# Joint analysis techniques for line intensity mapping, with application to the CO Mapping Array Project

Thesis by
Delaney Anne Dunne

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy in Astrophysics

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2026
Defended May 21, 2026

Delaney Anne Dunne
ORCID: 0000-0002-5223-8315

# ACKNOWLEDGEMENTS

I firmly believe that science is better when shared, and I have been so lucky in the people I was able to share the science in this thesis with. I have benefited immeasurably from the knowledge of my advisor Dr. Kieran Cleary, who has been an incredible supervisor and mentor, and has been nothing but supportive throughout my entire PhD. Kieran has taught me so much not just about science, but also invaluable and unquantifiable scientific skills like project management, mentorship, and leadership. He is extremely skilled at asking exactly the right questions about a project, something I hope I will be able to take into the rest of my career. I am so thankful to him for his guidance.

I have also had many other mentors and collaborators within COMAP, and I am deeply grateful to all of them. I would like to thank Prof. Dongwoo Chung for essentially taking me under their supervision—for always making time to answer my questions, whether about research or career advice, for always being willing to chat about science, and for being unquestionably on my side. Prof. Patrick Breysse was also an invaluable resource throughout my PhD. Thank you to Prof. Junhan Kim for babysitting me through an instrumentation project, and to Prof. Tony Readhead for writing me many letters of recommendation, sometimes on very late notice. I am also very grateful to Dr. Jonas Lunde and Dr. Nils-Ole Stutzer, who were my allies across the world throughout the PhD process. Our many shared adventures and science rabbit holes made conferences and collaboration meetings so much better, and it was an honor to be the COMAP graduate students together. Finally, thank you to all of the students I've worked with: Ashiria, James, Ella, Finneas, Megan, MinSu, and Hee Beom, for trusting me and being willing to take on my science, and being patient with me as I learned how to mentor.

Outside of COMAP, I have also been surrounded by many wonderful, kind scientists. Thank you to Prof. Gregg Hallinan, my faculty advisor, who always made time for me when I needed him, and to the rest of the Hallinan group, for welcoming me into their research meetings. Thank you to the HETDEX team, in particular Dr. Erin Mentuch Cooper for her unwavering support of my efforts within HETDEX. I am also so grateful to the other scientific communities I've found, particularly the ObsCos group at Caltech and the LIM community globally.

Community has been a theme throughout my time at Caltech. This started with all of the other graduate students in astronomy—the sense of community I felt among the graduate students here was part of what brought me to Caltech, and my PhD was so much more rewarding for having been surrounded with each of them. I have also had an incredible community through Dr. Cameron Hummels and the Astronomy Outreach program, where I had a lot of fun, saw lots of California, and critically learned a lot about communicating science. I learned so much about solidarity from each of the organizers involved in UAW 2478 (the graduate student and postdoc union at Caltech), and I'm so proud of what we achieved together during my time here. The Caltech Glee Club and Chamber Singers (and in particular Nancy Sulahian) gave me a demanding and rewarding artistic outlet that was so valuable throughout my PhD.

Finally, I am so, so grateful for the unwavering support system I have had in my friends and family. My first year here would have been extremely different without my cohort—Emily Silich, Katelyn Horstman, Jessie Miller, and Isabel Sands—my classes were so much more enjoyable, and I got so much more out of them, thanks to each of them. I owe so much to my roommates and best friends Dr. Sam Ponnada and Jakob Faber (and Cannoli), who very quickly became my California family. I do not have the words to thank my partner, Kelcey Wallens-Logan, who has made the past year one of the best of my life. On two different occasions, I realized the solution to a major scientific roadblock for the research in this thesis through rambling at her about what I was working on, and that's the least of what she has given me. Thank you to my brother Shane and my sister Kate, both of whom I have always known I can count on when I need to. Finally, none of this would have been possible without my parents, who have encouraged, supported, and loved me since the very beginning. I came into this PhD knowing all of the really important things, since they taught them to me. Everything I have learned since has been a bonus.

# ABSTRACT

Line intensity mapping (LIM) is an emerging observational technique with broad applications in cosmology and galaxy evolution. Rather than cataloging galaxies as discrete objects, LIM experiments measure the spatial fluctuations of galaxy populations across large three-dimensional volumes. Thus, LIM is able to measure the contribution to the overall spectral-line luminosity from all galaxies in a population, even the very faintest. LIM began with the 21 cm neutral hydrogen line but has recently been expanding to other tracers, such as CO, [CII], H$\alpha$, and Ly$\alpha$. The first generations of dedicated experiments targeting these tracers are beginning to come online. In order to draw science conclusions from their observations, these experiments must address challenges in the form of sensitivity and systematic errors.

In this thesis, I develop an analysis technique designed to mitigate these challenges by stacking LIM data on the positions of galaxies in an external galaxy catalog. I use a custom-designed simulation pipeline to produce simultaneous mock observations of a LIM experiment and galaxy survey covering the same galaxy population, and use these simulations to build a theoretical understanding of this stacking statistic. I find that cosmological clustering of galaxies around the objects in the reference galaxy catalog drives the stacked signal.

I also use the stack to draw conclusions about molecular gas in the period preceding the peak of cosmic star formation, using data from the CO Mapping Array Project (COMAP), the leading CO LIM experiment. I use eBOSS and DESI quasars, as well as Ly$\alpha$ emitters from HETDEX, as external reference catalogs. I do not detect any cosmic CO emission in the COMAP stacks, and I find a moderate tension between the resulting upper limits and the brightest models of the stacked emission. This tension could be attributable to non-linear radiative transfer or clustering properties of the galaxy catalogs, but it could also be due to models for cosmic CO overpredicting the average CO luminosity $z \sim 3$.

Finally, I search for spectral-line emission from axionic dark matter by stacking the DESI and HETDEX galaxies over a range of rest frequencies. I constrain the axion decay rate $\Gamma$ and the axion-photon coupling constant $|g_{a\gamma\gamma}|$. In the axion mass range probed by COMAP, I establish the tightest constraints available on an axion model of dark matter.

# PUBLISHED CONTENT AND CONTRIBUTIONS

Dunne, Delaney A. et al. (Mar. 2026). “COMAP Pathfinder – Season 2 results: IV. A stack on eBOSS/DESI quasars.” In: *Astronomy & Astrophysics* 707, A256, A256. DOI: 10.1051/0004-6361/202557864. arXiv: 2510.23568 [astro-ph.CO].

Dunne, Delaney A. et al. (Oct. 2025). “Three-dimensional stacking as a line intensity mapping statistic.” In: *Astronomy & Astrophysics* 702, A247, A247. DOI: 10.1051/0004-6361/202554780. arXiv: 2503.21743 [astro-ph.CO].

Mansfield, Ella M., Delaney A. Dunne, and Dongwoo T. Chung (Dec. 2025). “Clustering Confuses Spectro-photometry: An Investigation of 2D and 3D Forced Profile Matching for Stacking Line-intensity Mapping Data on Source Catalogues.” In: *arXiv e-prints (The Astrophysical Journal Accepted)*, arXiv:2512.22356. DOI: 10.48550/arXiv.2512.22356. arXiv: 2512.22356 [astro-ph.CO].

Dunne, Delaney A. et al. (Apr. 2024). “COMAP Early Science. VIII. A joint stacking analysis with eBOSS quasars.” In: *The Astrophysical Journal* 965.1, 7, p. 7. DOI: 10.3847/1538-4357/ad2dfc. arXiv: 2304.09832 [astro-ph.GA].

Delaney Dunne is the leading author on Dunne et al., 2024, Dunne et al., 2025, and Dunne et al., 2026c. Delaney conceived or was involved in the conception of each project, and developed the analysis pipelines to be applied to the existing reduced datasets. In Dunne et al., 2025, Delaney developed the mock observation simulation pipeline used, Delaney led the science analysis and writing of the manuscripts for each work. For Mansfield et al., 2025, Delaney was the second author, and co-supervised the leading author Ella Mansfield with Prof. Dongwoo Chung. Delaney was involved in the conception of the project, helped to develop the analysis pipeline used, and performed the bootstrapping analysis. Delaney was also involved in the rest of the subsequent science and the preparation of the manuscript.

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

*Chapter 1*

# INTRODUCTION

Progress in cosmology and galaxy formation requires detecting large numbers of faint galaxies over large cosmic volumes. This is a fundamental observational limitation, as any galaxy survey has an inherent luminosity limit below which galaxies cannot be detected, so the faintest galaxies are inevitably neglected in observations. When attempting to draw conclusions about the overall galaxy population, these faint galaxies' properties are either ignored or extrapolated over. This problem compounds further back in the Universe's history, due to the inverse-square relation of luminosity with distance. Beyond a redshift of $z \sim 2$, our understanding of the evolution of galaxies is largely based on detailed studies of individual bright objects; the large galaxy surveys that are driving modern-day astronomy do not extend that far back in time (e.g., DES Collaboration et al., 2022; DESI Collaboration et al., 2025c).

Over the past decade, an approach for addressing this 'completeness problem' has been emerging in the form of line intensity mapping (LIM). In this technique, galaxy populations are measured as three-dimensional (3D) fluctuation fields tracing a specific spectral line over a large cosmic volume, rather than as discrete cataloged objects (for reviews, see Kovetz et al., 2019; Chang and Lidz, 2026). A simulated LIM observation is shown in Figure 1.1, in comparison to the underlying large-scale structure (LSS) and a galaxy catalog covering the same cosmic volume. LIM fluctuations contain the integrated emission from all galaxies in a given cosmic volume, and LIM experiments are thus able to survey large volumes extremely efficiently, capturing emission from the faint objects that are not observable by traditional galaxy surveys. This has wide-reaching applications for galaxy formation studies, particularly for censuses of the bulk properties of some galaxy formation tracer across cosmic time, and for cosmology, where LIM has the potential to be a powerful tracer of the matter power spectrum either in autocorrelation or as a cross-correlation partner for some other observable (Karkare and Bird, 2018; Karkare et al., 2022a).

LIM initially developed in the context of the 21 cm neutral hydrogen line. Outside of the 21 cm field, LIM is a new but fast-growing field, with the first generations of

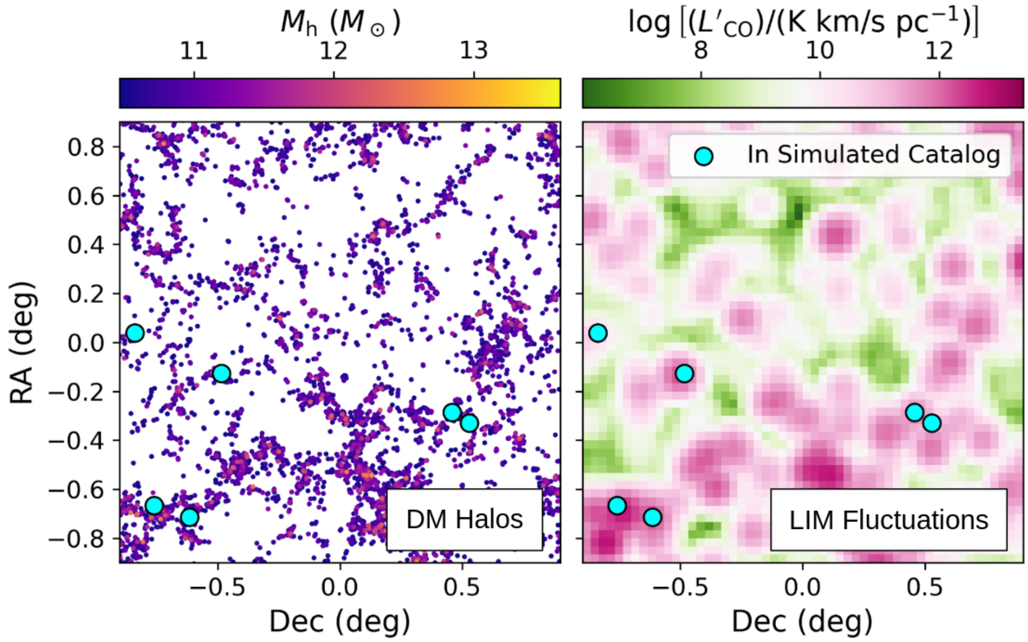


Figure 1.1: A simulated population of dark matter halos (left) with a mock of the resulting $z = 3$ LIM signal in CO(1–0) emission (center). The halos that would be observable by a mock galaxy survey are in blue. Figure adapted from Dunne et al. (2025).

dedicated experiments observing only now (Cleary et al., 2022; Staniszewski et al., 2014; Doré et al., 2014). In addition to the challenge inherent in drawing science conclusions from the (purposefully-)confused LIM datasets, LIM analyses are faced with two other primary challenges: (1) raw sensitivity (particularly for experiments targeting the CO and [CII] lines, which are faint compared to instrumental and atmospheric noise), and (2) systematic errors, which could have instrumental (Lunde et al., 2024; Chen et al., 2025) or astrophysical (Sun et al., 2021) origins.

This thesis comprises a series of efforts to address these challenges through combining LIM data from the CO Mapping Array Project (COMAP; Cleary et al., 2022) with external information from galaxy catalogs. The goals of this research are twofold: firstly, to draw conclusions about the assembly of star-forming material in galaxies at $z \sim 3$, and secondly to establish robust analysis strategies for the future of LIM (facilitating the broader set of science cases listed above). In the following sections, I provide the motivation behind this research, and describe the current state of the field.

## 1.1 Molecular gas at $z > 2$

In this section, I briefly review the current understanding about the global evolution of galaxies throughout the Universe's history. In particular, I focus on the molecular gas fueling star formation in the period $z \gtrsim 2$, which is the target of the research in this thesis. I use carbon monoxide (CO) emission as the primary tracer for molecular gas throughout my research, so I also introduce the physics behind CO emission from molecular gas regions in detail.

### The cosmic star formation rate history

The global star formation rate (SFR) density of the Universe has been well studied, primarily through infrared (IR) and ultraviolet (UV) galaxy surveys. These surveys typically take the form of catalogs of discrete objects, used to determine galaxy luminosity functions (LFs) measuring the number density of galaxies as a function of luminosity. Integrating these functions down to some limiting luminosity enables a measurement of the total SFR density. A compilation of several different measurements of the SFR density is shown in Figure 1.2, taken from the review in Madau and Dickinson (2014). It shows a clear global evolution with redshift: the SFR density rises steadily $3 \lesssim z \lesssim 8$, peaks at roughly $z \sim 2$, and then falls off towards the present day Figure 1.2. This period around this peak in global star formation is known as the Epoch of Galaxy Assembly (EoGA) or Cosmic Noon. Roughly half of the stellar mass in the modern-day Universe is thought to be formed during this time (Shapley, 2011).

I note that, while the existence of a peak in cosmic SFR is well-established, there are still some open questions about these measurements. In particular, the faint end slope of the IR LF has not been well constrained at $z \gtrsim 2$ (discussed by Madau and Dickinson, 2014), and the UV LFs are limited to rare, bright objects that may be biased by dust obscuration (these dust-obscured objects are beginning to be found with JWST, e.g., Harish et al., 2025, Stone et al., 2024). These limitations are currently being addressed by complementary LIM and LIM-adjacent surveys—tomographic measurements of the CIB through cross-correlation with galaxy catalogs have shown a significant ($\sim$ 60% at $z \sim 4$) contribution to the overall SFR that is obscured by dust (Chiang et al., 2025), although the broad evolutionary trends remain the same. Upcoming LIM constraints (in particular those from SPHEREx, e.g., Gong et al. 2017) also have the potential to circumvent these limitations more fully.

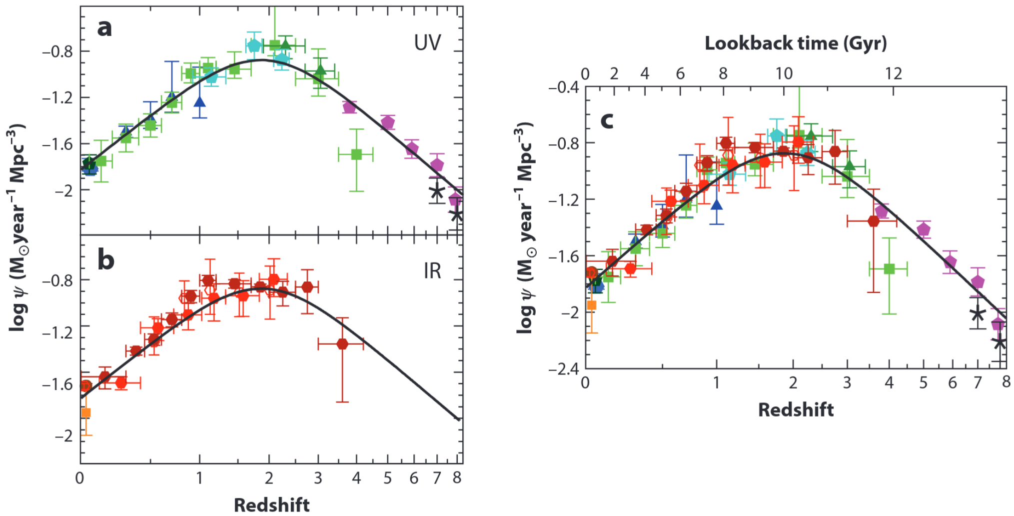


Figure 1.2: The evolution of the global SFR density of the Universe with time, as measured by ultraviolet (Panel a) and infrared (Panel b) galaxy surveys. The combined constraints are shown in Panel c. Figure from Madau and Dickinson (2014), with references enumerated in Table 1 of that work.

What is missing from this picture is a coherent model for the mechanisms driving this evolution in global SFR (e.g., Shapley, 2011). Star formation is a complicated and inefficient process even within a single galaxy, and the global SFR density is integrated over many different galaxy types. Any one of many factors including the accretion of gas onto galaxies, feedback from star formation, supernovae, and AGN, etc., could be evolving over time, and it is likely a balance between all of these factors that is driving the smooth evolution shown in Figure 1.2 (Madau and Dickinson, 2014). In order to study the interplay between these mechanisms, it is necessary to survey other galaxy properties—in particular the molecular hydrogen ($H_2$) gas that fuels star formation—in the lead-up to the EoGA. Efforts to do so systematically have primarily been in the form of small-area spectral line search galaxy surveys (Pavesi et al., 2018; Aravena et al., 2019; Lenkić et al., 2020) using facility instruments such as ALMA and NOEMA. These surveys tend to be limited by both completeness and cosmic variance, measuring $O(10)$ objects in a narrow range of DM halo masses. The properties of the molecular gas inside fainter galaxies typically have to be entirely extrapolated (e.g., Uzgil et al., 2019).

**CO as a tracer of star-forming fuel**

The primary star-forming fuel in galaxies is molecular hydrogen, $H_2$, located in dense giant molecular cloud (GMC) complexes (e.g., Kennicutt and Evans, 2012) with typical temperatures $10-50\,\mathrm{K}$ and densities $10^3-10^6\,\mathrm{cm}^{-3}$ (Draine, 2011). In this section, I review the physics behind the specific mechanism used throughout this thesis for measuring this star-forming molecular gas, which is observations of the CO(1–0) rotational transition. This discussion primarily follows the reviews in Carilli and Walter (2013), Scoville (2013), and Bolatto et al. (2013), with additional reference to Condon and Ransom (2016).

Molecular hydrogen is a symmetric molecule with no permanent electric dipole moment. As a result, the lowest-energy rovibrational transitions it emits are quadrupole transitions with high excitation temperatures (above 500 K, much higher than the temperatures available in GMCs, Carilli and Walter, 2013) that emit weakly in the IR. The preferred method for studying $H_2$ is instead to use polar tracer molecules that form and emit at similar pressures and temperatures to those inside GMCs. In this thesis, I will focus on $^{12}$CO, although lines such as $^{13}$CO, HCN, CS, and $CN^-$ are also used and can be analyzed in combination with $^{12}$CO to determine the properties of the observed molecular gas quantitatively.

A diatomic polar molecule with a moment of inertia $I$ exists at quantized rotational states $J$, with $J=0$ corresponding to the ground state with zero rotational energy. When falling from an excited state to the state below (the $J\rightarrow J-1$ transition), the molecule emits a photon of frequency

$$\nu = \frac{h}{4\pi^2 I}\cdot J. \tag{1.1}$$

The spectrum of this molecule is therefore a rotational 'ladder', with rungs at intervals of the base frequency $h/4\pi^2 I$. For $^{12}$CO (hereafter just 'CO'), this base frequency is 115.27 GHz. The focus of much of this thesis will be the $J=1\rightarrow 0$ transition, which I will hereafter refer to as CO(1–0).

