

COMAP Pathfinder – Season 2 results

IV. A stack on eBOSS/DESI quasars

D. A. Dunne^{1,*}, K. A. Cleary¹, J. G. S. Lunde², D. T. Chung³, P. C. Breyse⁴, N.-O. Stutzer², J. R. Bond^{5,6,7},
H. K. Eriksen², J. O. Gundersen⁸, G. A. Hoerning⁹, J. Kim¹⁰, E. M. Mansfield¹¹, S. R. Mason⁸,
N. Murray^{5,6,7}, T. J. Rennie^{12,9}, D. Tolgay^{5,6}, S. Valentine⁴, I. K. Wehus², and COMAP Collaboration

¹ California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA

² Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029, Blindern N-0315 Oslo, Norway

³ Department of Astronomy, Cornell University, Ithaca, NY 14853, USA

⁴ Department of Physics, Southern Methodist University, Dallas, TX 75275, USA

⁵ Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada

⁶ Department of Physics, University of Toronto, 60 St. George Street, Toronto, ON M5S 1A7, Canada

⁷ David A. Dunlap Department of Astronomy, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada

⁸ Department of Physics, University of Miami, 1320 Campo Sano Avenue, Coral Gables, FL 33146, USA

⁹ Jodrell Bank Centre for Astrophysics, Department of Physics & Astronomy, The University of Manchester, Oxford Road, Manchester M13 9PL, U.K.

¹⁰ Department of Physics, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon postcode34141, Republic of Korea

¹¹ Department of Physics, Cornell University, Ithaca, NY 14853, USA

¹² Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T 1Z1, Canada

Received 28 October 2025 / Accepted 28 January 2026

ABSTRACT

We present a stack of data from the second season of the CO Mapping Array Project (COMAP) on the positions of quasars from eBOSS and DESI. COMAP is a line intensity mapping (LIM) experiment targeting dense molecular gas via CO(1–0) emission at $z \sim 3$. Season 2 of COMAP represents a factor of three increase in map-level sensitivity over the previous early science data release. We do not detect any CO emission in the stack; instead, we find an upper limit of $10.0 \times 10^{10} \text{ K km s}^{-1} \text{ pc}^2$ at 95% confidence within an $\sim 18 \text{ cMpc}$ box. We compare this upper limit to models of the CO emission stacked on quasars and find a tentative ($\sim 3\sigma$) tension between the limit and the brightest stack models after accounting for a suite of additional sources of experimental attenuation and uncertainty, including quasar velocity uncertainty, pipeline signal loss, cosmic variance, and interloper emission in the LIM data. The COMAP-eBOSS/DESI stack is primarily a measurement of the CO luminosity in the quasars' wider environment, and is therefore potentially subject to environmental effects such as feedback. With our current simple models of the galaxy-halo connection, we are thus unable to confidently rule out any models of cosmic CO with the stack alone. Conversely, the stack's sensitivity to these large-scale environmental effects has the potential to make it a powerful tool for galaxy formation science, once we are able to constrain the average CO luminosity via the auto-power spectrum (a key goal of COMAP).

Key words. galaxies: evolution – galaxies: ISM – quasars: general – cosmology: observations – large-scale structure of Universe

1. Introduction

Despite major advances in instrumentation over the past decades, we do not yet have the observational capacity to resolve certain mysteries about the Universe's history of galaxy formation. For instance, the star formation density of the Universe is known to peak at $z \sim 3$, known as the epoch of galaxy assembly (EoGA; [Madau & Dickinson 2014](#)), but the mechanism driving that peak is still uncertain. The fuel for star formation, molecular hydrogen gas, traced by molecular lines such as carbon monoxide (CO), is expensive to observe at these redshifts, and galaxy surveys constraining its evolution are limited to small cosmic volumes (1–10 square arcmin) and bright galaxies (e.g. [Riechers et al. 2019](#); [Aravena et al. 2019](#); [Lenkić et al. 2020](#)). The faint end of the CO luminosity function is nearly entirely unconstrained during the EoGA, and thus empirically rooted models for the overall CO luminosity density span several orders of magnitude ([Li et al. 2016](#); [Padmanabhan 2018](#);

[Keating et al. 2020](#); [Yang et al. 2022](#)). Similarly, current models of star formation require feedback from active supermassive black holes (SMBHs) to disrupt star formation in high-mass galaxies in order to match observations, but the actual mechanism and extent of this feedback is poorly understood (e.g. [Sanders et al. 1988](#); [Hopkins et al. 2006](#)). Again, understanding this feedback requires tracing its effect on the fuel for future star formation in galaxies, and understanding this feedback is limited by selection effects and small sample sizes in CO observations of active galaxies ([Hill et al. 2019](#); [Muñoz-Elgueta et al. 2022](#)).

Line intensity mapping (LIM) is a rapidly maturing observational technique that measures galaxy populations as fluctuations in the line intensity field over a large 3D cosmic volume rather than as individual catalogued objects. LIM fluctuations contain the integrated emission from all galaxies in the measured volume, even the very faintest, and thus capture the faint halos not observable by traditional means. Although it is rapidly maturing, LIM is a field in its early stages, with pathfinder experiments both concluded and ongoing (e.g. [Keating et al. 2015](#),

* Corresponding author: ddunne@astro.caltech.edu

2020; CONCERTO Collaboration 2020; Cleary et al. 2022), and other experiments that are beginning to come online (e.g. Doré et al. 2014; Crites et al. 2014; Vieira et al. 2020; CCAT-Prime Collaboration 2023). The biggest challenge facing LIM experiments (outside of those targeting the 21 cm line) is raw sensitivity, but they are also susceptible to many forms of systematic error (e.g. caused by instrumental effects or interloping spectral line emission).

An emerging body of work is combining LIM observations with spectroscopic galaxy catalogues to improve sensitivity and to mitigate systematic error; as major systematic errors are primarily unique to one survey (e.g. Fronenberg & Liu 2024). This can be done as a real-space 3D co-addition of the LIM data on the position of the objects in the galaxy survey (called a stack) or as a real or Fourier-space cross-correlation. In their early stages, many experiments have been turning to the stack as a conceptually simpler analysis, with the potential to provide a detection in advance of the full auto-power spectrum analysis (Keenan et al. 2022; Dunne et al. 2024; MeerKLASS Collaboration & Barberi-Squarotti 2025; Chen et al. 2025). A stack on LIM data probes the larger-scale distribution of emission around the galaxy survey objects in addition to emission from the surveyed objects themselves, at scales of tens of megaparsecs. (Chen et al. 2025; Dunne et al. 2025). Spectroscopic galaxy catalogues that trace more massive dark matter (DM) halos (or more biased galaxies) thus yield a brighter stack as their surroundings contain more, and brighter, galaxies. Stacks on highly biased objects such as quasars should provide the brightest possible stack emission (Breyse & Alexandroff 2019; Dunne et al. 2024).

In this work we present a stack of LIM data from Season 2 of the CO Mapping Array Project (COMAP) Pathfinder telescope (Lunde et al. 2024; Stutzer et al. 2024; Chung et al. 2024a), on the positions of quasars from the extended Baryon Oscillation Spectroscopic Survey (eBOSS; Ahumada et al. 2020) and the Dark Energy Spectroscopic Instrument (DESI) quasi-stellar object (QSO) survey (DESI Collaboration 2025). COMAP targets molecular gas via CO(1-0) emission around the peak of cosmic star formation ($2.4 < z < 3.4$) and eBOSS and DESI together encompass the largest survey of quasars available in this redshift range. This work follows a similar analysis performed using data from COMAP Season 1 and eBOSS alone (Dunne et al. 2024), but this updated stack is more than three times more sensitive thanks to the improved COMAP Season 2 sensitivity. We have also considerably refined our methodology.

We introduce the three datasets in Sect. 2, and describe the simulated LIM and galaxy catalogue data to which we compare the stack in Sect. 3. The stack methodology is described in Sect. 4. In Sect. 5 we present the stack results, to which we compare models in Sect. 6. Throughout this work, we take the nine-year WMAP cosmology (Hinshaw et al. 2013) for consistency with the peak-patch simulations used in Sect. 3 and other COMAP works. This is a Λ CDM model with $\Omega_m = 0.286$, $\Omega_\Lambda = 0.714$, $\Omega_b = 0.047$, and $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ with $h = 0.7$. We assume base-10 logarithms unless otherwise stated.

2. Data

2.1. COMAP

The COMAP Pathfinder instrument is a 19-feed (single-polarization) focal plane array fielded on a 10.4-m antenna (Lamb et al. 2022) that has been performing LIM observations targeting CO(1-0) at $2.4 < z < 3.4$ since 2019 and is still con-

tinuing to observe. Thus far, two COMAP observing seasons have taken place. Season 1 encompassed 5200 on-sky observation hours taken over 13 months between May 2019 and August 2020 over three fields. In this work, we used the COMAP Season 2 dataset (Lunde et al. 2024; Stutzer et al. 2024; Chung et al. 2024a), which added 12 300 hours of observations collected between November 2020 and November 2023, for a total of 17 500 hours. This is an increase of 340% in raw data volume compared to the previous early science (ES) dataset. Combined with improvements to the data reduction pipeline that yield a much greater data efficiency (described in Lunde et al. 2024), COMAP Season 2 has seven times more effective integration time than the ES dataset, giving a per-voxel root-mean-square noise (RMS) in the final maps of 25–50 μK (Lunde et al. 2024; Stutzer et al. 2024).

The COMAP time-ordered data (TOD) were processed from raw detector readout into calibrated time-domain data in brightness temperature units feed-by-feed, using a series of filters designed to reduce correlated noise and systematic errors. The COMAP calibration uncertainty is $<2\%$, and we treated it as negligible here. The raw TOD were then binned into maps with $2' \times 2'$ spatial pixels (spaxels) and spectral channels with width 31.25 MHz using a simple nearest-neighbour mapmaker, as the filtered data were found to have negligible correlated noise. A map-domain PCA filter was used to remove lingering systematic effects. These steps are explained in full in Lunde et al. (2024). Each of these steps has the potential to remove signal as well as noise; we account for signal removed in the pipeline steps in Sect. 4.2. The end results of the COMAP pipeline are maps that are consistent with white (Gaussian) noise.

Finally, because COMAP is an experiment that scans across the sky to cover its fields, the RMS noise varies considerably across each COMAP field. The outer low-significance regions of the map contribute almost no sensitivity to the stack. We thus removed them for efficiency, and to reduce the chance of incorporating untreated systematic errors from these low-significance regions into the stack. In addition to the map-level cuts used by our autocorrelation pipeline (Lunde et al. 2024; Stutzer et al. 2024), we mask all voxels that exceed eight times the mean RMS of the 100 highest-significance voxels. We also clean out irregular regions of higher-significance voxels around the outskirts of the map (including some voxels that are totally surrounded by masked voxels). We masked all voxels which had fewer than four of their eight-connected neighbours unmasked. These added cuts masked 12.5% more voxels than the autocorrelation masking strategy across the three COMAP fields, but increased the stack uncertainty by only 2%.

2.2. eBOSS and DESI

We performed the stack on a combined catalogue of objects from the two largest spectroscopic quasar surveys available in the COMAP fields, eBOSS (part of the Sloan Digital Sky Survey, SDSS; Dawson et al. 2013; Ahumada et al. 2020; Lyke et al. 2020) and DESI (DESI Collaboration 2025). Quasars are ideal targets for a stacking analysis, as they trace high-mass DM halos which are very highly clustered. This, more than any other factor, leads to a bright stack signal (Dunne et al. 2025).

