. (2013) is based on a less empirical, more indirect set of assumptions to connect halo and galaxy properties, and more similar (both qualitatively and quantitatively) to our fiducial models or that of Li et al. (2016). Meanwhile, Padmanabhan (2018) shows $P(k)$ curves for both $f_{\rm duty} = 1$ and $f_{\rm duty} = 0.1$. We do not exclude the latter variation on this model in principle, although Figure 3 shows that this variation then predicts shot noise well below our UM+COLDz+COPSS model's expectation, as well as the COPSS measurement alone. Padmanabhan (2018) also notes that $f_{\rm duty} = 1$ is somewhat better supported in observational data. The tension between these two extremes (and their implications for the ratio between the clustering and shot-noise components of the power spectrum) could be feasibly bridged by a mass-dependent $f_{\rm duty}$ that falls from 1 with higher mass, as is the case for the empirical models of Yang et al. (2022).

3.2. Constraints on CO Power Spectrum Clustering and Shot-noise Amplitudes

At this early stage of the Pathfinder campaign, COMAP data will not yet place significant constraints on the parameters of our $L(M_h)$ model devised in Section 2. However, we show that the upper limit does place meaningful constraints on the integrated clustering and shot-noise amplitudes for the CO power spectrum. Furthermore, by leveraging our model priors from Section 2, we can obtain an upper limit on the mean temperature $\langle T \rangle$ at $z \sim 2.8$ from our clustering amplitude constraint, from which we derive limits on H_2 mass density in Section 3.3.

In real comoving space, we would model the power spectrum as

$$P(k) = A_{\rm clust} P_m(k) + P_{\rm shot}. \tag{11}$$

This is to say that the total $P(k)$ is the sum of a clustering component, the matter power spectrum $P_m(k)$ scaled by a clustering amplitude $A_{\rm clust}$, and a shot-noise component $P_{\rm shot}$. This neglects any possible scale-dependent bias or one-halo terms but is sufficient for our purposes.

We should then be able to consider likelihood contours and constraints for $A_{\rm clust}$ and $P_{\rm shot}$ based on our observational data, both in isolation and in combination with the COPSS $P(k)$ measurements from Keating et al. (2016). This mirrors the COPSS reanalysis performed by Keating et al. (2020a).

For the real-space $P(k)$, we would have $A_{\rm clust} = \langle Tb \rangle_2$, or the square of the mean line temperature–bias product across the LF:

$$\langle Tb \rangle \propto \int dM_h \, \frac{dn}{dM_h} L(M_h) b(M_h), \tag{12}$$

with appropriate conversions applied to convert luminosity density to brightness temperature. Without $b(M_h)$ in the integrand, the analogous integral would yield the mean CO brightness temperature $\langle T \rangle$; the line luminosity-averaged bias is then $b \equiv \langle Tb \rangle / \langle T \rangle$.

However, redshift-space distortions from the coherent infall of galaxies into large-scale structure (Kaiser 1987; Hamilton 1998) enhance the clustering component such that $A_{\rm clust} / \langle Tb \rangle^2 \approx 1 + 2/(3b) + 1/(5b^2)$ for small k (and $\Omega_m(z) \approx 1$, which is the case at $z \sim 3$). Furthermore, as explained in Section 2.3, line broadening introduces k-dependent attenuation, largely of the shot noise. In the context of $P(k)$, the parameter $v_{\rm eff}$ described there is sufficient to encapsulate the overall effect.

Given our limited knowledge of line bias and line broadening for CO at high redshift, we consider two different ways to present constraints on the power spectrum clustering and shot-noise amplitudes.

1. In the first method we carry out a b-agnostic, $v_{\rm eff}$-agnostic calculation of constraints on $A_{\rm clust}$ and $P_{\rm shot}$. We make no assumptions about values of b, instead constraining the overall observed amplitude $A_{\rm clust}$ that scales the matter power spectrum. We also ignore line broadening altogether and make no attempt to compensate for its effect on our data. Thus, we assume that the shot-noise component looks the same in real and redshift space (before transfer functions, for which we do compensate). This is closest to the analyses of Keating et al. (2020a) and Keenan et al. (2022), neither of which correct CO auto-spectra for line broadening or account for linear redshift distortions.
2. In the second method we constrain $\langle Tb \rangle_2$ and $P_{\rm shot}$ in a b-informed[22] and $v_{\rm eff}$-informed analysis, incorporating line broadening and expectations for line bias based on our UM+COLDz priors. From our UM+COLDz MCMC distribution, we obtain average values of b and $v_{\rm eff}$ across these dimensions[23] and define reasonable two-dimensional polynomial fits to those average values, as described in Appendix B. This allows us to directly calculate $P(k)$ including redshift-space distortions and line broadening, which should be appropriate to fit simultaneously to COPSS data and to the COMAP data that have been corrected in $(k_\parallel, k_\perp)$ space (before spherical averaging) to account for beam, filtering, and spectral effects.

We show results from both methods, using COMAP data and/or COPSS data, in Table 2. We also illustrate the COMAP–COPSS joint constraints graphically in Figure 4 for the b- and $v_{\rm eff}$-agnostic method and in Figure 5 for the b- and $v_{\rm eff}$-informed method.

[22] Despite making assumptions around b, we do not attempt to directly evaluate constraints in the 2D parameter space of $\langle T \rangle$–$P_{\rm shot}$. Such an approach will involve a scaling of $A_{\rm clust} / \langle T \rangle_2 = b_2 + 2b/3 + 1/5$ to estimate the clustering component of the power spectrum for a given value of $\langle T \rangle$, whereas our approach only involves scaling by $A_{\rm clust} / \langle Tb \rangle^2 \approx 1 + 2/(3b) + 1/(5b^2)$ for a given value of $\langle Tb \rangle$. Any unreliability in determining b will result in far greater relative error in the former than in the latter for plausible values of b in our models.

[23] We derive these values from the UM+COLDz priors to avoid double-counting any information from COPSS in our analysis, but the resulting fits hold equally well for the UM+COLDz+COPSS MCMC samples.

Table 2
Constraints on the CO Clustering and Shot-noise Amplitudes and on Derived Quantities

	b- and $v_{\rm eff}$-agnostic:		b- and $v_{\rm eff}$-informed:		b- and $v_{\rm eff}$-agnostic:		b- and $v_{\rm eff}$-informed:	
Data	$A_{\rm clust}$ (μK^2)	$P_{\rm shot}$ (10^3 μ K^2 Mpc^3)	$\langle Tb\rangle^2$ (μK^2)	$P_{\rm shot}$ (10^3 μ K^2 Mpc^3)	$\langle T\rangle$ (μK)	$\rho_{\rm H2}$ (10^8 $M_\odot$ Mpc^{-3})	$\langle T\rangle$ (μK)	$\rho_{\rm H2}$ (10^8 $M_\odot$ Mpc^{-3})
COPSS[a]	<630	$5.7^{+4.2}_{-3.6}$	<345	$12.1^{+7.5}_{-6.4}$	$<11.$	<7.4	<9.3	<6.4
COMAP Y1	<66	<19	<49	<24	<3.5	<2.4	<3.5	<2.5
COMAP Y1+COPSS	<69	$6.8^{+3.8}_{-3.5}$	<51	$11.9^{+6.8}_{-6.1}$	<3.5	<2.5	<3.6	<2.5

Notes. We use the terms "b- and $v_{\rm eff}$-agnostic/informed" to denote one of two methods used to infer and present constraints on the power spectrum component amplitudes, as discussed in Section 3.2. Bounds on the derived quantities $\langle T\rangle$ and $\rho_{\rm H2}$ depend on a priors-based assumption of $b > 2$ and other conversions discussed in Sections 3.2 and 3.3. Upper limits are 95% confidence; bounded intervals are 68% confidence.

[a] The $A_{\rm clust}$ constraint differs somewhat in our reanalysis from the reanalysis of Keating et al. (2020a), which found a 95% upper limit of 420 μK^2. We ascribe the discrepancy to differences in assumed cosmology, including in parameters not enumerated by Keating et al. (2020a) that determine $P_m(k)$. Our COPSS-based $P_{\rm shot}$ estimate uncorrected for line broadening, which does not depend on such parameters, corresponds to $2.0^{+1.4}_{-1.2} \times 10^3\,h^{-3}$ μK^2 Mpc^3, and is entirely consistent with the Keating et al. (2020a) estimate of $2000^{+1100}_{-1200}\,h^{-3}$ μK^2 Mpc^3.

With the agnostic method, COMAP data by themselves constrain $A_{\rm clust} \lesssim 70$ μK^2 at 95% confidence, with or without the COPSS data. For the COMAP–COPSS joint analysis, the accompanying shot-noise constraint is $P_{\rm shot} = 6.8^{+3.8}_{-3.5} \times 10^3$ μK^2 Mpc^3, around 78% of the total COPSS $P(k)$ measurement. Comparing this to the COPSS reanalysis by Keating et al. (2020a), which yielded a best estimate of $P_{\rm shot} = 2.0^{+1.1}_{-1.2} \times 10^3\,h^{-3}$ μK^2 Mpc^3 $= 5.8^{+3.2}_{-3.5} \times 10^3$ μK^2 Mpc^3 (around 66% of the total $P(k)$ measurement), shows that the limit placed by COMAP data on $A_{\rm clust}$ constrains how much of the COPSS signal could be ascribed to measuring clustering versus measuring shot noise.

We show COMAP–COPSS joint constraints from the informed method in Figure 5; with this model the COMAP data also drive a clustering constraint of $\langle Tb\rangle^2 \lesssim 50$ μK with or without COPSS data. The inferred actual $P_{\rm shot}$ value[24] of $1.2^{+0.7}_{-0.6} \times 10^4$ μK^2 Mpc^3 is significantly higher than from our first method and suggests that line broadening attenuates the COPSS measurement of shot noise by ≈40%; this is entirely consistent with the median expectation for the CO (1–0) $P(k)$ around $k \sim 1$ Mpc^{-1} from the simulations of Chung et al. (2021). That said, the upward correction merely reflects additional assumptions about line broadening rather than any added direct information. As a result, outside of this particular analysis we will not use this 40% difference to correct any measurements or derived constraints from either Keating et al. (2016) or Keating et al. (2020a) (e.g., as shown in Figures 6–9).

We also note a lack of sufficient sensitivity to further narrow our data-driven priors, even considering the clustering amplitude in isolation. Our upper limit for $\langle Tb\rangle^2$ corresponds to 13 times the value for the UM+COLDz+COPSS point estimate model, whereas the 68% credibility interval for $\langle Tb\rangle^2$ for any of our data-driven priors already spans less than an order of magnitude, as Figure 2 suggests.

Our two analyses arrive at either an $A_{\rm clust}$ constraint or a $\langle Tb\rangle^2$ constraint, but the two constraints are consistent with each other. Comparing the bottom panel of Figure 5 with the estimate of b as a function of $\langle Tb\rangle^2$ and $P_{\rm shot}$ in Appendix B, we can see that the parameter space preferred by the data tends to be associated with luminosity-averaged bias values of $b \sim 3$ (although specific points in that space, like our point estimate models, may have even higher b). Then, the COMAP–COPSS joint 95% upper limit from Table 2 of $\langle Tb\rangle^2 < 51$ μK should translate to an upper limit on the redshift-space clustering amplitude of $A_{\rm clust} \approx \langle T\rangle^2(b^2 + 2b/3 + 1/5) \approx 63$ μK^2. This is within 10% of the $A_{\rm clust}$ upper limit obtained from our previous method, with differences likely arising from our simplified treatment of line bias and signal distortions.

We are more conservative about b in deriving an upper limit on $\langle T\rangle$. For all of our priors, the sampled parameter sets all result almost entirely in $b > 2$; a value of $b = 2$ would be under the third percentile for "flat+COLDz" and under the first percentile for the others. Most models in the literature also favor superlinear $L(M_h)$ relations at lower mass (with possible exceptions being older models like those of Lidz et al. 2011 and Pullen et al. 2013, which had $L \propto M_h$) and thus fairly high values of b.

Combining the priors-based constraint of $b > 2$ with our first method's limit on $A_{\rm clust} = \langle T\rangle^2(b^2 + 2b/3 + 1/5)$, we would obtain $\langle T\rangle < 3.5$ μK. Combining $b > 2$ with our second method's limit of $\langle Tb\rangle^2 < 51$ μK^2 yields essentially the same limit (within 1%) of $\langle T\rangle < 3.6$ μK. In either case, this result—which the COMAP data primarily drive—is currently the best LIM clustering constraint on the CO (1–0) $\langle T\rangle$ at $z \sim 3$, outperforming by a factor of 3 the joint COPSS auto- and COPSS–galaxy cross-spectra analysis result of $\langle T\rangle < 10.9$ μK from Keenan et al. (2022). We illustrate this improvement, as well as the general history of constraints on $\langle Tb\rangle$—either from the CO auto-spectrum (via $\langle Tb\rangle^2$) or from a CO–galaxy cross-spectrum—in Figure 6.

3.3. Derived Constraints on Molecular Gas Abundance

The constraint on $\langle T\rangle$ directly translates into a constraint on the cosmic H_2 mass density $\rho_{\rm H2}$. The conversion $\alpha_{\rm CO}$ between H_2 mass (noting that here we do not deal with a gas mass density that includes heavier elements or atomic hydrogen) and CO luminosity is typically quoted with H_2 mass in intrinsic units of $M_\odot$ and CO luminosity in observer units of K km s^{-1} pc^2. Then, at redshift z, given $\alpha_{\rm CO}$ and the Hubble

[24] Unlike with the agnostic method, this value does not change significantly with the incorporation of COMAP data. The incorporation of line broadening into the informed method likely accounts for this fact. The COMAP data exclude very high values of $P_{\rm shot}$ that would be consistent with COPSS data on account of attenuation from line broadening at COPSS wavenumbers, but not with the COMAP data at lower k. This exclusion suppresses the inferred $P_{\rm shot}$ and cancels out the increase in inferred $P_{\rm shot}$ from clustering amplitude limits (which was the sole effect of COMAP data on COPSS interpretation with the agnostic method).