At the typical temperatures and densities of GMCs, CO(1–0) is above its critical density $n_* \sim 2.1\times10^3\,\mathrm{cm}^{-3}$ (Carilli and Walter, 2013)[1], and is thus both primarily collisionally excited and very bright. The surface brightness of the line is typically expressed as a brightness temperature in the Rayleigh-Jeans limit,

$$T_\mathrm{b} = \frac{2k_\mathrm{B}T_\mathrm{x}\nu^2}{c^2}, \tag{1.2}$$

[1] $n_* = A_{10}/(\sigma v)$, where $A_{10}$ is the Einstein A-coefficient for the transition, $\sim 7.2\times10^{-8}\,\mathrm{s}^{-1}$ (Schöier et al., 2010), $\sigma$ is the collisional cross-section of the CO molecule, $\sim 10^{-15}\,\mathrm{cm}^2$ (Condon and Ransom, 2016), and $v$ is the average velocity of the $H_2$ molecules, which are assumed to follow a Maxwell-Boltzmann distribution.

where $T_x$ is the CO excitation temperature and is assumed to be equal to the kinetic temperature of the $H_2$ molecules, $\nu$ is the rest frequency of the spectral line, $k_B$ is the Boltzmann constant, and $c$ is the speed of light. At high redshifts, this is typically converted into a 'line luminosity' in units of $\mathrm{K\,km\,s^{-1}\,pc^2}$,

$$L'_{\mathrm{CO}} = \frac{c^2}{2k_B}(S\Delta v)\frac{D_{\mathrm{L}}^2}{\nu_{\mathrm{obs}}^2(1+z)^3}, \tag{1.3}$$

where $D_{\mathrm{L}}$ is the luminosity distance to the galaxy, $\nu_{\mathrm{obs}} = \nu_{\mathrm{rest}}/(1+z)$ is the observed frequency, and $(S\Delta v)$ is the velocity-integrated flux,

$$(S\Delta v) = \frac{2\nu^2 k_B T_{\mathrm{b}}}{c^2}\Omega\Delta v_k \tag{1.4}$$

($\Omega$ is the solid angle subtended by the source on the sky, or the beam solid angle in the case of a point source, and $\Delta v$ is the linewidth in velocity units).

The optical depth ($\tau$) of GMCs to CO(1–0) emission is generally high. This ensures that the observable emission primarily comes from the $\tau = 1$ surface on the outskirts of the GMC, and sets the brightness temperature of the spectral line to the excitation temperature of this surface. This effect also provides a loose argument for quantitatively connecting the CO luminosity from a GMC to its $H_2$ mass $M_{\mathrm{H_2}}$ (as the cloud is assumed to be composed of primarily $H_2$): the CO luminosity of a cloud is its surface brightness integrated over its area. If a cloud is virialized, its area (particularly its radius) can be related to its mass by assuming a density distribution. This explanation is based on several assumptions that may or may not hold (particularly at high redshift), but an empirical connection between $L'_{\mathrm{CO}}$ and $M_{\mathrm{H_2}}$ has been measured in a variety of different ways (see Bolatto et al., 2013, for a review). This connection is what truly enables CO to be used as a tracer of the star-forming gas of a galaxy, and motivates its use in LIM. It is, however, dependent on a variety of astrophysical factors such as the metallicity and SFR of the galaxy and thus imparts an inherent uncertainty to any $M_{\mathrm{H_2}}$ measurements derived using CO. Throughout this thesis, I use a $L'_{\mathrm{CO}}$-$M_{\mathrm{H_2}}$ conversion factor of $\alpha_{\mathrm{CO}} = 3.6\,\mathrm{M_\odot}\,(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$, originally measured in galaxies at $z \sim 1.5$ (Daddi et al., 2010), for consistency with the CO galaxy surveys at $z \sim 3$ (Riechers et al., 2019; Decarli et al., 2021; Lenkić et al., 2020).

## 1.2 The current status of line intensity mapping

LIM evolved out of techniques designed for diffuse, fluctuating fields (primarily the cosmic microwave background and the 21 cm neutral hydrogen line before and during the epoch of reionization, when it was tracing the diffuse IGM; Madau et al., 1997; Tozzi et al., 2000). However, in this thesis I will focus on LIM experiments targeting spectral lines whose integrated emission is dominated by dense environments such as galaxies. In general I will treat these as a separate field from experiments performing intensity mapping of the 21 cm line, although there is considerable overlap in analysis techniques. This is a fast-growing landscape with many experiments actively observing or in production. Only a few experiments have achieved (primarily tentative) detections of their target spectral line (Keating et al., 2020; Lujan Niemeyer et al., 2026). Although I will focus on CO and the CO Mapping Array Project (COMAP; Cleary et al., 2022), the techniques I develop are broadly applicable to many different LIM experiments. As such, I will provide a brief overview of the LIM theoretical and experimental landscapes, as well as the current analysis techniques and challenges.

### Theoretical and analytical landscape

Because I am limiting the scope of this thesis to experiments targeting emission from within galaxies, I can follow much of the non-21 cm theoretical LIM literature (e.g., Righi et al., 2008; Lidz et al., 2011; Li et al., 2016; Padmanabhan, 2018; Chung et al., 2022) in modeling matter in the Universe as a three-dimensional distribution of dark matter (DM) halos, and luminosity originating only from these halos. More specifically, the spectral-line luminosity of each halo is modeled as directly connected to the mass of the host halo (and possibly its redshift; see Wechsler and Tinker, 2018 for a review) through a function $L(M)$. Each halo is treated as a point source; the $L(M)$ is integrated over the entire galaxy. There are typically several DM halos within a given three-dimensional LIM resolution element.

LIM experiments are intentionally confusing these discrete DM halos into 3D fluctuation fields, so galaxy populations must be analyzed in aggregate with summary statistics. The most commonly-used statistic is the auto-correlation power spectrum, $P(k)$. Although I do not directly measure any auto-correlation results in this thesis, it is often used as a comparison point so I introduce it here. $P(k)$ is typically broken into 'clustering' and 'shot' components,

$$P(k) = \langle T_{\text{line}} \rangle^2 b_{\text{line}}^2 P_m(k) + P_{\text{shot}}. \tag{1.5}$$

Here the clustering component, $\langle T_{\text{line}} \rangle^2 b_{\text{line}}^2 P_m(k)$ includes factors of the average line temperature $\langle T_{\text{line}} \rangle$, the (linear) line bias $b$, and the underlying matter power spectrum $P_m(k)$. The shot component, $P_{\text{shot}}$, is the variance in the mean line luminosity function and dominates over the clustering component at smaller spatial scales. In the current era of the field where many experiments are chasing a detection of any kind, a measurement of $P_{\text{shot}}$ would be an accomplishment. However, most of the cosmological information, and information about the $\langle T_{\text{line}} \rangle$ of a target spectral line, is contained within the clustering scales (e.g., Yue and Ferrara, 2019; Karkare et al., 2022a).

A similar formalism can be used for the cross-correlation power spectrum. In this thesis, I focus on cross-correlations between a LIM spectral-line tracer and a catalog of galaxies. If the galaxy catalog has bias $b_{\text{gal}}$, the cross-power spectrum becomes

$$P_{\times}(k) = \langle T_{\text{line}} \rangle b_{\text{line}} b_{\text{gal}} P_m(k) + C_{\text{LT}} \langle L_{\text{gal}} \rangle. \tag{1.6}$$

The cross-shot power in this instance is given by the mean CO luminosity of the galaxies in the galaxy catalog

$$\langle L_{\text{gal}} \rangle = \frac{1}{\langle n_{\text{gal}} \rangle} \int L \left. \frac{dn}{dL} \right|_{\text{gal}} dL, \tag{1.7}$$

converted to a line temperature using a conversion factor $C_{\text{LT}}$. More details are available in Breysse and Alexandroff (2019) or Wolz et al. (2017). I will develop a model for a third analysis technique, a stack, throughout this thesis (particularly in Chapter 3).

### Experimental landscape

There has been considerable investment in experiments using the LIM technique over the past decade. In pursuit of first detections, different LIM groups are targeting a wide variety of spectral lines, spatial scales, and redshift ranges, using diverse instrumentation techniques. A summary of the state of the field (as of 2022) is shown in Figure 1.3. There are also detailed reviews available in Kovetz et al. (2017), Kovetz et al. (2019), and Chang and Lidz (2026).

The first generation of LIM experiments used facility instruments to provide LIM constraints on smaller spatial scales. The CO Power Spectrum Survey (COPSS; Keating et al., 2015; Keating et al., 2016) used the Sunyaev-Zel'dovich Array (a subset of CARMA) to place constraints the shot-noise component of the CO power spectrum $z \sim 3$, as well as cross-power and stack constraints using a variety of spectroscopic galaxy catalogs (Keenan et al., 2022). Similarly, the millimeter-wave Intensity Mapping Experiment (mmIME; Keating et al., 2020) used a combination of ALMA and the ACA to place shot-power constraints (published as roughly $3\sigma$ detections) on the CO power spectrum.

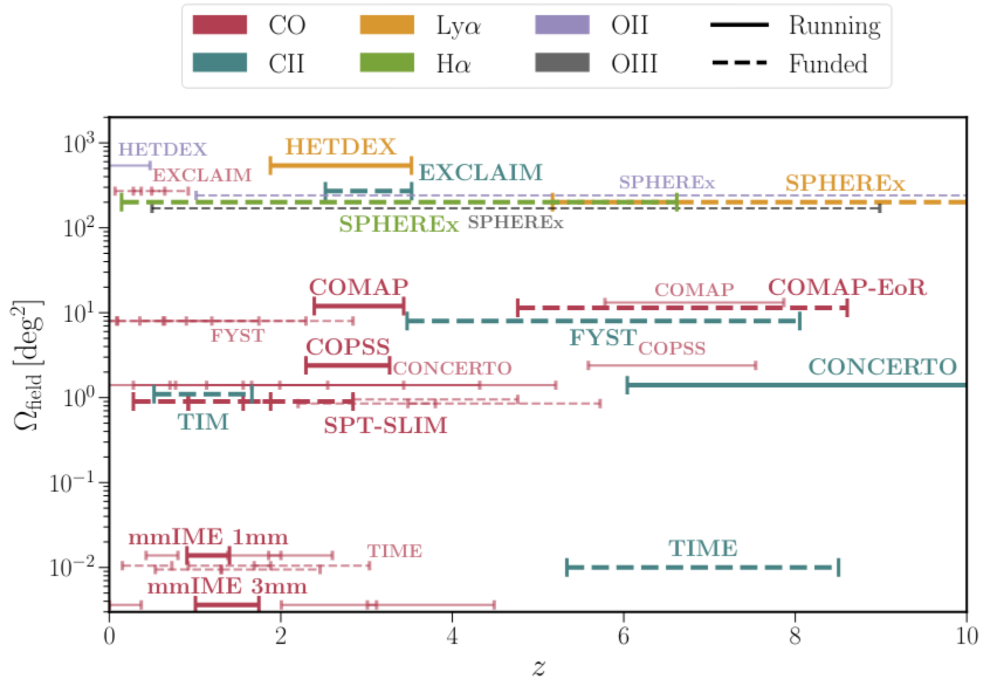


Figure 1.3: An overview of LIM experimentation, as of 2022. The field size and redshift range covered by each experiment is shown, colored by the primary spectral line being targeted. Since this time, SPHEREx has launched and begun science observations (Bock et al., 2026), and TIME and SPT-SLIM have entered commissioning phases. Figure from Bernal and Kovetz (2022a).

Since then, dedicated LIM instruments have begun to be constructed and to come online. In the cm-wave to sub-mm range, only COMAP (which I describe in considerably more detail in Section 1.3 below) has published CO science results. CONCERTO (Van Cuyck et al., 2023), now concluded, has produced constraints on a [CII] power spectrum severely limited by lower-redshift contamination from CO lines (Van Cuyck et al., 2023). TIME (Staniszewski et al., 2014), SPT-SLIM (Karkare et al., 2022b), EoR-Spec (CCAT-Prime Collaboration et al., 2023, listed as FYST in Figure 1.3), TIM (Vieira et al., 2020), TIFUUN (Kohno et al., 2024), and EXCLAIM (Oxholm et al., 2020) all have instrumentation in varying stages of completion and are planning to begin science observations in the next few years. TIME, SPT-SLIM, EoR-Spec, and TIFUUN are sub-mm spectrometer modules placed on shared dishes. TIM and EXCLAIM are balloon-bourne. These experiments are primarily targeting [CII] at different periods of the Universe's history, and will have lower-redshift CO lines that can be treated either as interloper contaminants for the [CII] or as science targets in their own right.

In the (observed) infrared-optical range, the two primary experiments are SPHEREx (Doré et al., 2014; Doré et al., 2016; Doré et al., 2018), and HETDEX (Gebhardt et al., 2021), both of which have completed significant fractions of their science observations. SPHEREx will observe multiple optical lines at different redshifts, including H$\alpha$, H$\beta$, Ly$\alpha$, [OII], [OIII], and broader PAH features (Cheng et al., 2025). HETDEX, while designed as a blind spectroscopic Ly$\alpha$ survey, can also perform intensity mapping of the Ly$\alpha$ line, and has published a cross-correlation detection between the discrete cataloged galaxies and the diffuse LIM emission in their dataset (Lujan Niemeyer et al., 2026).

**Joint analysis techniques**

In this early generation of LIM analysis, there are two main factors limiting the ability of experiments to produce autocorrelation science results. The first is simply raw sensitivity. COMAP's current auto-power constraints, for example, are mostly white-noise dominated (Lunde et al., 2024) and have not yet achieved the sensitivities required to rule out the best models of cosmic CO emission. This is also what forces some experiments to target primarily shot-power scales in their first stages (e.g., Keating et al., 2015; Sun et al., 2021), due to a tradeoff between achievable sensitivity in a given resolution element and survey area for an experiment with a finite observing time (the 'Knox criterion,' see Appendix C of Li et al., 2016 for a LIM-specific discussion). Detecting mostly shot-power scales makes estimation of parameters such as $\langle T_{\rm line} \rangle$ more difficult (Yue and Ferrara, 2019), and smaller-area surveys are much more limited by cosmic variance.

The other limiting factor is systematic error. This can be experimental—as in the CMB field, LIM experiments are sensitive to instrumental fluctuations across all spatial scales, and current experiments have all had to address instrumental systematics to some degree (e.g., Lunde et al., 2024; Lundgren et al., 2025). Systematic error can also be astrophysical. Many spectral lines can be redshifted into a given observing band, and they are very difficult to distinguish from the target line. This is a known issue, and many different techniques have been proposed to address it, including exploiting anisotropies in the projection of interloper lines at the power spectrum (e.g., Cheng et al., 2016) or map level (e.g., Moriwaki et al., 2020), masking with an external catalog (e.g., Sun et al., 2018; Karoumpis et al., 2024), or tomographic techniques involving fitting multiple spectral lines at the same time (e.g., Cheng et al., 2020). Each of these have their own limitations.

It is for these reasons that techniques combining LIM experiments with external surveys of galaxies have become increasingly valued. These analyses improve sensitivity by introducing priors on the location of the signal (e.g., Pullen et al., 2013b; Chung et al., 2019), and reduce the set of systematic errors to those that are correlated between the two datasets (Fronenberg and Liu, 2024)[2]. Joint analyses also allow for exploration of the interplay between multiple tracers of the complex process of galaxy formation (e.g., Breysse and Alexandroff, 2019). The efficacy of these techniques has already been proven in the 21 cm field, where signal is beginning to be detected. Almost all of the initial 21 cm results are joint analyses involving galaxy surveys (e.g., Pen et al., 2009; Chang et al., 2010; Masui et al., 2013; Wolz et al., 2022; MeerKLASS Collaboration et al., 2025; Chen et al., 2025; Amiri et al., 2023; Amiri et al., 2024).

Much of the LIM joint analysis theory and forecasting focuses on the cross-correlation (e.g., Lidz et al., 2011; Pullen et al., 2013b; Breysse and Alexandroff, 2019; Chung et al., 2019; Bracks et al., 2026), and some LIM experiments are already publishing cross-correlation results (Keenan et al., 2022; Lujan Niemeyer et al., 2026). Cross-correlation makes use of the large-scale clustering present in both the LIM and galaxy datasets, increasing the available information content. Ultimately, this is the target joint-analysis technique for the LIM field.

However, cross-correlations are sensitive to the existing angular structures in a survey strategy, and they suffer from complicated mode-mixing effects when the observational strategy of the LIM experiment is correlated with that of the galaxy catalog. These issues are tractable, but require careful characterization of the observational properties of both experiments to overcome. It is for these reasons that stacking (coadding the LIM data on the positions of the objects in a galaxy catalog, returning the average luminosity of the galaxies in the catalog and their surroundings) has the potential to be extremely value as a conceptually-simple, early-adoption analysis technique for LIM. This technique is the focus of my thesis. Outside of my own work with COMAP, stacking has been used successfully for COPSS (Keenan et al., 2022) and the 21 cm experiment MeerKLASS (Chen et al., 2025).

[2]This is not quite true—while *additive* systematics will be removed, *multiplicative* systematics will persist. This is a particular problem for 21 cm experiments, which tend to struggle with extremely bright continuum foregrounds taking on spectral variation due to the behavior of the beam with frequency. Systematics will also inflate the variance of the statistic. Fronenberg and Liu (2024) discuss this in detail in the context of the cross-correlation.

## 1.3 The CO Mapping Array Project

Although the stacking techniques developed in this thesis are applicable to any LIM experiment, I will apply them to COMAP (Cleary et al., 2022) throughout. As such, I provide a brief overview of the instrument and autocorrelation results here.

### The COMAP Pathfinder instrument and data reduction

The COMAP Pathfinder (described in detail in Lamb et al., 2022) is a 19-feed focal plane array receiver mounted on a 10.4 m Leighton Cassegrain dish (Leighton, 1977) at the Owens Valley Radio Observatory (OVRO) in California (Figure 1.4f). Each feed ('pixel') is located within a cryostat at the secondary focus of the dish (Figure 1.4b,c,e). A summary of the receiver chain is shown in Figure 1.4a: light at $26 - 34$ GHz (corresponding to coverage of the redshifted 115.27 GHz CO(1–0) transition from $z = 2.4 - 3.4$) enters the feedhorns, is downcoverted twice and split into four 2 GHz-wide sidebands, and then read out by ROACH-2 spectrometers every 20 ms. The native spectral resolution is 2 MHz, making the telescope an $R \sim 10,000$ spectrometer; data are binned to 31.25 MHz spectral channels for science. The median system temperature ($T_{\rm sys}$) is 44 K.

The telescope observes three cosmological fields (each $\sim 1.5^\circ \times 1.5^\circ$) by scanning over them using constant-elevation scans. Each feed has its own receiver chain and is effectively an independent total-power instrument. Each feed has a $4.5'$ beam FWHM at 30 GHz, and beams are separated by $12'$ on-sky. We calibrate the telescope using a 'vane' made of microwave absorber material which can be moved into the path of the feedhorns between scans (Figure 1.4e), providing an ambient-temperature blackbody load for determination of the system temperature. Each scan lasts roughly 5 minutes. COMAP began observing in 2019, and has since logged more than 17,500 observing hours on the cosmological fields. For analysis purposes, we split these observations into two observing 'Seasons': Early Science (05/2019–08/2020, 5,200 hours) and Season 2 (11/2020–11/2023; 12,300 hours). The details of the observing strategy are given in Foss et al. (2022) and Lunde et al. (2024).