As we had previously performed a COMAP stacking analysis using eBOSS QSOs (Dunne et al. 2024), and the first season of COMAP data, we refer to that work for the details of the eBOSS catalogue reduction. We took DESI data from their first data release (DR1; DESI Collaboration 2025). DESI is still ongoing; there will be more objects in the COMAP fields

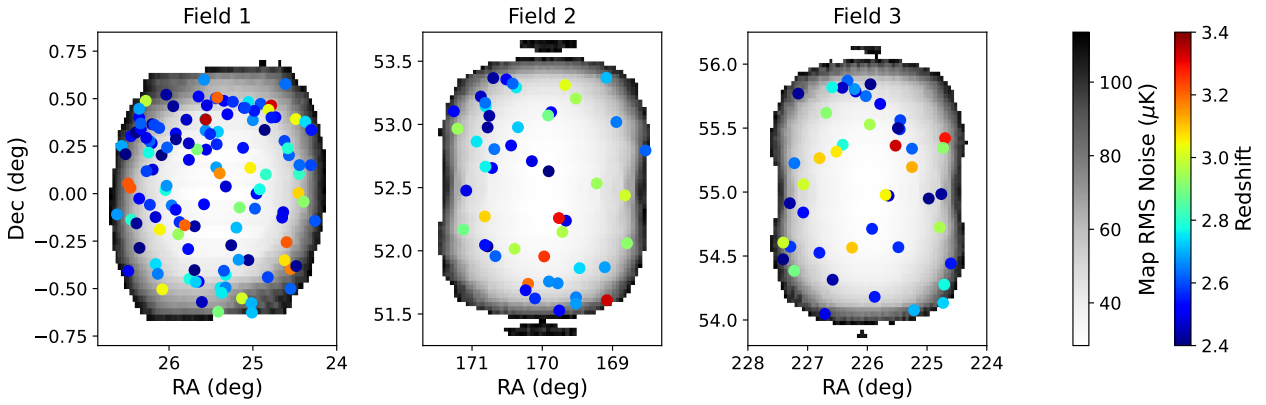


Fig. 1. Spatial distribution of the eBOSS/DESI quasars compared to the footprint of the COMAP data. The COMAP footprint is coloured by the RMS in each voxel, averaged across all spectral channels, and the quasars are coloured by their redshift.

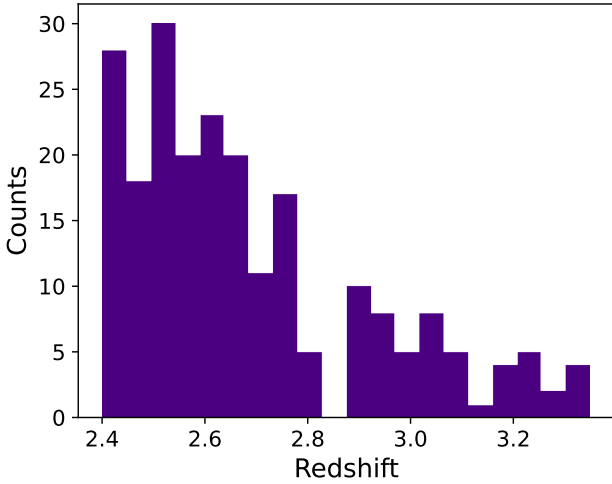


Fig. 2. Redshift distribution of the eBOSS and DESI quasars included in the stack catalogue.

in the future as the survey continues. The DESI target selection pipeline does not account for whether a QSO was previously observed with eBOSS, so there is significant overlap between the two samples. We treated any DESI object within $1.3'$ (which is the average SDSS seeing full width at half maximum, [Delubac et al. 2017](#)) and $\Delta z = 0.012$ (which is the eBOSS+DESI uncertainty on the redshifts, and roughly corresponds to the 1185 km s^{-1} redshift offset we observe in quasars) of an eBOSS object as a duplicate. This is a similar cut to that of a comparable analysis performed by the DESI team, who treated any objects within $1'$ as a match ([Adame et al. 2025](#)). Where duplicates were present, we used the DESI position and redshift.

Additionally, we applied a systematic velocity offset of 1185 km s^{-1} to each QSO in the combined eBOSS-DESI catalogue. This value was determined in [Dunne et al. \(2024\)](#) from comparing optically determined and CO-based redshift values for a selection of $z \sim 3$ QSOs. Astrophysically, it originates from the optical spectral lines used by eBOSS/DESI tracing bulk in- and outflows in QSOs, whereas the CO emission traces dense gas at the centre of the galaxy. We note that varying the magnitude of this offset does not meaningfully affect our results. This effect also results in a large ($\sim 1660 \text{ km s}^{-1}$, or several COMAP

spectral channels) scatter in redshift offsets, which we address in Sect. 6.2.2.

Between the two surveys, there are 231 QSOs in the COMAP footprint. Of these, 95 are in eBOSS alone, and 136 are either in DESI or in both surveys. The angular distribution of the galaxies across the COMAP fields is shown in Figure 1, and their redshift distribution in Figure 2. We note that there are fewer QSOs in this analysis than the 243 eBOSS QSOs used for COMAP stacking in [Dunne et al. \(2024\)](#), even after the additional DESI objects are included. This is because the ES maps did not incorporate the same RMS cuts, and thus covered areas on the sky that in practice added no sensitivity. The median redshift of the quasars included in this sample is 2.62.

3. Simulations

We generated a suite of simulated LIM data cubes and associated galaxy catalogues for comparison with the real COMAP and quasar data. These were generated using the `joint_limlam_mock`¹ code described in [Dunne et al. \(2025\)](#), based on peak-patch DM halo simulations ([Bond & Myers 1996](#); [Stein et al. 2019](#)). We refer the reader to [Dunne et al. \(2025\)](#) for the methodology behind this simulation code. The suite of simulation parameters in [Dunne et al. \(2025\)](#) was designed to emulate a stack of COMAP data on the positions of Lyman- α emitters from the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX; [Hill et al. 2024, 2021](#)). In the modelling for this work, we primarily adopted the same parameters for the CO LIM map, but altered the parameters to match the quasar catalogue. We additionally tested a suite of models for CO luminosity as a function of halo mass, $L(M)$. For all other parameters, we defaulted to those used in [Dunne et al. \(2025\)](#).

3.1. CO modelling

We defaulted to the COMAP fiducial (UM+COLDz+COPSS) CO model from [Chung et al. \(2022\)](#), which takes the form

$$\frac{L'_{\text{CO}}(M_{\text{h}})}{\text{K km s}^{-1} \text{ pc}^2} = \frac{C}{(M_{\text{h}}/M)^A + (M_{\text{h}}/M)^B}. \quad (1)$$

A , B , C , and M are free parameters. Here, we adopted values determined using priors from UNIVERSEMACHINE

¹ https://github.com/delaneydunne/joint_limlam_mock

(Behroozi et al. 2019), conditioned on luminosity function constraints from the CO Luminosity Density at High Redshift survey (COLDz; Riechers et al. 2019) and the power spectrum measurement from the CO Power Spectrum Survey (COPSS; Keating et al. 2015). The resulting parameter values were $A = -2.85$, $B = -0.42$, $\log C = 10.63$, and $\log \frac{M}{M_\odot} = 12.3$. We applied a log-normal scatter with $\sigma_{\text{CO}} = 0.42$ dex. As the model space for CO at $z \sim 3$ remains extremely uncertain, we additionally tested several other models, which are shown in Figure 3. These are

- An observationally motivated model that determines halos’ CO luminosity through a simulated star formation rate (SFR), which is in turn generated from their infrared luminosities L_{IR} ,

$$\log L'_{\text{CO}} = \frac{1}{\alpha} [\log L_{\text{IR}} - \beta], \quad (2)$$

$$L_{\text{IR}} = \frac{\text{SFR}}{\delta_{\text{MF}}} \times 10^{10}. \quad (3)$$

This is based on Li et al. (2016). However, we took the Keating et al. (2020) extension to this model, which uses the more up-to-date parameter values from Kamenetzky et al. (2016) rather than Carilli & Walter (2013). These values have the additional advantage of connecting the CO luminosities directly to L_{IR} , without the SFR intermediary. The Kamenetzky et al. (2016) parameter values are $\delta_{\text{MF}} = 1.0$, $\alpha = 1.27$, $\beta = -1.0$. We used a scatter of 0.3 dex, and took

$$\frac{L_{\text{CO}}}{L_\odot} = 4.9 \times 10^{-5} \frac{L'_{\text{CO}}}{\text{K km s}^{-1} \text{ pc}^2}. \quad (4)$$

Hereafter we refer to this model as L16-K20.

- A CO(1-0) model generated from fits to the Santa Cruz semi-analytic model (e.g. Somerville et al. 2015) from Yang et al. (2022). This takes a redshift-dependent double-power-law form,

$$L = 2N(z)M \left[\left(\frac{M}{M_1(z)} \right)^{-\alpha(z)} + \left(\frac{M}{M_1(z)} \right)^{\beta(z)} \right]^{-1}, \quad (5)$$

with a duty cycle

$$f_{\text{duty}} = \frac{1}{1 + \left(\frac{M}{M_2(z)} \right)^{\gamma(z)}}. \quad (6)$$

We took values for $M_1(z)$, $N(z)$, $\alpha(z)$, and $\beta(z)$, as well as $M_2(z)$ and $\gamma(z)$ for CO(1–0) at $1.0 \leq z < 4.0$ from Table 1 of Yang et al. (2022). We refer to this model as Y22.

- An observationally driven model abundance-matched onto a collection of $z = 3$ CO observations by Padmanabhan (2018). This is another redshift-dependent double-power-law model,

$$L_{\text{CO}}(M_{\text{h}}, z) = \frac{2N(z)M_{\text{h}}}{\left(\frac{M_{\text{h}}}{M_1(z)} \right)^{-b(z)} + \left(\frac{M_{\text{h}}}{M_1(z)} \right)^{y(z)}}, \quad (7)$$

with $M_1(z)$, $N(z)$, $b(z)$, and $y(z)$ each given by Equation 24 of Padmanabhan (2018). We also adopted the Padmanabhan (2018) duty cycle factor $f_{\text{duty}} = 0.1$. We refer to this model as P18.

- The flat+COLDz model from C22. This follows the same double power law functional form as the fiducial model above, but is conditioned on luminosity function constraints

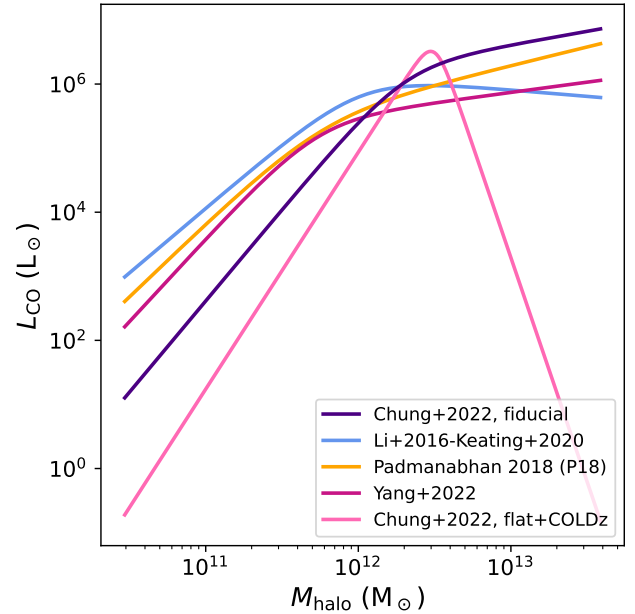


Fig. 3. Models for CO luminosity as a function of halo mass used to generate simulated stacks. We defaulted to the C22 fiducial model.

from the COLDz blind CO galaxy survey (Riechers et al. 2019) alone, with an otherwise flat prior. The coefficients that result are $A = -3.7$, $B = 7.0$, $\log C = 11.1$, and $\log \frac{M}{M_\odot} = 12.5$, and it has a lognormal scatter of 0.36 dex. We refer to this model as C22, flat+COLDz.