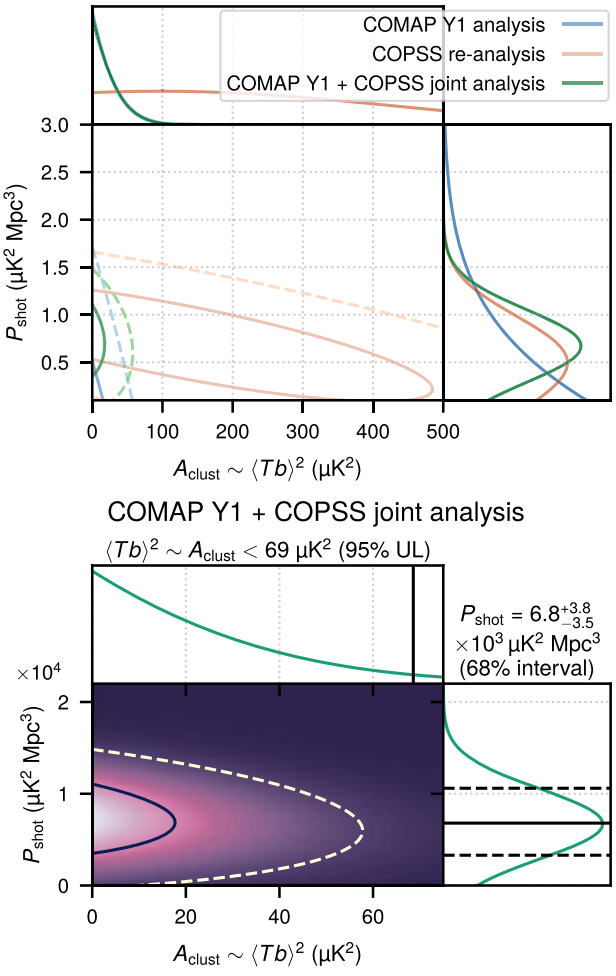


Figure 4. Top: likelihood contours (largest panel) and marginalized probability distributions (smaller panels) for the clustering and shot-noise amplitudes of the CO power spectrum, based on different data sets. The solid and dashed 2D contours respectively represent $\Delta\chi^2 = \{1, 4\}$ relative to the minimum χ^2 obtained in the parameter space, corresponding to 1σ and 2σ for 2D Gaussians. Bottom: likelihood distribution conditioned jointly on COMAP and COPSS $P(k)$ measurements, with the corresponding contours reproduced from the top panel and marginalized constraints on each of $A_{\rm clust}$ and $P_{\rm shot}$ shown.

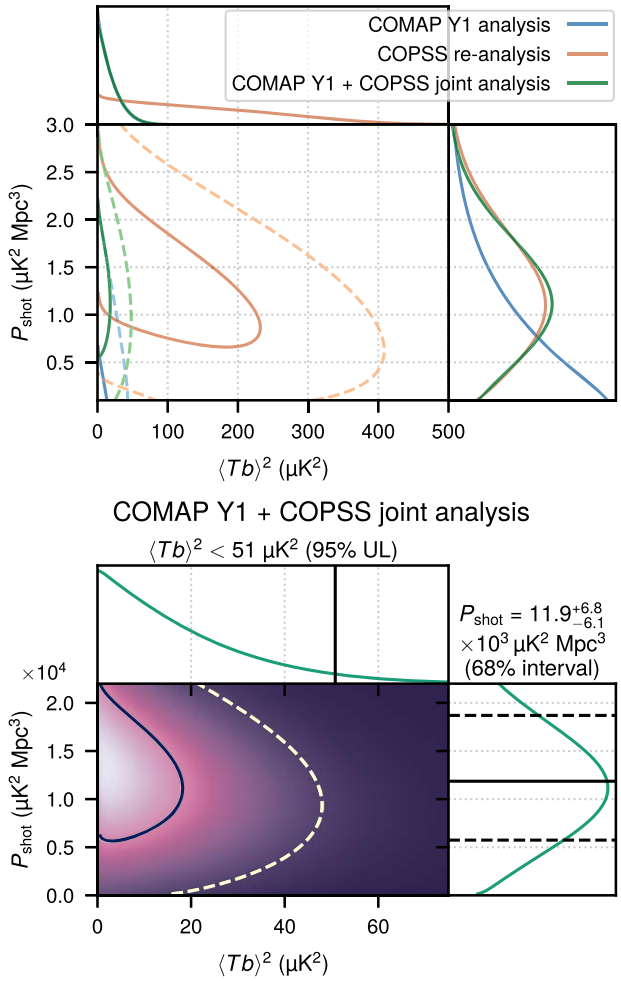


Figure 5. Same as Figure 4, but with strong assumptions around line bias and line broadening as discussed in the main text. These assumptions also allow us to claim a constraint on $\langle Tb \rangle^2$ itself rather than the redshift-space clustering amplitude $A_{\rm clust}$ as considered near the start of Section 3.2.

parameter $H(z)$,

$$\rho_{\rm H2} = \frac{\alpha_{\rm CO} \langle T \rangle H(z)}{(1+z)^2}. \tag{13}$$

At the COMAP central redshift of $z \approx 2.8$, our upper limit of $\langle T \rangle < 3.6$ μK thus translates to an upper limit of $\rho_{\rm H2} < 2.5 \times 10^8\, M_\odot$ Mpc^{-3} given $\alpha_{\rm CO} = 3.6\, M_\odot$ (K km s^{-1} pc^2)$^{-1}$, which we use for easy comparison with other works that use the same conversion, such as Decarli et al. (2020), Lenkić et al. (2020), and Riechers et al. (2019). We show our upper limit alongside these other works in Figure 7.

Since the COMAP upper limit for $\langle Tb \rangle^2$ was less stringent than our data-driven (COLDz-based) priors for the clustering power spectrum, we do not expect our upper limit on $\rho_{\rm H2}$ to be more constraining than our priors either. Indeed, the 90% interval for $\rho_{\rm H2}$ for our UM+COLDz+COPSS priors is given by $\log[\rho_{\rm H2}/(M_\odot\ {\rm Mpc}^{-3})] = 7.58^{+0.23}_{-0.25}$, slightly higher than the COLDz stand-alone calculation from Riechers et al. (2019) owing to both UM and COPSS favoring a higher abundance of molecular gas, respectively, through preference for a steep faint-end LF slope and through simply a higher measurement as shown in Figure 7. Therefore, the 95th percentile value of $\rho_{\rm H2} = 9.5 \times 10^7\, M_\odot$ Mpc^{-3} from our fiducial priors sits at less than one-half of the COMAP upper limit.

On the other hand, the upper limit is still notable in relation to the other constraints shown in Figure 7. For one, we obtained this limit across a much wider area—on the order of square degrees—compared to the other surveys, which all operate across patches of $\sim$ square arcminutes. The small volumes of these surveys can result in substantial cosmic variance and systematic biases not necessarily currently accounted for by their analyses, although the COPSS survey design (which spans multiple fields distributed widely across the sky, vs. ASPECS and COLDz spanning one or two fields) should be less susceptible to these effects (Keenan et al. 2020).

Since the upper limit is within a factor of between two and three of the upper edge of our priors, the final COMAP Pathfinder measurement should indeed have constraining power beyond our priors, which we explore in Section 4.2. By making use of up to 69 times more science-quality integration time (which would correspond to a map noise level lower by more than 8 times) than even the Field 1 CES-only

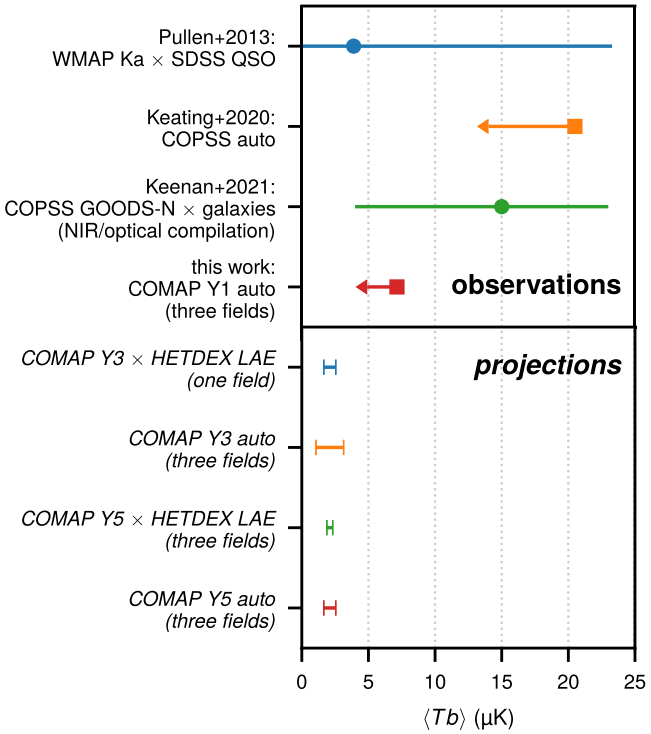


Figure 6. $\langle Tb \rangle$ constraints from previous observational analyses—the broadband cross-correlation of Pullen et al. (2013), the auto-spectrum of Keating et al. (2016), and the 3D cross-correlation of Keenan et al. (2022)—alongside our current upper limit. We also show projections for future results based on COMAP auto- and CO–galaxy cross-spectra in subsequent years (see Section 4.3).

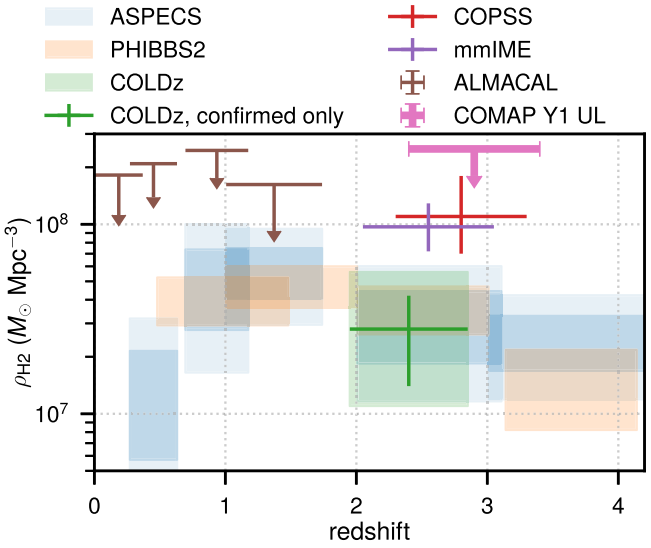


Figure 7. Current COMAP constraint on ρ_{H2} (thick bar with downward arrow) in relation to past CO-based results from ASPECS (Decarli et al. 2020), PHIBBS2 (Lenkić et al. 2020), COLDz (Riechers et al. 2019, from which we show results based on either all line candidates or only those that have confirmatory independent spectroscopic measurements), COPSS (Keating et al. 2016), mmIME (Keating et al. 2020a), and ALMACAL (Klitsch et al. 2019). All results use $\alpha_{CO} = 3.6\,M_\odot$ (K km s^{-1} pc^{-2})$^{-1}$ except COPSS, which uses a conversion of $\alpha_{CO} = 4.3\,M_\odot$ (K km s^{-1} pc^{-2})$^{-1}$.

results (which dominate our coadded CES-only sensitivity), 5 yr results from the COMAP Pathfinder should be on par with the other results shown in Figure 7 and should act as an independent check on those measurements of $z \sim (2\text{–}3)\rho_{H2}$. We will discuss the expected 5 yr constraint on ρ_{H2} in more quantitative detail later in this work (Section 4).

As the present COMAP constraint and future expected constraints derive from directly measuring $\langle Tb \rangle$ as opposed to reconstructing $\langle T \rangle$ from individual detections or shot-noise measurements, they will serve the community as a strongly complementary probe of cosmic molecular gas density at $z \sim 3$. In particular, we note that the results of Keating et al. (2020a) depend strongly on models of the multiple overlapping CO lines encompassed by ALMA observing frequencies.

Incidentally, our upper limit also compares favorably to the ALMACAL upper limits of Klitsch et al. (2019) derived for $z \sim 0$–2, from a blind search for CO absorption lines against background ALMA calibrators. The survey design for ALMACAL enables >1500 hr of integration time spanning a wide sky area—unusual for a community instrument and enabled only by the use of calibrator source observations. However, the ALMACAL approach cannot extend beyond $z \sim 2$ owing to the nature of ALMA calibrators, the majority of which appear to lie below $z \sim 1.5$ with a small tail of the redshift distribution stretching out to $z \sim 3$ (Bonato et al. 2018). Thus, while molecular gas surveys not limited by cosmic variance are possible with ALMA through absorption-line searches, these will not be able to survey the same redshifts as LIM or emission-line searches.

4. Expectations for COMAP Pathfinder Future Science Results

As Foss et al. (2022) note in their Section 4.2, future observing seasons should improve the rate at which we acquire science-quality integration time through a combination of improvements in hardware, observing efficiency, and analysis. This implies that by the end of Year 5 of the Pathfinder campaign (Y5), sensitivity relative to the current Y1 power spectrum results of Ihle et al. (2022) will improve not by a factor of 5 but by as much as a factor of 69 over the Field 1 Y1 result (which, as noted above, accounts for much of the current sensitivity). Of interest is how this final Pathfinder sensitivity will enable exclusion or detection not only of our fiducial UM +COLDz+COPSS model but also of other models previously considered in the literature.

We first briefly discuss the expected raw detection sensitivity in Section 4.1 and then simulate how this sensitivity will enable inferences about $z \sim 3$ CO in Section 4.2. Finally, in Section 4.3 we touch on possible science gains between now and Y5 results through cross-correlation with the Hobby-Eberly Telescope Dark Energy eXperiment (HETDEX; Hill et al. 2008, 2021; Gebhardt et al. 2021).

4.1. Current Predictions for Detection Significance

We show current and expected Pathfinder sensitivities (with the latter based on the aforementioned improvements forecast by Foss et al. 2022) in Figure 8, alongside several models of the CO power spectrum. As our current sensitivity already excludes some of the models shown, as already considered in Section 3.1, we will not make signal-to-noise ratio (S/N) forecasts for those models.

We expect Y5 COMAP Pathfinder results to yield confident detections across multiple k-bins of other models yet to be excluded, including our own fiducial model, which would be detectable with an all-k S/N of 9 (excluding sample variance).

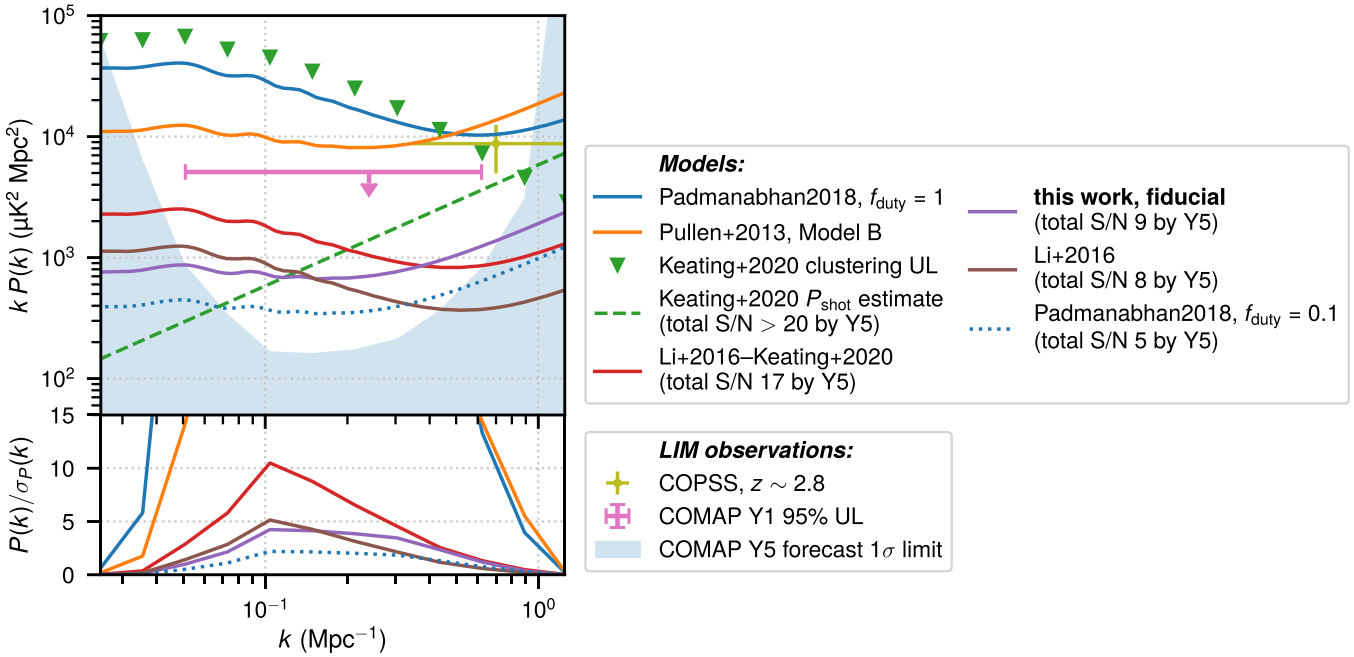


Figure 8. Top panel: the same models and COPSS interpretations from Figure 3 shown in relation to our Y5 Pathfinder sensitivity forecast (blue shaded area). The legend also indicates the expected S/N with which we would reject the null hypothesis (i.e., excluding sample variance from the calculation). Bottom panel: S/N per k-bin of width $\Delta[\log(k\ \mathrm{Mpc})] \approx 0.16$ for each $P(k)$ prediction shown in the top panel, accounting for attenuation from line widths in the presence of a beam with FWHM of 4.′5 on sky.