Data are processed using a filter-and-bin mapmaker (Foss et al., 2022). Raw time-ordered data (spectra at each time sample) are passed through a series of filters (to remove signals correlated with local pointing, continuum signals, residual atmospheric variations, and standing waves) and then masked and calibrated. The TOD are then binned into three-dimensional datacubes with 3D pixel ('voxel') sizes of

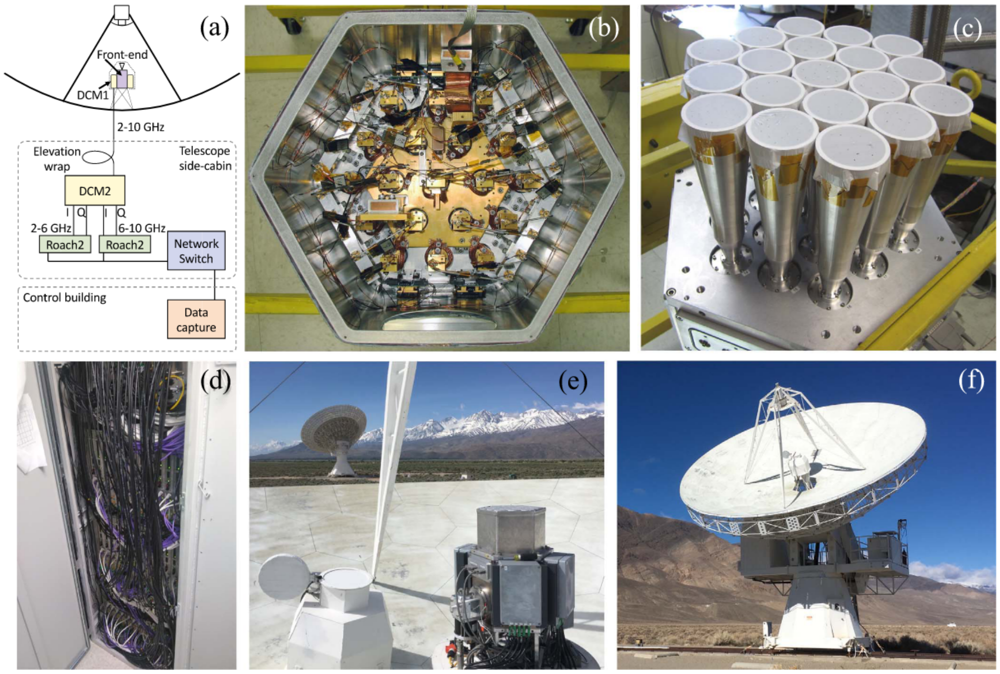


Figure 1.4: The COMAP Pathfinder instrument. (a) A simplified version of the receiver chain for a single COMAP feed. (b) Interior and (c) exterior of the cryostat, showing the low-noise amplifier and feedhorn, respectively, for each of the 19 feeds. (d) ROACH-2 spectrometers in the digital rack. (e) Receiver package at the secondary focus of the telescope, with the weather enclosure removed (the calibration vane is visible on the weather enclosure). (f) 10.4 m COMAP dish at OVRO. Figure from Cleary et al. (2022).

$2' \times 2' \times 31.25$ MHz, based on a pointing matrix. For the Season 2 dataset, we apply a map-level principal component analysis (PCA) filter (Lunde et al., 2024). Each of these steps are done separately for each feed. For map-level analyses such as the stack, we then coadd the feeds together into a single datacube for each field. For the autocorrelation power spectrum, pairs of feeds are cross-correlated to remove residual systematics correlated between feeds (Ihle et al., 2022; Stutzer et al., 2024).

**COMAP autocorrelation results**

The COMAP Season 2 autocorrelation power spectrum is shown in Figure 1.5a. This is currently the most up-to-date COMAP constraint; it is also the state-of-the-art for measurements of cosmic CO emission $z \sim 3$. Different models for $z \sim 3$ CO emission are also shown (from Pullen et al., 2013b; Li et al., 2016; Padmanabhan, 2018; Keating et al., 2020; Yang et al., 2022), as well as the constraint from COPSS (Keating et al., 2016) and COMAP Early Science (Ihle et al., 2022; Chung et al., 2022). No signal has yet been detected in the clustering regime[3]. While COMAP is ruling out the brightest models, there remains nearly an order of magnitude between CO models. This is an artifact of the lack of observational constraints described in Section 1.1 above – each of these models is nearly equally viable in the context of available observations.

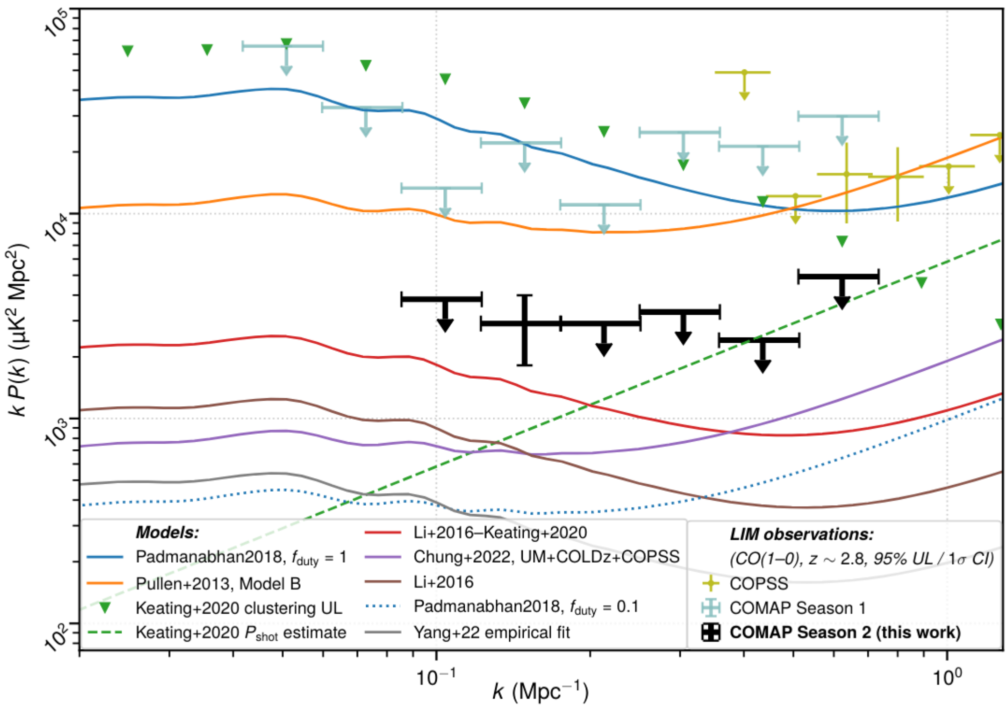


Figure 1.5: The autocorrelation power spectrum measured from COMAP's Season 2 dataset (the dark black limits), in comparison to the model space for $z \sim 3$ CO and other constraints from COPSS (Keating et al., 2016) and COMAP Early Science (Season 1, Ihle et al., 2022). Figure from Chung et al. (2024b).

Using the halo modeling described in Section 1.2, the power spectrum constraint can be converted into a constraint on the mean CO temperature-bias product, $\langle Tb \rangle$. We do this in Chung et al. (2024b), finding an upper limit $\langle Tb \rangle < 4.8\,\mu$K. By assuming a (conservative) prior on the bias $b > 2$, we can break the degeneracy between $T$ and $b$ in the autocorrelation power spectrum on its own, and constrain the average CO line temperature to a 99% upper limit of $\langle T \rangle < 2.4\,\mu$K. This corresponds to an average molecular gas density $\rho_{H2} < 1.7 \times 10^8\,M_\odot\,\mathrm{Mpc}^{-3}$, which is roughly a factor of 3 – 4 above the extrapolated values from resolved CO galaxy surveys

[3]One of the COMAP $k$-bins is plotted as a detection; this is because the value in that bin was measured to be just above 2× the uncertainty. After considering the look-elsewhere effect it is not statistically significant (Stutzer et al., 2024).

such as ASPECS (Aravena et al., 2019) and PHIBSS2 (Lenkić et al., 2020) (see Section 1.1).

## 1.4 Thesis outline

This main body of this thesis is split into five research papers. In Chapter 2, I present my first attempt at a COMAP stack, stacking the COMAP Early Science data on the positions of 243 quasars from the extended Baryon Oscillation Spectroscopic Survey (eBOSS). I introduce the stacking framework as I apply it to COMAP (contained in the pipeline `lim_stacker`; Dunne et al., 2026a), and place a limit on the average CO luminosity of the quasars and on the average molecular gas density of the Universe in their surroundings, without making any attempt to differentiate between the two theoretically.

In Chapter 3, I develop a more thorough theoretical understanding of stacking analyses that I apply throughout the rest of the thesis. This is done through simulations using the custom pipeline `joint_limlam_mocker` (Dunne et al., 2026b), which I developed for this application. I find that the stack signal is dominated by cosmological clustering rather than the galaxies actually included in the reference catalog – almost all of the luminosity in a typical stack average originates from halos neighboring the reference halo. I also establish a set of parameters that optimize the S/N ratio of a stack, given a finite observing time.

Chapters 4 and 5 apply the knowledge gained from Chapter 3 to stacks using COMAP's Season 2 dataset, which is roughly 3.4× more sensitive than the Early Science dataset. Chapter 4 stacks again on quasars, including those from eBOSS and additional detections from DESI. Chapter 5 stacks on Ly$\alpha$ emitting galaxies (LAEs) from the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX; Gebhardt et al., 2021). In both cases, I do not detect any cosmological CO emission in the stack, and in both cases there is a mild to moderate tension observed between the upper limit obtained from the data and models of CO stacks on similar mock catalogs. I test a wide variety of experimental attenuation factors as possible explanations for this tension and find that it still holds.

I present a novel application of the stacking framework in Chapter 6. I vary the rest frequency at which the stack is centered across a spectrum of rest frequencies around 100 GHz, thus carrying out a spectral line scan. I use this scan to search for two-photon decay from axionic dark matter, placing constraints on the axion decay rate $\Gamma$ and the axion-photon coupling constant $|g_{a\gamma\gamma}|$. In the mass range to which it is sensitive, COMAP is able to provide the first constraints on axion dark matter decay.

Finally, I summarize the conclusions of each work in Chapter 7, and discuss prospects for future joint analyses with COMAP and other LIM experiments. These include near-future COMAP instrument improvements, changes to the stacking methodology designed to improve S/N and refine science conclusions, and ultimately the step forward to full cross-correlations.

*Chapter 2*

# COMAP EARLY SCIENCE: VIII. A JOINT STACKING ANALYSIS WITH EBOSS QUASARS

Dunne, Delaney A. et al. (Apr. 2024). "COMAP Early Science. VIII. A joint stacking analysis with eBOSS quasars." In: *The Astrophysical Journal* 965.1, 7, p. 7. DOI: 10.3847/1538-4357/ad2dfc. arXiv: 2304.09832 [astro-ph.GA].

## 2.1 Chapter abstract

We present a new upper limit on the cosmic molecular gas density at $z = 2.4 - 3.4$ obtained using the first year of observations from the CO Mapping Array Project (COMAP). COMAP data cubes are stacked on the 3D positions of 243 quasars selected from the Extended Baryon Oscillation Spectroscopic Survey (eBOSS) catalog, yielding a 95% upper limit for flux from CO(1-0) line emission of 0.129 Jy km/s. Depending on the balance of the emission between the quasar host and its environment, this value can be interpreted as an average CO line luminosity $L'_{\rm CO}$ of eBOSS quasars of $\leq 1.26 \times 10^{11}$ K km pc$^2$ s$^{-1}$, or an average molecular gas density $\rho_{\rm H_2}$ in regions of the universe containing a quasar of $\leq 1.52 \times 10^8$ M$_\odot$ cMpc$^{-3}$. The $L'_{\rm CO}$ upper limit falls among CO line luminosities obtained from individually-targeted quasars in the COMAP redshift range, and the $\rho_{\rm H_2}$ value is comparable to upper limits obtained from other Line Intensity Mapping (LIM) surveys and their joint analyses. Further, we forecast the values obtainable with the COMAP/eBOSS stack after the full 5-year COMAP Pathfinder survey. We predict that a detection is probable with this method, depending on the CO properties of the quasar sample. Based on the achieved sensitivity, we believe that this technique of stacking LIM data on the positions of traditional galaxy or quasar catalogs is extremely promising, both as a technique for investigating large galaxy catalogs efficiently at high redshift and as a technique for bolstering the sensitivity of LIM experiments, even with a fraction of their total expected survey data.

## 2.2 Introduction

Spectral Line Intensity Mapping (LIM) is an emerging observational technique with the potential to enhance our understanding of the universe by constraining the global properties of galaxies over cosmic time. LIM surveys do not aim to resolve individual galaxies, but instead measure 3-dimensional fluctuations in the integrated emission from many galaxies, allowing for efficient mapping of galaxies across large cosmic volumes (see Kovetz et al., 2019, for a review). Because LIM measures integrated emission, it is sensitive to the faintest galaxies, which are nearly impossible to detect in traditional surveys but nevertheless make up the bulk of the galaxy population at any given cosmic time. Observationally, the field is still in its early stages, with dedicated LIM instruments only beginning to publish early autocorrelation constraints (e.g., COMAP, Cleary et al. 2022; the Hydrogen Epoch of Reionization Array, HERA, Abdurashidova et al. 2022; and MeerKAT, Paul et al. 2023).

Currently, the results being released from these first LIM surveys are primarily upper limits (for COMAP, Chung et al., 2022) and are not yet sufficient to achieve the measurements described above. This is partially a consequence of the newness of the field – strong autocorrelation detections of a given spectral line are required to measure its cosmological fluctuations, and most LIM surveys have simply not been integrating long enough to achieve the needed sensitivities (for example, COMAP projects a secure autocorrelation detection after 5 years; Chung et al., 2022). However, cosmic line emission may be detectable even with present-day LIM datasets, if careful signal processing techniques making use of external datasets are applied.

In particular, averaging together intensity-map voxels (3D pixels) that are known to contain galaxies (by comparison with some other survey) in a stacking analysis has the potential to make a detection of the average CO line temperature associated with catalog galaxies, even with an intensity map insufficiently sensitive for other analyses (Silva et al., 2021). Stacking, or coadding, is an established technique for improving sensitivity in traditional, targeted galaxy surveys (e.g., Stanley et al., 2019; Jolly et al., 2021; Romano et al., 2022; Lujan Niemeyer et al., 2022), and has recently been extended to LIM observations as well (Keenan et al., 2022). The efficacy of this technique will depend on factors such as the number of traditional survey objects that fall into the LIM survey footprint, and the redshift accuracy in the traditional catalog (Chung et al., 2019; Silva et al., 2021), as well as the chosen weighting scheme (Sinigaglia et al., 2022).

In this paper, we use LIM data from the CO Mapping Array Project (COMAP; Cleary et al., 2022), the first survey to use a purpose-built instrument (the COMAP Pathfinder, described in detail in Lamb et al., 2022), to impose direct constraints on the clustering-scale CO power spectrum at high redshifts. We use COMAP

Season 1 data, taken during the first 13-month observing season of the project. COMAP observations cover a frequency range of 26–34 GHz (redshifts of $z = 2.4$–$3.4$) and encompass three $\sim 4\,\mathrm{deg}^2$ fields. The data reduction and map-making processes for these Season 1 data are described in Foss et al., 2022, while power spectrum estimation techniques and constraints are described in Ihle et al., 2022 and Chung et al., 2022. Observations are continuing with the Pathfinder to complete the nominal 5-year survey, at the end of which a detection of the CO auto-power spectrum is forecast. The COMAP fields were selected to overlap with the HETDEX survey of Ly-$\alpha$ emitters (Gebhardt et al., 2021), allowing us to perform stacking and cross-correlation analyses of these two datasets.

As a preliminary step, we investigate in this work the potential for the BOSS/eBOSS quasar sample (Dawson et al., 2016) to provide additional constraints and inform our understanding of the relationship between active galaxies and star formation. The eBOSS catalog was developed with the intention of studying the Ly$\alpha$ forest and therefore covers a redshift range overlapping with that of COMAP – the eBOSS DR16 catalog includes spectroscopic observations of $> 239,000$ quasars in the range $z > 2.1$. Depending on the assumptions made, a stack of COMAP data on the positions of eBOSS quasars can be treated as a measurement of either the average CO luminosity of the quasars themselves or the average molecular gas density at quasar positions.

In addition to its extensive size, which allows for a high-sensitivity stack, the combination of COMAP with the eBOSS catalog enables the study of the CO properties of a large sample of high-redshift quasars. Quasar feedback, particularly in relation to the molecular gas traced by CO transitions, is poorly understood on a statistical scale. Current studies mainly involve individual objects, using the CO rotational ladder in the host galaxies of Active Galactic Nuclei (AGN) to study the effects of AGN on their surroundings. These are complex measurements requiring extremely CO-bright objects, and thus the pool of available galaxies for study is small and biased, especially at high redshifts. LIM measurements will be able to provide a more complete picture of the wider population of AGN and their host galaxies (e.g., Breysse and Alexandroff, 2019).

In this work, we will investigate the CO properties of the 243 eBOSS quasars in the COMAP fields through stacking, thus providing constraints on the molecular gas content of a large sample of high-redshift quasars for the first time. We will additionally use this analysis to explore the viabilty of stacking as a LIM technique. As such, we will forecast the sensitivity of the COMAP Pathfinder survey to CO emission from the eBOSS quasars after five years of observing. Section 2.3 introduces both the COMAP Pathfinder (§2.3) and eBOSS surveys (§2.3), and Section 2.4 describes the stacking methodology we use in our analysis. Section 2.5 describes verification of the stacking analysis. We present results in Section 2.6, including forecasts of the results of stacking the full Pathfinder survey (§2.6). We discuss the implications of these results in Section 2.7.

We assume a ΛCDM cosmology based broadly on nine-year WMAP (Hinshaw et al., 2013) results throughout, with $\Omega_{\rm m} = 0.286$, $\Omega_\Lambda = 0.714$, $\Omega_{\rm b} = 0.047$, and $H_0 = 100h\,{\rm km\,s^{-1}\,Mpc^{-1}}$ with $h = 0.7$. These are the same values used in all previous COMAP works. Unless otherwise indicated, distances are given as proper values.

## 2.3 Data

### COMAP

The COMAP Season 1 data consist of 13 months of observations using the COMAP Pathfinder instrument, a single-polarization 19-feed spectrometer array fielded on a 10.4 m telescope located at the Owens Valley Radio Observatory (OVRO) in California (the Pathfinder is described in detail in Lamb et al. 2022). The Pathfinder instrument observes between 26 and 34 GHz, and is therefore sensitive to the CO(1–0) rotational transition ($\nu_0 = 115.27$ GHz) emitted in a redshift range of $z = 2.4{-}3.4$. These observations were towards three $\sim$ 4-deg$^2$ cosmological fields (shown in Table 2.1).

Table 2.1: The locations of the three fields used in the COMAP Pathfinder survey.

| Field | RA (J2000) | Dec (J2000) |
|---|---|---|
| Field 1 | $01^{\rm h}\,41^{\rm m}\,44$ | $00°\,00'\,00''$ |
| Field 2 | $11^{\rm h}\,20^{\rm m}\,00$ | $52°\,30'\,00''$ |
| Field 3 | $15^{\rm h}\,04^{\rm m}\,00$ | $55°\,00'\,00''$ |

During observations, the telescope is positioned at the leading edge of the field and performs a scanning motion while the cosmological field drifts through. This usually takes 3–10 minutes, after which the telescope is pointed to the leading edge of the field again and the procedure is repeated. This means that over time, we are scanning the fields from several different directions leading to a smoothly varying noise distribution in the final maps, with low noise in the central regions of the field, and gradually increasing levels of noise towards the edges of the fields.

Raw time-ordered data (TOD) from the telescope are processed by applying a series of filters with the goal of removing correlated noise, standing waves, continuum foregrounds, ground contamination and other systematic effects from the TOD. This process and the COMAP scanning strategies are described in detail by Foss et al., 2022. The output of this pipeline is a set of three calibrated three-dimensional intensity maps, each of angular size $\sim 4\,\mathrm{deg}^2$ with 31.25 MHz spectral resolution and a 4.9–4.4 arcmin beam FWHM. At small scales these maps are dominated by uncorrelated, Gaussian noise (Foss et al., 2022). The maps are somewhat elongated in the right ascension direction, giving us appreciable coverage in a 3D region with dimensions (in comoving coordinates) roughly 300 Mpc $\times$ 200 Mpc in directions perpendicular to the line of sight (at redshift $z = 2.9$) and about 1000 Mpc along the line of sight, for a total volume of the order $6{\times}10^7$ Mpc$^3$ comoving for each field.