Overall, the COMAP fiducial model is brightest for high-mass halos ($M_{\text{h}} \gtrsim 10^{12} M_\odot$), and the L16-K20 model for most other halo masses. The Y22 model is the faintest on average.

3.2. Quasar modelling

Because of the resolution of COMAP, the stack integrates over so many DM halos (roughly 250 over $18 \times 18 \times 9$ cMpc) that the host galaxies of the actual quasars included in the catalogue provide only minor contributions to the total stack luminosity (Dunne et al. 2025). Conversely, this means that the stack is extremely sensitive to the number and CO luminosity of DM halos in the vicinity of the quasars. A successful mock catalogue, then, will recover the relationship between the (observed) quasars and the underlying large-scale structure (LSS), with the CO properties of the quasars themselves only a secondary effect.

As the quasar luminosity function (QLF) has considerably more observational backing from large surveys $z \sim 3$ than the CO models (e.g. Ahumada et al. 2020; Palanque-Delabrouille et al. 2016), we followed Dunne et al. (2025) in generating modelled quasar luminosities by abundance-matching onto observed QLFs. We show the parametrizations of the QLF we use in this work in Figure 4. The default model we used is based on a Ly α luminosity function generated from quasars in the Physics of the Accelerating Universe survey (PAU; Torralba-Torregrosa et al. 2024), and follows a Schechter (1976) form,

$$\frac{dn}{dL} = \frac{\phi^*}{L^*} \left(\frac{L}{L^*} \right)^\alpha e^{-L/L^*}. \quad (8)$$

We took the lowest-redshift bin, which is centred slightly above the $z = 2.62$ median redshift of our combined eBOSS-DESI

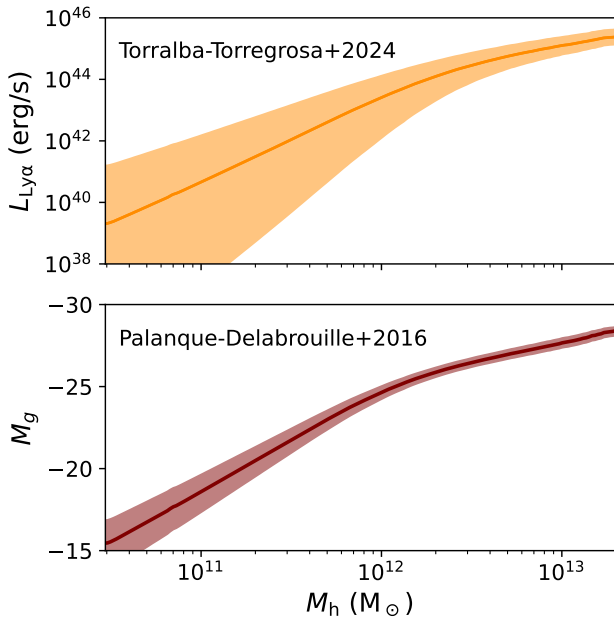


Fig. 4. Models for quasar luminosity used to generate simulated catalogues, created from abundance-matching onto observed luminosity functions. The top panel shows the PAU model from (Torralba-Torregrosa et al. 2024) and the bottom panel shows the Palanque-Delabrouille et al. (2016). We defaulted to the Torralba-Torregrosa et al. (2024) model.

catalogue, yielding $\log(L^*/\text{erg s}^{-1}) = 44.89$, $\log(\phi^*/\text{Mpc}^{-3}) = -6.48$, and $\alpha = -1.44$. Through propagating forward the uncertainty on these best-fit values, we found an average scatter at the high-mass end ($M_{\text{halo}} > 5 \times 10^{12} M_{\odot}$) of $\sigma_{\text{QSO}} = 0.39$ dex. We took the limiting luminosity $L_{\text{cut}} = 10^{44} \text{ erg s}^{-1}$ (Torralba-Torregrosa et al. 2024).

Additionally, we tested the QLF from Palanque-Delabrouille et al. (2016), generated from eBOSS and used for target selection for DESI. The luminosities treated in this work are absolute g -band magnitudes, normalized to $z = 2$. The luminosity function follows a double power law form,

$$\Phi(M_g, z) = \frac{\Phi^*}{10^{0.4(\alpha+1)(M_g - M_g^*)} + 10^{0.4(\beta+1)(M_g - M_g^*)}} \quad (9)$$

with several redshift-dependent modifications modulating Φ^* , M_g^* , and α . We followed these modifications, using the luminosity and density evolution (LEDE) functional form that the authors suggest for $z > 2.2$. These modifications are

$$\log[\Phi^*(z)] = \log[\Phi^*(z_p)] + c_{1a}(z - z_p) + c_{1b}(z - z_p)^2, \quad (10)$$

$$M_g^*(z) = M_g^*(z_p) + c_2(z - z_p), \quad (11)$$

and

$$\alpha(z) = \alpha(z_p) + c_3(z - z_p). \quad (12)$$

In this high-redshift limit, there are eight fitted parameters ($M_g^*(z_p)$, $\Phi^*(z_p)$, α , β , c_{1a} , c_{1b} , c_2 , and c_3). We took the best-fit values for each from Table 4 of Palanque-Delabrouille et al. (2016). We calculated the luminosity function at $z = 2.62$ (the median redshift of the combined eBOSS-DESI catalogue), and

used this across the entire COMAP redshift range. We determined the log-normal scatter in this relation from the uncertainty on the best-fit parameters, propagating them into the $L(M)$ relation and taking the median scatter at the high-mass end ($M_{\text{halo}} > 5 \times 10^{12} M_{\odot}$). This yields a scatter of 0.11 dex.

The actual magnitude limits of the DESI quasar catalogue are defined in the r band. We used the Chaussidon et al. (2023) definition of a (relative) magnitude limit of $r = 23.0$, and corrected this to the g band using the BOSS composite spectrum of Harris et al. (2016), which provides a $g - r$ value of 0.08 at our median redshift of 2.62. The correction stays below $g - r = 0.2$ over our entire redshift range. We thus defined a relative magnitude limit of $g < 23.08$ for this model.

In both cases, once quasar luminosities were assigned to DM halos, we generated a mock catalogue in two steps. First, we removed from the mock catalogue all objects with luminosities (or magnitudes) below the cut-off established for each model. These are objects which would not be observable due to sensitivity limits. We then selected the goal number of objects (N_{obj}) randomly from the remaining simulated quasars, weighting linearly by the quasar luminosity. Each model has its own simulated sensitivity limit, relationship between quasar luminosity and DM halo mass, and scatter in that relationship. Because of these differences, the average mass of the halos selected to be included in the mock catalogue varies between models, and thus the degree of clustering of other DM halos around the catalogued objects also varies.

4. Stacking methodology

The stack is a coaddition of the COMAP data at the position of each quasar. To accomplish this we primarily followed the methodology established in Dunne et al. (2024), and we refer the reader to that work for the full details. Here, we briefly summarize the methodology and describe the modifications we made to account for pipeline signal loss.

4.1. Coaddition

We first converted the three-dimensional COMAP maps for each field from brightness temperature (T_b) units into CO line luminosity values (L'_{CO}) following Solomon et al. (1997),

$$L'_{\text{CO}} = \frac{c^2}{2k_B} (S \Delta v) \frac{D_L^2}{v_{\text{obs}}^2 (1+z)^3}, \quad (13)$$

where

$$(S \Delta v) = \frac{2v^2 k_B T_b}{c^2} \Omega_{\text{vox}} \Delta v. \quad (14)$$

Here $(S \Delta v)$ is the integrated CO line flux in the voxel, D_L is the luminosity distance of the catalogue object, v_{obs} is the observed frequency and z is the redshift. The parameter Ω_{vox} refers to the solid angle of the sky subtended by the voxel. We then extracted from the COMAP data cubes a 3D ‘cubelet’ centred around the spatial position and observed CO(1-0) frequency of each catalogue object. This centring is performed only at the level of the COMAP voxel. We explore more advanced sub-voxel centring in Mansfield et al. (2025). The cutout cubelets were $31 \times 31 \times 81$ voxels ($62' \times 62' \times 2.53 \text{ GHz}$, or $113 \text{ cMpc} \times 113 \text{ cMpc} \times 337 \text{ cMpc}$), with the catalogue object itself located within the central voxel of the cubelet. For context, the full COMAP cubes are $120 \times 120 \times 256$ voxels. The cubelets were

purposefully much larger than the expected extent of the stack signal in order to account for larger-scale structure.

Finally, we combined the COMAP data cubelets associated with each catalogue object voxel-by-voxel, weighting by the voxel RMS using an inverse-variance scheme to account for the inhomogeneous noise response across the COMAP data cubes. We did this separately for each field. We thus obtained a single stacked cubelet, containing the average three-dimensional profile of the CO emission surrounding the catalogued objects, for each field. We refer to this average cubelet as \mathbf{s} . For presentation purposes, we summed the averaged cubelet across its central 5×5 spatial voxels or 3 spectral voxels to generate 1D spectra or 2D images, respectively. These are shown in Figure 6.

4.2. Extraction of line luminosity

The goal output value of the stack is the average CO luminosity of the quasars in the eBOSS/DESI catalogue and their surroundings over some given cosmic volume, $\langle L'_{\text{CO}} \rangle$. In Dunne et al. (2024), we obtained this value by integrating the stacked cubelet data \mathbf{s} over a central aperture 3×3 spatial pixels and seven spectral channels wide. However, the COMAP data reduction pipeline, intended to take the data from raw TOD to a map free of systematic errors, will act on the signal as well as the noise, and modify both the absolute magnitude and 3D distribution of the input ground-truth signal. In order to account for this, we changed the methodology by which we extract the $\langle L'_{\text{CO}} \rangle$ from the stacked datacube. We determined $\langle L'_{\text{CO}} \rangle$ by scaling a 3D model \mathbf{r} , containing our best approximation of the predicted stack signal, by an amplitude a fit to the real stacked data \mathbf{s} . We then integrated over the scaled model to yield $\langle L'_{\text{CO}} \rangle$. We did this separately for each field, as the different on-sky locations of the fields meant that the pipeline normalization filter, which applies to the TOD, has a different behaviour in RA and Dec for each field.

4.2.1. Extraction methodology

We extracted the signal in four steps, performing each step individually for each COMAP field. Firstly, we generated a 3D model of the signal to be stacked using simulations generated with `joint_limlam_mock`. Then, we processed the raw simulated signal \mathbf{r} by the COMAP analysis pipeline, to ensure the effects of the filters on the real data are accounted for in the model, by injecting the simulated signal into subsets of the COMAP TOD and passing these injection simulations through the pipeline. We generated stacks of the simulated signal, both before and after processing by the pipeline. We refer to the resulting unprocessed model as \mathbf{r} and the pipeline-processed model as \mathbf{t} (these are discussed in more detail below in Sect. 4.2.2, and shown in Figure 5). Thirdly, we used the processed template \mathbf{t} as a matched filter, and determined the amplitude a that would scale it to the real stacked datacube \mathbf{s} ,

$$a = (\mathbf{t}^T \mathbf{t})^{-1} \mathbf{t}^T \mathbf{s}. \quad (15)$$

We then determined the input stacked signal best-explained by the actual data by multiplying a into the original, unprocessed stack model \mathbf{r} . This step also corrected for signal loss incurred by the pipeline filters. In the final step, we calculated the corrected output stack luminosity $\langle L'_{\text{CO}} \rangle$ by integrating over the scaled template signal \mathbf{ar} in three dimensions,

$$\langle L'_{\text{CO}} \rangle = a \sum_{k=-1}^{+1} \sum_{i,j=-2}^{+2} r_{i,j,k}, \quad (16)$$

where r is the voxel of the model cube \mathbf{r} indexed by i and j in the two spatial dimensions and k in the spectral (\mathbf{r} is in units of line luminosity, so the area and spectral width of each voxel are accounted for). In this work, we integrated over the central 5×5 spaxels (in contrast to the 3×3 used in Dunne et al. 2024) because the actual stack signal is dominated by cosmological clustering and is thus considerably more broadened than a point source. Additionally, we summed over 3 spectral channels, rather than 7, and accounted for signal spread out of this central aperture due to uncertainties on the systematic redshifts of the quasars using a scaling factor f_{VI} (see Sect. 6.2.2).