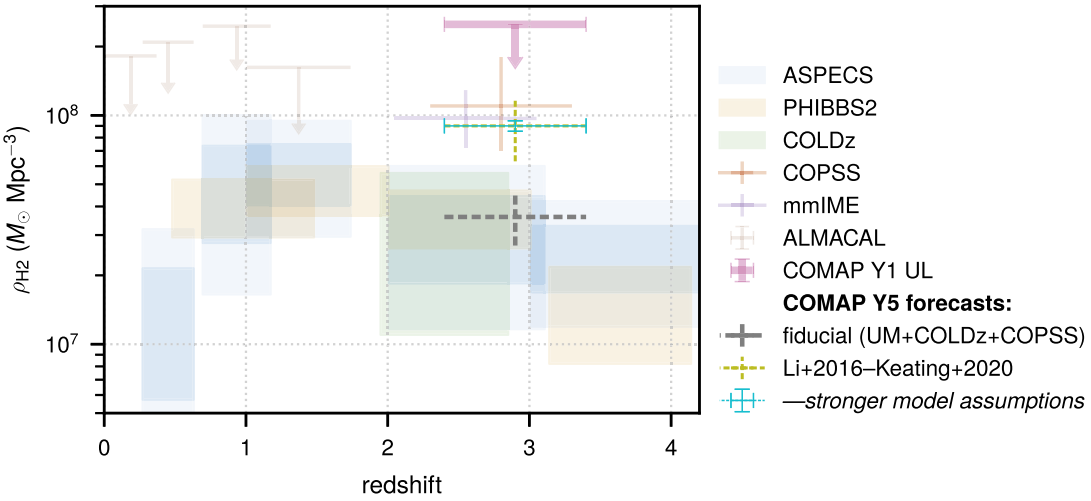


Figure 9. Forecast constraints on ρ_{H2} for COMAP Y5 power spectrum analysis, alongside the same constraints from Figure 7. We show forecasts for both the fiducial model in this work and a more optimistic model from Breysse et al. (2022a), which is more consistent with current LIM measurements. For the latter we also show a forecast constraint with stronger line bias priors, better mirroring highly informed analyses like that of Keating et al. (2020a).

This level of sensitivity will allow COMAP data to discriminate clearly between several of the models shown.

Molecular gas constraints. In Section 4.2 we rigorously consider how this detection, combined with characterization of the VID, will enable inference of model parameter constraints and of the CO LF. Before we do this, however, we consider a quick Fisher forecast of expected constraints on $\langle T \rangle$ and thus on ρ_{H2}.

In addition to our fiducial model, which, as we noted toward the end of Section 2.2, is a conservative estimate by the very nature of data-driven priors based on direct-detection measurements, we also consider the signal estimate derived from the empirical CO model of Keating et al. (2020a). This model, which we label "Li et al. (2016)–Keating et al. (2020a)" to distinguish it from the COPSS-based shot-noise estimate also calculated by Keating et al. (2020a), is also one of the primary models that Breysse et al. (2022a) use for COMAP forecasts beyond the Pathfinder. The model borrows the general approach of Li et al. (2016), which composes the simulation- and data-driven halo mass–SFR connection from Behroozi et al. (2013a, 2013b) with an empirical IR–CO luminosity fit but uses newer (albeit exclusively local) IR–CO correlation fits from Kamenetzky et al. (2016). The predicted CO (1–0) $\langle T \rangle$ at the COMAP central redshift of $z = 2.8$ is 1.3 μK, which is several times higher than our fiducial COLDz-driven conservative prediction of 0.5 μK, due to significant differences in the faint end of the $L(M_h)$ relation and thus the faint-end slope of the CO LF. Under this model, a Y5 power spectrum analysis would reject the null hypothesis at an all-k S/N of 17.

We run a Fisher analysis across the parameters $\{\langle T\rangle, b, P_{\rm shot}, v_{\rm eff}\}$, imposing loose Gaussian priors around the central line bias and $v_{\rm eff}$ values with width $\sigma[b] = 1$ and $\sigma[v_{\rm eff}] = 120\ {\rm km\ s^{-1}}$ (mostly to keep both away from negative values). The applicable central parameter values for our fiducial model are $\{0.52\ \mu{\rm K},\ 4.0,\ 1.9\times10^3\ \mu{\rm K^2\ Mpc^3},\ 330\ {\rm km\ s^{-1}}\}$, and those for the Li et al. (2016)–Keating et al. (2020a) model are $\{1.3\ \mu{\rm K},\ 2.7,\ 9.7\times10^2\ \mu{\rm K^2\ Mpc^3},\ 210\ {\rm km\ s^{-1}}\}$.

The forecast suggests that the primary parameter being constrained in this exercise is $\langle T\rangle$, with expectations of $0.52\pm0.14\ \mu{\rm K}$ for the conservative fiducial model and $1.3\pm0.4\ \mu{\rm K}$ for the more optimistic Li et al. (2016)–Keating et al. (2020a) model. Through the same conversion as we used in Section 3.3, these $\langle T\rangle$ constraints respectively translate into $\rho_{\rm H2}$ constraints of $(3.6\pm0.9)\times10^7\ M_\odot\ {\rm Mpc^{-3}}$ and $(9.0\pm2.7)\times10^7\ M_\odot\ {\rm Mpc^{-3}}$, as shown in Figure 9.

However, our priors around the CO model are fairly loose, whereas some real-world analyses like those of Keating et al. (2020a) or some Fisher analyses like those of Breysse et al. (2022a) make stronger assumptions about the shape of the $L(M_h)$ relation—which then completely determines at least the line bias—and constrain only the overall normalization of $L(M_h)$. In our Fisher forecast's parameter space this would be equivalent to imposing very narrow priors on b. If we keep the same prior width for $v_{\rm eff}$ but narrow the width for the bias prior to $\sigma[b] = 0.1$, we would obtain constraints around the Li et al. (2016)–Keating et al. (2020a) model of $\langle T\rangle = (1.3\pm0.07)\ \mu{\rm K}$ and $\rho_{\rm H2} = (9.0\pm0.5)\times;10^7\ M_\odot\ {\rm Mpc^{-3}}$. We show the latter also in Figure 9.

Finally, as these forecasts use the CO power spectrum alone, additional information from the VID and even from cross-correlations would further improve these constraints.

4.2. Simulated Inferences

In Section 2 we developed a new parameter set to describe the halo–CO connection, estimated a set of priors for these parameters, and discussed an accurate method to take into account the effect of CO line widths. Now also equipped with predictions for Y5 sensitivities, we can go on to the question of how we could use our model methods to infer constraints from the COMAP experiment.

Following Ihle et al. (2019), but using the model developed in Section 2, we run an MCMC inference from simulated data to forecast constraints on astrophysical observables like the LF, $\phi(L)$, as well as posterior distributions of our parameters from Section 2, $\theta = \{A, B, \log C, \log(M/M_\odot), \sigma\}$. This inference uses both the CO $P(k)$ and the VID in a joint analysis that accounts for covariance between all observables, as first considered by Ihle et al. (2019).

We focus here on the results of the simulated MCMC inference but provide further details on the MCMC setup, including the exact priors and survey parameters assumed, in Appendix C. Broadly speaking, the noise level assumed corresponds to Y5 sensitivity projections already discussed in Section 4.1, and the signal simulation uses the fiducial point estimate model (UM+COLDz+COPSS) defined in Table 1. The results shown here are from one MCMC run (i.e., one signal and noise realization) and will change somewhat from realization to realization.

Figure 10 shows the posterior distribution of all the individual model parameters resulting from one MCMC simulated inference run. Comparing the posterior (black curves) to the prior (green dashed curves), we see modest but clear shifts and tightening of the distributions. The simulated COMAP data constrain the power-law slopes, with 95% limits of $A < -2.1$ and $A > -0.78$, bounding $L(M_h)$ from above in both cases. The data also tighten the probability distributions projected in the σ–$\log(M/M_\odot)$ and in the σ–$\log C$ planes. This would appear to chiefly reflect information from the VID on the high-luminosity end of the LF, as the anticorrelation of σ with both $\log(M/M_\odot)$ and $\log C$ largely affects predicted abundances of CO emitters beyond the knee of the LF. Overall, the comparison between posterior and prior distributions shows that even when including COLDz and COPSS detections in the prior, COMAP improves the constraints on the model.

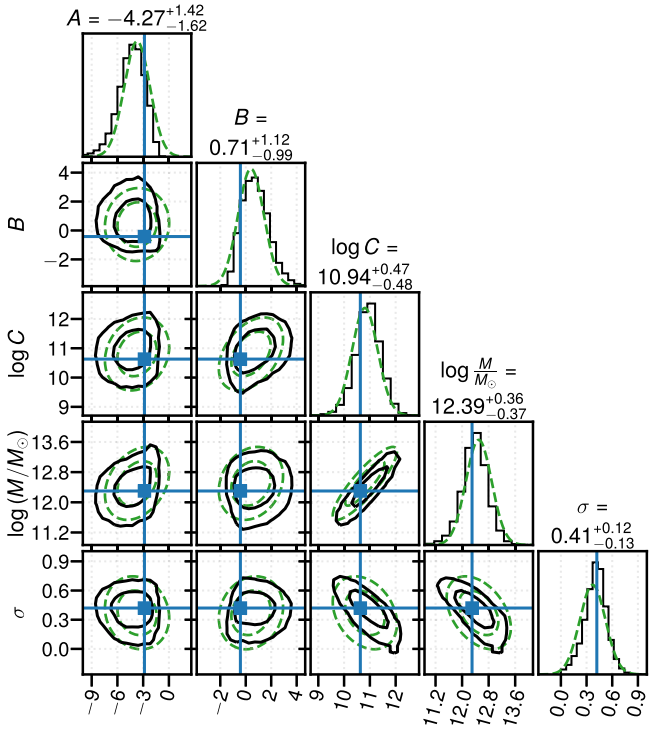


Figure 10. Forecasted posterior distributions for a single realization of the COMAP 5 yr experiment, given the UM+COLDz+COPSS model parameters from Table 1. Blue points in cross hairs show the model point estimates used for the simulated input signal. Black curves outline the posterior distributions, while the green dashed lines correspond to the (Gaussian) priors used for the MCMC. The two curves of each color correspond to 68% and 95% credibility regions, respectively. The numbers on top of each column correspond to the 68% credibility interval for each parameter.

The LF constraints, in Figure 11, show that even though the improvement of the parameter constraints appeared modest, the LF is significantly more constrained using COMAP compared to the prior (based on COLDz and COPSS), especially at the high-luminosity end. This in turn will correspond to significantly improved measurements of integrated and derived quantities like the previously discussed $\rho_{\rm H2}$.

4.3. HETDEX Cross-correlation Expectations

We have considered prospects for cross-correlation between CO intensity maps from COMAP and Lyα emitter (LAE) data from HETDEX in other works by Chung et al. (2019) and Silva et al. (2021). However, Chung et al. (2019) presented cross-spectrum forecasts well before we could characterize real-world performance of the COMAP Pathfinder instrument and data

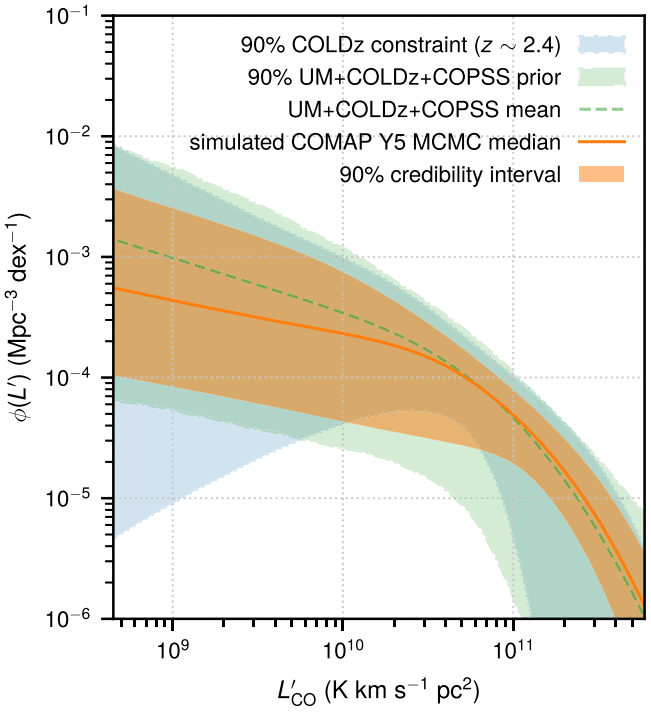


Figure 11. Constraints on the LF from the same single realization of the COMAP 5 yr experiment as used for Figure 10. The orange shaded area and solid curve represent the MCMC 90% credibility interval and median, respectively. We also show 90% intervals given by the COLDz constraints of Riechers et al. (2019) at $z \sim 2.4$ (shaded area bounded by dotted curves) and by the UM+COLDz+COPSS data-driven prior when applied to our $z \sim 2.8$ simulations (shaded area bounded by dashed–dotted curves). The green dashed line corresponds to the ensemble mean LF of the UM+COLDz+COPSS model producing the input signal.

pipeline, and Silva et al. (2021) consider a very detailed LAE model but solely in the context of voxel-level analyses.