As described by Foss et al. (2022), Ihle et al. (2022) and Rennie et al. (2022), the maps are calibrated to the total power entering the telescope. About 72 % of this power comes from the main beam, with another roughly 10 % of the power in the near sidelobes (within about 30 arcmin). To take this scale-dependence into account Ihle et al. (2022) use a model of the COMAP beam to construct a beam transfer function for the power spectrum. For the purposes of our stacking analysis, since we are mostly interested in the smallest scales in the map, we simply divide the map (and the corresponding uncertainty map) by a single overall factor of 0.72 to get the right calibration at scales corresponding to the main beam. We mask any map voxel (3-D pixel) with an integration time below 1000 seconds (50000 'hits', a stricter cut than is taken for the overall COMAP pipeline; Foss et al. 2022).

**eBOSS**

The Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013) and its extension (eBOSS; Ahumada et al. 2020) together encompass eleven years of observations using the BOSS spectrograph at Apache Point Observatory (Gunn et al., 2006; Smee et al., 2013). We use the large ($\sim$ 240000-object) eBOSS spectroscopic sample of Quasi-Stellar Objects (QSOs) at $z \geq 2.4$, which was intended to enable studies of the Ly$\alpha$ forest at $z \sim 1$. These observations targeted objects that were selected based on WISE/SDSS DR13 imaging data. We use the eBOSS DR16 superset catalog (Lyke et al., 2020; Ahumada et al., 2020), released in 2020 – the final iteration of the BOSS/eBOSS catalog. While principally composed of traditional quasars, the superset applied no quasar-selection pipeline and thus contains a small percentage (3.3%, in the overlap with COMAP fields) of bright galaxies, broad absorption line quasars, and damped Ly$\alpha$ systems.

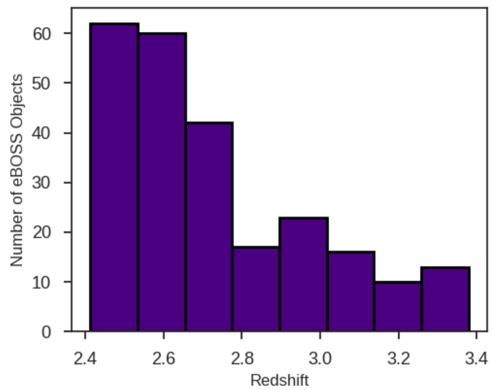


Figure 2.1: The redshift distribution of the eBOSS objects used for this stacking analysis.

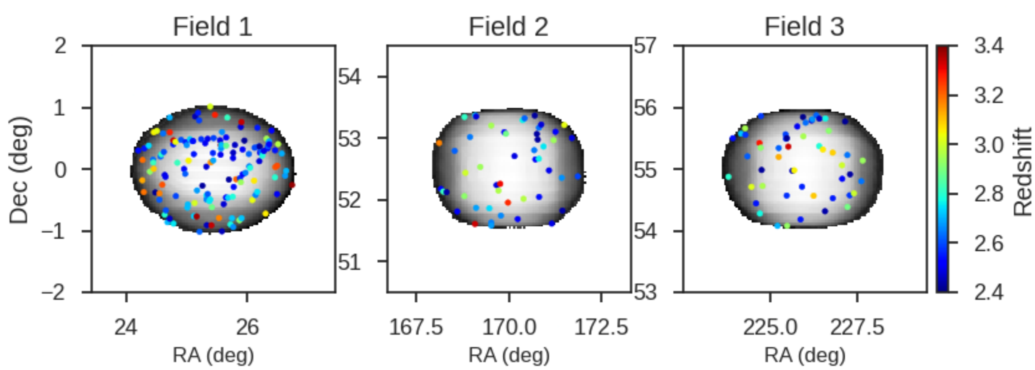


Figure 2.2: The distribution in space and redshift of the eBOSS objects used for this stacking analysis, plotted on the COMAP Season 1 median map RMS for each field (see Foss et al., 2022).

Objects in the eBOSS catalog are identified and their redshifts are determined simultaneously, through the fitting of templates spanning galaxies, quasars, and stars to the eBOSS spectra (stepping over redshift; Bolton et al., 2012). Spectra are optical, and at the redshifts of interest the templates to be fit typically include some combination of Ly$\alpha$, H$\alpha$ and H$\beta$, and a range of optical and NUV forbidden metal lines. The fit with the minimum reduced-$\chi^2$ value is used, provided $\chi_r^2 > 0.01$ and there is no other similarly confident fit determination greatly offset in redshift from the best-fit value. Visual followup is used to confirm the least confident identifications, and to test the algorithmic determination[1]. Tests of the classification algorithm on blank-sky spectra showed a $< 2\%$ false positive rate – the eBOSS redshift determinations are generally considered highly trustworthy. Based on the pixel size of the detector, the uncertainty in the eBOSS redshifts is at most $\sim$ 207 km s$^{-1}$, or $|\Delta z| \leq 0.01$ (Bautista et al., 2017) (approximately 1 COMAP channel). We note that this uncertainty does not account for the systematic gas in- or outflows that may be present in quasars; we discuss this additional consideration in §2.5.

We remove duplicates and any object with an SDSS `NEGATIVE_EMISSION` warning from the eBOSS catalog, and then cut to the COMAP fields. In total, 243 eBOSS objects lie in the COMAP fields after these cuts. The average redshift of these objects is $z = 2.73$, with their redshift distribution as shown in Figure 2.1. Their spatial distribution is shown in Figure 2.2, against the COMAP Season 1 sensitivity.

## 2.4 Stacking methods

Broadly, the stack is a three-dimensional coaddition of the COMAP maps on the spatial and spectral positions of the quasars in the eBOSS catalog: we extract a three-dimensional 'cubelet' centered around the spatial position and redshifted CO(1-0) frequency of each source, and then combine them together into a single stacked cubelet containing the average luminosity of each source.

### Determining Physical Quantities from the COMAP Data Cubes

The COMAP data cubes are provided in brightness temperature units, $T_{\mathrm{b},i,j,k}$, where $i$ and $j$ index voxels in directions perpendicular to the line of sight (spatial axes) and $k$ indexes voxels parallel to the line of sight (spectral axis). The stacking pipeline begins by converting $T_{\mathrm{b},i,j,k}$ into velocity-integrated line flux, $(S\Delta v)_{i,j,k}$:

$$(S\Delta v)_{i,j,k} = \frac{2\nu^2 k_B T_{\mathrm{b},i,j,k}}{c^2}\Omega_{\mathrm{vox}}\Delta v_k. \tag{2.1}$$

In this case, $T_{\mathrm{b},i,j,k}$ is the brightness temperature in the voxel, $\Omega_{\mathrm{vox}}$ is the solid angle of the sky subtended by the voxel (which is constant across the map), and $\Delta v_k$ is the channel width in km/s in the $k^{\mathrm{th}}$ frequency channel,

$$\Delta v_k = \frac{\Delta \nu}{\nu_{\mathrm{obs},k}}c. \tag{2.2}$$

[1]Other redshift values, in particular those based on specific emission lines, are also available. For a full list of the various methods for redshift determination, see: `https://data.sdss.org/datamodel/files/BOSS_QSO/DR16Q/DR16Q_Superset_v3.html`

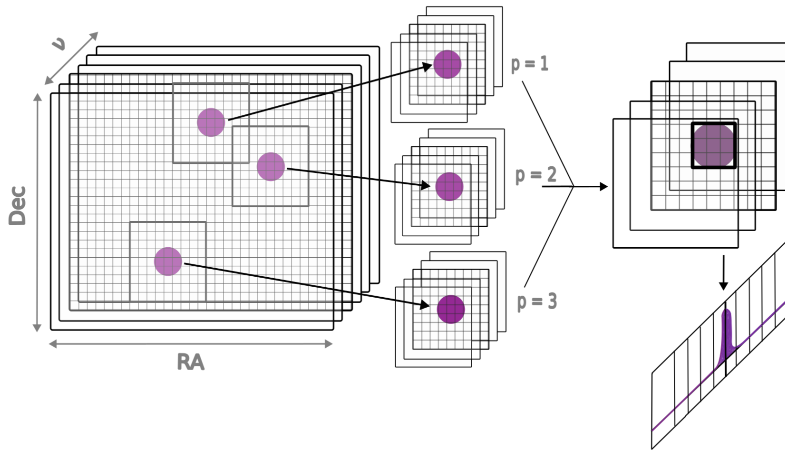


Figure 2.3: A diagram displaying the methodology used for this stacking analysis. From the full three-dimensional COMAP data cube (left), smaller 3-D cubelets are cut out centered on the position of each eBOSS object (center). These cubelets are then averaged into a single 3-D stack (top right), which is used to determine the average COMAP spectrum of the eBOSS objects (bottom right).

We follow Solomon et al., 1997 to calculate $L'_{\rm CO}$ in each voxel,

$$L'_{\mathrm{CO},i,j,k} = \frac{c^2}{2k_B}(S\Delta v)_{i,j,k}\frac{D^2_{\mathrm{L},k}}{\nu^2_{\mathrm{obs},k}(1+z_k)^3}, \tag{2.3}$$

where $\nu_{\mathrm{obs},k}$ is the observed frequency of the $k^{\mathrm{th}}$ frequency channel, $z_k$ is the redshift required to place the 115.27 GHz CO(1-0) line into that channel's frequency range, and $D_{L,k}$ is the luminosity distance associated with that redshift. We perform this conversion before stacking because several of these values vary across the COMAP data cube.

### Extraction of Cubelets

For each eBOSS quasar, we extract from the converted COMAP $L'_{\rm CO}$ cubes a three-dimensional 'cubelet' centered around the spatial position and redshifted CO(1-0) frequency of the quasar. These cubelets have dimensions 42 arcmin × 42 arcmin × 1.28 GHz and contain CO line luminosity values $L'_{\mathrm{CO},l,m,n,p}$, where $l$ and $m$ are in spatial directions, $n$ is along the spectral axis, and $p$ indexes each cubelet. These are purposefully large compared to the COMAP spatial and spectral resolution, so they can be visually searched for potential large-scale fluctuations. If any cubelet has its central voxel masked, or more than half of the voxels immediately adjacent to the central voxel in all three directions masked, it is excluded from the catalog of cubelets to be stacked.

**Stacking**

We then stack the CO emission by directly combining the full three-dimensional cubelets associated with each individual galaxy, voxel-by-voxel. This is done to allow for additional testing of the fidelity of any potential detections. To account for the inhomogeneous noise response across the map (§2.3), values are weighted by their RMS noise using inverse-variance weighting. Thus, the line luminosity in each voxel of the stacked cubelet is

$$L'_{\mathrm{CO},l,m,n} = \frac{\sum_p L'_{\mathrm{CO},l,m,n,p}/\sigma^2_{l,m,n,p}}{\sum_p 1/\sigma^2_{l,m,n,p}}, \tag{2.4}$$

for cubelet voxel line luminosities $L'_{\mathrm{CO},l,m,n,p}$ with RMS noise $\sigma_{l,m,n,p}$. The uncertainty in each voxel of the resulting stacked cubelet is then

$$\sigma_{l,m,n} = \sqrt{\frac{1}{\sum_p 1/\sigma^2_{l,m,n,p}}}. \tag{2.5}$$

We additionally attempt to lessen any interactions between asymmetries in the position of catalog galaxies in the data cubes and the inhomogenous COMAP noise response by rotating each cubelet randomly in its spatial axes by either 0°, 90°, 180°, or 270° before performing the coaddition.

**Obtaining luminosity and density limits**

Once a full three-dimensional stack cubelet is obtained, we calculate the average line luminosity of the included quasars in two steps. Firstly, we calculate the line luminosity in each frequency channel, $L'_{\mathrm{CO},n}$, by summing over the central 3 spatial pixels (spaxels) of the stacked cubelet (the region indicated with a black box in Figures 2.3 and 2.4),

$$L'_{\mathrm{CO},n} = \sum_{l,m=-1}^{+1} L'_{\mathrm{CO},l,m,n} \tag{2.6}$$

which leaves us with a one-dimensional frequency spectrum. This $3 \times 3$ region corresponds to a 6′ square aperture, roughly 1.5 times the size of the COMAP main beam, meaning that most of the signal from any catalog object located anywhere in the central spaxel will not be spread outside this aperture by the beam. Thus, we avoid excluding any potential signal from the spectrum, while keeping the aperture as small as possible to mitigate the effects of beam dilution and maximize our signal-to-noise ratio (S/N).

Finally, to determine the average CO luminosity of the stacked catalog objects $L'_{\rm CO}$, we sum the central seven frequency channels of the stacked spectrum,

$$L'_{\rm CO} = \sum_{n=-3}^{+3} L'_{{\rm CO},n}. \tag{2.7}$$

The seven-channel width corresponds to 218.75 MHz, or 2065 km/s at the average redshift of the galaxy catalog objects. This chosen spectral width allows for velocity offset between the CO emission and the optical lines used to determine the eBOSS redshifts, which can be quite significant (see §2.5). Instrumental aliasing between adjacent COMAP science channels is negligible (Lamb et al., 2022), so broadening considerations in the spectral axes are almost entirely astrophysical.

We use the average line luminosity to determine an average molecular gas density in the stack. We use the Bolatto et al. (2013) CO-to-$H_2$ conversion factor of 3.6 $M_\odot$(K km s$^{-1}$ pc$^2$)$^{-1}$ (calculated for the Milky Way) to determine the molecular gas mass associated with our stacked line luminosity, and then convert this mass into a molecular gas density $\rho_{\rm H_2}$ by dividing by the *comoving* volume of the stack aperture.

This is the simplest possible version of a stack, ignoring other available information such as the significance of the catalog objects' detections with eBOSS, or the optical luminosity of the catalog objects, both of which could also be used for weighting (see, for example, Sinigaglia et al., 2022). Functionally, this means assuming a $\delta$-function for the luminosity function. While we do plan to explore refinements to this methodology in the future, we assume for now that the noise in the Season 1 COMAP data cube is much greater than any variations across the eBOSS catalog.

## 2.5 Data verification

### Bootstrapped error analysis

While we believe the noise in our maps to be Gaussian-distributed ('white' noise: see §2.3, also Foss et al. 2022) over the entire map, this stacking analysis introduces a new sampling strategy, and it is possible that the white noise assumption does not hold under this new sampling. We therefore perform a spatial 'bootstrap' test on each stack to confirm the noise distribution in our maps, and to determine the uncertainty in each of the stacked values.

This is done by binning the actual eBOSS catalog by COMAP field and by redshift (into three redshift bins), and then generating an artificial catalog with the same overall redshift and field distribution as the real catalog but with random 3D positions. Variations in catalog number density between field, as well as instrumental effects such as the changing main beam with frequency and potential systematic errors correlating to on-sky position, will thus appear both in the bootstrapped stacks and the real stack. The LIM map is then stacked on the artificial catalog. This is repeated for 10000 different random artificial catalog iterations for each real stack run, to fully characterize the stack noise response.

**Constant-Luminosity Simulation Tests**

**Simulation Methodology**

As an end-to-end test of the stacking pipeline, we create mock COMAP observations of simulated cosmic CO signal, generate mock galaxy catalogs associated with these cubes, and perform the stacking analysis on these simulated data. Mock CO emission is generated by painting CO luminosities onto a dark matter halo catalog from peak-patch $N$-body simulations (Stein et al., 2019). Because we are looking to check the efficacy of the stacking pipeline rather than make realistic cosmological predictions, we paint each halo with the same luminosity regardless of its mass. We artificially broaden the CO emission line in each halo by an effective velocity $v_{\rm eff}$ of 319 km/s (see §2.5 below), again to more easily test the effects of the stacking pipeline. To generate mock COMAP data cubes, we additionally beam-smooth the emission in the spatial axes by convolving the mock cubes with a 2D Gaussian kernel with a 4-arcminute FWHM to approximate the COMAP main beam.

The artificial CO emission cube generated from these simulations is then converted into a timestream, scaled by a factor of 1000 to simulate an (unrealistically) strong detection, and added into real time-ordered data from the first COMAP observing season (Stutzer et al., *in prep*). The positions of the simulated halos should have no correlation with any existing real cosmic structure, so this is an excellent way to simulate observations with the actual noise structure that COMAP observes. This mock timestream is passed through the COMAP pipeline, so any effects of the several pipeline filters (Foss et al., 2022) will show up in the mock data cube.

In addition to the mock COMAP data cubes, we generate an artificial galaxy catalog by randomly selecting 1000 dark matter halos from the most massive 30% of halos in the peak-patch catalog (dark-matter masses between $9\times10^{13}$ M$_\odot$ and $4\times10^{12}$ M$_\odot$), and taking these halos to be the objects emitting brightly enough in some other galaxy tracer to be detected in a traditional galaxy survey. Only the 3D positions of the halos are required for the stacking analysis, so we do not calculate any other mock parameters for the galaxy survey.

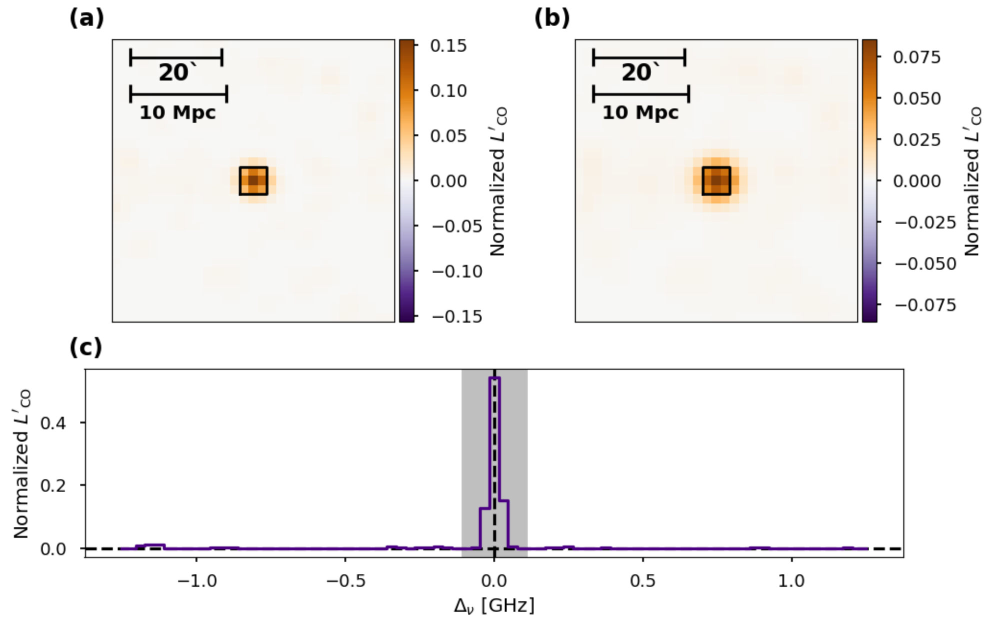


Figure 2.4: A noiseless simulated stacking detection, normalized to the expected integrated luminosity. (a) The central three frequency channels coadded to provide a view of the spatial distribution of signal. The black box in the center indicates the smaller aperture used to calculate single stack values. (b) The same spatial representation of the stack, smoothed with a 4′ Gaussian kernel to approximate the COMAP main beam. (c) The frequency spectrum of the stack. The value in each spectral channel is determined by summing over the spatial aperture indicated in (a). The three frequency channels making up the stack aperture are highlighted in grey. While the simulation is noiseless, neighboring halos may show up in some objects' cutouts, appearing as small fluctuations in the stacked spectra.

**Spatial attenuation due to COMAP beam**

We then run the stacking pipeline on ten different realizations of the mock data cube and galaxy catalog. This is done principally to confirm that the COMAP pipeline is not attenuating the stacking signal in any unexpected ways, and to determine the size of the stacking aperture discussed in §2.4. An example of a resulting simulated stack, shown in Figure 2.4, confirms that the $6' \times 6'$ spatial aperture we selected indeed encloses the majority of the stacked flux, without incorporating much empty space. The line luminosity output by this stack is 88.5% of the input value, meaning the input signal is attenuated by a factor of 1.13. This attenuation is due to the

COMAP main beam spreading signal outside of the stack's $3 \times 3$ spatial aperture (Figure 2.4). Instead of broadening the aperture further and incorporating more empty space into the stack, which would affect our sensitivity, we multiply the values output by the stack by 1.13 to correct for this attenuation.