This is a simplified version of the full adaptive matched filtering analysis that we present in Mansfield et al. (2025). There are two advantages to this strategy: extracting $\langle L'_{\text{CO}} \rangle$ using a template weights most heavily the voxels which contain the most signal, and it enables us to efficiently account for signal loss due to the analysis pipeline. Finally, this methodology does introduce a model-dependence to the output $\langle L'_{\text{CO}} \rangle$. We describe how we accounted for this in Sect. 6.3.1.

4.2.2. Template generation

In order to perform the matched filtering correctly, we needed to use as accurate an estimation of the raw signal output from the COMAP stack as possible. Thus, the templates for each field needed to incorporate any modification of the signal incurred during signal processing. To generate an accurate template \mathbf{t} , we thus applied the pipeline filters to the simulated signal. We did this by injecting simulated signal into a subset of the raw COMAP TOD, and passing it through the full pipeline and mapmaker. Because we are not detecting CO signal in either autocorrelation or this stacking analysis, and because the simulated halo distribution will not line up with the real one in any case, the raw data functioned as noise (we verify this in Appendix A). Using signal injected into raw data allowed us to estimate signal loss using the actual COMAP noise distribution, thus including its angular and frequency shape and any low-level systematics.

We created 100 different realizations of a simulated CO fluctuation cube and associated QSO catalogue using the `joint_limlam_mock` code and default parameters described in Sect. 3. We used 50 different peak-patch realizations, generating luminosity values using two different random seeds for each (for a total of 100 realizations). This is increased from the ten unique realizations used by Lunde et al. (2024) for the autocorrelation analysis, in order to account more thoroughly for the effects of cosmic variance (see Sect. 6.3.2). We then injected each simulation realization into a unique subset of the full TOD (using a hundredth of the total data volume) for COMAP Fields 1 and 2, boosting the injected signal by a factor of 50. This boost had to be chosen carefully; we required a high-significance template for effective scaling, but the signal still needed to be small enough to not cause non-linearities in any of the pipeline steps. We discuss our choice of boost in more detail in Appendix A. Each of the 100 TOD subsets with injected signal was passed through the full COMAP analysis pipeline and mapmaker, resulting in 100 independent mock LIM cubes for each field, each with accurate on-sky footprints and filtering.

For each field of each injected realization, we then performed a mock stack. For efficiency, we used the template generated for Field 2 for both Field 2 and Field 3, as both fields traverse the sky in a similar fashion, making their scan pattern almost identical in Equatorial coordinates. This produces very similar pipeline effects, and they are thus expected to have very similar templates. We reproduced the quasar density in the real COMAP-

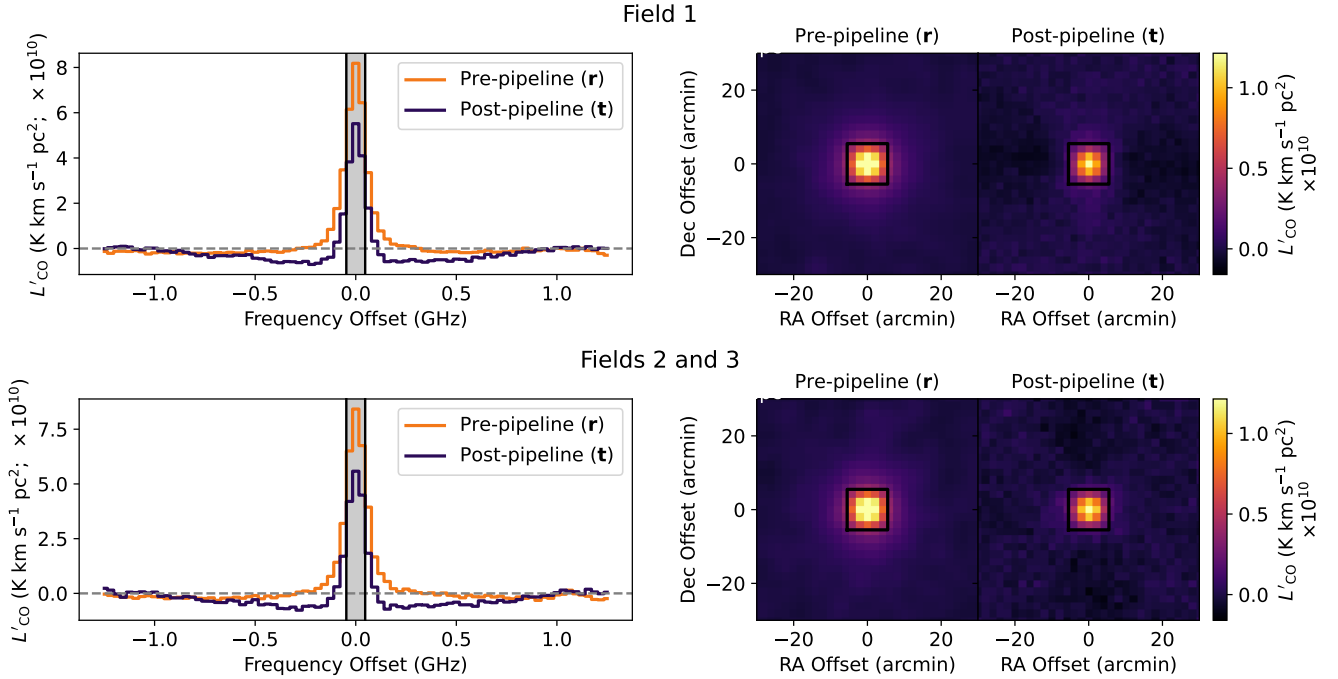


Fig. 5. Averaged simulated stacks used to account for pipeline signal loss in the stack, shown here for COMAP Field 1 above and Field 2 (which is also used for Field 3) below. We show the stacked spectra on the left (averaging across the central 5×5 spaxels for presentation purposes) and the stacked images (averaging across the central three spectral channels for presentation purposes) on the right. The apertures over which we sum are shown in black. The spatial and spectral effects of the COMAP analysis pipeline, and the different effects for each field, are visible.

eBOSS/DESI stack by stacking on the same number of mock quasars as are present in each real field (the injections recover the same hit patterns as the real COMAP fields, so the cosmic volumes covered by each field were also accurate). We rounded the number of quasars in Fields 2 and 3 to 50 (From 53 and 47). We simulated the uncertainty in the quasar systemic redshifts by offsetting the spectral position of each mock quasar by a random value pulled from a normal distribution with standard deviation $\sigma_v = 1660 \text{ km s}^{-1}$ (Sect. 2.2). This resulted in a 3D stacked cubelet for each realization. The 131-object Field 1 stacks with a $50\times$ signal boost had S/N values of 6 on average in the central voxels, and the 50-object stacks for Field 2 and 3 had S/N values of 4, calculated using the map-level RMS noise propagated from the TOD. Finally, we averaged together the stacked cubelets for each of the 100 different realizations, voxel-by-voxel, using inverse-variance weighting. The resulting averaged cubelet is our template \mathbf{t} , shown in Figure 5. When averaged over 100 iterations, the S/N of the central voxel of the template rises to 67 for Field 1 and 49 for Fields 2 and 3.

4.3. Variance

While it is possible to determine the uncertainties in each stack from propagating through the RMS noise from the TOD, this methodology doesn't capture any potential residual correlated noise in the maps, or any of the noise processing by the filters in the pipeline. Thus, in order to robustly calculate the variance in the output stacks, we bootstrapped stacks on randomized catalogues and measured the distribution of their measured line luminosity values. We randomized the catalogues using the same method as was used in Dunne et al. (2024), offsetting the coordinates of each object in the stack catalogue by some value (between one and ten spaxels in the angular directions, and between one and ten channels in the spectral direction. This

works well for the eBOSS/DESI catalogues, as both are fairly spatially homogeneous. We ran the stack 10 000 times, each time using a different random catalogue realization generated using this method. We calculated $\langle L'_{\text{CO}} \rangle$ from the resulting stacked cubelets using the methodology outlined in Sect. 4.2. We then fit a Gaussian distribution to the output $\langle L'_{\text{CO}} \rangle$ values and took the standard deviation of this Gaussian to be the uncertainty on the stack. This was done separately for each field, because both the average RMS noise and the number of quasars included in the stack varies between fields.

As demonstrated in Chen et al. (2025), this bootstrapped strategy only accounts for variance due to instrumental noise (both correlated and uncorrelated), and does not account for sample variance, which also affects the stack considerably. However, sample variance only applies to actual signal, and in this work we achieve only an upper limit, so the sample variance does not affect the limit directly. Instead, we counted sample variance as an uncertainty on results output from our modelling; we discuss this in Sect. 6.3.2.

5. Results

We show the final COMAP-eBOSS/DESI stacks for each field in Figure 6, noting that the extended 2D image and 1D spectrum representations of the stack are uncorrected for any pipeline signal loss effects. We calculated the scale factor a for each field, using its respective template \mathbf{t} , finding $a = [0.406, 0.026, 0.027]$ for Fields [1,2,3] respectively. The scaled templates $a\mathbf{t}$ are also shown in Figure 6. We then determined $\langle L'_{\text{CO}} \rangle$ for each field using its respective model \mathbf{r} , finding $\langle L'_{\text{CO}} \rangle = [8.44, 0.56, 0.59] \times 10^{10} \text{ K km s}^{-1} \text{ pc}^2$ for Fields [1,2,3].