Detailed models of the CO–LAE cross-spectrum are beyond the principal scope of our early science papers, which concern themselves with detection and interpretation of CO intensity mapping observations by themselves. However, we note that, based on our fiducial CO model and current expectations of LAE bias and number density, we expect to reach an all-k S/N of 7 on the CO–LAE cross-spectrum even with only Year 3 (Y3) data in hand for only Field 1 (whereas we will need all data through Y5 to achieve a similar S/N for the CO auto-spectrum). Fisher forecasts (in the style of, e.g., Breysse & Alexandroff 2019) suggest that even with relaxed priors on CO bias and line broadening compared to our assumptions from earlier, this Y3 single-field cross-correlation detection should allow for a constraint of $\langle Tb \rangle_{\rm CO} = 2.1 \pm 0.4$—a $\approx 5\sigma$ result. This contrasts with an upper limit of $\langle Tb \rangle_{\rm CO} < 5\ \mu$K with only the CO auto-$P(k)$ in the same field, or a marginal 2σ result of $\langle Tb \rangle_{\rm CO} = 2.1 \pm 1.0$ coadding auto-$P(k)$ measurements across all three fields.

The constraints from cross- and auto-$P(k)$ would respectively improve to $\langle Tb \rangle_{\rm CO} = 2.1 \pm 0.2$ and $\langle Tb \rangle_{\rm CO} = 2.1 \pm 0.5$ with the full three-field Y5 data and completely overlapping HETDEX LAE survey coverage in hand. That said, our forecasts suggest that HETDEX data would enable strong constraints on the CO clustering amplitude in advance of Y5, as we illustrate graphically in Figure 6 alongside current observational constraints. We refer the reader to Appendix D for further details on these simple forecasts.

Furthermore, as previous intensity mapping works have shown (Switzer et al. 2013; Keenan et al. 2022), cross-correlation constraints show strong robustness against systematics present in intensity mapping data. Whether this may relax our data selection requirements in the context of cross-correlation analyses will be the subject of future work, in which we also hope to mirror the more detailed LAE model of Silva et al. (2021) in larger cosmological simulations like the peak patch simulations used for Section 4.2.

5. Conclusions

This paper synthesizes model updates and early COMAP Pathfinder data to answer the following key questions:

1. *What inferences do our early science verification data enable about the $z \sim 3$ CO (1–0) power spectrum and molecular gas abundance?* Our current result of $\langle Tb \rangle^2 \lesssim 50\ \mu$K^2 already excludes certain models directly in the clustering regime and places a much stronger upper limit on the clustering amplitude of the CO power spectrum than COPSS. In addition, our upper limit is consistent with and readily complements existing constraints on $\rho_{\rm H2}$ at $z \sim 3$.
2. *Given early science sensitivities and updated $z \sim 3$ models, what are our present expectations for constraints on these same quantities, and others like the CO LF, at the end of 5 yr of COMAP Pathfinder observations?* We expect a detection of the $z \sim 3$ CO power spectrum to enable clear discrimination between different models from existing literature that predict different degrees of contribution of faint emitters to the total $P(k)$. For our conservative fiducial data-driven model we forecast an all-scale S/N of 9. Such a firm detection would also enable significant constraining power on the CO LF beyond our priors that conventional direct-detection surveys have not been able to offer, as well as a measurement of cosmic molecular gas abundance that will be a strong independent check on results from other surveys.

These promising early results are possible owing to the quality of the COMAP Pathfinder data at the present time, which are entirely consistent with uncorrelated white noise with any systematics successfully suppressed below white noise through data cuts. With further integration time we fully expect the COMAP Pathfinder to detect an excess power spectrum over white noise. The key question is whether this excess will be uncharacterized contamination or we will be able to attribute it to the CO signal we are targeting, which Pathfinder Y5 sensitivities should be sufficient to detect and even distinguish between many currently viable CO models as shown in Section 4. We can only be confident in the interpretation of such an excess through continued technical improvements, not only in mapmaking and power spectrum derivation but also in forward models of the signal.

With future work we also hope to present significant improvements not only in confidence of interpretation of the COMAP data but also in qualitative range of possible constraints through cross-correlation with external data sets, through both simple power spectrum cross-correlation and voxel-level analyses that will provide high information content around redshift evolution of CO emission and molecular gas content (Silva et al. 2021).

This material is based on work supported by the National Science Foundation under grant Nos. 1518282, 1517108, 1517598, 1517288, and 1910999; by the Keck Institute for Space Studies under "The First Billion Years: A Technical Development Program for Spectral Line Observations"; and by a seed grant from the Kavli Institute for Particle Astrophysics and Cosmology.

D.T.C. is supported by a CITA/Dunlap Institute postdoctoral fellowship. The Dunlap Institute is funded through an endowment established by the David Dunlap family and the University of Toronto. The University of Toronto operates on the traditional land of the Huron-Wendat, the Seneca, and, most recently, the Mississaugas of the Credit River; D.T.C. is grateful to have the opportunity to work on this land. P.C.B. is supported by the James Arthur Postdoctoral Fellowship at New York University. K.C., J.W.L., A.C.S.R., B.D.U., and D.P.W. acknowledge support from NSF awards 1518282 and 1910999. Work at the University of Oslo is supported by the Research Council of Norway through grants 251328 and 274990 and from the European Research Council (ERC) under the Horizon 2020 Research and Innovation Program (grant agreement No. 819478, COSMOGLOBE). Parts of the work were carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration, and funded through the internal Research and Technology Development program. H.P. acknowledges support from the Swiss National Science Foundation through Ambizione grant PZ00P2_179934. J.O.G. acknowledges support from the Keck Institute for Space Science, NSF AST-1517108, and the University of Miami. S.E.H. acknowledges support from an STFC Consolidated grant (ST/P000649/1). L.C.K. was supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 885990. J.K. is supported by a Robert A. Millikan Fellowship from Caltech.

We thank Riccardo Pavesi for access to the COLDz ABC posterior sample used in this work. This research made use of NASA's Astrophysics Data System Bibliographic Services. Some of the computing for this project was performed on the Sherlock cluster. D.T.C. would like to thank Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to these research results. The Scientific color maps `acton` and `tokyo` (Crameri 2021) are used in this study to prevent visual distortion of the data and exclusion of readers with color-vision deficiencies (Crameri et al. 2020).

Finally, we would like to thank an anonymous referee whose comments and suggestions significantly improved this manuscript.

Software: `hmf` (Murray et al. 2013); Matplotlib (Hunter 2007); `corner.py` (Foreman-Mackey 2016); Astropy, a community-developed core Python package for astronomy (Astropy Collaboration et al. 2013).

Appendix A
Details of CO Model Prior Formulation

Throughout this appendix, we examine potential ways to inform our CO model priors. First, we consider what information we can incorporate from the model papers of Li et al. (2016) and Behroozi et al. (2019); then, we consider $z \sim 2$–3 CO (1–0) observations in the past several years and how they can further refine our priors.

Note that for this section only we use a slightly different cosmology for consistency with Behroozi et al. (2019), which uses the cosmology of Planck Collaboration et al. (2016), such that $\Omega_m = 0.307$, $\Omega_\Lambda = 0.693$, $\Omega_b = 0.047$, $h = 0.678$, $\sigma_8 = 0.823$, and $n_s = 0.96$. Differences in cosmological quantities like $H(z)$ and comoving distance are around or less than 1% at COMAP redshifts, and while the higher Ω_m will likely result in a $\sim$10% difference in predicted halo abundances versus our fiducial cosmology, this is a much smaller relative uncertainty than many of our other model uncertainties, including the uncertainties surrounding some observational constraints.

A.1. Initial Prior Setup from Previous Models

The new parameters $\{A, B, C, M\}$ in the parameterization of Section 2.1 are expressible in terms of the parameters used in the scaling relations that have gone into this functional form (again, under the approximation of $\alpha \approx 1$ or at least $\alpha \ll 1$ and $\alpha \gg 1$):

$$A = 0.3\alpha_{\rm UM}/\alpha; \tag{A1}$$

$$B = 0.3\beta_{\rm UM}/\alpha; \tag{A2}$$

$$\log C = (10 - \log \delta_{\rm MF} - \beta + \log \epsilon)/\alpha; \tag{A3}$$

$$\log(M/M_\odot) = \log M_{200\,{\rm km/s}^{-1}} + (10/3)\log(V/200). \tag{A4}$$

Then, we can propagate through the above equations the priors on α, β, and $\delta_{\rm MF}$ from Li et al. (2016)—$\alpha = 1.17 \pm 0.37$, $\beta = 0.21 \pm 3.74$, and $\log \delta_{\rm MF} = 0.0 \pm 0.5$—and the 68% interval around the best-fit values of the other parameters from Behroozi et al. (2019). (We used the best-fit model from the Early Data Release; we do not consider the changes between this and the official Data Release 1 large enough to recalculate our priors.) The model of Behroozi et al. (2019) is redshift dependent, but here we fix $z = 2.4$, to match the median redshift of the COLDz survey. There should be relatively little evolution in CO abundances and thus the power spectrum between $z = 2.4$ and the COMAP central redshift of $z = 2.8$ (certainly little more than a factor of 2 or so, less than the current level of uncertainty in models of the signal).

The resulting initial priors on $\{A, B, C, M\}$ are

$$A = -1.66 \pm 2.33, \tag{A5}$$

$$B = 0.04 \pm 1.26, \tag{A6}$$

$$\log C = 10.25 \pm 5.29, \tag{A7}$$

$$\log(M/M_\odot) = 12.41 \pm 1.77. \tag{A8}$$

The central values for these priors do not change significantly across the COMAP redshift range, at least compared to the widths of the priors. We also set an initial prior of $\sigma = 0.4 \pm 0.2$ (dex), which takes the central value from the 0.37 dex total scatter in the Li et al. (2016) fiducial model and assumes a slightly broader prior than that model would have prescribed.

We consider several alternate sets of initial priors on the model parameters, depending on how confident we think we can be in various pieces of information in the literature. Thus we have, as in Table 3,

1. a conservative set of "flat," uninformative priors;
2. an informed set of priors (used for the fiducial model) deriving from empirical models of the galaxy–halo connection as described above;

Table 3
Initial Parameter Priors for the CO (1–0) Model

Prior Prefix	Initial Priors on:				
	A	B	$\log C$	$\log(M/M_\odot)$	σ
"flat"	$\mathcal{U}(-18, 9)$	$\mathcal{U}(-18, 9)$	$\mathcal{U}(5, 15)$	$\mathcal{U}(10, 15)$	$\mathcal{U}(0, 1)$
"UM"	$\mathcal{N}(-1.66, 2.33)$	$\mathcal{N}(0.04, 1.26)$	$\mathcal{N}(10.25, 5.29)$	$\mathcal{N}(12.41, 1.77)$	$\mathcal{N}(0.4, 0.2)$
"P18"	$\mathcal{N}(-2.29, 0.52)$	$\mathcal{N}(-0.57, 0.36)$	$\mathcal{N}(10.59, 0.70)$	$\mathcal{N}(11.79, 0.64)$	$\mathcal{N}(0.4, 0.2)$

3. and an extrapolation-heavy set of priors that derive from calculating the best-fit parameters and errors of the Padmanabhan (2018) model at $z = 2.4$ ("P18"), which builds in a range of $z = 0$–3 data including LF constraints from COPSS.

The names of these initial priors act as prefixes for our data-driven priors, as they represent information unconditioned on observational data.

A.2. Observational Constraints on High-redshift CO (1–0)

As reviewed by Carilli & Walter (2013), CO observations at high redshift in general are not especially novel, with hundreds of detections at $z \gtrsim 1$. However, a complication is that many of these detections—certainly the "main sequence" or "normal" star-forming galaxies surveyed by Daddi et al. (2010) and Tacconi et al. (2010)—are in CO (2–1) or CO (3–2) (if not higher-J CO lines), whereas we want to specifically consider CO (1–0) emission. In any case, we have already folded information from all the detections reviewed by Carilli & Walter (2013) owing to the fact that their values for α and β are one of four results incorporated into the Li et al. (2016) priors on these parameters.

While the CO LF was not constrained beyond $z = 0$ at the time of the Carilli & Walter (2013) review, several major projects have taken place to directly measure the CO LF at redshifts that COMAP will survey. We consider each of these and our rationale for incorporating or not incorporating them into our priors.

ASPECS.—As a molecular line scan survey, ASPECS searches for CO line emitters in a deep interferometric data cube without external preselection. The latest iteration is a Large Programme (LP) on ALMA (González-López et al. 2019) covering 4.6 arcmin^2—roughly five times the area of its pilot precursor (Walter et al. 2016)—and the observations in ALMA Band 3 (84–115 GHz) cover CO (3–2) emission at $z \sim 2.0$–3.1, as well as lower-J (or higher-J) CO lines at lower (or higher) redshift.

While ASPECS LP does constrain the CO LF at COMAP redshifts, we choose not to incorporate these results into our priors for the simple reason that the observations at COMAP redshifts are in CO (3–2) and not CO (1–0). Initial inferred CO (1–0) LF estimates presented by Decarli et al. (2019) relied on specific assumptions about CO line excitation, including a line luminosity ratio of $L'_{\mathrm{CO}(3-2)}/L'_{\mathrm{CO}(1-0)} = 0.42 \pm 0.07$ taken from Daddi et al. (2015), which averaged line ratios from three near-IR-selected "normal" star-forming galaxies at $z = 1.5$. While the uncertainties around this ratio were incorporated into the inference of the CO (1–0) LF, in hindsight the quoted uncertainties are severe underestimates of the probable error of the nominal value with respect to the global ratio at $z \sim 2.5$. Of the four CO (3–2) detections from González-López et al. (2019), three were observed robustly in CO (1–0) in Very Large Array (VLA) data by Riechers et al. (2020), and the line ratios were found to be closer to 0.8–1.1. Further work by Boogaard et al. (2020) yielded CO excitation models that favored an average line luminosity ratio of 0.80 ± 0.14—almost twice the original fiducial value used—that was then used for updated LF constraints by Decarli et al. (2020). The revised value resulted in estimates of luminosity densities and thus molecular gas abundances at roughly half of what was presented by Decarli et al. (2019).

Such significant changes in the presentation of the ASPECS LP results in the span of 2 yr strongly demonstrate both the uncertainty and possible variance in CO excitation across the population of high-redshift galaxies being surveyed. Due to this large uncertainty, we forgo using inferred constraints on $z \sim 2$–3 CO (1–0) from ASPECS.

COLDz.—The CO Luminosity Density at High-z (COLDz) survey (Pavesi et al. 2018) is also a molecular line survey, but it is in the COSMOS and GOODS-N fields and uses Ka-band VLA observations at $z = 2.0$–2.9, altogether covering almost 60 arcmin^2. The measurement is more directly applicable to our context, as it measures CO (1–0) line emission rather than a higher-J CO line. While the survey only identifies four secure (independently confirmed) line candidates across both fields at $z \sim 2$–3, the LF calculation also incorporates a catalog of line candidates that have not been independently confirmed, many of which do not have a spatially coincident counterpart in optical or near-infrared imagery.

The possibility of spurious line detections should in principle only discourage interpreting each line candidate individually (which Pavesi et al. 2018 explicitly do when presenting their nonsecure line candidates). However, the understanding of what line candidates should be considered "reliable" and which should not continues to evolve. For instance, in the case of ASPECS, between the pilot and large surveys, the requirement on the fidelity of a source (essentially the probability that the source is a genuine line detection rather than a noise peak) to be considered for analysis evolved from 60% to 90%. However, of the eight sources (across all CO lines and redshifts) identified by the pilot survey (Walter et al. 2016) in the overlapping area between the pilot and large surveys, only the four sources with identified optical or near-IR counterparts had above 90% fidelity (simply because having a counterpart meant that the fidelity was 100%). The other four sources had no counterparts, had below 90% fidelity, and were not recovered by the large survey. Therefore, whether 90% fidelity is a sufficient threshold to exclude spurious detections remains an open question.