### Spectral attenuation due to offsets between optical and CO redshifts

Although the cataloged eBOSS redshifts are considered very precise, high-redshift quasars often contain large systematic outflows/inflows that cause velocity offsets between optical emission lines (such as those used by eBOSS), and the CO emission, which traces the cold gas of the galaxy itself (Banerji et al., 2017; Herrera-Camus et al., 2020; Bischetti et al., 2021). This is an important consideration, as even small redshift uncertainties can cause significant attenuation in the signal of LIM/galaxy catalog cross-correlation analyses (Chung et al., 2019; Chung et al., 2021). The shot component of the cross-power spectrum (*i.e.*, the cross power at small scales, the most direct analogue to a stacking analysis) is where this effect is at its worst, although COMAP is most sensitive to larger spatial scales than those where attenuation due to redshift uncertainty is truly catastrophic.

To quantify systemic offsets between the cataloged redshift for our eBOSS sample ($z_{\rm opt}$) and the redshift of the CO emission ($z_{\rm CO}$), we assemble a sample of quasars that have (clean) spectroscopic optical redshifts in the COMAP redshift range from SDSS, and have been individually detected in a molecular gas tracer line. These come from three studies:

- Eight objects from Hill et al. (2019), who targeted CO(3-2) in 13 QSOs selected from the Keck Baryonic Structure Survey (KBSS) for additional cosmic web study with SCUBA-2. CO measurements used NOEMA (the NOrthern Extended Millimeter Arary; Chenu et al., 2016). The $z_{\rm opt}$ values for these objects are specifically from eBOSS.
- Two objects from the sample of nine QSO MUSEUM quasars targeted for CO(6-5) and CO(7-6) observations with the SEPIA180 (Belitsky et al., 2018) receiver on APEX in Muñoz-Elgueta et al. (2022; ME22). These objects also have optical redshifts from eBOSS.
- Two quasars from Riechers et al. (2011; R11), who targeted CO(1-0) in five quasar host galaxies using both the Extended Very Large Array (EVLA), and the Green Bank Telescope. Optical redshifts are from SDSS for these quasars.

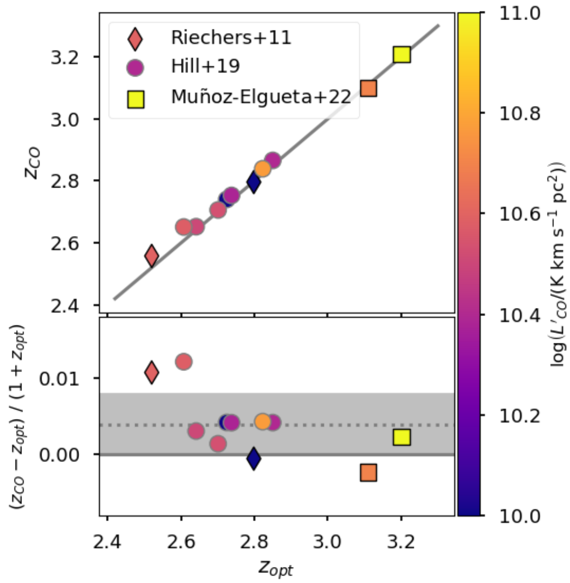


Figure 2.5: The optically-determined redshift values compared with CO-based redshifts for a selection of quasars individually detected in CO in the COMAP redshift range. The colormap is the extrapolated CO(1-0) luminosity of each object, and the line where $z_{\rm opt} = z_{\rm CO}$ is shown for reference. In the bottom panel we indicate the mean offset of the sample with a dotted line, and the $1\sigma$ scatter in that scatter with the shaded region.

In each case, we exclude objects with no available eBOSS/SDSS redshift or that fall far outside the COMAP redshift range. We also exclude objects with SDSS `NEGATIVE_EMISSION` warnings, for consistency with our own generated eBOSS catalog. Over this entire compiled catalog, the quasars have an average molecular gas emission-line FWHM of 319 km/s.

We plot $z_{\rm opt}$ against $z_{\rm CO}$ for each of these quasars in Figure 2.5. We find that $z_{\rm opt}$ is systematically smaller than $z_{\rm CO}$ with a mean value of $(z_{\rm CO} - z_{\rm opt})/(1 + z_{\rm opt}) = 0.00397 \pm 0.00408$. This offset does not appear to be correlated with either redshift or extrapolated CO(1-0) line luminosity. At the average redshift of our eBOSS quasar catalog, it corresponds to a negative frequency offset of 122 MHz, or 1185 km/s ($\sim$ 4 COMAP spectral channels). This is in good agreement with observed bulk velocity offsets in quasars, which are typically $\sim 1000 - 2000$ km/s (Orellana et al., 2011; Tytler and Fan, 1992).

In addition to exploring corrections in the bulk velocity offset between quasar gas phases, we also investigate the scatter in the relation. The standard deviation in $(z_{\rm CO} - z_{\rm opt})/(1 + z_{\rm opt})$ in the individually-detected quasar sample is $\sigma_{\delta z} = 0.00408$ from this effect alone, which is large enough to move a proportion of the signal from quasars in the stack outside of our defined 7-channel frequency aperture. We note that the variety of CO transitions used in the molecular gas measurements of the plotted quasars could create a larger scatter than a pure CO(1-0) analysis, as they trace slightly different phases of (dense) gas (e.g., Muñoz-Elgueta et al., 2022).

We perform a separate analysis to determine signal attenuation due to velocity offsets in quasars between the optical lines measured by eBOSS and the CO(1-0) line. This is addressed in detail in Appendix 2.10. We find that after this effect is accounted for, the input signal is attenuated by an additional factor of 1.58. As above, we multiply the values output by the stack by this factor to correct for this source of signal attenuation. Combined with the spatial attenuation discussed above, this is an overall attenuation factor of 1.79.

## 2.6 Results

### CO(1-0) Stack Results

The final COMAP/eBOSS stack is shown in Figure 2.6. We find, at our current sensitivity level, no detection of CO emission associated with the 243 eBOSS quasars in the COMAP fields. The results of the COMAP/eBOSS bootstrap test (discussed in §2.5) are shown in Figure 2.7. As expected, the maps remain consistent with Gaussian noise centered at $T_{\rm b} = 0$ K even in this stacking analysis. As this is a more robust calculation of the distribution of stack values expected from random noise, we use the 95% confidence region calculated from a Gaussian fit to these bootstrapped $L'_{\rm CO}$ values as the uncertainty in our stack.

We set a 95% upper limit on the CO frequency-integrated line flux of $S\Delta v \leq 0.129$ Jy km/s, corresponding to a limit on the CO line luminosity of $L'_{\rm CO} \leq 1.26 \times 10^{11}$ K km pc$^2$ s$^{-1}$ and a limit on the average molecular gas density in regions containing a quasar of $\rho_{\rm H_2} \leq 1.52 \times 10^8$ M$_\odot$ Mpc$^{-3}$.

Additionally, as the quasar velocity offset discussed in Section 2.5 is large enough to move any flux from our stack outside our chosen stack aperture, we correct for it by applying a 122 MHz frequency offset to our stack. We again find no significant CO emission, with a new upper limit of $L'_{\rm CO} \leq 8.69 \times 10^{10}$ K km pc$^2$ s$^{-1}$ (Figures

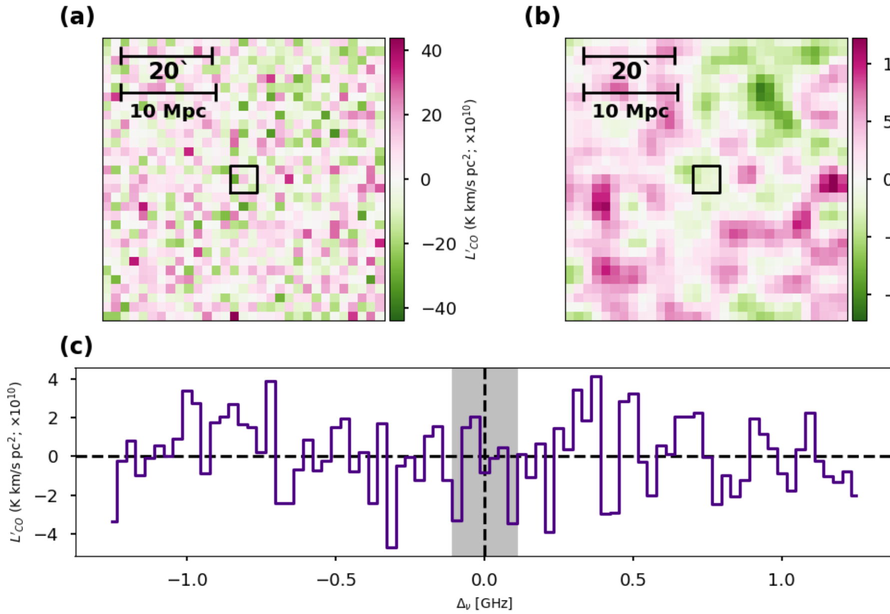


Figure 2.6: The result of stacking the COMAP data cubes on the 3D positions of eBOSS quasars. (a) The central seven frequency channels coadded to provide a view of the spatial distribution of signal. The black box in the center indicates the smaller aperture used to calculate single stack values. (b) The same spatial representation of the stack, smoothed with a 4′ Gaussian kernel to approximate the COMAP main beam. (c) The frequency spectrum of the stack. The value in each spectral channel is determined by summing over the spatial aperture indicated in (a). The seven frequency channels making up the 3D stack aperture are highlighted in grey. All distance scales are given as proper values.

2.10 and 2.11). Accounting for this frequency offset thus does not meaningfully change our results, at our current level of sensitivity.

#### Foreground Lines

A major concern for LIM experiments is contamination of signal at the redshift of interest by other (primarily foreground) spectral lines, which may also be redshifted into the frequency range of the experiment by coincidence (Cheng et al., 2016; Lidz and Taylor, 2016). CO LIM experiments suffer less from this type of contamination than LIM experiments targeting other spectral lines, because the (primarily molecular) lines that fall into our frequency range are expected to be faint compared to CO(1-0) (Cleary et al., 2022). Joint analyses such as stacking, which require emission in two different spectral lines to be detected from the target sources, are also expected to be fairly robust against foreground contamination (Chung et al.,

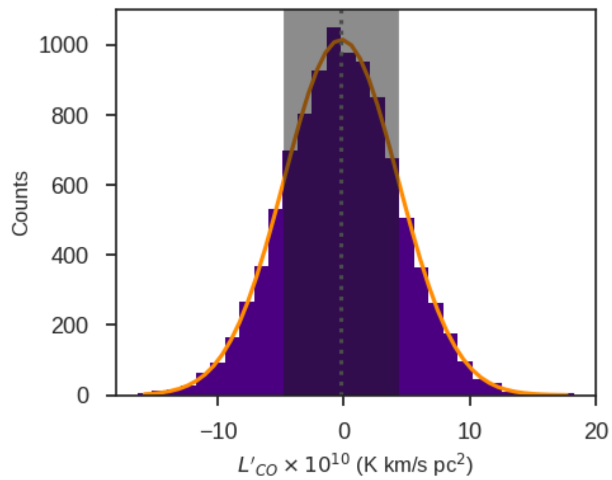


Figure 2.7: A histogram of the $L'_{\mathrm{CO}}$ values returned from 10000 stacks performed on random 3D positions in the COMAP data cubes, with the same spectral and field distribution of objects as the real eBOSS catalog. A Gaussian fit to the histogram is shown in orange. The 66% confidence interval of the fitted Gaussian (*i.e.*, 1 standard deviation about the fitted mean) is shown in dark gray, and the mean value is shown as a dotted gray line (see §2.4).

Table 2.2: 95% upper limits obtained by stacking the COMAP data cube on the CO(1-0) line and expected foreground emission lines.

| Spectral Line | # of Objects | $\nu_{\mathrm{rest}}$ (GHz) | Redshift Range | $S\Delta v$ 95% Limit (Jy km s$^{-1}$) | $L'$ 95% Limit (K km pc$^2$ s$^{-1}$) |
|---|---|---|---|---|---|
| $^{12}$CO(1-0) | 243 | 115.27 | 2.4 – 3.4 | $\leq 0.129$ | $\leq 1.26 \times 10^{11}$ |
| $^{12}$CO(1-0) with offset[a] | 244 | 115.15 | 2.4 – 3.4 | $\leq 0.270$ | $\leq 8.69 \times 10^{10}$ |
| HCN(1-0) | 595 | 88.63 | 1.6 – 2.4 | $\leq 0.196$ | $\leq 6.21 \times 10^{10}$ |
| CS(2-1) | 479 | 91.98 | 1.7 – 2.5 | $\leq 0.285$ | $\leq 1.18 \times 10^{11}$ |
| $^{13}$CO(1-0) | 311 | 110.20 | 2.2 – 3.2 | $\leq 0.236$ | $\leq 1.14 \times 10^{11}$ |
| CN$^{-}$(1-0) | 269 | 113.49 | 2.3 – 3.4 | $\leq 0.316$ | $\leq 1.29 \times 10^{11}$ |

[a]Shifted by the mean velocity offset of $z_{\mathrm{eBOSS}}$ from $z_{\mathrm{CO}}$ for the individually-detected quasars in §2.5 (see Figure 2.5).

2019). Because joint analyses are targeted toward specific emission lines, however, they present a unique opportunity for testing theoretical predictions that there will be minimal foreground emission present. Additionally, non-CO molecular emission lines are also interesting tracers of molecular gas in their own rights, and worth investigating.

Thus, we perform the stacking analysis on four other molecular emission lines which will likely fall into the COMAP frequency range, in addition to the $^{12}$CO(1-0) spectral line for which COMAP was designed. These lines are the $^{13}$CO isotope's $J = 1 \rightarrow 0$ rotational transition, the HCN $J = 1 \rightarrow 0$ transition, the CS $J = 2 \rightarrow 1$ transition, and the $CN^-$ radical's $J = 1 \rightarrow 0$ rotational transition. These lines are discussed in more detail in Appendix 2.12, or in Breysse et al. (2015). We make no detection of any foreground line. The resulting upper limits are shown in Table 2.2, and representations of the stacked cubelets are shown in Appendix 2.12.

**Forecasts**

As we are working with only the first year of data from COMAP's planned multi-year Pathfinder survey, we also project the stack sensitivity after five years of observing. The eBOSS survey has finished, so we only forecast improvements to the stack sensitivity arising from deeper COMAP data, instead of other possible stack improvements such as including more catalog objects or refining their redshift or positional accuracy. Under this assumption, the stack will improve by an identical factor to the sensitivity of the COMAP data cubes themselves.

Detailed projections were made by Foss et al., 2022 to forecast improvements to COMAP's sensitivity. Primarily, this improvement will be due to the obvious increase in total integration time by the end of the survey, but the Pathfinder's first season of operation was partially used for commissioning and refining the survey design, so we also expect significant improvements to the overall percentage of usable data. The effect of these improvements on COMAP's power spectrum sensitivity specifically are broken down in Table 2 of Foss et al. (2022), and collated in Equation 26. Almost all of them apply to the stacking analysis, but some are power-spectrum specific:

- The Foss et al., 2022 forecasts are made by projecting the sensitivity of a single COMAP field to the sensitivity of all three fields. As we are already using the full COMAP on-sky area in the stack analysis, we do not include the factors of $\sqrt{3}$ corresponding to this projection.

- Improvements to the filtering and map-making stages of the COMAP pipeline should result in these stages removing roughly 10% less actual signal from the maps (this is the transfer function efficiency, $E_{\mathrm{TF}}$), but will mostly act on the large angular scales to which the stack is not sensitive. We therefore do not multiply by $E_{\mathrm{TF}}$ when forecasting for the stacks.

- The process of calculating the power spectrum from the completed intensity map involves generating many jackknifed cross-spectra and only keeping those that pass quality tests. The retention efficiency of these steps, parameterized by Foss et al. (2022) as $E_{\chi^2_{P(k)}}$ and $E_{\mathrm{split}}$, will likely improve in future COMAP seasons as the earlier pipeline stages become more robust. Since these data cuts use statistics calculated from already-generated intensity maps, they only affect the power spectrum itself and we do not include them here.

All other efficiency improvements apply at the map level, and thus will apply to the stack. The forecast stacking sensitivity for the full COMAP Pathfinder survey then becomes

$$\sigma^{\mathrm{5yr}}_{\mathrm{stack}} = \frac{\sigma^{\mathrm{S1}}_{\mathrm{stack}}}{\sqrt{D^{\mathrm{5yr}}_{\mathrm{map}}}}, \tag{2.8}$$

where

$$D^{\mathrm{5yr}}_{\mathrm{map}} \equiv \frac{T^{\mathrm{S1}} E^{\mathrm{S1}}_{\mathrm{tot}} + (5 \cdot 365 - T^{\mathrm{S1}}) E^{\mathrm{proj}}_{\mathrm{tot}}}{T^{\mathrm{S1}} E^{\mathrm{S1}}_{\mathrm{tot}}}. \tag{2.9}$$

$T^{\mathrm{S1}}$ is the calendar duration of Season 1, 440 days. Both instances of $E_{\mathrm{tot}}$ indicate total observing efficiency, which combines bare on-sky observing efficiency $E_{\mathrm{obs}}$ with the fraction of data retained after all pipeline steps $E_{\mathrm{data}}$. For the exact breakdown of these quantities, see Foss et al., 2022. $E^{\mathrm{S1}}_{\mathrm{tot}}$ is the actual Season 1 efficiency, and $E^{\mathrm{proj}}_{\mathrm{tot}}$ is the projected value after five years. Adjusting the values calculated by Foss et al., 2022 to exclude the power spectrum-only factors discussed above, these are 9.15% and 33.2%, respectively. We thus project that the sensitivity of the COMAP/eBOSS stack will improve by a factor of $D^{\mathrm{5yr}}_{\mathrm{map}} = 25.2$, corresponding to a CO(1-0) flux sensitivity $\sigma^{\mathrm{5yr}}_{\mathrm{stack}}$ of 0.026 Jy km s$^{-1}$ at the 95% level.

While we have chosen in this work to focus specifically on the CO properties of quasars, we note that more general high-redshift spectroscopic catalogs can be used for stacking if the goal is to extract the CO properties of the universe as a whole. In particular, we have already made projections (Chung et al., 2019; Chung et al., 2022) for the cross-correlation of COMAP data with the HETDEX (the

Hobby-Eberly Telescope Dark Energy EXperiment; Gebhardt et al. 2021) catalog of Lyman-Alpha Emitting galaxies (LAEs). We plan in future to perform a stacking analysis on the same catalog. To first order, the stack sensitivity increases with the number of objects $N$ in the spectroscopic catalog by a factor of $\sqrt{N}$, so we project significant sensitivity increases with this larger catalog. We note, however, that different galaxy tracers are subject to different biases, and any anticorrelation between the brightness of the tracers used and the CO brightness of a given object will affect this prediction; we plan to quantify this effect in future works.

## 2.7 Discussion

In general, how we interpret the results of LIM/galaxy survey stacking analyses depends on the relative contribution of the CO emission associated with the catalog galaxy compared to that of other sources within the COMAP beam. We can bracket the range of possible interpretations by considering the two extreme cases: i) the galaxies traced by our chosen survey dominate the brightness temperature of their surroundings (making the stack simply a measurement of the average CO luminosity, $L'_{\mathrm{CO}}$, of the survey objects), and ii) the cataloged galaxies are unspectacular themselves, but instead tend to fall into overdense regions of the large-scale structure (making the stack an measurement of the average CO luminosity of bright regions of the universe, which can be converted into their molecular gas density, $\rho_{\mathrm{H}_2}$).

Due to the dearth of CO observations of quasars at cosmic noon, it is not yet clear which of these cases is closer to the truth; we discuss the available evidence below. Neither assumption, however, invalidates the other in the context of the upper limits we present here. Quasar companions will not be distinguishable from the objects themselves at COMAP resolutions, meaning any nearby objects will cause a source confusion effect pushing the measured $L'_{\mathrm{CO}}$ value upwards. Conversely, abnormal amounts of CO emission from the quasars themselves will drive the local $\rho_{\mathrm{H}_2}$ upwards. In both cases, therefore, the stack values should likely be treated as upper limits even once a confident detection is made at this level of analysis.