We calculated the variance for each field using bootstrapping, as described in Sect. 4.3. We combined the values extracted for each field using inverse-variance weighting, using the out-

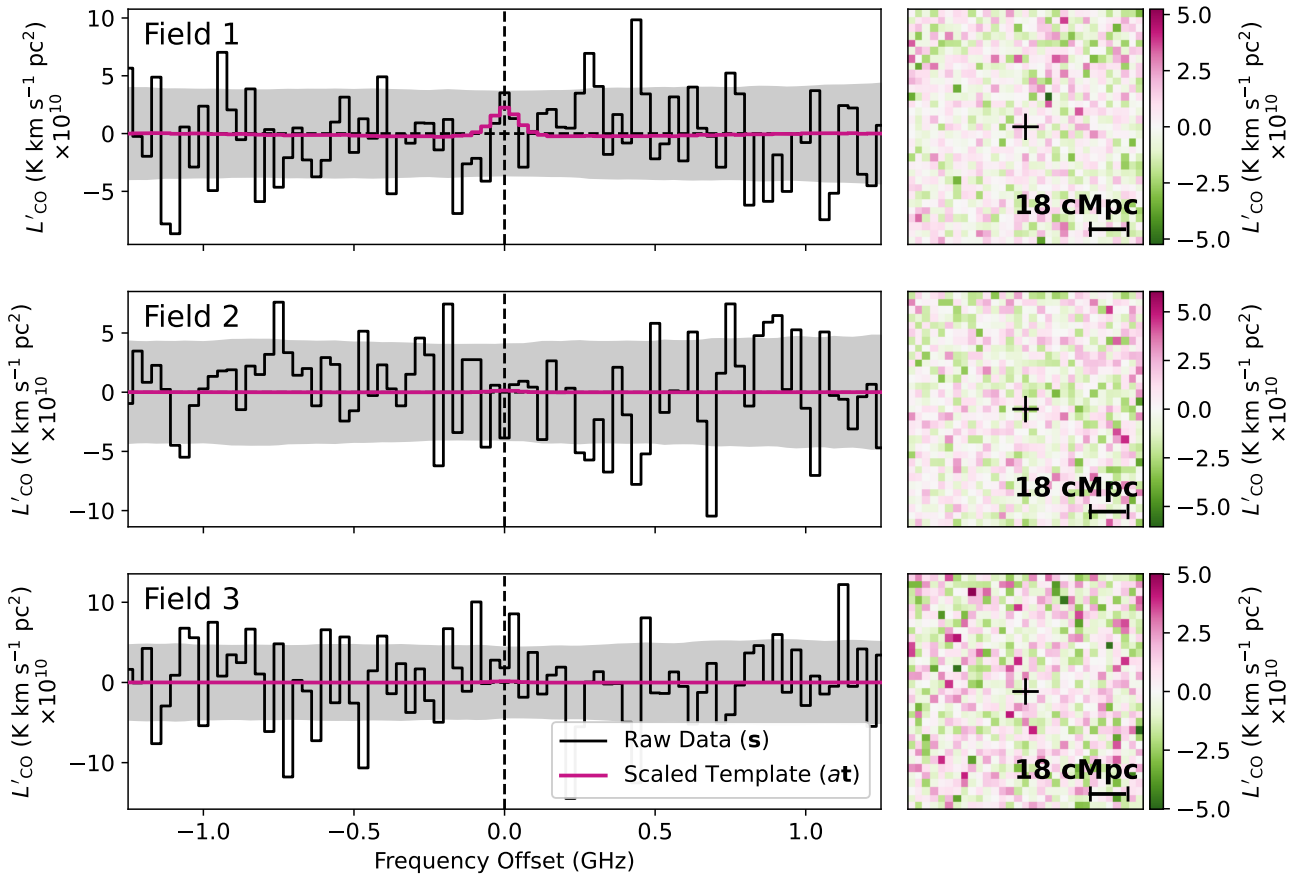


Fig. 6. Stacked spectra (left) and 2D images (right) for COMAP data in each of the three COMAP fields at the positions of the eBOSS/DESI QSOs. The per channel RMS values are shown as grey shaded regions in the stacked spectrum, and the scaled template at is overlaid. The stack is not corrected for any pipeline signal-loss effects, and the templates have the same signal loss present.

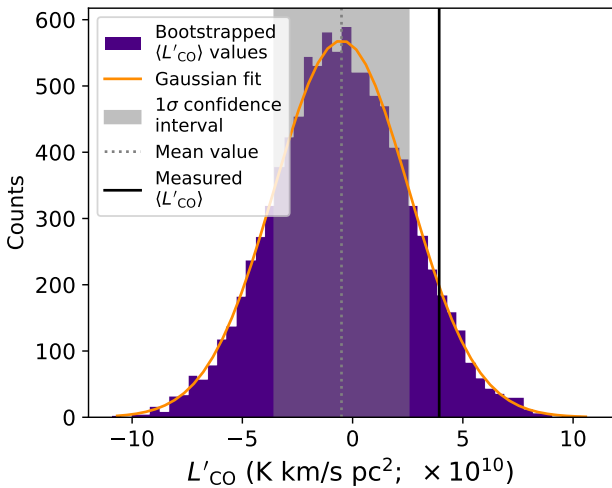


Fig. 7. Bootstrapped $\langle L'_{\text{CO}} \rangle$ values determined from stacks on randomized eBOSS/DESI catalogues, with a Gaussian fit. The mean and 1σ confidence interval of the Gaussian are shown in grey.

puts from the bootstraps in each field to determine the variances. We found an overall output stack luminosity for the COMAP-eBOSS/DESI stack of $\langle L'_{\text{CO}} \rangle = 4.15 \times 10^{10} \text{ K km s}^{-1} \text{ pc}^2$, with an

uncertainty $\sigma_{\langle L'_{\text{CO}} \rangle} = 2.94 \times 10^{10} \text{ K km s}^{-1} \text{ pc}^2$. The stack therefore has a significance of 1.41σ , and is a non-detection. The 2σ (95%) upper limit is $\langle L'_{\text{CO}} \rangle \leq 10.0 \times 10^{10} \text{ K km s}^{-1} \text{ pc}^2$. A histogram showing the bootstrapped overall output luminosity values is shown in Figure 7, along with fits to a Gaussian distribution. Averaging the scale factors for each field together using the same inverse-variance weighting, we found an overall scale factor for the stack of $a = 0.20$. Physically, this means that the scaling of the fiducial model that best explains the data has a fifth of the overall luminosity of our fiducial model.

6. Comparison to models

To put this upper limit value into context, we compared it directly to models for CO emission stacked on QSO positions at $z \sim 3$. While these models innately include many factors which could affect the output signal (such as astrophysical line broadening of CO, instrumental beam smearing, and the relationship to the LSS of both the CO emission and the QSOs), there remain several experimental and modelling factors that could serve to either attenuate the stacked signal or add variance, and thus must be accounted for when comparing the real stack luminosity to modelled versions. These are in addition to signal loss from the COMAP analysis pipeline, which is accounted for when extracting $\langle L'_{\text{CO}} \rangle$. We discuss these factors below. The effects of each on the simulated signal are shown in Table 2 and in Figure 9.

Table 1. Stack $\langle L'_{\text{CO}} \rangle$ values for each of the CO and galaxy catalogue models tested.

Model	TT24 ^a $\langle L'_{\text{CO}} \rangle$ (K km s ⁻¹ pc ²) $\times 10^{10}$	PD16 ^b $\langle L'_{\text{CO}} \rangle$ (K km s ⁻¹ pc ²) $\times 10^{10}$
<i>Before corrections</i>		
Fiducial ^c	26.87 ± 2.43	22.32 ± 1.88
L16-K20	30.53 ± 2.23	26.36 ± 1.70
Y22	15.83 ± 1.17	13.45 ± 0.91
flat+COLDz ^c	14.86 ± 1.14	13.25 ± 1.17
P18	12.34 ± 0.97	10.38 ± 0.77
<i>After corrections</i>		
Fiducial ^c	21.19 ± 2.15	17.55 ± 1.71
L16-K20	24.36 ± 2.25	20.81 ± 1.64
Y22	12.59 ± 1.15	10.61 ± 0.86
flat+COLDz ^c	11.79 ± 1.06	10.49 ± 1.09
P18	9.79 ± 0.91	8.17 ± 0.69

Notes. Top: Values are shown before correcting for any attenuation due to experimental parameters, with uncertainties from raw sample variances. Bottom: Values are shown after applying all corrections to both the signal and the variance. ^aThe quasar model from Torralba-Torregrosa et al. (2024). ^bThe quasar model from Palanque-Delabrouille et al. (2016). ^cFrom C22.

6.1. Modelled luminosity values

We tested the five CO models and two QLF models listed in Sect. 3.1. In each case, we determined the modelled $\langle L'_{\text{CO}} \rangle$ by averaging 50 simulated stack realizations. Each simulated stack consisted of three independent mock COMAP fields (simulated using one of 25 peak-patch realizations each with two random seeds for determining luminosity values), each with the real COMAP hit pattern and masking for that field. Unlike the injection simulations presented in Sect. 4.2.2, we did not pass these mock maps through the COMAP pipeline, instead keeping them as the true CO distribution unadulterated by the pipeline filters (i.e. as $\mathbf{r}_{\text{model}}$ cubes). In each field, we generated a mock catalogue cut to the same number of objects as that field in the real stack ($N_{\text{obj}} = 131, 53, \text{ and } 47$ for Fields 1, 2, and 3), and stacked over that catalogue. This is because quasars tend to occupy more massive DM halos in a region, which are rare (e.g. Richardson et al. 2012; Hopkins et al. 2008; Chowdhary & Chatterjee 2025). If either the cosmic volume covered by the LIM experiment or the number of catalogued objects is much altered, the mock halo selection will go farther into the lower-mass end than is realistic, affecting the output values, despite the somewhat randomized selection process described in Sect. 3.2. The $\langle L'_{\text{CO}} \rangle$ values for each model are listed in Table 1, both before any corrections for experimental sources of signal attenuation and with uncertainties from raw sample variance (Sect. 6.3.2), and after the corrections are made. The uncorrected spectra are shown in Figure 8.

6.2. Sources of signal attenuation

As we have no signal to compare to in the data themselves, we applied any factor that could attenuate signal directly to the models. In each case, we treated the attenuation as a simple multiplicative factor f . The effects are shown on our fiducial CO model in Figures 9.

Table 2. Experimental sources of signal attenuation and variance, and their contribution to the stack.

Description	Factor	Value
<i>Signal attenuation</i>		
Pointing uncertainties	f_{PO}	1.05
Redshift uncertainties	f_{VI}	1.21
<i>Variance</i>		
Model-dependence	d_{MD}	1.07
LIM interloper emission	d_{LI}	1.04

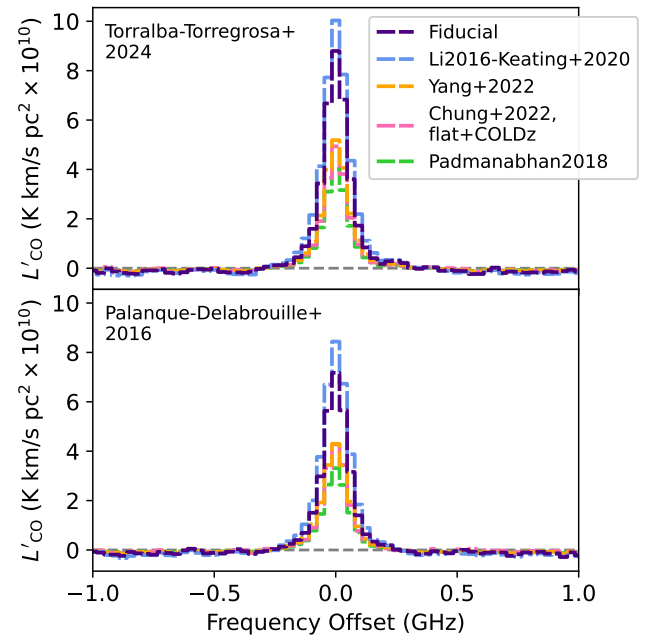


Fig. 8. Spectra for each CO model choice, shown for a quasar catalogue generated using the Ly α -based Torralba-Torregrosa et al. (2024) model (top) and the eBOSS g -band luminosity function Palanque-Delabrouille et al. (2016) (bottom). Spectra are shown before correcting for any of the attenuation factors discussed in Sect. 6.2.

6.2.1. Pointing uncertainty

There is a known issue with pointing in the COMAP Pathfinder telescope leading to systematic offsets between real and assumed source locations. Using a series of bright radio calibrators (including Taurus A, Cassiopeia A, and Cygnus A), these offsets were determined to be systematically $<1'$ in both the azimuthal (az) and elevation (el) directions, with an additional scatter of $<1'$. The az and el offsets are dependent on the absolute az-el being observed. For Field 1 (which is at 0° declination), this translates to slight (sub-arcmin) offset in RA and Dec, with $\sim 0.8'$ scatter. For the other two fields, which are at higher declination, offsets occur in different directions when the fields are rising and setting, spreading the RA-Dec scatter along the RA axis. In each case, any systematic offset and the majority of the scatter is contained within a single COMAP spaxel. We verified this by stacking on radio continuum sources which appear in continuum-leakage maps (Appendix B), finding no systematic offset in the continuum sources and very little additional spread in the signal (Table 2).