Given the complexities in understanding which sources identified by a molecular line scan like ASPECS or COLDz are spurious, this might discourage using even statistical LF constraints from these surveys. However, it is worth noting that even if the ASPECS-Pilot analysis incorporated spurious sources as a significant fraction of its statistical sample, its $z \sim 2.6$ CO (3–2) LF measurement (Decarli et al. 2016) is

actually largely consistent with the ASPECS LP measurement (Decarli et al. 2019). Therefore, as purity (along with completeness and other various sources of error and uncertainty) is given due accounting in these analyses, we treat the COLDz measurement of the LF (Riechers et al. 2019) as a reliable one, even if not all of the individual sources in the statistical sample are individually reliable.

COPSS.—The work of Keating et al. (2016) represents the first attempt at a dedicated CO (1–0) LIM survey, targeting the same redshifts as COMAP. Following an analysis of Sunyaev–Zel'dovich Array (SZA) archival data (Keating et al. 2015), the same interferometer carried out observations specifically designed to measure the CO power spectrum at $z \sim 3$. The result was a constraint of $P(k) = (3.0 \pm 1.3) \times 10^3\ h^{-3}\ \mu\mathrm{K}^2\ \mathrm{Mpc}^3$, or $(8.7 \pm 3.8) \times 10^3\ \mu\mathrm{K}^2\ \mathrm{Mpc}^3$, at $k \sim 1\ h\ \mathrm{Mpc}^{-1} = 0.7\ \mathrm{Mpc}^{-1}$. Theoretical models, including our own, suggest that this should predominantly be a measurement of the shot-noise component of the power spectrum.

Keating et al. (2020a) recently reinterpreted the COPSS results to allow for the possibility that the clustering component contributes to the COPSS $P(k)$ value, reporting an estimate of $P_{\mathrm{shot}} = 2.0^{+1.1}_{-1.2} \times 10^3\ h^{-3}\ \mu\mathrm{K}^2\ \mathrm{Mpc}^3$. However, significant modification of P_{shot} away from the original COPSS value requires $\langle Tb \rangle^2 \gg 10\ \mu\mathrm{K}^2$, which we consider to be unlikely based on our models; we thus use the original constraint from Keating et al. (2016), rather than the revised constraint from Keating et al. (2020a).

mmIME.—The design of mmIME combines archival data and LIM observations on community instruments across a wide range of frequencies to probe CO line emission at high redshift, with Keating et al. (2020a) announcing results from ALMA observations. Using a combination of ASPECS data and ALMA Compact Array observations, Keating et al. (2020a) find a nonzero shot power, which they attribute to a combination of CO lines from different redshifts. Based on a CO model consistent with (although not constrained by) the total shot power measured, they expect CO (2–1) at $z \sim 1.3$ and CO (3–2) at $z \sim 2.5$ to contribute the bulk of this; using assumed line luminosity ratios (again from Daddi et al. 2015), the decomposition can be translated into an estimate of the CO (1–0) shot-noise power spectrum at $z \sim 2.5$.

We do not incorporate this measurement into our priors because, in addition to the complications reviewed previously surrounding CO line ratios and excitation, the mmIME estimate of CO at $z \sim 2.5$ relies on decomposing the total shot power appropriately into the contributions from different CO lines. Since Keating et al. (2020a) assume a specific model to do this, the $z \sim 2.5$ CO (1–0) P_{shot} estimate could change significantly depending on the model parameters; accounting for these additional uncertainties is beyond the scope of this work.

PHIBBS2.—The principal design of PHIBBS2 (Freundlich et al. 2019) is not as a molecular line scan survey, but as targeted observations of CO (2–1), CO (3–2), and CO (6–5) emission from $z = 0.5$–0.8, $z = 1$–1.6, and $z = 2$–3 "main sequence" star-forming galaxies. However, Lenkić et al. (2020) were able to identify serendipitous CO line emission from secondary sources in 110 observations of primary PHIBBS2 targets and constrain the CO LF across $z \sim 0.6$–3.6.

As with ASPECS, the measurements at COMAP redshifts are of CO (3–2) or higher-J CO lines. While we thus do not incorporate PHIBBS2 results into our priors either, we note that the ASPECS LP, COLDz, and PHIBBS2 results are all reasonably consistent with each other—at worst in slight tension—when translated to CO (1–0) LF constraints.

A.3. Data-driven Priors Constrained by Observational Results

To incorporate information from COLDz into our priors and thus generate refined "flat/UM/P18+COLDz" priors for each set of initial priors, we run an MCMC with initial priors on the five parameters $\{A, B, \log C, \log(M/M_\odot), \sigma\}$ as outlined above. At each step of the MCMC, we convert halo masses from a snapshot of the Bolshoi–Planck simulation (as used by Behroozi et al. 2019) at $a = 0.293560$ into CO luminosities given the sampled model parameters and calculate the resulting CO LF. Then, to determine the likelihood, we fit a Schechter function to the CO LF and compare the resulting Schechter parameter values to the posterior distribution of Schechter parameters from the COLDz approximate Bayesian computation (ABC). (In a minority of cases, the fitting procedure fails to produce a reasonable result; we find that including or excluding these cases does not significantly influence the posterior distribution).

We also run an MCMC using UM priors that incorporates the COPSS power spectrum measurement into the likelihood as well. This is done by calculating the expected shot-noise power spectrum from the LF as

$$P_{\mathrm{shot}} = \left[\underbrace{\frac{c^3 (1+z)^2}{8\pi k_B \nu_{\mathrm{rest}}^3 H(z)}}_{\equiv C_{\mathrm{LT}}} \right]^2 \int dL\ \frac{dn}{dL}\ L^2, \tag{A9}$$

which is then compared to the COPSS measurement of $(3.0 \pm 1.3) \times 10^3\ h^{-3}\ \mathrm{Mpc}^3\ \mu\mathrm{K}^2$. This UM+COLDz+COPSS MCMC will provide our fiducial data-driven prior.

The posterior distribution of this MCMC should then incorporate both our initial priors of Table 3 and the constraints from COLDz and COPSS. Thus, this distribution ("UM +COLDz+COPSS" in particular) should be a suitable prior distribution for COMAP analysis, and one that provides a new fiducial model for the CO (1–0) power spectrum at the COMAP redshifts.

In all MCMCs, we do not force $A < B$ while the chain is run, but we do apply the prior for A to the smaller of the two and the prior for B to the larger, and in analyzing the chain after completion, we always take the smaller value of the two at each sample to be A, and the larger to be B.

While the resulting posterior distributions are highly complex with all kinds of degeneracies, we show them in Figure 12. When using these as data-driven priors for COMAP analysis, we approximate them as multivariate Gaussian distributions based on the means and covariances.

Looking at the posterior distributions of the predicted LFs plotted in Figure 13, we find that they are largely consistent with COLDz constraints, which is exactly as expected. However, one quirk is that the LFs from our MCMC runs tend to have negative faint-end slope, whereas the COLDz constraints do not favor either negative or positive faint-end slope values. This is to be expected based on the fact that the procedure of Riechers et al. (2019) makes no assumptions about the CO emitters beyond the statistical sample from the survey, whereas we have the implicit assumption of the halo mass function, which approximately follows $dn/dM_h \sim M_h^{-2}$ at the low-mass end. Thus, at the faint end of the LF, we expect

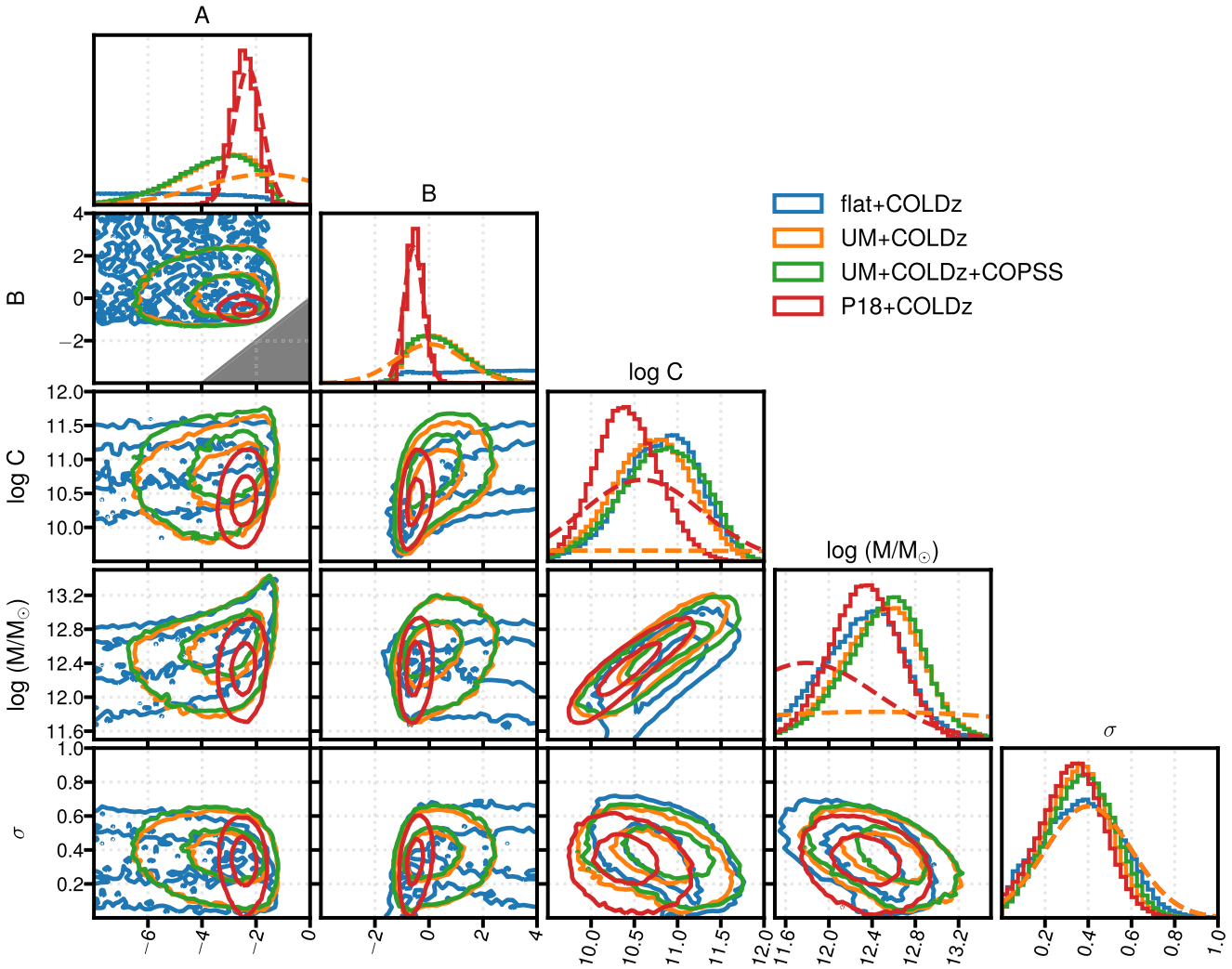


Figure 12. Parameter posterior (or data-driven prior) distributions from the MCMC combining our initial priors (dashed lines in marginalized posterior plots) with a likelihood based on the COLDz ABC constraints. Contours represent 39% and 86% mass levels, corresponding to 1σ and 2σ levels for 2D Gaussians. The legend indicates the colors for both initial and data-driven prior distribution curves. The dark-gray triangle in the A vs. B plot indicates the forbidden parameter space where $A > B$.

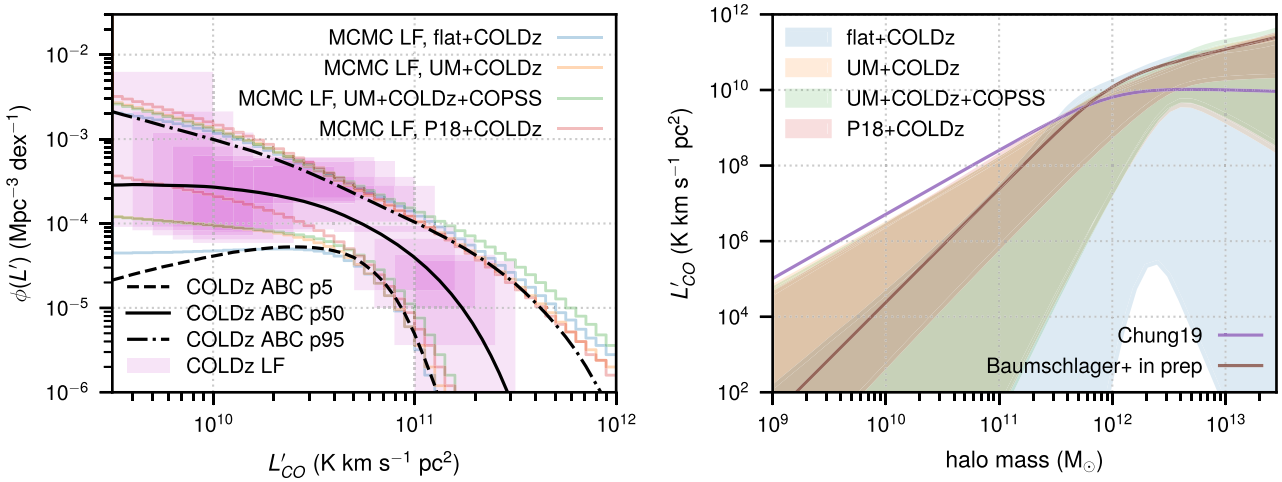


Figure 13. Left: $z \sim 2.4$ CO LF posterior distributions calculated from the MCMC results. We show 90% intervals for the MCMC (step plots indicated in legend), the COLDz direct constraints (shaded rectangles), and the COLDz ABC constraints on a Schechter LF (black dashed, solid, and dashed–dotted showing 5th, 50th, and 95th percentiles, respectively). Right: 90% intervals from each of the MCMC results for the $L(M_h)$ relation. For reference we also show $L(M_h)$ relations from Chung (2019)—which closely ties to the Li et al. (2016) model—as well as the `TNG300_2` model from B. Baumschlager et al. (2022, in preparation).