### CO in Quasars

The response of the molecular gas in quasar host galaxies to the extremely active SMBH at their centers is likely extremely complicated, especially at cosmic noon. The abundances of both quasars and gas-rich, sub-mm-bright galaxies (SMGs) seem to peak during this epoch, suggesting a link between the two (Simpson et al., 2012). Indeed, both theoretical and observational arguments have suggested that the SMBH

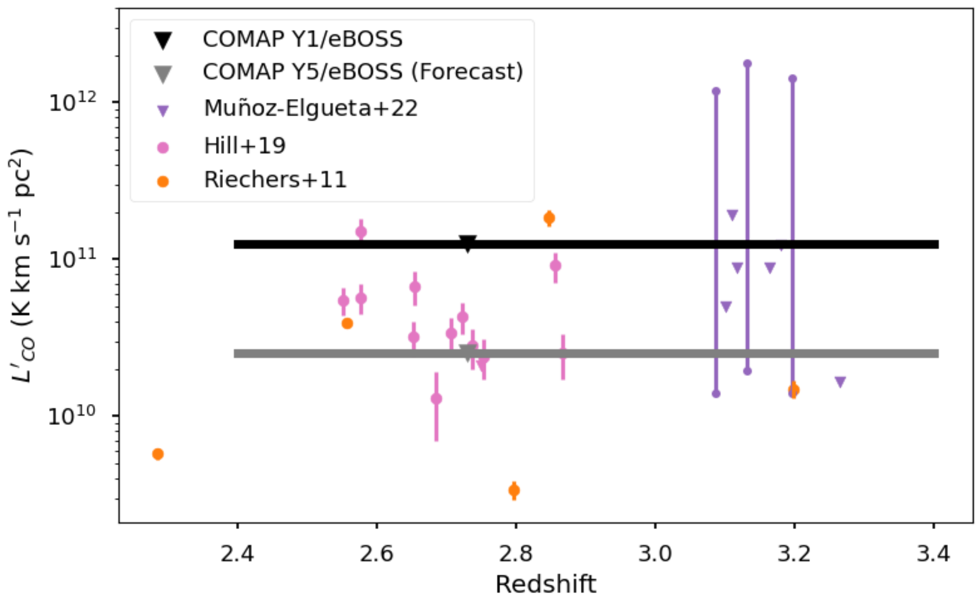


Figure 2.8: The COMAP/eBOSS 95% upper limit on $L'_{\rm CO}$ (the black bar), plotted against quasars detected individually in CO from Riechers et al. (2011), Hill et al. (2019), and Muñoz-Elgueta et al. (2022). Because Muñoz-Elgueta et al. (2022) report a range of molecular gas mass values $M_{\rm H_2}$ from SED fitting of multiple lines, we show the $L'_{\rm CO}$ values determined from those values as a range using vertical bars. Upper limits are shown as inverted triangles. The extent of the COMAP/eBOSS bar indicates the redshift range to which COMAP is sensitive, and the marker is located at the mean redshift of the eBOSS objects on which we stack. A forecast for the stack sensitivity at the end of the full COMAP Pathfinder Survey is plotted in gray (see §2.6).

activity driving quasars is triggered from normal SMGs by mergers rich in cold and cool gas (Sanders et al., 1988b; Hopkins et al., 2008; Brusa et al., 2018; Herrera-Camus et al., 2020; Bischetti et al., 2021), suggesting that, at least in their initial phases, quasars should be extremely rich in cold gas.

Whether that cold gas survives the subsequent barrage of heat and pressure from its central SMBH, however, is still uncertain. SMG studies of molecular gas have shown that the presence of an AGN does not seem to significantly affect the gas properties of its host (Bothwell et al., 2013; Banerji et al., 2017), and observed gas masses in quasars themselves follow typical SMG values closely (Hill et al., 2019), but other observations have shown that CO Spectral-Line Energy Distributions (SLEDs) in quasar hosts tend towards higher excitation states than normal (e.g., Brusa et al., 2018). At least for the purposes of COMAP, which measures the lowest CO excitation state, this would read as damped CO emission in quasar hosts. Additionally, many separate quasar studies have observed large-scale outflows of all gas phases from the host galaxy, although a fraction of this gas may remain in the molecular phase, and thus be observable with COMAP (Bothwell et al., 2013; García-Burillo et al., 2014; Feruglio et al., 2015; Brusa et al., 2018).

In the limiting case that much of the CO luminosity contributing to the LIM maps at the quasar positions comes from the quasars themselves, then the upper limit on $L'_{\rm CO}$ determined from our COMAP/eBOSS stack becomes a direct upper limit on the average CO line luminosity of the 243 eBOSS quasars included in the stack. Under this analysis, the $L'_{\rm CO}$ value can only ever be an upper limit, because the cosmic volume covered by the stack is large (9 Mpc $\times$ 9 Mpc $\times$ 2065 km/s, in proper units) compared to an individual quasar, so it is unlikely that the quasars are the only sources contributing to the CO luminosity of the region. Because we are reporting an upper limit, however, this remains valid.

The measured average $L'_{\rm CO}$ value is shown in Figure 2.8, plotted against other measures of quasar line luminosity in a similar redshift range, for comparison. These include the sample of 13 QSOs from KBSS, detected in CO(3-2) with NOEMA by Hill et al., 2019; the sample of 9 objects from the QSO MUSEUM, surveyed with APEX in CO(6-5) and CO(7-6) by Muñoz-Elgueta et al., 2022; and the sample of 5 QSOs detected in CO(1-0) by Riechers et al. (2011) using the Extended Very Large Array and the Green Bank Telescope.

The Hill et al. (2019) CO(3-2) measurements are converted into CO(1-0) line luminosities using the conversion factor $L'_{\rm CO(3-2)}/L'_{\rm CO(1-0)} = 0.97 \pm 0.19$ from Carilli and Walter (2013). In several cases, a single KBSS object was found to be associated with two different CO sources slightly offset on-sky. For each of these objects, we sum both CO sources together to obtain the plotted $L'_{\rm CO}$ value, as both sources would fall well into the COMAP beam. Muñoz-Elgueta et al. (2022) converted their higher-$J$ CO measurements into molecular gas masses directly, by fitting to CO SLEDs, incorporating also the [CI] luminosity. In order to determine CO(1-0) line luminosities associated with the objects, we extrapolate their calculated $M_{\rm H_2}$ values to $L'_{\rm CO}$ values using the Milky-Way $\alpha_{\rm CO}$ conversion factor of 3.6 $\rm M_\odot$ (K km s$^{-1}$ pc$^{-2}$)$^{-1}$ (Bolatto et al., 2013), for consistency with our own previous analyses. These objects are mostly nondetections, and the SLED fitting returned a range of $M_{\rm H_2}$ values, so these ranges are what we plot in Figure 2.8. Riechers et al. (2011) report CO(1-0) line luminosities directly.

Figure 2.8 shows that the COMAP/eBOSS stacking upper limit is already comparable with the line luminosities of (the brightest) individually surveyed quasars in its redshift range, even using only the first season of COMAP data. In some cases, the COMAP value is actually a stricter limit, particularly for the APEX objects (Muñoz-Elgueta et al., 2022). This illustrates a powerful potential application for LIM – directly detecting CO in individual objects becomes prohibitively expensive very quickly at these redshifts. A statistically-significant survey would require integration times that are not practical on high-demand community instruments. LIM-based stacking analyses, able to efficiently survey large samples of galaxies, could thus act as an important complement to these individual detection-based surveys by placing constraints on the ensemble emission from many galaxies. The minimum spatial scales accessible from a LIM instrument are large compared to galaxy sizes, meaning such analyses would have to remain upper limits even if a detection were made, but these still have the potential to be extremely enlightening.

Under this interpretation, our forecasted COMAP Pathfinder sensitivity to CO luminosity from a stack on the same 243 eBOSS quasars after five years of observing falls below the mean luminosity of the individually-detected quasars in Figure 2.8, meaning it has the potential to be genuinely constraining—we will likely be able to tell with COMAP Y5 if these three samples are representative of the eBOSS quasar population as a whole. If we are able to constrain the $L'_{\mathrm{CO}}$ of eBOSS quasars to this degree, it will provide significant insight into the complex feedback cycles present in quasars, and thus into star formation processes as a whole at cosmic noon.

**Cosmic Molecular Gas Density**

The other limiting case—that quasars trace large-scale structure but do not themselves dominate the local CO emission—is perhaps better-supported, as it follows from hierarchical structure formation. SMBH mass and galaxy mass have been shown several times to be correlated, meaning quasars will preferentially be found in massive galaxies that are likely to be at the center of large dark-matter halos (e.g., Gebhardt et al., 2000; Hopkins et al., 2008). This argument is additionally supported by clustering measurements of both BOSS and SDSS Stripe 82 quasars (White et al., 2012; Timlin et al., 2018), as well as observations of the individual objects shown in Figure 2.8: a quarter of the Hill et al. (2019) objects, for example, have nearby CO-bright companions (also Banerji et al., 2017; Banerji et al., 2018; Decarli et al., 2021; Bischetti et al., 2021). The argument that quasars require major

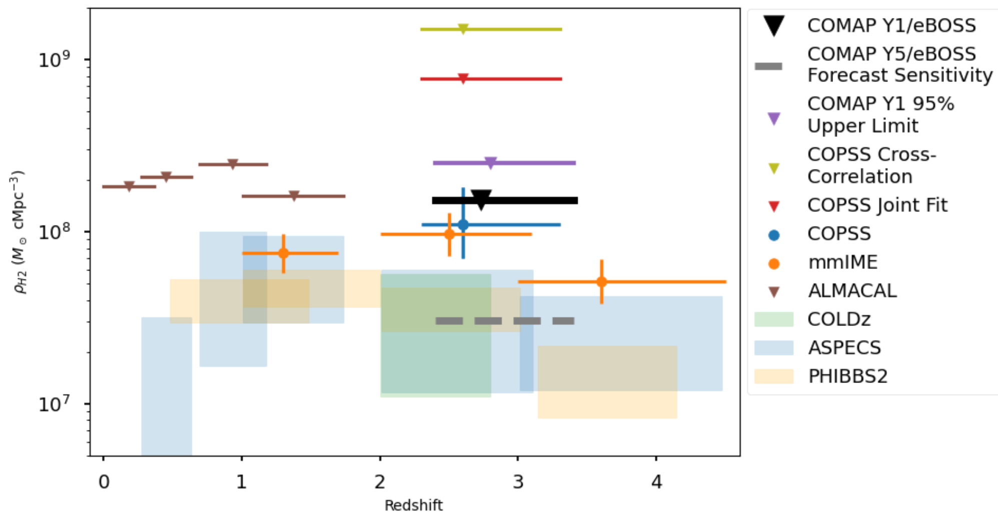


Figure 2.9: The upper limit on the average molecular gas density determined from the COMAP/eBOSS stack (black line), compared to other $\rho_{\mathrm{H_2}}$ measurements made using different survey techniques. The upper limit from the COMAP Season 1 power spectrum analysis is shown in purple. Also shown is the COPSS LIM survey's measured value (Keating et al., 2016), as well as upper limits from COPSS joint analyses with galaxy surveys (Keenan et al., 2022). Shaded boxes are measurements made by large-scale traditional surveys COLDz (Riechers et al., 2019), ASPECS (González-López et al., 2019), and PHIBSS2 (Lenkić et al., 2020). We show the forecast $2\sigma$ sensitivity level of the COMAP survey after five years stacked on eBOSS as a grey upper limit.

mergers to ignite also necessitates at least one nearby massive galaxy in their recent past.

In this limit, where the quasar host itself is contributing negligibly to a local overdensity of CO emission, we can use the stack's upper limit on $T_\mathrm{b}$ to calculate the average molecular gas density $\rho_{\mathrm{H_2}}$ in bright regions at the COMAP redshifts. As the stack is only sensitive to a single, $\sim 9$ (comoving) Mpc spatial scale (the size of our chosen aperture), this value is most similar to a random 'shot' power component of the $\rho_{\mathrm{H_2}}$ determined from a power spectrum analysis. We plot this value in Figure 2.9, alongside several other measurements of the same value made using a variety of techniques, including COMAP's own auto-spectrum based early science constraint (Chung et al., 2022). These other measurements include:

- COPSS (Keating et al., 2016), a LIM survey, targeting the same CO(1-0) transition as COMAP. COPSS used the Sunyaev-Zeldovich Array, and is subject to interferometric insensitivity to large scales. The $\rho_{H_2}$ value from COPSS is thus based mainly on the shot-noise component of the CO power spectrum, converted to a molecular gas density using the Milky-Way $\alpha_{CO}$ value of 3.6 $M_\odot$ (K km $s^{-1}$ $pc^{-2}$)$^{-1}$ (the same value we adopt in this work). Two other COPSS-determined values (Keenan et al., 2022) are plotted, both determined using joint analyses combining COPSS data with (after correcting for main beam weighting) 145 spectroscopically-confirmed cataloged galaxies. Keenan et al. (2022) cross-correlated the COPSS data with a data cube obtained by gridding the catalog objects in 3D space (yielding one limit on $\rho_{H_2}$) and fit the resulting cross-power spectrum jointly with both the auto-power spectrum of COPSS alone and the galaxy-galaxy power spectrum of their galaxy catalog (yielding a second $\rho_{H_2}$ limit).

- ALMACAL (Klitsch et al., 2019) used archival ALMA calibrators as background sources to search blindly for CO absorption lines. As these calibrator sources must be observed frequently, ALMACAL boasts $>$ 1500 hours of observations on a community instrument over a wide enough area to not be limited by cosmic variance. However, these calibrators do not extend past $z \sim 2$.

- The Millimeter-wave Intensity Mapping Experiment (mmIME) used a combination of archival and targeted ALMA and Atacama Compact Array (ACA) observations at frequencies tracing multiple CO transitions (Keating et al., 2020). As with COPSS, these constraints are primarily constraints on the spectral shot power.

- Finally, the shaded boxes in Figure 2.9 correspond to $\rho_{H_2}$ measurements made using traditional (*i.e.*, resolved and targeted) galaxy surveys. These include ASPECS (Uzgil et al., 2019; González-López et al., 2019) and PHIBBS2 (Lenkić et al., 2020), each tracing multiple CO rotational transitions, as well as COLDz (Riechers et al., 2019), which traces exclusively CO(1-0) at $z \sim 3$. While each of these traditional surveys provide $\rho_{H_2}$ values that seem like much better constraints than any of the LIM-based numbers, they cover much smaller survey volumes and are thus extremely subject to cosmic varience effects. Additionally, as traditional surveys, they are insensitive to any emission below their detection limits, and so will miss any contribution to the total gas density from fainter objects, likely underestimating the total gas density.

Clearly, these $\rho_{H_2}$ measurements were made using a diverse range of different techniques, so some of the scatter in Figure 2.9 can be attributed simply to the biases associated with each strategy. There is also likely additional scatter present due to variations in the CO-to-$H_2$ ratio, both astrophysically and in the conversion factor $\alpha_{CO}$ chosen for each analysis, as well as the relative intensity of the different CO line transitions, which will also change based on environment.

In our case, while we are indeed subject to the assumptions associated with our choice of $\alpha_{CO}$ and the bias of CO emission as a galaxy tracer, we are also only investigating the regions of space associated with eBOSS quasars, so the stacking analysis necessarily introduces a secondary bias factor. As discussed above, regions of this size surrounding quasars are likely to be more dense than the universe as a whole, biasing the stacked $\rho_{H_2}$ value upwards because of clustering. The upper limit we report here is therefore likely an overestimate—an analysis which fully explored clustering effects would be more constraining. Because we report an upper limit we find this to be acceptable. We will explore the effects of clustering more thoroughly in future work.

Due to this additional bias, the $\rho_{H_2}$ value from the eBOSS stack is more similar to a cross-correlation analysis, such as the early WMAP cross-correlations with SDSS and BOSS quasars (Pullen and Hirata, 2013a; Pullen et al., 2013b). Unlike these cross-correlation analyses, the stack's $\rho_{H_2}$ value is only sensitive to a single spatial scale (the scale of the 3D stack aperture, $\sim$ 9 Mpc). We note that, because of our definition of a voxel in the COMAP data cubes, this scale does vary across the cube, and the $\rho_{H_2}$ values being stacked are thus probing variable volumes. We explore this more thoroughly in Appendix 2.11.

Even with these caveats, however, our predicted value is comparable to the values returned from the different COPSS joint analyses, reinforcing the viability of stacking as a LIM tool introduced by Keenan et al., 2022. More promising is our forecast sensitivity for the COMAP Y5/eBOSS stack, which actually falls into the regions reported by traditional galaxy surveys. As these galaxy-survey values are likely underestimates of the actual cosmic $H_2$ density, our stacking analysis should detect emission from even very pessimistic models of cosmic CO, provided our forecasts hold true. As we already predict that our autocorrelation analysis will allow us to discriminate between CO models by COMAP Y5 (Chung et al., 2022), stacking will provide another valuable tool in characterizing the amount and properties of molecular gas emission at cosmic noon.

Additionally, quasar catalogs are far from the only spectroscopic information available on astrophysical objects at our redshifts. As discussed in §2.6, increasing the number of objects in our galaxy catalog will increase the sensitivity of our stack, but introducing new objects detected using different tracers will also provide new opportunities for analysis. It will be very interesting to compare, for example, the CO properties of the LAEs in the HETDEX survey (Gebhardt et al., 2021) to the CO properties of the eBOSS quasars. Galaxy evolution is an inherently multitracer problem, so there will be significant analysis benefits to collating as many different surveys (whether traditional resolved surveys or LIM experiments) of the same region of space as possible. As the LIM field progresses, there will be more and more opportunities to do so.

## 2.8 Summary and conclusions

By stacking CO line-intensity maps from COMAP's first year of science operations on the 3D positions of quasars from the SDSS eBOSS survey, we have obtained a new upper limit on the cosmic abundance of molecular hydrogen gas. We describe the methodology behind the stack in detail, including the COMAP RMS-based weighting scheme we use when coadding and the bootstrapped error analysis technique we use to quantify uncertainties. We use the stack to search additionally for interloper emission from foreground galaxies in four molecular lines, finding no evidence of a detection in any case. We can interpret our stacked upper limit as a constraint on the CO emission from either the cataloged quasars themselves, or the regions of the universe surrounding those quasars. Likely, the most realistic interpretation is some combination of both cases. Additionally, we forecast stacking results for the COMAP Pathfinder survey after five years of observing.

In the limit that any potential signal in the stack is dominated by CO emission from the eBOSS quasars themselves, we treat our measurement as the average CO luminosity of these objects. We compare this average luminosity to resolved CO measurements of quasars in the COMAP redshift range, and find that our upper limit already probes CO luminosities fainter than those of the brightest objects observed to date. The quasar studies to which we compare typically required hours of observations using large community facilities to detect CO in only a handful of objects. Determining the CO properties of a large sample of objects such as our subset of the eBOSS catalog would thus be prohibitively expensive through traditional means. LIM stacks, therefore, could potentially be extremely useful as a tool for studying large samples of high-redshift galaxies and quasars.

Conversely, in the limit that the eBOSS quasars do not significantly contribute to the integrated CO emission of their surroundings, the stacked flux measurement can be converted to a measurement of the cosmic molecular gas density. While we compare this value directly to other $\rho_{\mathrm{H}_2}$ measurements from various sources, its interpretation is somewhat more complicated, as the molecular gas density in the regions traced by the stack depends heavily on the bias of the cataloged quasars towards large-scale structure. Additionally, the stack probes only a single spatial scale, unlike the more conventional power spectrum-based LIM measurements. These differences make stacking an excellent complement to other LIM analyses.

We propose, therefore, that stacking analyses with existing galaxy catalogs are a promising addition to the LIM analysis toolbox, especially when using LIM to approach as complex and multi-tracer a problem as galaxy evolution. To take full advantage of their potential benefits, stacking analyses should be performed on catalogs using as many different galaxy tracers as possible, to probe this phase space more fully. We aim to investigate stacking as a galaxy analysis tool more fully in future works, including by stacking on the extensive HETDEX catalog of Lyman-$\alpha$ emitters (Gebhardt et al., 2021).

## 2.9 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant Nos. 1518282, 1517108, 1517598, 1517288, 1910999 and 2206834; 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.

DD acknowledges support from NSF Award 2206834. KC acknowledges support from NSF Awards 1518282, 1910999 and 2206834. PCB is supported by the James Arthur Postdoctoral Fellowship. DC 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. HI, HKE, and IW acknowledge support from the Research Council of Norway through grant 251328. JG acknowledges support from the Keck Institute for Space Science, NSF AST-1517108 and University of Miami and Hugh Medrano for assistance with cryostat design. JK is supported by a Robert A. Millikan Fellowship from Caltech. HP's research is supported by the Swiss National Science Foundation via Ambizione Grant PZ00P2_179934.