This effect can be ignored in the COMAP autocorrelation, but the stack is predicated on the ability to centre the LIM signal on regions of sky containing catalogue objects, and thus will be attenuated by any imprecision in the instrumental pointing. We determined the magnitude of this attenuation by convolving kernels modelling the RA-Dec pointing scatter described above into the maps for each field, noting that these are upper limits on the magnitude of the pointing errors. We found that the effect is indeed negligible. Signal is attenuated by a factor of 1.04 in Field 1 and by a factor of 1.07 in Fields 2 and 3, using these upper limit values. Combining these factors weighted by the number of objects in each field yields an overall attenuation of 1.05.

6.2.2. Catalogue redshift uncertainty

Redshift uncertainty in the galaxy catalogue can attenuate the stacked signal by spreading luminosity outside of the central aperture where signal is expected (Dunne et al. 2025). We adopted the magnitude of the scatter calculated in Dunne et al. (2024): a redshift uncertainty of 1669 km s^{-1} (in addition to a systematic offset, which is accounted for above). We treated the redshift uncertainty directly for each model, by repeating the stack described above (Sect. 6.1), offsetting the velocity of each catalogue offset by a value randomly drawn from a Gaussian distribution with zero mean and standard deviation equivalent to the redshift uncertainty. The signal is attenuated on average across models by $f_{\text{VI}} = 1.21$. Instead of simply dividing the $\langle L'_{\text{CO}} \rangle$ output from each model by this value when accounting for redshift uncertainty, we used the actual $\langle L'_{\text{CO}} \rangle$ value from the stack with the introduced redshift scatter.

6.3. Error budget

Certain experimental and modelling parameters serve not to attenuate potential signal directly, but instead to increase the uncertainty on either the data (beyond the instrumental noise calculated using bootstraps in Sect. 5) or model values. Again, we treated each of these as multiplicative factors, labelling them as d .

6.3.1. Model dependence

Our method of signal extraction makes one key assumption: that the spatial and spectral profile of the real stacked signal matches that of the model. The stack is not a point source, and the shape can vary significantly with the CO and quasar model assumed. In Dunne et al. (2025) we found that the stack profile is fairly well-fit by a double-Gaussian model in the spatial axes, and that the width of each can change by 5–10% with the CO model used, and much more with galaxy catalogues targeting different galaxy types (up to 30% between a $z \sim 3$ LAE model and a $z \sim 3$ quasar model, although there would be considerably less variation between the fairly well-characterized quasar models explored here). It is thus critical to know how the output stack luminosity changes with the adopted model.

The injection simulations are computationally expensive, prohibiting us from determining a template \mathbf{t} for each CO model and applying it directly to the real data. Instead, we tested for model-dependence by scaling the pre-pipeline stack on the fiducial model \mathbf{r}_{fid} to the pre-pipeline model stack for each other model $\mathbf{r}_{\text{model}}$, and then extracting the $\langle L'_{\text{CO}} \rangle$ from the scaled fiducial model $a_{\text{model}} \mathbf{r}_{\text{fid}}$. The output model $\langle L'_{\text{CO}} \rangle$ values calculated using this method differed by at most 6.9%, with the C22 (flat+COLDz) model the most different. We account for this

effect as a multiplicative factor $d_{\text{MD}} = 1.069$ applied directly to the uncertainty on the data value.

6.3.2. Sample variance

In addition to the variance from noise in the COMAP data, which is determinable through bootstrapping and is driving the level of the upper limit we report in Sect. 5, there is an inherent variance in the level of the expected signal in our models. This arises from several sources, including basic cosmic variance, scatter in the CO brightnesses of DM halos, and scatter in the number and mass of DM halos in the vicinity of objects included in galaxy catalogues. This is the difference between the shuffled covariance and true mock covariance discussed by Chen et al. (2025). This variance applies directly to the luminosity of the signal itself, and thus is impossible to account for in our stack value, which detects no luminosity. However, it does apply to the luminosity values returned from our modelling, and will bring down the significance of any discrepancy between the models and our upper limit.

To characterize sample variance for each of the models we test, we calculated the mean absolute deviation (MAD) between the stack $\langle L'_{\text{CO}} \rangle$ value returned from each of 50 independent simulation realizations. The sample variance is sensitive to the amount of cosmic volume probed by the LIM cubes, another reason for our recreation of the exact volume probed by the stack for each simulation realization (including the masking and hit patterns of the actual COMAP data). Regions with fewer hits are downweighted and contribute less to the overall stack luminosity, effectively decreasing the cosmic volume probed by the stack. The final variances are listed as the uncertainties on each model luminosity in Table 1. In this particular stacking configuration, these uncertainty values are of similar magnitude to those from the random noise in the COMAP stack data.

6.3.3. LIM interloper emission

While $z \sim 3$ CO(1–0) emission is fairly isolated in the COMAP bandwidth, interloper spectral line emission from higher- or lower-redshift ensembles of galaxies can still disrupt the CO signal. In the case of COMAP, this interloper emission will come primarily from $z \sim 8$ galaxies emitting in CO(2–1). The exact degree of contamination from this emission is unknown (because CO at high redshifts is largely unconstrained), but it is estimated to be roughly 30% of the luminosity of the $z \sim 3$ CO(1–0) in map space (Chung et al. 2024b). The interloper contamination is heavily dependent on the number of objects being stacked. We found a roughly 20% scatter from a 10% interloper (on a stack on an LAE catalogue) in a 100-object stack in Dunne et al. (2025). We thus repeated this analysis of Dunne et al. (2025) using the exact quasar parameters: we performed 99 simulation realizations with a mock galaxy catalogue generated using the Torralba-Torregrosa et al. (2024) QLF model and 10% LIM interloper contamination, taking 231-object stacks on each. We found that the LIM interloper contamination should add 3.8% scatter. This number is considerably less than the scatter from comparable 100-object stacks in Dunne et al. (2025), due to a combination of the brighter signal from the quasar stack and the smaller aperture being summed over. We thus took $d_{\text{LI}} = 1.038$.

7. Interpretations

In this work, we present stacks of CO emission from COMAP on the positions of quasars from eBOSS and DESI. The stack is a

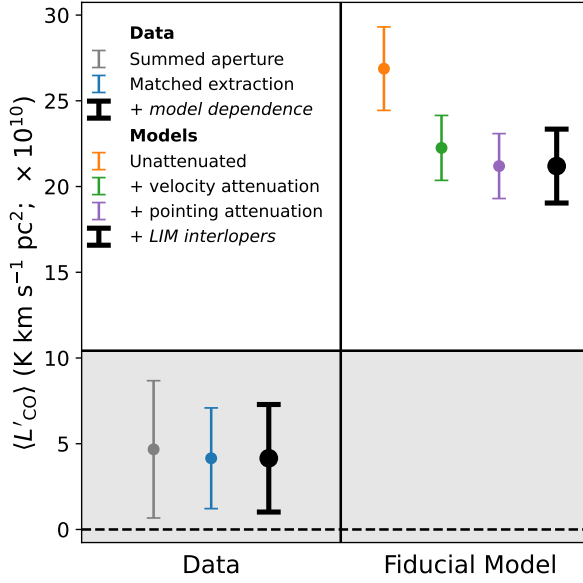


Fig. 9. Factors attenuating signal or increasing variance, shown for the upper limit from the COMAP-eBOSS/DESI stack and the fiducial model from C22. The corrections applied to the other models are qualitatively similar. The shaded region indicates the 2σ upper limit from the data alone. The various corrections are outlined in Sect. 6.

biased tracer of the overall CO luminosity, probing (in this case) the average CO luminosity of an $18 \times 18 \times 9$ cMpc region around surveyed quasars. We do not detect any cosmic CO in the stack, and the upper limit we place on the stacked signal falls below the modelled stack luminosity for simulations incorporating several different CO and quasar models. We have attempted a comprehensive census of possible sources of signal loss and uncertainty (Figure 9), which brings the modelled luminosity values closer to the limit placed by the COMAP-eBOSS/DESI stack, but still leaves a 3.3σ disagreement between the two using our fiducial modelling parameters (the UM+COLDz+COPSS model from Chung et al. 2022 and the Torralba-Torregrosa et al. (2024) quasar model). We show the disagreement as a simple S/N for this model and our suite of other models in Figure 10. Stacks on the L16-K20 CO model are discrepant with our upper limit to $>3\sigma$ for both choices of quasar luminosity function.

Because the stack is not a straightforward tracer of cosmic CO emission, interpretation of this result is somewhat complicated. The stack does not directly probe the average CO emission in the universe as a whole. It also does not directly probe the average CO emission from the galaxies hosting the quasars. In general the contribution of the objects surrounding the catalogued objects to the stacked luminosity is far greater than the stacked luminosity of the objects themselves (Dunne et al. 2025). For higher-mass DM halos such as those containing quasars this difference is not as marked as it is for lower-mass halos (roughly a factor of 6 for our fiducial modelling), but the host galaxies are still a minor factor. For the purposes of discussion, we also assume that all experimental considerations affecting $\langle L'_{\text{CO}} \rangle$ have been accounted for in Sect. 6, although there is always a possibility that our result could be indicative of an unaccounted-for experimental systematic error. We break the remaining potential interpretations into three categories: those affecting the cosmological clustering around the quasars, environmentally dependent baryonic physics that are not accounted for in the simplistic modelling we use, and fainter average CO emission.

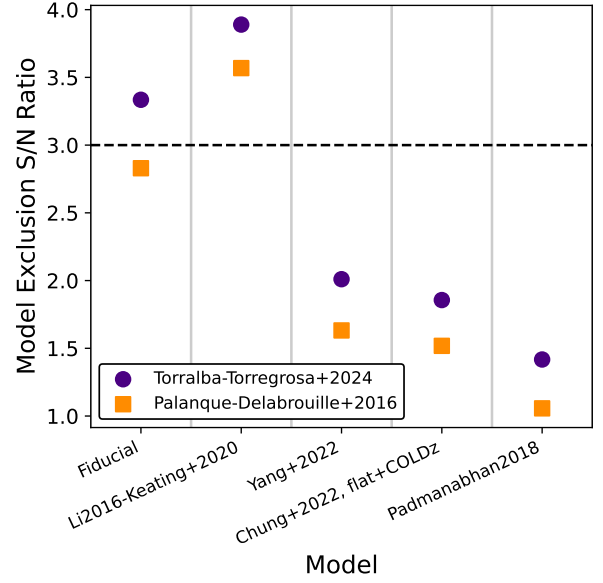


Fig. 10. S/N at which models are discrepant with the COMAP-eBOSS/DESI upper limit for various models of CO emission, including all potential sources of uncertainty and signal attenuation listed in Sect. 6.

7.1. Clustering of halos

Because the stack luminosity is dominated by clustering effects, it is extremely sensitive to the number and mass (which translates nearly directly to luminosity in our models) of the DM halos neighbouring those containing the quasars. In our modelling, this effect is treated through a combination of the QLF, onto which we abundance-match our simulated halos, and the observational cuts (in particular the cutoff luminosity L'_{cut}) used to generate the actual mock catalogue being stacked upon. With our default parameter choices, we find roughly 250 DM halos fall into the stack aperture, in addition to the actual quasar host. The QLF is fairly well-characterized (e.g. Shenhav et al. 2020; Chaussidon et al. 2023), and statistics other than the QLF have also shown quasars to be highly clustered (White et al. 2012; Timlin et al. 2018), so it is unlikely that our modelling is drastically incorrect. Still, small differences in the QLF can significantly affect the modelled stack luminosity (there is an average 18% discrepancy between the two luminosity functions we test here). The model exclusion by our upper limit is tentative, and could vanish entirely with different choices of quasar modelling.