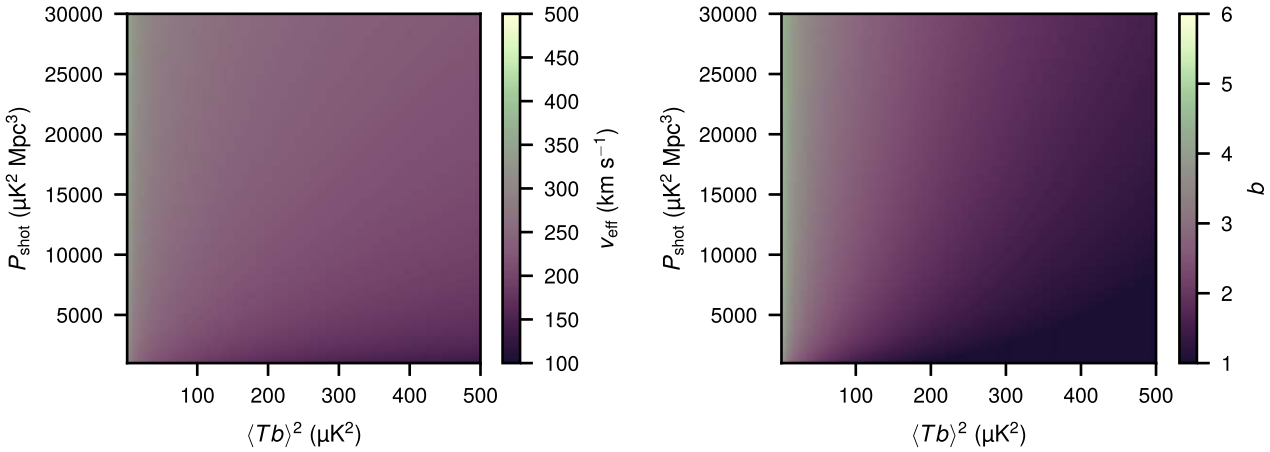


Figure 14. Polynomial fits of $v_{\rm eff}$ (left) and b (right) with respect to $\langle Tb\rangle^2$ and $P_{\rm shot}$, as detailed in the main text. Note the general trends of co-correlation of both variables with $P_{\rm shot}$, as well as anticorrelation of both with $\langle Tb\rangle^2$.

$\phi(L') = (dn/dM_h)/(d(\log L')/dM_h) \sim L^{1/A}$, and $A < 0$ being strongly favored means that a negative power-law slope at the faint end is also strongly favored.

The posterior distributions of the model parameters can be summarized as a posterior distribution of the $L(M_h)$ relation, as shown in the right panel of Figure 13. Our "flat+COLDz" prior-likelihood combination does not meaningfully constrain anything other than the turnabout scale, but the other data-driven priors tend to additionally favor a relatively flat bright-end slope and a faint-end power law in the M_h^2–M_h^6 range.

Appendix B
Average Values of Line Bias and Effective Line Width for the UM+COLDz and UM+COLDz+COPSS MCMC Posterior Distributions

We find that the behavior of b and $v_{\rm eff}$ with changing $\langle Tb\rangle$ and $P_{\rm shot}$ is relatively smooth across the UM+COLDz MCMC distribution. This allows us to devise the following fits:

$$\begin{aligned} b \approx & -2.85 - 2.30\log\left(\frac{\langle Tb\rangle}{\mu\mathrm{K}}\right) \\ & + 4.16\log\left(\frac{P_{\rm shot}}{\mu\mathrm{K}^2\,\mathrm{Mpc}^3}\right) + 2.18\log^2\left(\frac{\langle Tb\rangle}{\mu\mathrm{K}}\right) \\ & - 0.888\log^2\left(\frac{\langle Tb\rangle}{\mu\mathrm{K}}\right)\log\left(\frac{P_{\rm shot}}{\mu\mathrm{K}^2\,\mathrm{Mpc}^3}\right) \\ & + 0.0144\log^2\left(\frac{\langle Tb\rangle}{\mu\mathrm{K}}\right)\log^2\left(\frac{P_{\rm shot}}{\mu\mathrm{K}^2\,\mathrm{Mpc}^3}\right) \\ & - 0.557\log^2\left(\frac{P_{\rm shot}}{\mu\mathrm{K}^2\,\mathrm{Mpc}^3}\right) \\ & + 0.421\log\left(\frac{\langle Tb\rangle}{\mu\mathrm{K}}\right)\log^2\left(\frac{P_{\rm shot}}{\mu\mathrm{K}^2\,\mathrm{Mpc}^3}\right) \\ & - 1.42\log\left(\frac{\langle Tb\rangle}{\mu\mathrm{K}}\right)\log\left(\frac{P_{\rm shot}}{\mu\mathrm{K}^2\,\mathrm{Mpc}^3}\right); \end{aligned} \tag{B1}$$

$$\begin{aligned} \frac{v_{\rm eff}}{\mathrm{km\ s}^{-1}} \approx & 241. - 219.\log\left(\frac{\langle Tb\rangle}{\mu\mathrm{K}}\right) + 32.5\log\left(\frac{P_{\rm shot}}{\mu\mathrm{K}^2\,\mathrm{Mpc}^3}\right) \\ & + 21.8\log\left(\frac{\langle Tb\rangle}{\mu\mathrm{K}}\right)\log\left(\frac{P_{\rm shot}}{\mu\mathrm{K}^2\,\mathrm{Mpc}^3}\right). \end{aligned} \tag{B2}$$

Residuals versus these fits are mostly confined to 10%–30% relative error, against both the UM+COLDz and UM+COLDz +COPSS samples. This level of error is sufficient for our purposes given the large uncertainties associated with the observational data. We plot these fits in Figure 14, although note that the MCMC posterior samples only span parts of the parameter space being plotted, largely toward lower values of both $P_{\rm shot}$ and $\langle Tb\rangle$.

Appendix C
Details of MCMC Inference Simulations

C.1. Survey Simulations

We use a simplified COMAP experimental setup with a sensitivity corresponding roughly to the Y5 sensitivity forecast previously mentioned in Section 4.1. The experimental parameters are summarized in Table 4. We assume a uniform noise distribution in three cosmological fields each covering 4 deg^2, with 256 frequency bins covering the full 26–34 GHz range. Following Ihle et al. (2019), we choose a pixel size (4×4 arcmin^2) comparable to the instrumental beamwidth of 4.5 arcmin (FWHM), which gives us a 30×30 pixel grid for each field.

Our signal simulations are based on mock DM halo catalogs generated using the peak patch approach (Bond & Myers 1996; Stein et al. 2019). We associate CO luminosities with each of the DM halos using the model presented above. Luminosities are converted to equivalent brightness temperature and then separated by virial velocity before adding up the contributions to each voxel in a high-resolution comoving grid. The maps corresponding to the different virial velocity are convolved with the appropriate Gaussian line width, as discussed above, before they are added together and convolved with the angular

Table 4
Simplified COMAP Experimental Parameters for MCMC Simulated Inference

Parameter	Value
System temperature (K)	44
Number of feeds	19
Beam FWHM (arcmin)	4.5
Frequency band (GHz)	26–34
Channel width (MHz)	31.25
Number of fields	3
Field size ($\deg^2$)	4
Number of pixels per field	30×30
Noise per voxel[a] (μK)	17.8

Note.
[a] This value corresponds to the Y5 sensitivity forecast discussed at the start of Section 4.

beam. Finally, we degrade the map to the low resolution used for the main analysis.

We use 161 independent light cones, each covering 9.6×9.6 $\deg^2$, and divide them into smaller angular pieces to correspond to the size of our cosmological fields. This way we get a large number of semi-independent cosmological realizations to use for generating covariance matrices.

C.2. Observables and Covariances

Ihle et al. (2019) showed that using a combination of the power spectrum, $P(k)$, and the VID, $\mathcal{P}(T)$, is a good way to capture different parts of the information in a set of line intensity maps in an efficient manner. We use the same approach here.

The spherically averaged power spectrum, $P(k)$, is calculated from the (discrete) 3D Fourier components, $f_{\boldsymbol{k}}$, of the temperature map

$$P(k) = \frac{V_{\rm vox}}{N_{\rm vox}} \langle |f_{\boldsymbol{k}}|^2 \rangle \approx \frac{V_{\rm vox}}{N_{\rm vox} N_{\rm modes}} \sum_{j=1}^{N_{\rm modes}} |f_{\boldsymbol{k}_j}|^2 \equiv P_{k_i}, \quad \text{(C1)}$$

where $P_{k\rm i}$ is the estimated power spectrum in bin number i, $V_{\rm vox}$ is the voxel volume, $N_{\rm vox}$ is the total number of voxels in the map, and $N_{\rm modes}$ is the number of Fourier components with wavenumber $|\boldsymbol{k}_j| \approx k_i$ (i.e., in the bin corresponding to wavenumber $k = k_i$).

The most natural observable related to the VID, $\mathcal{P}(T)$, is the temperature bin count

$$\langle B_i \rangle = N_{\rm vox} \int_{T_i}^{T_{i+1}} \mathcal{P}(T) dT, \quad \text{(C2)}$$

where B_i is the number of voxels with a temperature in the ith temperature bin.

We combine both observables into a data vector

$$d_i = (P_{k_i}, B_i). \quad \text{(C3)}$$

If all the components of d_i were independent, they would have the following variance, which we denote as the *independent variance*:

$$\mathrm{Var}_{\rm ind}(P_{k_i}) = \langle P_{k_i} \rangle^2 / N_{\rm modes}, \quad \text{(C4)}$$

$$\mathrm{Var}_{\rm ind}(B_i) = \langle B_i \rangle. \quad \text{(C5)}$$

This assumes that the Fourier modes $f_{\boldsymbol{k}}$ of the maps are independent Gaussians and that the total number of voxels is much larger than $\langle B_i \rangle$.

Since there typically are correlations between the different elements of the data vector, we can take this into account using a full covariance matrix

$$\xi_{ij} = \mathrm{Cov}(d_i, d_j). \quad \text{(C6)}$$

We now have all the ingredients we need to build up a likelihood. We assume a Gaussian likelihood of the form (up to a constant)

$$-2 \ln P(d|\theta) = \frac{N_s}{N_s + 1} \sum_{ij} [d_i - \langle d_i \rangle] (\xi^{-1})_{ij} \times [d_j - \langle d_j \rangle] + \ln|\xi|, \quad \text{(C7)}$$

where $\langle d \rangle(\theta)$ and $\xi(\theta)$ are the mean values and covariance matrix of the observables d_i for specific parameters θ. N_s is the number of simulations used to estimate $\langle d \rangle$, and the factor $N_s/N_s + 1$ takes into account the effect of the uncertainty in the estimate of $\langle d \rangle$. We refer the reader to Ihle et al. (2019) for further details on the mock DM catalogs, the simulation procedure, and how the covariance matrices are estimated.

C.3. Mock MCMC Setup

The posterior distribution for our model parameters, $\theta = \{A, B, \log C, \log(M/M_\odot), \sigma\}$, is given by Bayes's theorem,

$$P(\theta|d) \propto P(d|\theta) P(\theta), \quad \text{(C8)}$$

where $P(\theta)$ is the prior on the model parameters, θ. As a prior we approximate the constraints derived in Section A.3 as a multidimensional Gaussian distribution, for computational efficiency. The prior parameters are given in Table 5.

To sample from the posterior distribution, we use the `emcee` package (Foreman-Mackey et al. 2013) implementing an affine-invariant ensemble MCMC with 60 walkers. As the "data," d, for the MCMC forecast we use a single (three-field) cosmological realization of the UM+COLDz+COPSS point estimate from Table 1, our default model. At each step in the MCMC we estimate the mean observables, $\langle d \rangle(\theta)$, using 10 simulations in order to evaluate the posterior at the current point in parameter space. We estimate the mean CO LF from the 10 signal realizations each step in the MCMC and use this as an estimate of the LF at this point in parameter space. This way we obtain a large number of samples of the LF sampled according to the posterior distribution of the model parameters, giving us a simple way to get posterior constraints on the LF.

We use a burn-in period of 500 samples out of 5940 and treat the subsequent samples as valid samples from the posterior. We use the Gelman–Rubin and Geweke convergence diagnostics as implemented by `ChainConsumer` (Hinton 2016), which both suggest the MCMC converges. The Geweke statistic in particular suggests that we could even shorten the burn-in period considerably to the first 30 samples, but we do not presume that these statistics have great sensitivity in identifying chain convergence, although they may have great specificity. Therefore, regardless of the convergence statistics, we take a conservative approach and continue to treat the first 500 MCMC samples as the burn-in period.

Table 5
Mean and Covariance Matrix for Gaussianized UM+COLDz+COPSS Priors on Model Parameters

Parameter	Mean	Covariance Matrix				
		A	B	$\log C$	$\log \frac{M}{M_\odot}$	σ
A	−3.71	2.26	0.0651	0.143	0.185	−0.0305
B	0.41	0.0651	1.03	0.207	0.0805	0.0239
$\log C$	10.8	0.143	0.207	0.251	0.17	−0.0227
$\log \frac{M}{M_\odot}$	12.5	0.185	0.0805	0.17	0.151	−0.0243
σ	0.371	−0.0305	0.0239	−0.0227	−0.0243	0.0228

Appendix D
Details of COMAP–HETDEX Fisher Forecasts

The quantitative forecasts of Chung et al. (2019) for COMAP–HETDEX cross-spectra are somewhat esoteric at the present time for several reasons:

1. COMAP Pathfinder parameters such as field sizes and observation efficiencies have substantially evolved.
2. The LAE model, while in principle matching LAE counts in the literature, did not correctly account for Lyα emission duty cycles of high-redshift galaxies, effectively setting the LAE fraction to 100%. Thus, the model of Chung et al. (2019) overestimates the mean CO luminosity of LAE samples and thus overestimates the cross shot noise, while also overestimating the LAE bias.
3. The forecast S/N values were never translated to parameter constraints.

It is entirely out of scope for this paper to provide a full-fledged Lyα emission model in order to forecast COMAP–HETDEX cross-correlation analyses. We can, however, forecast the observable auto- and cross-spectra without having to devise such a model.

D.1. Observables and Parameters

Adapting Breysse & Alexandroff (2019), in real (comoving) space we would have

$$P_{\rm CO}(k) = \langle Tb\rangle_{\rm CO}^2 P_m(k) + P_{\rm shot,CO}; \tag{D1}$$

$$P_{\rm LAE}(k) = b_{\rm LAE}^2 P_m(k) + \bar{n}_{\rm LAE}^{-1}; \tag{D2}$$

$$P_{\rm CO\times LAE}(k) = \langle Tb\rangle_{\rm CO} b_{\rm LAE} P_m(k) + P_{\rm shot,CO\times LAE}. \tag{D3}$$

Normally forward models would link parameters like our fiducial model's $\{A, B, C, M, \sigma\}$ to quantities like $\langle Tb\rangle_{\rm CO}$ and $P_{\rm shot}$. However, our approach for these simple forecasts will focus on directly constraining the "derived" quantities, much as we directly constrained $\langle Tb\rangle$ and $P_{\rm shot}$ with the early science data.

Apart from the matter power spectrum $P_m(k)$, five quantities fully define the real-space power spectra: the CO clustering amplitude $\langle Tb\rangle_{\rm CO}$, the CO shot noise $P_{\rm shot,CO}$, the LAE bias $b_{\rm LAE}$, the LAE number density $\bar{n}_{\rm LAE}$, and finally the cross shot noise $P_{\rm shot,CO\times LAE}$, which encodes the mean CO luminosity of the LAE population:

$$\begin{aligned} P_{\rm shot,CO\times LAE} &= C_{\rm LT}\langle L_{\rm CO|LAE}\rangle \\ &= \frac{C_{\rm LT}}{\bar{n}_{\rm LAE}} \int L_{\rm CO} \left.\frac{dn}{dL_{\rm CO}}\right|_{\rm LAE} dL_{\rm CO}, \end{aligned} \tag{D4}$$

where $C_{\rm LT}$ is the same as defined in Equation (A9).