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, Observatário 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 made use of NASA's Astrophysics Data System Bibliographic Services. 

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

## 2.10 Appendix: Signal attenuation due to finite spectral aperture size

As discussed in Section 2.5, quasars have inherently large velocity offsets between different redshift tracers, due to inflowing and outflowing of different phases of the gas in the galaxy. In the stack, this manifests as an offset of the centroid of the stacked emission in the spectral axis, and a broadening of this emission. This will serve to reduce the signal-to-noise ratio, as emission is spread across spectral channels, and also attenuate the stack luminosity, as these velocity offsets can be wide enough to move signal out of the 7-channel aperture used to calculate the stack luminosity. This effect is compounded by other sources of spectral broadening in the stack, such as the inherent FWHM of the CO line. We defer the full details of the effects of spectral broadening on a stacking analysis to a future work, but we here explore the specific case of eBOSS quasars.

We use the previous studies of molecular gas in individual eBOSS quasars shown in Figure 2.5 to determine (1) an average CO FWHM for these objects of 319 km/s, (2) an average offset between the molecular redshift and the eBOSS redshift of $(z_{\rm CO} - z_{\rm opt})/(1 + z_{\rm opt}) = 0.00397$, and (3) a scatter in that offset of 0.00408. We note that these correction factors are illustrative values, limited by the small number of eBOSS quasars detected in CO. We would require many more CO measurements of individual quasars to determine a more confident calibration, and these values are not available at the moment. Additionally, we note that the assembled literature measurements use a variety of tracers which may trace slightly different phases of molecular gas, and thus this scatter may be enhanced compared to a single-tracer study such as COMAP. We deal with the average offset by simply shifting the stack centroid in the spectral axis by this amount. This remains a nondetection (as discussed in Section 2.5); the resulting stack is shown in Figures 2.10 and 2.11.

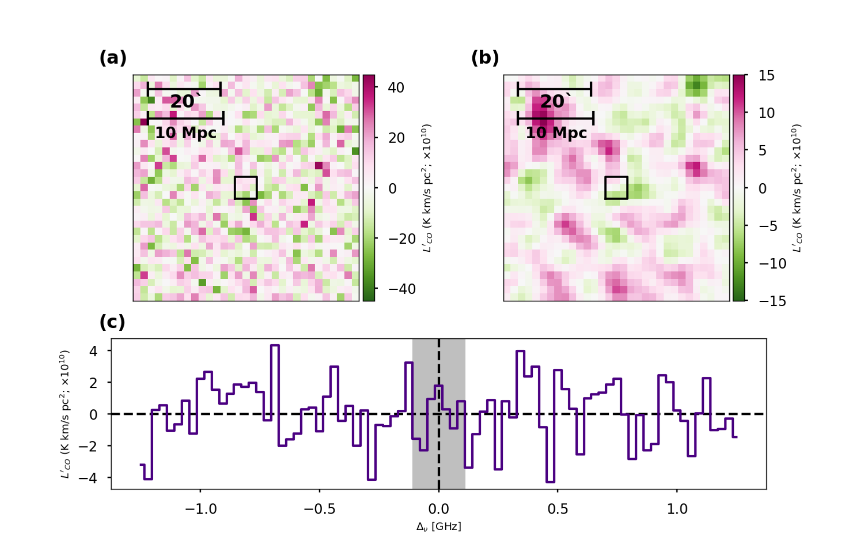


Figure 2.10: The 3D stack performed on the eBOSS catalog using the $^{12}$CO(1-0) emission wavelength offset by the frequency offset found in individually-detected quasars (see §2.5). See Figure 2.6 for panel descriptions.

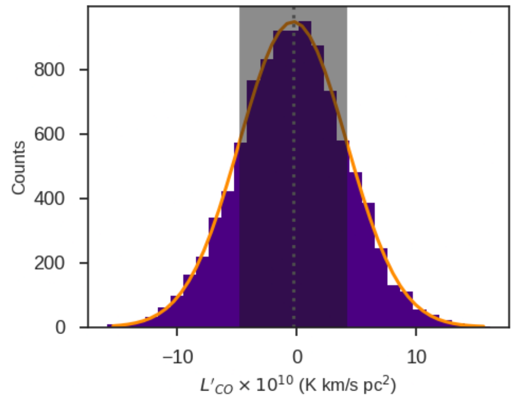


Figure 2.11: The bootstrapped uncertainty in $L'_{\mathrm{CO}}$ for the offset $^{12}$CO(1-0) stack (see Figure 2.7).

To address the scatter in the velocity offset between tracers (functionally, this introduces a redshift uncertainty in the stack), we use the constant luminosity simulations discussed in Section 2.5 as a framework. We artificially set the linewidth of each CO emitter to be 319 km/s, as measured in Section 2.5. We test the attenuation from various values of scatter in $(z_{\mathrm{CO}} - z_{\mathrm{opt}})/(1 + z_{\mathrm{opt}})$ by offsetting the spectral centroid of each halo individually by a randomly-selected amount pulled from a normal distribution with standard deviation set to the $(z_{\mathrm{CO}} - z_{\mathrm{opt}})/(1 + z_{\mathrm{opt}})$ scatter. This is done before stacking. We then stack the simulated data cube as normal. We perform 10 different iterations of this analysis for each value of $(z_{\mathrm{CO}} - z_{\mathrm{opt}})/(1 + z_{\mathrm{opt}})$ we test, using the same underlying DM halo distribution as a baseline and varying only the halo velocities. The resulting spectra for the scatter value we measure from individual eBOSS quasars are shown in Figure 2.12. In Figure 2.13, we show how this scatter attenuates signal compared to an un-broadened stack. Using the values determined from individual eBOSS quasars, we find that the returned signal is 63.1% of the actual value, corresponding to an attenuation factor of 1.58.

## 2.11 Appendix: Cosmic volume variation across the stack

One constraint with our current stacking methodology is that the physical volume of each cubelet is redshift-dependent – the COMAP data cubes stretch from redshift 2.4 to redshift 3.4, so the $6' \times 6' \times 218.75$ MHz aperture we sum to determine quantities such as $L'_{\mathrm{CO}}$ and $\rho_{\mathrm{H_2}}$ varies in size from 3.0 Mpc $\times$ 3.0 Mpc $\times$ 2519 km/s at low redshifts to 2.7 Mpc $\times$ 2.7 Mpc $\times$ 1929 km/s at high redshifts. We investigate here how this affects the stack. Ideally, this would be addressed by rebinning the data

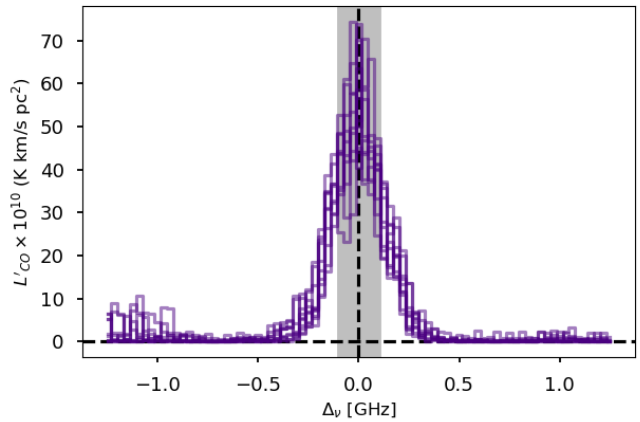


Figure 2.12: Spectra of simulated stacks, offset to account for the velocity difference between eBOSS-determined redshifts and CO-determined redshifts $(z_{\rm CO} - z_{\rm opt})/(1 + z_{\rm opt}) = 0.00408$). Ten different iterations are shown. The fluctuations in the fringes of the spectrum are due to neighboring objects falling into some cutouts.

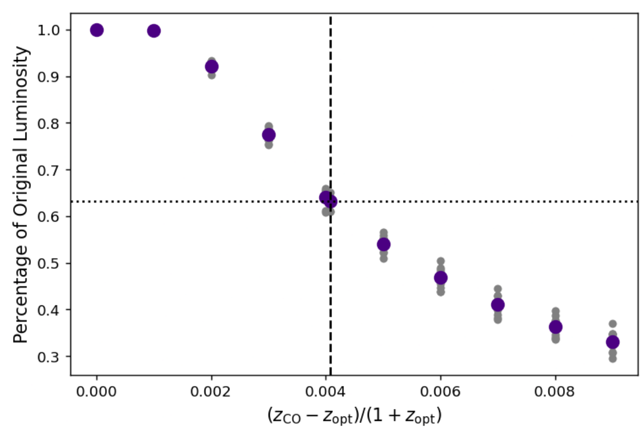


Figure 2.13: The percentage of the input luminosity returned by a constant-luminosity simulated stack as a function of an induced scatter between $z_{\rm opt}$ and $z_{\rm CO}$ in the individual halos making up the stack. The average scatter calculated from the independently-detected quasars presented in Section 2.5, taken to be the average of our eBOSS quasar stack, is indicated with the vertical black dashed line. We repeat the test ten times – the average values are shown as large purple dots.

at the map level to create voxels that are constant physical sizes, and we plan to implement this in future work.

As in Appendix 2.10, we investigate this effect using the constant-luminosity simulations from Section 2.5, again artificially imposing a linewidth of 319 km/s for each CO emitter. We generate three different 1000-halo simulations in which emitters are only located in a specific redshift range: a ‘low-z’ simulation with halos

$2.39 < z < 2.42$, a 'mid-z' simulation with halos $2.82 < z < 2.86$, and a 'high-z' simulation with halos $3.89 < z < 3.43$. Each of these redshift ranges corresponds to 12 spectral channels in a COMAP data cube, meaning that the density of sources is the same in each simulated cube. We create three different iterations of each cube at each redshift.

In the spatial axes, the size of the voxel is varying across the stack at a similar rate to the size of the main beam, which is also defined in on-sky coordinates (Lamb et al., 2022). At the COMAP spatial resolution, nearly any astrophysical object is a point source, so the beam defines the spatial distribution of the emission. The distribution of the emission between map voxels is therefore extremely similar across the entire redshift range – we are stacking together point sources with a similar measured distribution, even if their physical distributions vary.

In the spectral axis, any line broadening is from astrophysical sources and not instrumental sources, as our spectral aliasing is near-zero. At low redshifts, this is $1.15\times$ the width of a map channel, whereas at high redshifts it is $0.9\times$ the width of a channel. An identical CO emitter will thus (unphysically) appear more strongly peaked at high redshifts than at low redshifts (Figure 2.14). However, the same amount of emission is still present (spread between more spectral channels) and our 7-channel aperture is more than enough to encompass this spread, so the returned $L'_{\rm CO}$ value remains the same.

While the line luminosity is unaffected by the change in cosmic volume across the stack, the resulting molecular gas density, $\rho_{\rm H_2}$, does vary across the stack, because the volume being averaged over varies across the stack. This is because, in our simulations, all of the input luminosity is coming from a single source, but being spread over variable volumes across the stack due to the main beam. Volumes vary by roughly 37% across the full range of the stack—the same percentage difference we observe in the $\rho_{\rm H_2}$ values output from these simulations.

## 2.12 Appendix: Foreground stacks

As discussed in Section 2.6, there are several other molecular spectral lines whose emission may fall into the COMAP frequency range. These spectral lines could potentially contaminate the expected CO(1-0) LIM signal. Conversely, these spectral lines are important tracers of galaxy properties in their own rights, and any signal would be interesting and informative. We present stacks on each of these foregrounds lines below.

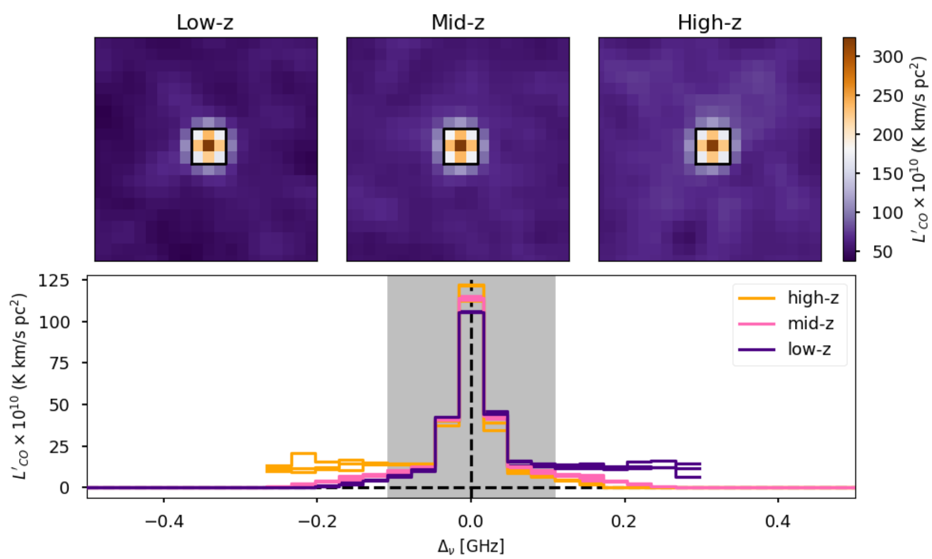


Figure 2.14: A comparison of constant-luminosity, constant-linewidth simulations at three different redshifts spanning the width of the COMAP data cubes. The spatial distribution of (stacked) emission at each redshift is shown in the top three panels, and the bottom panel shows the spectrum of the stacked cubelets. While the high-redshift stacks are more peaked than the low-redshift stacks, each stack returns the same CO line luminosity. The low-level fringes in the spectra are background emission from objects other than the one being stacked appearing in the spectrum – asymmetries appear due to the low-redshift and high-redshift cutoffs.

**HCN(1-0)**

HCN(1-0) is a tracer of dense ($n(H_2) \geq 3 \times 10^4$ cm$^{-3}$) molecular gas in galaxies, associated primarily with the star-forming cores of Giant Molecular Clouds (Gao and Solomon, 2004). HCN luminosity correlates strongly with CO luminosity, although in a given galaxy $L'_{\rm HCN}$ is roughly an order of magnitude less than $L'_{\rm CO}$. HCN has a rest frequency of 88.63 GHz, and 595 eBOSS quasars with this rest frequency fall into the COMAP volume.

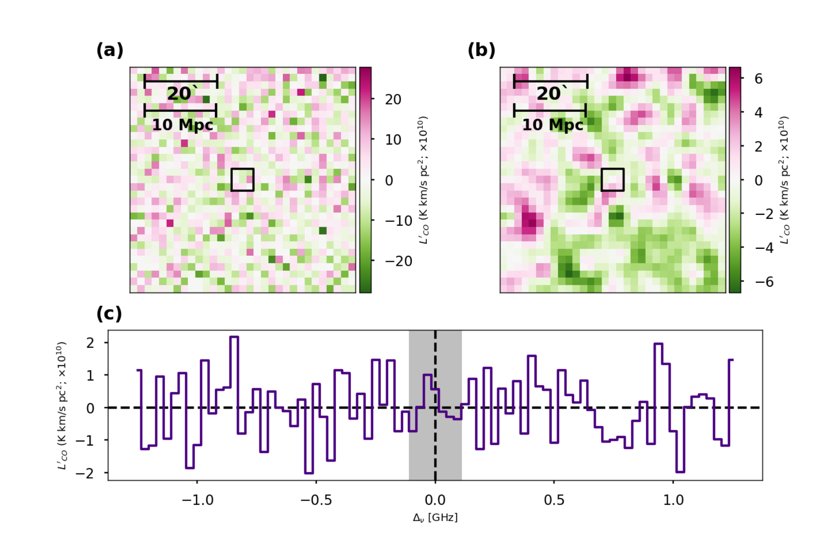


Figure 2.15: The 3D stack performed on the eBOSS catalog using the HCN(1-0) emission wavelength. See Figure 2.6 for panel descriptions.

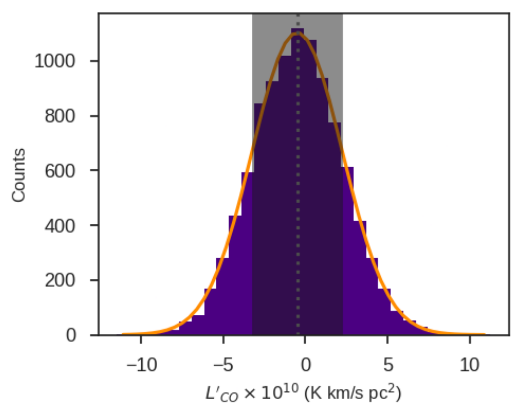


Figure 2.16: The bootstrapped uncertainty in $L'_{\mathrm{CO}}$ for the HCN(1-0) stack (see Figure 2.7).

### CS(2-1)

CS is interesting as a dense gas tracer independent of HCN/$CN^-$ chemistry, although it is quite faint (Wang et al., 2011; Zhang et al., 2014, e.g.). With a rest frequency of of 91.98 GHz, there are 479 eBOSS objects included in this stack.

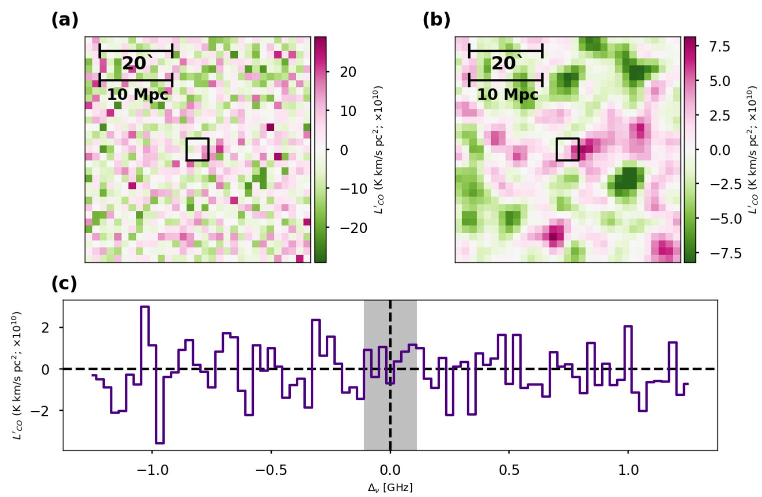


Figure 2.17: The 3D stack performed on the eBOSS catalog using the CS(2-1) emission wavelength. See Figure 2.6 for panel descriptions.

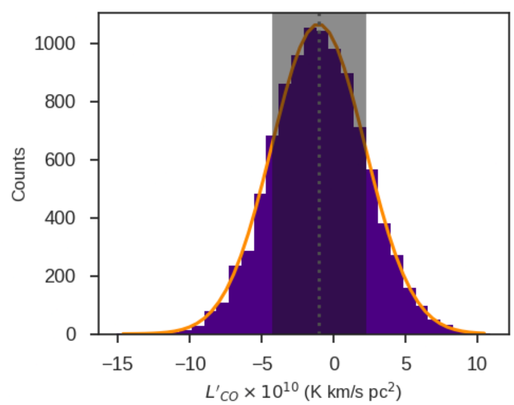


Figure 2.18: The bootstrapped uncertainty in $L'_{\rm CO}$ for the CS(2-1) stack (see Figure 2.7).

### $^{13}$CO(1-0)

$^{13}$CO, as a rarer and slightly heavier isotopologue of $^{12}$CO, has a similar rest frequency (110.20 GHz) and traces a similar phase of gas, although with a higher critical density. As an isotopologue, $^{13}$CO is of particular interest in determining the CO-$H_2$ conversion factor in galaxies (e.g., Israel, 2020). 311 eBOSS objects are included in this stack.

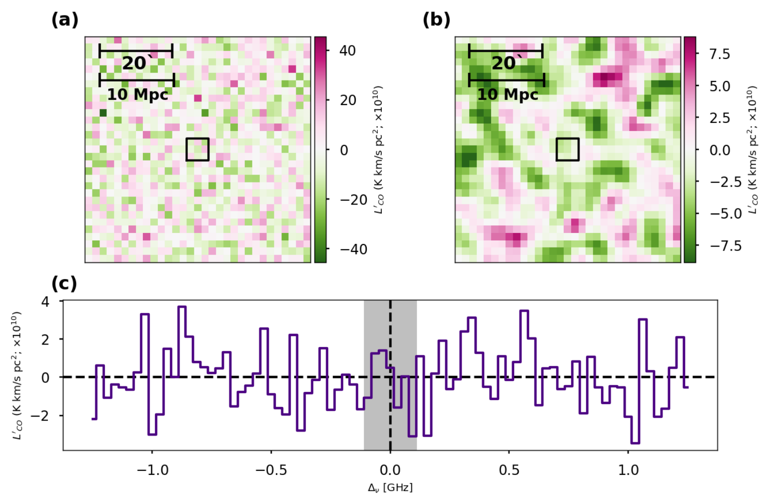


Figure 2.19: The 3D stack performed on the eBOSS catalog using the $^{13}$CO(1-0) emission wavelength. See Figure 2.6 for panel descriptions.