Modelling errors in our determination of the quasar velocity offsets could also change the number of halos falling into the stack aperture, by changing their systemic velocities. While we have attempted to account for the large bulk motions of gas into and out of quasars (Sect. 6.2.2), CO observations of quasars $z \sim 3$ are difficult to obtain, and our determination of the scatter in the relationship between the CO and optically measured velocities is based on only 12 objects, measured using multiple CO lines. If the scatter is being underestimated, that would drive the modelled luminosities down and reduce the tension between the models and the data.

7.2. Feedback and environmental effects

The models we use to compare with the stack luminosity in this work are extremely simple. They assume a very basic galaxy-halo connection, taking only the DM mass of the halo (and pos-

sibly its redshift) into consideration when assigning luminosities for both the CO and the quasar tracers. Any other physics is ignored. In the context of a discrepancy between stack and model luminosities, any ignored physics which could result in galaxies in dense environments, or galaxies around quasars, having decreased CO emission, could explain our results.

Such environmental dependence is plausible. Galaxy properties such as SFR are well-known to correlate with their local density, potentially in a manner independent of their halo mass (e.g. Park et al. 2007; Peng et al. 2010). Traditionally, this environmental-dependence has been thought to matter most at small scales, but dependence on scales of 1–10 Mpc has been observed up to $z \sim 1$. (Park et al. 2007; Wang et al. 2018). Quasars preferentially occupy dense environments; if environmental quenching is occurring in these environments, it would bring our modelled stack luminosities down. Additionally, quasar-specific feedback effects such as jets (which could extend up to several megaparsecs and have been observed at redshifts as high as $z \sim 5$; Oei et al. 2024; Gloude-mans et al. 2025) could quench galaxies and disrupt CO emission around the eBOSS/DESI objects.

This sensitivity to large-scale environmental effects is a unique strength of the stack. While there are currently multiple degenerate explanations for the discrepancy between stack models and data, these degeneracies can be broken in the future with ancillary information. Once the CO autocorrelation power spectrum is measured, for example, this could be used to constrain the CO luminosity directly and any remaining discrepancy between models and data could be explained by environmental factors.

7.3. Average CO luminosity

The final potential explanation for the discrepancy between the data and models is that some of the models we present for cosmic CO are simply too bright. We note that this is counter to some tentative existing LIM-based evidence which has been presented favouring the L16-K20 model (suggesting bright cosmic CO, Keating et al. 2020; Chung et al. 2024a).

While considerable effort has been dedicated to realistically modelling the CO luminosity of DM halos as a function of their mass, only a very narrow mass range of halos have actually been observed in CO at $z \sim 3$ (e.g. Riechers et al. 2019), and the faint end of the luminosity function has had to be largely inferred from prior empirical models connecting halo properties to star formation and stellar mass. As a result, the models considered here make stronger assumptions than the observations from which they derive. Being in tension with the models thus does not necessitate being in tension with the basic observations.

For the two CO models that are in tension with the observed COMAP stack, it is likely true that the COMAP stack is not in tension with the observations on which the models are based. The C22 fiducial CO model is based on the COLDz luminosity function, which is fit to observations that are complete only to $L'_{\text{CO}} \sim 10^{10} \text{ K km s}^{-1} \text{ pc}^2$ at $z \sim 2.5$ (Pavesi et al. 2018). There remains a large uncertainty in the faint-end slope. If this uncertainty is causing an overestimate of the luminosity of the faint-end objects, this would lead in turn to an overestimate in the stack flux without causing any discrepancy with the actual COLDz observations. The Kamenetzky et al. (2016) observations, on which the L16-K20 model is based, are more complete (down to $L'_{\text{CO}(1-0)} \sim 10^8 \text{ K km s}^{-1} \text{ pc}^2$), but are also primarily of objects in the local universe, at much later times than those being observed by COMAP ($z \lesssim 0.1$). In this model, the CO(1–0) luminosity is extrapolated from an assumed FIR luminosity

using a relationship based on these observations, so a tension between observed stack luminosity and the L16-K20 stack luminosity could be explained by an evolution in the CO-FIR relationship with redshift. We note that existing evidence suggests such an evolution may be mild (Genzel et al. 2010; Greve et al. 2014), although observations at $z \sim 2-3$ are extremely lacking.

The remaining three models we consider are not formally in tension with our stack result: Y22, C22 flat+COLDz, and P18. We note that these either assume little contribution to the overall CO luminosity of the Universe by the faint end of the luminosity function (Y22, C22 flat+COLDz), or invoke a duty cycle that assumes that only a percentage of halos which could emit in CO actually are emitting (P18).

Overall, the tension that we observe between stack models and the COMAP quasar stack could be explained by current models of CO emission over-estimating the contribution from faint galaxies, which are extremely difficult to observe in quantity with traditional galaxy survey techniques at $z \sim 3$. Sensitivity to these faint objects and their abundance (or lack thereof) is a unique strength of LIM in general. However, as enumerated above, many degenerate effects other than a fainter faint end may also cause the stack luminosity to be fainter than expected. Each of these would need to be further constrained by observations in order to use the stack to draw conclusions about the faint-end CO luminosity.

8. Summary and conclusions

In this work we have presented an update to the COMAP-eBOSS stack of Dunne et al. (2024). In addition to the factor of three improvement in sensitivity in the stack that results from the increase in COMAP data since the ES data release, we have refined the stack methodology considerably, through a template-based matched filter. We placed a 2σ upper limit on the CO emission surrounding 231 quasars from eBOSS and DESI of $\langle L'_{\text{CO}} \rangle = 10.0 \times 10^{10} \text{ K km s}^{-1} \text{ pc}^2$. We compared this to models of the stack signal, varying both the prescription for CO emission and the QLF. We find that the COMAP-eBOSS/DESI upper limit excludes our brightest model of CO emission stacked on quasars (L16-K20/Torralba-Torregrosa et al. 2024) to $>3\sigma$ even after accounting for experimental attenuation and variance from factors including sample variance, pipeline and experimental signal loss, velocity attenuation, model-dependence of the $\langle L'_{\text{CO}} \rangle$ extraction, and interloper emission in the LIM data. The COMAP fiducial/Torralba-Torregrosa et al. (2024) model is also strongly disfavoured at 3.3σ .

However, a wide variety of disparate quasar and galaxy properties affect the stack in addition to the cosmic CO luminosity itself. This is because the stack is not a measure of the true average CO luminosity of the COMAP fields, instead probing only the CO luminosity of the regions surrounding quasars observable by eBOSS or DESI. While we rule out models for the stack emission, which combine CO and quasar modelling, we are thus not able to confidently rule out any model for cosmic CO emission from the stack alone. To truly isolate the cause of the observed discrepancy between data and models will require more detailed modelling and a much broader suite of LIM statistics to work from, including stacks on different galaxy tracers subject to independent environmental considerations, access to correlation data on more spatial scales from a cross-correlation power spectrum between the LIM data and the galaxy surveys, and ultimately a detection of cosmic CO in a full autocorrelation power spectrum as a comparison point. With these ancillary data points,

the stack will be a powerful probe of large-scale environmental effects affecting galaxies.

Acknowledgements. This material is based upon work supported by the National Science Foundation under Grant Nos. 1517108, 1517288, 1517598, 1518282, 1910999, and 2206834, as well as by the Keck Institute for Space Studies under ‘The First Billion Years: A Technical Development Program for Spectral Line Observations’. JG acknowledges support from the Keck Institute for Space Science, NSF AST-1517108 and University of Miami. GAH acknowledges the funding from the Dean’s Doctoral Scholarship by the University of Manchester. SM acknowledges support from the U.S. Department of Defense through the Science, Mathematics, and Research for Transformation (SMART) Scholarship-for-Service Program. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00340759). This research was funded by the 2025 KAIST-U.S. Joint Research Collaboration Open Track Project for Early-Career Researchers, supported by the International Office at the Korea Advanced Institute of Science and Technology (KAIST). This research made use of the SDSS-IV eBOSS survey. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is www.sdss4.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics | Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University. This research used data obtained with the Dark Energy Spectroscopic Instrument (DESI). DESI construction and operations is managed by the Lawrence Berkeley National Laboratory. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High-Energy Physics, under Contract No. DE-AC02-05CH11231, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract. Additional support for DESI was provided by the U.S. National Science Foundation (NSF), Division of Astronomical Sciences under Contract No. AST-0950945 to the NSF’s National Optical-Infrared Astronomy Research Laboratory; the Science and Technology Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Humanities, Science and Technology of Mexico (CONAHCYT); the Ministry of Science and Innovation of Spain (MICINN), and by the DESI Member Institutions: www.desi.lbl.gov/collaborating-institutions. The DESI collaboration is honored to be permitted to conduct scientific research on I’oligam Du’ag (Kitt Peak), a mountain with particular significance to the Tohono O’odham Nation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. National Science Foundation, the U.S. Department of Energy, or any of the listed funding agencies.