We can define fiducial values easily for four of the five quantities. For CO, given the UM+COLDz+COPSS point model (which is our fiducial model) we have $\langle Tb\rangle_{\rm CO} = 2\ \mu$K (with $b_{\rm CO} = 4$) and $P_{\rm shot,CO} = 1.9\times10^3\ \mu{\rm K}^2\,{\rm Mpc}^3$. For HETDEX, references in Gebhardt et al. (2021) suggest $b_{\rm LAE} = 1.8$–2.2, so we can take a central value of 2. Taking into account the fact that HETDEX sparsely samples most of its survey footprint at 1/4.6-fill, Gebhardt et al. (2021) quote an expected number density of $\bar{n}_{\rm LAE} = 1.1\times10^5\ {\rm Gpc}^{-3}$.

For the cross shot noise, we assume an LAE fraction and give our best estimate for the mean CO luminosity for the LAE population based on that assumption. Using various subsamples of UV-selected galaxies and deep data from the Multi-Unit Spectroscopic Explorer (MUSE), Kusakabe et al. (2020) found the LAE fraction to range between 4% and 30%, rising weakly with redshift and not evolving significantly with absolute rest-frame UV magnitude. Based on the LAE fractions found in their Table 1, we consider $X_{\rm LAE} = 0.05$ a reasonable LAE fraction to assume.

Proceeding from this assumption that 5% of galaxies are LAEs, with no significant dependence on UV luminosity—and thus, one might assume, no significant dependence on SFR or halo mass—we assume that 5% of DM halos host LAEs. To reach a number density of $\bar{n}_{\rm LAE} = 1.1\times10^5\ {\rm Gpc}^{-3}$ with 1/4.6-fill, which is to say $\bar{n}_{\rm LAE} = 5\times10^5\ {\rm Gpc}^{-3}$ without sparse sampling, the number density of halos that could *ever* possibly host an LAE would have to be $\bar{n}_{\rm LAE}/X_{\rm LAE} = 10^7\ {\rm Gpc}^{-3}$. This number density can be achieved by selecting all DM halos with a halo mass above $M_h > 9.3\times10^{10}\,M_\odot$, a halo population with an average halo bias of 2.2 roughly consistent with the $b_{\rm LAE}$ central expectation. For this population the average CO luminosity is $4.9\times10^4\,L_\odot$ under our fiducial model, which we multiply by $C_{\rm LT}$ to obtain an estimate of $P_{\rm shot,CO\times LAE} = 51\ \mu{\rm K}\,{\rm Mpc}^3$.

However, the redshift-space observables are more complicated than the real-space power spectra, as the CO intensity field is subject to line broadening and large-scale redshift-space distortions manifest in all observed clustering. As in the early science analyses in the main text, here we are only concerned with the simply spherically averaged monopole power spectrum, so we approximate the effect of line broadening with a single Gaussian filter with associated effective line width $v_{\rm eff}$. On top of these astrophysical redshift-space effects, we must also consider the COMAP transfer function and the VIRUS instrumental resolution.

Although the early science analysis in the main text uses debiased sensitivities already corrected for the transfer function, for this section we compare the uncorrected signals against raw sensitivities. We can approximate the transfer function for COMAP CES data, in the space of the transverse and line-of-sight wavevector components $k_\perp$ and $k_\parallel$, as a

combination of sigmoid and Gaussian functions:

$$\mathcal{T}(k_\perp, k_\parallel) = \frac{0.4\exp[-(3.5\ \mathrm{Mpc}\cdot k_\perp)^2]\exp[-(1.8\ \mathrm{Mpc}\cdot k_\parallel)^2]}{[1+\exp(5-100\ \mathrm{Mpc}\cdot k_\perp)][1+\exp(5-200\ \mathrm{Mpc}\cdot k_\parallel)]}. \quad \text{(D5)}$$

We can combine this with an estimated noise power spectrum of $P_N \sim 10^6\ \mu\mathrm{K}^2\ \mathrm{Mpc}^3$ for Field 1 CES data, as well as Fourier mode counts expected from a $60\times60\times256$ voxel grid spanning $2^\circ\times2^\circ\times8$ GHz in angular and spatial extent (mirroring the actual COMAP pixelization in all dimensions). The resulting noise limit and transfer function are both within 1/3 of ground truth across a majority of the range of k values. Scaling the resulting noise limit (P_N divided by the number of modes) down by 69 to obtain our Y5 sensitivity estimate and applying the approximate $\mathcal{T}$ and line broadening to the fiducial CO model power spectrum, we obtain a forecast all-k S/N of 8 for the uncorrected signal versus the raw sensitivity. This estimate is not at all far from the value of 9 forecast in the main text comparing the debiased noise limit against a line-broadened signal.

Working from Chung (2019), we can define the pseudo-power auto- and cross-spectra as functions of $(k_\perp, k_\parallel)$:

$$\tilde{P}_{\mathrm{CO}}(k_\perp, k_\parallel) = \mathcal{T}(k_\perp, k_\parallel)\exp(-k_\parallel^2\sigma_{\mathrm{eff}}^2) \times [\langle Tb\rangle_{\mathrm{CO}}^2(1+\beta_{\mathrm{CO}}k_\parallel^2/(k_\parallel^2+k_\perp^2))^2P_m(k)+P_{\mathrm{shot,CO}}]; \quad \text{(D6)}$$

$$\tilde{P}_{\mathrm{LAE}}(k_\perp, k_\parallel) = \exp(-k_\parallel^2\sigma_{\mathrm{VIRUS}}^2) \times [(b_{\mathrm{LAE}}+fk_\parallel^2/(k_\parallel^2+k_\perp^2))^2P_m(k)+\bar{n}_{\mathrm{LAE}}^{-1}]; \quad \text{(D7)}$$

$$\tilde{P}_{\mathrm{CO}\times\mathrm{LAE}}(k_\perp, k_\parallel) = \mathcal{T}^{1/2}(k_\perp, k_\parallel)\exp\left(-\frac{k_\parallel^2(\sigma_{\mathrm{eff}}^2+\sigma_{\mathrm{VIRUS}}^2)}{2}\right) \times [\langle Tb\rangle_{\mathrm{CO}}(1+\beta_{\mathrm{CO}}k_\parallel^2/(k_\parallel^2+k_\perp^2))(b_{\mathrm{LAE}} + fk_\parallel^2/(k_\parallel^2+k_\perp^2))P_m(k)+P_{\mathrm{shot,CO}\times\mathrm{LAE}}]. \quad \text{(D8)}$$

Note that we define σ_{eff} in terms of v_{eff}:

$$\sigma_{\mathrm{eff}} = \frac{(1+z)}{H(z)}\frac{v_{\mathrm{eff}}}{\sqrt{8\ln2}}. \quad \text{(D9)}$$

We may then average the pseudo-power spectra in cylindrical k-space into k-bins, weighting by inverse variance, to yield spherically averaged pseudo-signal spectra $\tilde{P}_i(k)$. For the biased signals this is equivalent to simply weighting by the mode count in each pair of $k_\perp$- and $k_\parallel$-bins, thus ending up with a simple arithmetic average over some number of modes $N_m(k)$ for each k-bin. Since the field being autocorrelated or Fourier transformed is entirely real, only half the Fourier modes are independent; we count modes so that $N_m(k)$ already includes this halving.

D.2. Fisher Forecasts

For a Fisher analysis in the style of Breysse & Alexandroff (2019), we first define the covariance matrix of the (distorted) CO intensity and LAE overdensity fields:

$$C_{ij}(k) = \begin{bmatrix} \tilde{P}_{\mathrm{CO}}(k)+P_N & \tilde{P}_{\mathrm{CO}\times\mathrm{LAE}}(k) \\ \tilde{P}_{\mathrm{CO}\times\mathrm{LAE}}(k) & \tilde{P}_{\mathrm{LAE}}(k) \end{bmatrix}. \quad \text{(D10)}$$

Unlike Breysse & Alexandroff (2019), we should not divide P_N by any kind of window function, as we already fold applicable transfer functions into $\tilde{P}_i(k)$.

To forecast constraints for parameters $\{x_i\}$, we would obtain the covariance of those parameters by inverting the Fisher matrix:

$$F_{ij} = \sum_k N_m(k)\mathrm{Tr}\left[\frac{\partial C}{\partial x_i}C^{-1}\frac{\partial C}{\partial x_j}C^{-1}\right]. \quad \text{(D11)}$$

Note that we discard the factor of 1/2 from Equation (13) of Breysse & Alexandroff (2019), as we use $N_m(k)$ to denote the number of *independent* Fourier modes in each k-bin, whereas Breysse & Alexandroff (2019) appear to write Equation (13) under the assumption that N_m denotes twice this. Corroborating this is the fact that Equation (12) of Chung et al. (2019) divides by $2N_m(k)$—using an equivalent definition of $N_m(k)$ as in this section—when calculating cross-spectrum variance, but Equation (19) of Breysse & Alexandroff (2019) divides only by $\sqrt{N_m(k)}$ when calculating cross-spectrum error.

We project constraints for six parameters: $\langle Tb\rangle_{\mathrm{CO}}$ (in units of μK), β_{CO}, $p_{\mathrm{shot,CO}} \equiv P_{\mathrm{shot,CO}}/(10^3\ \mu\mathrm{K}^2\ \mathrm{Mpc}^3)$, v_{eff} (in units of $\mathrm{km\ s^{-1}}$), b_{LAE}, and $p_{\mathrm{shot,CO}\times\mathrm{LAE}} \equiv P_{\mathrm{shot,CO}\times\mathrm{LAE}}/(10^2\ \mu\mathrm{K\ Mpc}^3)$. We do not attempt to project constraints on $\bar{n}_{\mathrm{LAE}}$, as the HETDEX data by themselves will constrain this extremely finely. We also note the addition of two parameters not in our real-space parameter set: β_{CO}, which must be defined separately from $\langle Tb\rangle_{\mathrm{CO}}$ to fully describe redshift-space distortions; and v_{eff}, which we use to describe line broadening. We base the central values for these parameters on our fiducial model and thus obtain fiducial values for all six parameters in our Fisher forecast:

$$\langle Tb\rangle_{\mathrm{CO}} = 2.1\ [\mu\mathrm{K}]; \quad \text{(D12)}$$

$$\beta_{\mathrm{CO}} = 0.24; \quad \text{(D13)}$$

$$p_{\mathrm{shot,CO}} = 1.9; \quad \text{(D14)}$$

$$v_{\mathrm{eff}} = 330\ [\mathrm{km\ s^{-1}}]; \quad \text{(D15)}$$

$$b_{\mathrm{LAE}} = 2; \quad \text{(D16)}$$

$$p_{\mathrm{shot,CO}\times\mathrm{LAE}} = 0.51. \quad \text{(D17)}$$

We also continue to assume $\bar{n}_{\mathrm{LAE}} = 1.1\times10^{-4}\ \mathrm{Mpc}^{-3}$ to calculate $\tilde{P}_{\mathrm{LAE}}(k)$.

Since a significant degeneracy exists between $\langle Tb\rangle_{\mathrm{CO}}$ and β_{CO}, and also between $p_{\mathrm{shot,CO}}$ and v_{eff}, we impose Gaussian prior distributions of $\beta_{\mathrm{CO}} = 0.24\pm0.15$ and $v_{\mathrm{eff}} = 330\pm165$ (again, in units of $\mathrm{km\ s^{-1}}$). These priors are conservative; the prior width on v_{eff} is defined to keep $v_{\mathrm{eff}} > 0$ in the vast majority of cases, while the prior width on β_{CO} derives from the main text's assertion that $b_{\mathrm{CO}} > 2$.

We run Fisher analyses for two survey sensitivity scenarios:

1. *Y3:* We assume that, at minimum, by Y3 we will have a (sparsely sampled) HETDEX LAE catalog that covers Field 1, and that we will have accumulated 15 times the integration time that we currently have in this field (entirely consistent with our Y5 forecast). This means we scale the estimated current $P_N = 10^6\ \mu\mathrm{K}^2\ \mathrm{Mpc}^3$ down by a factor of 15 but assume we only have Fourier modes available in this one field for cross-correlation. For the CO auto-spectrum we consider sensitivities for both one field and for all three fields, keeping P_N the same but tripling the number of modes available in the survey for the latter case.

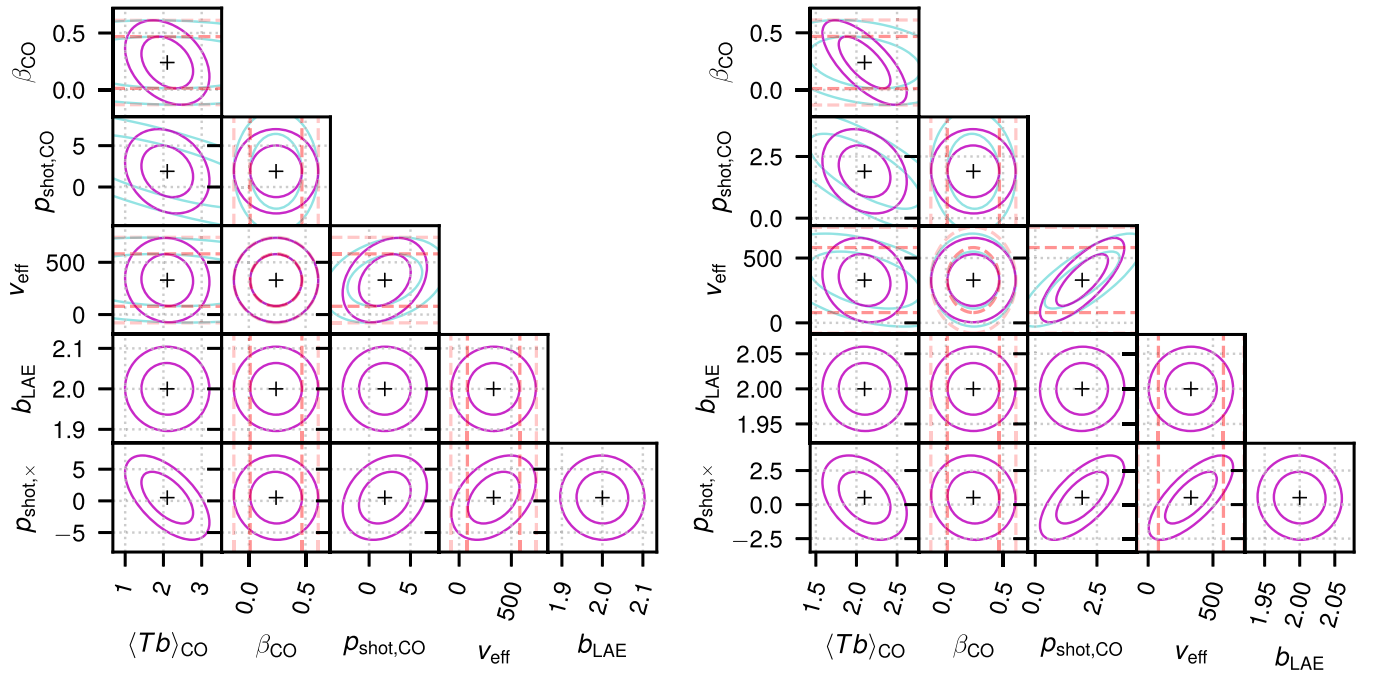


Figure 15. 68% and 95% ellipses from the Fisher analyses described in the text for the Y3 one-field (left) and Y5 three-field (right) scenarios. Parameters are dimensionless except $\langle Tb\rangle_{CO}$ (in units of μK) and v_{eff} (in units of km s^{-1}). Faint cyan ellipses show constraints expected from the CO auto-spectrum only, while the solid magenta ellipses show joint constraints expected from CO and LAE data. We also show priors for β_{CO} and v_{eff} (red dashed) applied in the Fisher analyses.