References

- Adame, A. G., Aguilar, J., Ahlen, S., et al. 2025, *JCAP*, **2025**, 017
- Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, *ApJS*, **249**, 3
- Aravena, M., Decarli, R., González-López, J., et al. 2019, *ApJ*, **882**, 136
- Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, *MNRAS*, **488**, 3143
- Bond, J. R., & Myers, S. T. 1996, *ApJS*, **103**, 1
- Breyse, P. C., & Alexandroff, R. M. 2019, *MNRAS*, **490**, 260
- Carilli, C., & Walter, F. 2013, *ARA&A*, **51**, 105
- CCAT-Prime Collaboration (Aravena, M., et al.) 2023, *ApJS*, **264**, 7
- Chaussidon, E., Yèche, C., Palanque-Delabrouille, N., et al. 2023, *ApJ*, **944**, 107
- Chen, Z., Cunningham, S., Pourtsidou, A., et al. 2025, *ApJS*, **279**, 19
- Cheng, C., Parsons, A. R., Kolopanis, M., et al. 2018, *ApJ*, **868**, 26
- Chowdhary, A., & Chatterjee, S. 2025, *ApJ*, **992**, 21
- Chung, D. T., Breyse, P. C., Cleary, K. A., et al. 2022, *ApJ*, **933**, 186
- Chung, D. T., Breyse, P. C., Cleary, K. A., et al. 2024a, *A&A*, **691**, A337
- Chung, D. T., Chluba, J., & Breyse, P. C. 2024b, *Phys. Rev. D*, **110**, 023513
- Cleary, K. A., Borowska, J., Breyse, P. C., et al. 2022, *ApJ*, **933**, 182
- CONCERTO Collaboration (Ade, P., et al.) 2020, *A&A*, **642**, A60
- Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, *AJ*, **115**, 1693
- Crites, A. T., Bock, J. J., Bradford, C. M., et al. 2014, in *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*, eds. W. S. Holland, & J. Zmuidzinas, *SPIE Conf. Ser.*, **9153**, 91531W
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, **145**, 10
- Delubac, T., Raichoor, A., Comparat, J., et al. 2017, *MNRAS*, **465**, 1831
- DESI Collaboration (Abdul-Karim, M., et al.) 2025, arXiv e-prints [arXiv:2503.14745]
- Doré, O., Bock, J., Ashby, M., et al. 2014, arXiv e-prints [arXiv:1412.4872]
- Dunne, D. A., Cleary, K. A., Breyse, P. C., et al. 2024, *ApJ*, **965**, 7
- Dunne, D. A., Cleary, K. A., Breyse, P. C., et al. 2025, *A&A*, **702**, A247
- Fronenberg, H., & Liu, A. 2024, *ApJ*, **975**, 222
- Genzel, R., Tacconi, L. J., Gracia-Carpio, J., et al. 2010, *MNRAS*, **407**, 2091
- Glouemans, A. J., Sweißen, F., Morabito, L. K., et al. 2025, *ApJ*, **980**, L8
- Greve, T. R., Leonidaki, I., Xilouris, E. M., et al. 2014, *ApJ*, **794**, 142
- Harris, D. W., Jensen, T. W., Suzuki, N., et al. 2016, *AJ*, **151**, 155
- Hill, R., Chapman, S. C., Scott, D., et al. 2019, *MNRAS*, **485**, 753
- Hill, G. J., Lee, H., MacQueen, P. J., et al. 2021, *AJ*, **162**, 298
- Hill, G. J., Lee, H., Good, J. M., et al. 2024, in *Ground-based and Airborne Instrumentation for Astronomy X*, eds. J. J. Bryant, K. Motohara, & J. R. D. Vernet, *SPIE Conf. Ser.*, **13096**, 130960F
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *ApJS*, **208**, 19
- Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, *ApJS*, **163**, 1
- Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, *ApJS*, **175**, 356
- 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., & Keenan, R. P. 2020, *ApJ*, **901**, 141
- Keenan, R. P., Keating, G. K., & Marrone, D. P. 2022, *ApJ*, **927**, 161
- 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, T. Y., Wechsler, R. H., Devaraj, K., & Church, S. E. 2016, *ApJ*, **817**, 169
- Lunde, J. G. S., Stutzer, N. O., Breyse, P. C., et al. 2024, *A&A*, **691**, A335
- Lunde, J. G. S., Breyse, P. C., Chung, D. T., et al. 2025, *A&A*, submitted, [arXiv:2510.23502]
- Lyke, B. W., Higley, A. N., McLane, J. N., et al. 2020, *ApJS*, **250**, 8
- Madau, P., & Dickinson, M. 2014, *ARA&A*, **52**, 415
- Mansfield, E. M., Dunne, D. A., & Chung, D. T. 2025, *ApJ*, submitted, [arXiv:2512.22356]
- MeerKLASS Collaboration, A., Barberi-Squarotti, M., et al. 2025, *MNRAS*, **537**, 3632
- Muñoz-Elgueta, N., Arrigoni Battaia, F., Kauffmann, G., et al. 2022, *MNRAS*, **511**, 1462
- Oei, M. S. S. L., Hardcastle, M. J., Timmerman, R., et al. 2024, *Nature*, **633**, 537
- Padmanabhan, H. 2018, *MNRAS*, **475**, 1477
- Palanque-Delabrouille, N., Magneville, C., Yèche, C., et al. 2016, *A&A*, **587**, A41
- Park, C., Choi, Y.-Y., Vogeley, M. S., et al. 2007, *ApJ*, **658**, 898
- Pavesi, R., Sharon, C. E., Riechers, D. A., et al. 2018, *ApJ*, **864**, 49
- Peng, Y.-J., Lilly, S. J., Kovač, K., et al. 2010, *ApJ*, **721**, 193
- Richardson, J., Zheng, Z., Chatterjee, S., Nagai, D., & Shen, Y. 2012, *ApJ*, **755**, 30
- Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, *ApJ*, **872**, 7
- Sanders, D. B., Soifer, B. T., Elias, J. H., et al. 1988, *ApJ*, **325**, 74
- Schechter, P. 1976, *ApJ*, **203**, 297
- Shen, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2020, *MNRAS*, **495**, 3252
- Solomon, P. M., Downes, D., Radford, S. J. E., & Barrett, J. W. 1997, *ApJ*, **478**, 144
- Somerville, R. S., Popping, G., & Trager, S. C. 2015, *MNRAS*, **453**, 4337
- Stein, G., Alvarez, M. A., & Bond, J. R. 2019, *MNRAS*, **483**, 2236
- Stutzer, N. O., Lunde, J. G. S., Breyse, P. C., et al. 2024, *A&A*, **691**, A336
- Timlin, J. D., Ross, N. P., Richards, G. T., et al. 2018, *ApJ*, **859**, 20
- Torralba-Torregrosa, A., Renard, P., Spinoso, D., et al. 2024, *A&A*, **690**, A388
- Vieira, J., Aguirre, J., Bradford, C. M., et al. 2020, arXiv e-prints [arXiv:2009.14340]
- Wang, H., Mo, H. J., Chen, S., et al. 2018, *ApJ*, **852**, 31
- White, M., Myers, A. D., Ross, N. P., et al. 2012, *MNRAS*, **424**, 933
- Yang, S., Popping, G., Somerville, R. S., et al. 2022, *ApJ*, **929**, 140

Appendix A: Signal injection boost strength

We determine pipeline signal loss and extract the stack $\langle L'_{\text{CO}} \rangle$ using injection simulations, where we inject simulated CO LIM cubes into subsets of the full COMAP data, boosting the injected signal as we do. As discussed in §4.2, we use a boost strength of 50. This yields high S/N ratios in the injected signal when stacked and averaged over 100 iterations (S/N of 49-67 in the central voxel, depending on field). Here, we justify that this boost is also small enough not to cause non-linearities in any of the pipeline steps, or to interact with existing signal in the data.

Ensuring that the boosted injected signal will not affect the TOD pipeline is relatively simple. We find that the brightest voxel of the simulated emission is still roughly four orders of magnitude below the system temperature even after boosting by a factor of 50. The simulated TOD signal will thus be well below the noise, and the TOD pipeline will not be affected by the injection of this signal. The other potentially non-linear pipeline step is the PCA filter applied after map-making (Lunde et al. 2024). The map S/N is greatly increased compared to the S/N of the TOD, so it is possible for the signal to be in the non-linear regime of this filter even when it is well below the noise in the TOD. Lunde et al. (2024) show how the filter behaves with the voxel-level S/N of the maps, $\langle \text{SNR}_{\text{vox}} \rangle$. They find that the map-level PCA filter behaves linearly with respect to the CO signal up to $\langle \text{SNR}_{\text{vox}} \rangle = 0.02$. The boosted, signal-injected data subsets we use in this work have $\langle \text{SNR}_{\text{vox}} \rangle \sim 0.005$, so we are well within this linear regime. We also note that our best estimates for the real signal will be well within the linear regime of this filter. The unboosted fiducial CO model would be at a $\langle \text{SNR}_{\text{vox}} \rangle = 2 \times 10^{-4}$ in each subset of the full COMAP S2 dataset, and this rises to $\langle \text{SNR}_{\text{vox}} \rangle = 0.002$ after averaging the 100 subsets together. We therefore do not expect either the signal loss analysis or the stack itself to be affected by pipeline nonlinearities. As an additional check that there is no signal being removed by the map-level PCA filter, we stack on the PCA modes removed by the filter. We find no signal.

We also note that signal-injection simulations have been shown to give inaccurate results when there are interactions between real and injected signal (Cheng et al. 2018). This is not an issue when the injected signal is considerably stronger than the real signal in the data. While we do not know the actual strength of the signal in the COMAP data, this work suggests that it is likely weaker than the L16-K20 model, and may be weaker than the COMAP fiducial model. Assuming the fiducial model is correct, the injected signal will be on average 50× brighter than the real signal in the data cubes. Stacking introduces an additional prior on the position of the signal, so even if the two were of equal brightness the real signal would be at least 9.5× fainter than the injected signal in a stack on the injection catalogue. This means that overall, the real signal must be at least 500× fainter than the injected, so we do not expect biasing due to the injection-based simulation strategy.

Appendix B: Pointing verification using continuum radio sources

We verify the COMAP pointing and the stack coordinate system by performing a stack on known radio continuum sources from the NRAO VLA Sky Survey (NVSS; Condon et al. 1998). The COMAP pipeline is designed to reject continuum, and does so to a high degree of significance (Lunde et al. 2024). However, continuum-leakage maps can be generated by purposefully accounting for the system temperature spikes incorrectly dur-

ing time-domain processing (Lunde et al. 2025). These maps can then be averaged across the frequency axis to give continuum. While the flux densities in these maps are extremely difficult to calibrate, the spatial distribution of any emission present is accurate. Visible in these maps are clear signatures of the Cosmic Microwave Background, as well as several bright radio continuum sources (Figure B.1).

There are 22 sources from NVSS that fall within the COMAP fields. We perform a stack on these 22 sources, using the frequency-averaged map. The resulting angular profile is shown in Figure B.2. We find a source profile that is somewhat more spread-out than the 4.5' Gaussian to be expected from stacking on sources exactly centred in their spaxels and broadened only by the primary beam, but which still fits well within $3' \times 3'$. This confirms the COMAP pointing accuracy.

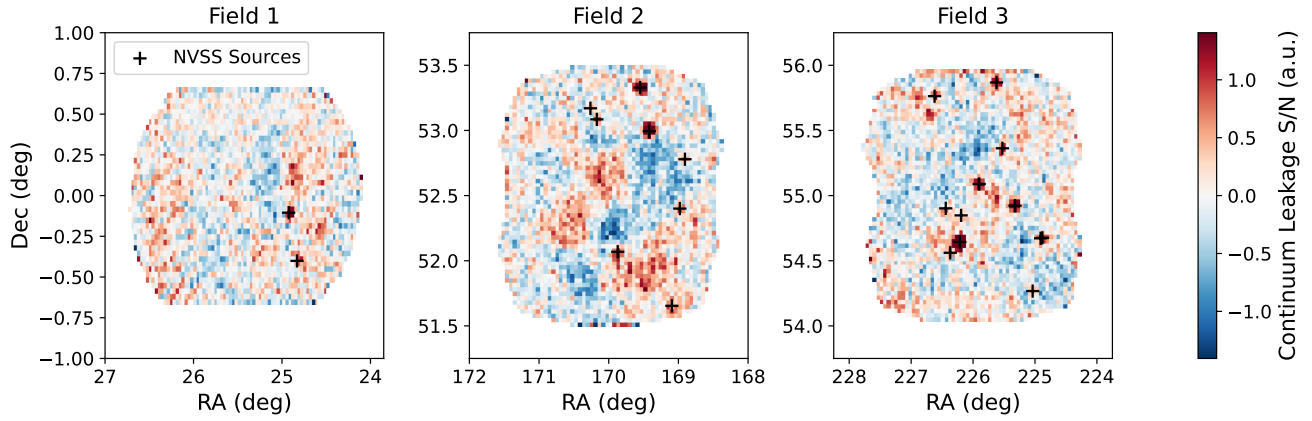


Fig. B.1. Continuum-leakage maps of the three COMAP fields, with radio continuum sources from NVSS (Condon et al. 1998) overplotted.

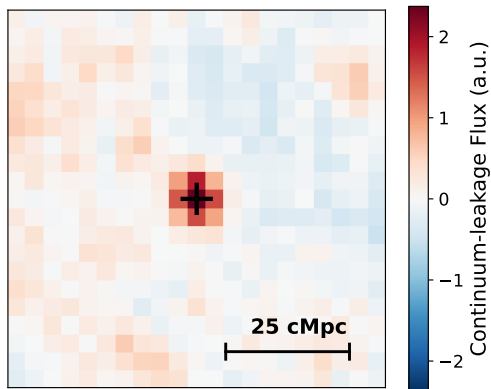


Fig. B.2. Stack of 22 NVSS radio sources on the COMAP continuum-leakage maps. The resulting image is roughly beam-sized, indicating good pointing accuracy. The scale bar is shown for $z = 2.62$ (the median redshift of the eBOSS/DESI sources). It corresponds to $\sim 14'$.