Table 6
Fisher Forecasts for Clustering Constraints from COMAP Auto- and COMAP–HETDEX Cross-spectra

		$\langle Tb\rangle_{CO}/\sigma[\langle Tb\rangle_{CO}]$				
Model	$\langle Tb\rangle_{CO}$	COMAP Y3 ×HETDEX LAE (One Field)	COMAP Y3 Auto (One Field)	COMAP Y3 Auto (Three Fields)	COMAP Y5 ×HETDEX LAE (Three Fields)	COMAP Y5 Auto (Three Fields)
UM+COLDz+COPSS	2.1 μK	4.7	1.2	2.0	10.	4.6
Li+2016–Keating+2020	3.5 μK	8.0	3.7	6.1	13.	12.

Note. The model labels "UM+COLDz+COPSS" and "Li+2016–Keating+2020" respectively denote the fiducial model from this work (derived from the namesake data-driven prior) and the Li et al. (2016)–Keating et al. (2020a) model, both discussed in the main text.

2. *Y5:* We assume that we integrate deep enough to achieve a noise power spectrum of $10^6/40 = 2.5 \times 10^4\ \mu\mathrm{K}^2\,\mathrm{Mpc}^3$ in all fields and that a sparsely sampled HETDEX LAE catalog covers all fields as well. This scenario is designed such that the net auto-spectrum sensitivity gain of $40\sqrt{3} = 69$ relative to the current Field 1 limit is consistent with the improvement forecast for Y5 in the main text.

The assumed HETDEX LAE data availability for Y3 and Y5 does not reflect potential proprietary periods for HETDEX data before they are shared with either the general community or the COMAP collaboration specifically. However, given the expectation of full-fill sampling of the HETDEX zero-decl. field (which overlaps with COMAP Field 1) and the current estimated HETDEX survey completion date of 2024 quoted by Gebhardt et al. (2021), we believe we have a reasonable guess of how much data HETDEX would have available internally.

We show the resulting error ellipses in Figure 15. While we only resolve some parameter degeneracies through priors, note that we significantly reduce the degeneracy between $\langle Tb\rangle_{CO}$ and $p_{shot,CO}$—which is to say that we can better disambiguate CO clustering from CO shot noise—through cross-correlation.

The main parameter we would meaningfully constrain in the Y3 scenarios is $\langle Tb\rangle_{CO}$, so we tabulate the constraining power in Table 6 as the ratio between the central $\langle Tb\rangle_{CO}$ value and the Fisher forecast uncertainty $\sigma[\langle Tb\rangle_{CO}]$. From Field 1 data alone we expect $\sigma[\langle Tb\rangle_{CO}] = 0.445$ from a joint analysis of the CO auto- and CO–LAE cross-spectra (the latter detected with an all-k S/N of 7). This would be a significant improvement over the early science result of $\langle Tb\rangle^2_{CO} < 51\ \mu$K (or $\langle Tb\rangle_{CO} < 7\ \mu$K) and would still be better than an analysis of the Field 1 CO auto-spectrum by itself, which would only yield an upper limit of $\langle Tb\rangle_{CO} < 5\ \mu$K. (Note that the priors on v_{eff} and line bias applied in the Fisher forecasts are much looser than the axiomatic assumptions applied in the main text's analysis, so we should not expect this forecast upper limit to improve on the early science result by a factor of $\sqrt{15} \approx 4$.) Even if we multiply the number of modes available by a factor of three to simulate an auto-spectrum-based constraint that uses data from all three COMAP fields, we forecast a marginal 2σ result, as the predicted uncertainty is $\sigma[\langle Tb\rangle_{CO}] = 1.04$.

With Y5 data in hand, the CO–LAE cross-spectra detection should continue to improve to an all-k S/N of 19, and the uncertainty on the CO clustering amplitude from joint analysis of auto- and cross-spectra should also improve by approximately a factor of 2 to $\sigma[\langle Tb\rangle_{\rm CO}] = 0.209$. The CO auto-spectrum alone will now be securely detected, as we have forecast previously in this work, and achieve $\sigma[\langle Tb\rangle_{\rm CO}] = 0.455$, on par with the Y3 Field 1 cross-correlation result.

We also repeat these analyses for the Li et al. (2016)–Keating et al. (2020a) model that we considered in Section 4.1, recalculating central values for all parameters except $b_{\rm LAE}$:

$$\langle Tb\rangle_{\rm CO} = 3.5\ [\mu{\rm K}]; \quad \text{(D18)}$$

$$\beta_{\rm CO} = 0.36; \quad \text{(D19)}$$

$$p_{\rm shot,CO} = 0.97; \quad \text{(D20)}$$

$$v_{\rm eff} = 210\ [{\rm km\ s^{-1}}]; \quad \text{(D21)}$$

$$p_{\rm shot,CO\times LAE} = 1.06. \quad \text{(D22)}$$

We continue to use Gaussian priors for β and $v_{\rm eff}$ with the same widths, though the central values are changed. The stronger auto-spectrum detection forecast for this model, as forecast previously in Section 4.1, means that Y5 results are similar between auto- and cross-spectrum analyses. However, cross-correlation still provides a significant advantage in intermediate stages of data acquisition, as we show in tabulations alongside the fiducial predictions in Table 6. We also recall the point raised toward the end of Section 4.3 about the advantages of LIM–galaxy cross-correlation against systematics as discussed by other works (e.g., Switzer et al. 2013; Keenan et al. 2022).

In all cases, HETDEX observations beyond sparse sampling that fully fill in all COMAP patches would benefit S/N by lowering HETDEX shot noise. The quantitative improvement would depend on relative contribution of HETDEX shot noise versus COMAP thermal noise to the uncertainties, but the improvement predicted by Chung et al. (2019) was around 50%.

We do not detect the cross shot noise in any of the scenarios considered above. However, even if the cross-spectrum yields only an upper limit on the mean CO luminosity of LAEs, this can be combined with voxel-level analyses as proposed by Silva et al. (2021) and should still provide key insights into galaxy and IGM properties at $z \sim 3$.

ORCID iDs

Dongwoo T. Chung https://orcid.org/0000-0003-2618-6504
Patrick C. Breysse https://orcid.org/0000-0001-8382-5275
Kieran A. Cleary https://orcid.org/0000-0002-8214-8265
Håvard T. Ihle https://orcid.org/0000-0003-3420-7766
Hamsa Padmanabhan https://orcid.org/0000-0002-8800-5740
Marta B. Silva https://orcid.org/0000-0003-0209-4816
J. Richard Bond https://orcid.org/0000-0003-2358-9949
Delaney A. Dunne https://orcid.org/0000-0002-5223-8315
Hans Kristian Eriksen https://orcid.org/0000-0003-2332-5281
Marie Kristine Foss https://orcid.org/0000-0001-8896-3159
Stuart E. Harper https://orcid.org/0000-0001-7911-5553
Andrew I. Harris https://orcid.org/0000-0001-6159-9174
Brandon Hensley https://orcid.org/0000-0001-7449-4638
Laura C. Keating https://orcid.org/0000-0001-5211-1958
Junhan Kim https://orcid.org/0000-0002-4274-9373
Timothy J. Pearson https://orcid.org/0000-0001-5213-6231
Liju Philip https://orcid.org/0000-0001-7612-2379
Anthony C. S. Readhead https://orcid.org/0000-0001-9152-961X
Thomas J. Rennie https://orcid.org/0000-0002-1667-3897
Nils-Ole Stutzer https://orcid.org/0000-0001-5301-1377
Bade D. Uzgil https://orcid.org/0000-0001-8526-3464
Marco P. Viero https://orcid.org/0000-0003-0545-4872
Duncan J. Watts https://orcid.org/0000-0002-5437-6121
Risa H. Wechsler https://orcid.org/0000-0003-2229-011X
Ingunn Kathrine Wehus https://orcid.org/0000-0003-3821-7275

References

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013a, ApJL, 762, L31
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013b, ApJ, 770, 57
Bonato, M., Liuzzo, E., Giannetti, A., et al. 2018, MNRAS, 478, 1512
Bond, J. R., & Myers, S. T. 1996, ApJS, 103, 1
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., Chung, D. T., Cleary, K. A., et al. 2022a, ApJ, 933, 188
Breysse, P. C., Yang, S., Somerville, R. S., et al. 2022b, ApJ, 929, 30
Calzetti, D. 2013, in Secular Evolution of Galaxies, ed. J. Falcón-Barroso & J. H. Knapen (Cambridge: Cambridge Univ. Press), 419
Carilli, C. L., & Walter, F. 2013, ARA&A, 51, 105
Casey, C. M., Narayanan, D., & Cooray, A. 2014, PhR, 541, 45
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., Viero, M. P., Church, S. E., et al. 2019, ApJ, 872, 186
Cleary, K. A., Borowska, J., Breysse, P. C., et al. 2022, ApJ, 933, 182
Crameri, F. 2021, Scientific colour maps, v7.0.0, Zenodo, doi: 10.5281/zenodo.1243862
Crameri, F., Shephard, G. E., & Heron, P. J. 2020, NatCo, 11, 5444
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
Foreman-Mackey, D. 2016, JOSS, 1, 24
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Foss, M. K., Ihle, H. T., Borowska, J., et al. 2022, ApJ, 933, 184
Freundlich, J., Combes, F., Tacconi, L. J., et al. 2019, A&A, 622, A105
Gebhardt, K., Mentuch Cooper, E., Ciardullo, R., et al. 2021, ApJ, 923, 217
Gong, M., Ostriker, E. C., Kim, C.-G., & Kim, J.-G. 2020, ApJ, 903, 142
González-López, J., Decarli, R., Pavesi, R., et al. 2019, ApJ, 882, 139
Hamilton, A. J. S. 1998, in The Evolving Universe, ed. D. Hamilton (Dordrecht: Kluwer), 185
Hill, G. J., Gebhardt, K., Komatsu, E., et al. 2008, in ASP Conf. Ser. 399, Panoramic Views of Galaxy Formation and Evolution, ed. T. Kodama, T. Yamada, & K. Aoki (San Francisco, CA: ASP), 115
Hill, G. J., Lee, H., MacQueen, P. J., et al. 2021, AJ, 162, 298
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hinton, S. R. 2016, JOSS, 1, 00045
Hunter, J. D. 2007, CSE, 9, 90
Ihle, H. T., Borowska, J., Cleary, K. A., et al. 2022, ApJ, 933, 185
Ihle, H. T., Chung, D., Stein, G., et al. 2019, ApJ, 871, 75
Inoguchi, M., Hosokawa, T., Mineshige, S., & Kim, J.-G. 2020, MNRAS, 497, 5061
Kaiser, N. 1987, MNRAS, 227, 1
Kamenetzky, J., Rangwala, N., Glenn, J., Maloney, P. R., & Conley, A. 2016, ApJ, 829, 93
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. 2020a, ApJ, 901, 141
Keating, L. C., Richings, A. J., Murray, N., et al. 2020b, MNRAS, 499, 837
Keenan, R. P., Keating, G. K., & Marrone, D. P. 2022, ApJ, 927, 161
Keenan, R. P., Marrone, D. P., & Keating, G. K. 2020, ApJ, 904, 127
Kennicutt, R. C. J. 1998, ApJ, 498, 541

Klitsch, A., Péroux, C., Zwaan, M. A., et al. 2019, MNRAS, 490, 1220
Kovetz, E., Breysse, P. C., Lidz, A., et al. 2019, BAAS, 51, 101
Kovetz, E. D., Viero, M. P., Lidz, A., et al. 2017, arXiv:1709.09066
Kusakabe, H., Blaizot, J., Garel, T., et al. 2020, A&A, 638, A12
Lamb, J. W., Cleary, K. A., Woody, D. P., et al. 2022, ApJ, 933, 183
Lenkić, L., Bolatto, A. D., Förster Schreiber, N. M., et al. 2020, AJ, 159, 190
Li, Q., Narayanan, D., Davè, R., & Krumholz, M. R. 2018, ApJ, 869, 73
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
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Madden, S. C., Cormier, D., Hony, S., et al. 2020, A&A, 643, A141
Moradinezhad Dizgah, A., & Keating, G. K. 2019, ApJ, 872, 126
Murray, S. G., Power, C., & Robotham, A. S. G. 2013, A&C, 3, 23
Padmanabhan, H. 2018, MNRAS, 475, 1477
Pavesi, R., Sharon, C. E., Riechers, D. A., et al. 2018, ApJ, 864, 49
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15
Rennie, T. J., Harper, S. E., Dickinson, C., et al. 2022, ApJ, 933, 187
Riechers, D. A., Boogaard, L. A., Decarli, R., et al. 2020, ApJL, 896, L21
Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, ApJ, 872, 7
Righi, M., Hernández-Monteagudo, C., & Sunyaev, R. A. 2008, A&A, 489, 489
Seifried, D., Haid, S., Walch, S., Borchert, E. M. A., & Bisbas, T. G. 2020, MNRAS, 492, 1465
Silva, M. B., Baumschlager, B., Cleary, K. A., et al. 2021, arXiv:2111.05354
Smit, R., Bouwens, R. J., Franx, M., et al. 2012, ApJ, 756, 14
Solomon, P. M., & Vanden Bout, P. A. 2005, ARA&A, 43, 677
Stein, G., Alvarez, M. A., & Bond, J. R. 2019, MNRAS, 483, 2236
Sun, G., Hensley, B. S., Chang, T.-C., Doré, O., & Serra, P. 2019, ApJ, 887, 142
Switzer, E. R., Masui, K. W., Bandura, K., et al. 2013, MNRAS, 434, L46
Tacconi, L. J., Genzel, R., Neri, R., et al. 2010, Natur, 463, 781
Uzgil, B. D., Carilli, C., Lidz, A., et al. 2019, ApJ, 887, 37
Visbal, E., & Loeb, A. 2010, JCAP, 11, 16
Visbal, E., Trac, H., & Loeb, A. 2011, JCAP, 8, 10
Walter, F., Decarli, R., Aravena, M., et al. 2016, ApJ, 833, 67
Wang, R., Carilli, C. L., Neri, R., et al. 2010, ApJ, 714, 699
White, M. 2001, A&A, 367, 27
White, M. 2002, ApJS, 143, 241
Yang, S., Popping, G., Somerville, R. S., et al. 2022, ApJ, 929, 140
Yang, S., Somerville, R. S., Pullen, A. R., et al. 2021, ApJ, 911, 132