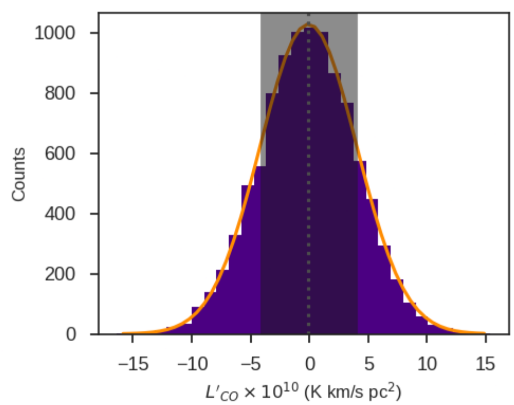


Figure 2.20: The bootstrapped uncertainty in $L'_{\mathrm{CO}}$ for the $^{13}$CO(1-0) stack (see Figure 2.7).

### CN⁻(1-0)

$CN^-$(1-0) traces a very similar phase of gas to HCN (*i.e.*, dense molecular cores; §2.12), but has a critical density which is lower by roughly a factor of 5, and likely additionally prefers gas layers affected by stellar UV radiation (Riechers et al., 2007). The $CN^-$ rest frequency of 113.49 GHz is very close to CO(1-0), and 269 eBOSS quasars fall into the COMAP volume at this frequency.

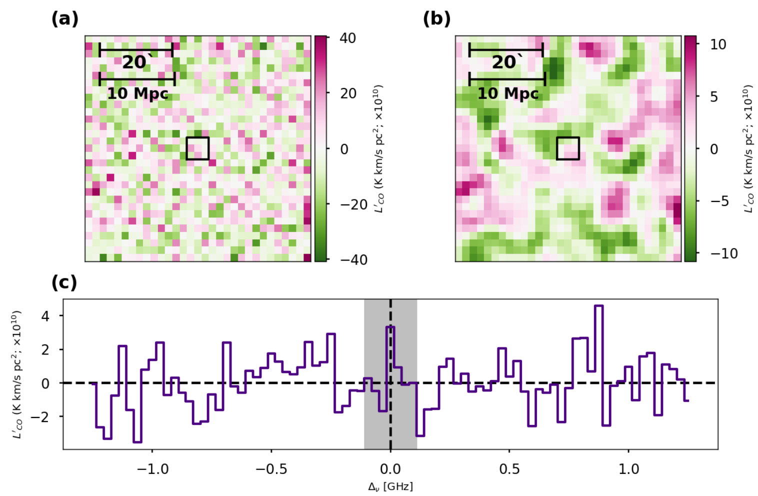


Figure 2.21: The 3D stack performed on the eBOSS catalog using the CN$^-$(1-0) emission wavelength. See Figure 2.6 for panel descriptions.

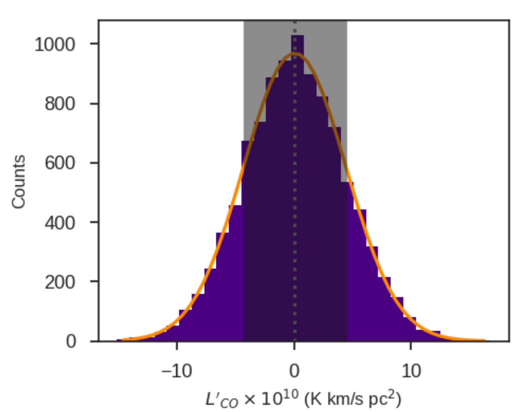


Figure 2.22: The bootstrapped uncertainty in $L'_{\rm CO}$ for the CN$^-$(1-0) stack (see Figure 2.7).

*Chapter 3*

# THREE-DIMENSIONAL STACKING AS A LINE INTENSITY MAPPING STATISTIC

Dunne, Delaney A. et al. (Oct. 2025). "Three-dimensional stacking as a line intensity mapping statistic." In: *Astronomy & Astrophysics* 702, A247, A247. DOI: 10.1051/0004-6361/202554780. arXiv: 2503.21743 [astro-ph.CO].

## 3.1 Chapter abstract

Line intensity mapping (LIM) is a growing technique that measures the integrated spectral line emission from unresolved galaxies over a three-dimensional region of the Universe. Although LIM experiments ultimately aim to provide powerful cosmological constraints via auto-correlation, many LIM experiments are also designed to take advantage of overlapping galaxy surveys, thus enabling joint analyses of two datasets. We introduce a flexible simulation pipeline that can generate mock galaxy surveys and mock LIM data simultaneously for the same population of simulated galaxies. Using this pipeline, we explore a simple joint analysis technique: three-dimensional co-addition (stacking) of LIM data on the positions of galaxies from a traditional galaxy catalog. We test how the output of this technique reacts to changes in experimental design of both the LIM experiment and the galaxy survey, its sensitivity to various astrophysical parameters, and its susceptibility to common systematic errors. We find that an ideal catalog for a stacking analysis targets as many high-mass dark matter halos as possible. We also find that the signal in a LIM stacking analysis originates almost entirely from the large-scale clustering of halos around the catalog objects rather than the catalog objects themselves. While stacking is a sensitive and conceptually simple way to achieve a LIM detection, thus providing a valuable way to validate a LIM auto-correlation detection, it will likely require a full cross-correlation to achieve further characterisation of the galaxy tracers involved, as the cosmological and astrophysical parameters we explore here have degenerate effects on the stack.

## 3.2 Introduction

Line intensity mapping (LIM) is an observational technique designed to measure the integrated spectral line emission from all galaxies in a population while generating a three-dimensional fluctuation map of a given spectral line with intentionally broadened spatial resolution. Specifically, this technique is sensitive to the integrated emission from even the faintest end of the luminosity function and is thus a possible solution to the completeness problem in galaxy surveys, particularly those at high redshift (for a review, see Kovetz et al., 2019). LIM is still an emerging field, and no experiment has yet made a confirmed auto-correlation detection of their target emission line (e.g., Staniszewski et al., 2014; Keating et al., 2015; Keating et al., 2020; Cleary et al., 2022; Fasano et al., 2022; CCAT-Prime Collaboration et al., 2023; Paul et al., 2023) due to the faintness of the target signal, the need for exquisite control of instrumental systematic errors, and (in some cases) the presence of interloper spectral line emission.

Because of these difficulties, analysis techniques that combine LIM data with external tracers of the matter density (such as other intensity mapping experiments or more traditional galaxy surveys) will be an extremely valuable way of validating potential auto-correlation detections as the field matures. These joint analyses up-weight regions of the LIM map that are most likely to contain a signal and can therefore mitigate the effects of instrumental and astrophysical systematic errors, as these should primarily be unique to one of the two tracers included in the analysis.

The most commonly studied technique for combining intensity mapping data with a galaxy survey is full three-dimensional cross-correlation (e.g., Pullen et al., 2013b; Chung et al., 2019; Keenan et al., 2022; Breysse and Alexandroff, 2019), as it allows for the study of the interaction between the two tracers as a function of spatial scale. However, this approach is sensitive to the existing angular structures in a survey strategy, and it suffers from complicated mode-mixing effects when the observational strategy of the LIM experiment is correlated with that of the galaxy catalog. These problems are tractable but complex to manage, particularly in the early stages of an experiment. Because of this, several LIM experiments supplement their cross-correlation analyses with an analysis that is simpler conceptually, such as a three-dimensional stack (or co-addition) of the LIM data on the positions of objects in the galaxy survey (e.g., Keenan et al., 2022; Dunne et al., 2024; Chen et al., 2025). Functionally, this is equivalent to the zero-lag region of the two-point correlation function, or a real-space version of a cross-correlation on the smallest spatial scale. Although this analysis does not probe the larger-scale modes that the cross-spectrum does, and thus loses some overall sensitivity compared to the cross-spectrum, it is easy to implement and is extremely useful for addressing systematic errors such as those listed above while still constraining the overall amplitude of the signal from a LIM tracer (Chen et al., 2025).

While the stack is conceptually simple, in practice it is generally unclear how astrophysical and experimental factors will affect its output. Previous work has shown that, for example, astrophysical line broadening has a dampening effect on cross-correlations that is worse in the higher-k bins, which are where the stack sensitivity lies (e.g., Chung et al., 2021). These effects need to be explored in detail with simulations. Therefore, in this work, we devise a simple yet robust simulation pipeline that can generate a realistic population of multiple astrophysical tracers as well as parametrize the interaction between the two. Using this pipeline, we aim to answer the following questions:

1. What experimental design (of both the LIM and catalog experiment) is optimal for a stacking analysis?
2. Where does the signal in a stacking analysis originate?
3. What can the stack tell us about the properties of the galaxies contributing to the signal?

We have structured this work as follows: In §3.3, we introduce the joint simulation pipeline, which is an extension to the existing `limlam_mocker`[1] simulation pipeline known as `joint_limlam_mocker`.[2] We briefly summarize the stacking methodology in §3.4 and attempt to parametrize the stack analytically. We then test the effects of various instrumental and astrophysical parameters on the outcome of the stacking analysis in §3.5. In §3.6, we discuss the implications of these results.

We assumed base ten logarithms unless otherwise stated, and where necessary, we took a ΛCDM cosmology with $\Omega_{\mathrm{m}} = 0.286$, $\Omega_{\Lambda} = 0.714$, $\Omega_{\mathrm{b}} = 0.047$, and $H_0 = 100h\mathrm{km\ s^{-1}\ Mpc^{-1}}$ with $h = 0.7$. This matches the choice of parameters used in the peak-patch simulations in §3.3 and is based on the nine-year WMAP results (Hinshaw et al., 2013). All cosmological distances are presented as comoving quantities.

[1] `https://github.com/georgestein/limlam_mocker`
[2] `https://github.com/delaneydunne/joint_limlam_mocker`

## 3.3 Simulated multi-tracer observations

Because LIM targets cosmological scales but is also sensitive to variations in spectral line emission at the level of an individual dark matter (DM) halo, observational LIM data are difficult to simulate. For joint analyses such as the stack being explored in this work, we are additionally including a second spectral line tracer, and thus need to realistically simulate the distribution across halo masses of both tracers, as well as the interaction between the tracers at the individual halo level.

For this work we follow previous efforts to simulate LIM data, adding the capability to generate a galaxy catalog using a different spectral line tracer. The resulting pipeline, known as `joint_limlam_mocker`, is publicly available[2]. The overall flow of the pipeline is outlined in Figure 3.1, and is as follows: we begin with a DM halo catalog generated from peak-patch $N$-body simulations (§3.3). We then assign to each halo simulated luminosities corresponding to each of two different emission lines calculated based on the halo's DM mass—one line associated with the LIM map (§3.3) and one associated with the galaxy catalog (§3.3). We add scatter to each luminosity value, correlated between the two tracers on a per-halo basis (§3.3). We then generate a synthetic line intensity map (§3.3) and a synthetic galaxy catalog (§3.3).

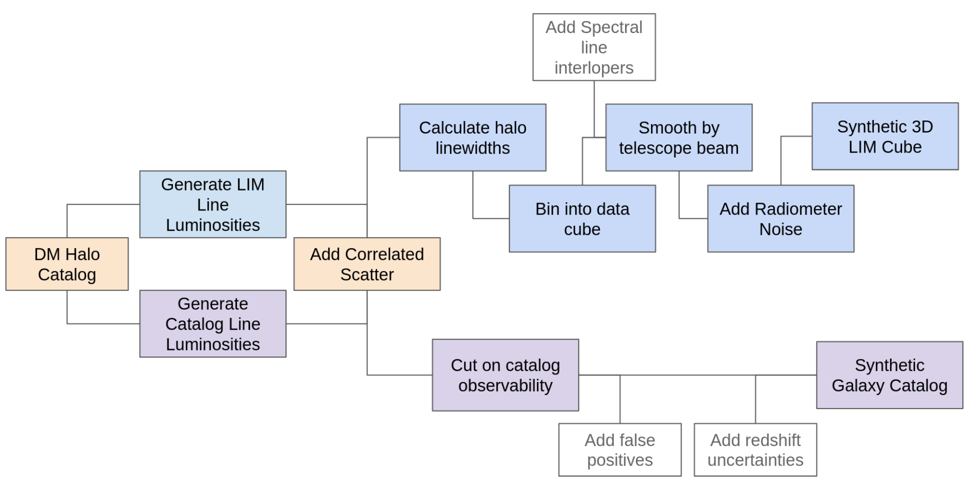


Figure 3.1: Flowchart depicting the multi-tracer simulation pipeline. Orange boxes indicate steps that affect both the galaxy catalog and LIM data, blue boxes are actions on the simulated LIM data, and purple boxes are actions on the simulated galaxy catalog. Steps with white boxes are optional.

The outputs of the simulation pipeline are multi-survey and multi-tracer synthetic observations of the same population of simulated DM halos. The synthetic observations include a rudimentary treatment of the correlation between the tracers due to galaxy-scale astrophysics, as well as other complicating instrumental or astrophysical parameters, such as spectral line broadening, interloper emission in the LIM map, and redshift uncertainties in the galaxy catalog. Although the treatment of these astrophysical and instrumental parameters is basic, we can vary the prescription used for any of the steps in this process, in order to explore the importance of each step to the stacking analysis.

As an illustrative example, we based the instrumental and spectral line modeling parameters for the LIM experiment on the CO Mapping Array Project (COMAP) Pathfinder experiment (Cleary et al., 2022) and the parameters for the galaxy catalog on the Hobby-Eberly Telescope Dark Energy eXperiment (HETDEX, Gebhardt et al., 2021), meaning that the two spectral lines we treat are the (1–0) transition of carbon monoxide (CO), and Hydrogen Lyman-$\alpha$ (Ly$\alpha$), respectively. We thus refer to the simulated luminosity values generated for each halo as the 'CO luminosity' ($L_{\rm CO}$), which is used to generate the mock LIM data cube, and the 'Ly$\alpha$ Luminosity' ($L_{\rm Ly\alpha}$), which is used to generate the resolved mock galaxy catalog. We note, however, that this framework should be generic to any joint LIM-galaxy analysis. We explain the choices of the various parameters chosen for the default simulation setup below and summarize them in Table 3.2.

**Peak-patch $N$-body simulations**

The basis of the synthetic observations is a set of $N$-body simulations generated using the peak-patch method (Bond and Myers, 1996; Stein et al., 2019). These simulations provide a realistic catalog of the three-dimensional (3D) positions and masses of DM halos, onto which luminosities can be painted. The minimum halo mass we include is $2.9\times10^{10}\ M_\odot$, and the maximum mass is $9.1\times10^{13}\ M_\odot$, following Chung et al. (2022). A cutout of the resulting distribution of DM halos is shown in Figure 3.2.

**LIM tracer luminosity modeling: CO(1–0)**

The prescriptions we used to determine the CO luminosity of the simulated halos are the same as those used for our COMAP auto-correlation analyses, to enable direct comparison. Each of these models calculates the CO luminosity of a DM halo directly from its mass. They are:

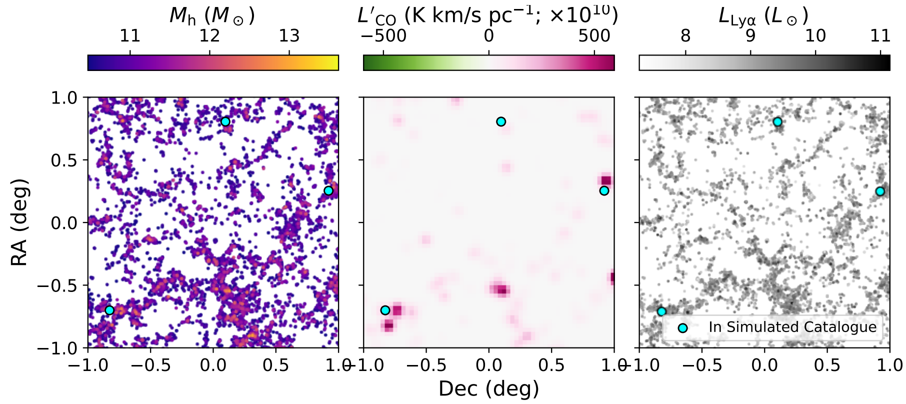


Figure 3.2: Zoomed-in frequency slices of the simulated population of DM halos and resulting mock observations. *Left:* DM halos, colored by their halo mass. *Centre:* Mock LIM fluctuation map of the CO emission (with no noise added). *Right:* Ly$\alpha$ luminosity of each DM halo. The halos that would actually be detected by the mock survey and thus included in the galaxy catalog are shown as larger cyan circles in all three panels.

- The data-driven 'COMAP fiducial' model Chung et al., 2022, hereafter C22, which we treated as the default for these simulations. It follows the functional form
$$\frac{L'_{\mathrm{CO}}(M_{\mathrm{h}})}{\mathrm{K\ km\ s^{-1}\ pc^2}} = \frac{C}{(M_{\mathrm{h}}/M)^A + (M_{\mathrm{h}}/M)^B}, \tag{3.1}$$
where $A$, $B$, $C$, and $M$ are all free parameters. Chung et al. (2022) determine fiducial values for these free parameters based on the UniverseMachine (UM) framework of Behroozi et al. (2019) and conditioned on data from COLDz (Riechers et al., 2019) and COPSS (Keating et al., 2016), which are listed in Table 3.2.

- The CO model from Padmanabhan, 2018, which was abundance-matched onto CO observations at $z = 3$ from Keating et al., 2016. The model takes the form of a double power law with redshift-dependent parameters:
$$L_{\mathrm{CO}}(M_{\mathrm{h}}, z) = \frac{2N(z)M_{\mathrm{h}}}{\left(\frac{M_{\mathrm{h}}}{M_1(z)}\right)^{-b(z)} + \left(\frac{M_{\mathrm{h}}}{M_1(z)}\right)^{y(z)}}. \tag{3.2}$$

Here, the parameters $M_1(z)$, $N(z)$, $b(z)$, and $y(z)$ all contain a constant term for $z \sim 0$ and a term that evolves with redshift:

$$\begin{aligned} \log M_1(z) &= \log M_{10} + M_{11} z/(z+1) \\ N(z) &= N_{10} + N_{11} z/(z+1) \\ b(z) &= b_{10} + b_{11} z/(z+1) \\ y(z) &= y_{10} + y_{11} z/z(z+1). \end{aligned} \quad (3.3)$$

We took the best-fitting parameters from Table 1 of Padmanabhan, 2018. This model includes an additional factor quantifying the duty cycle of DM halos (the percentage of which are actually emitting in CO), $f_{\rm duty}$. Here, we took $f_{\rm duty} = 0.1$. We refer to this model hereafter as P18.

- The CO model from Li et al., 2016, which determines the CO luminosity of a given DM halo via its infrared luminosity calculated through a modeled star formation rate (SFR),

$$\log L'_{\rm CO} = \frac{1}{\alpha} \left[\log L_{\rm IR} - \beta\right] \quad (3.4)$$

where

$$L_{\rm IR} = \frac{\rm SFR}{\delta_{\rm MF}} \times 10^{10}. \quad (3.5)$$

The coefficient $\delta_{\rm MF}$ depends on the initial mass function, and we calculated SFR values following Behroozi et al., 2013a; Behroozi et al., 2013b. As in Li et al., 2016, we took $\delta_{\rm MF} = 1$. For CO(1–0),

$$\frac{L_{\rm CO}}{L_\odot} = 4.9 \times 10^{-5} \frac{L'_{\rm CO}}{\rm K\ km\ s^{-1}\ pc^2}. \quad (3.6)$$

We refer to this model as L16.

For each of these different models, we show the CO luminosity $L(M_{\rm h})$ as a function of halo mass $M_{\rm h}$, in Figure 3.3. We note that there is more than an order of magnitude variation between the models at most halo masses – this is a reflection of the large uncertainty that exists in this model space.

**Catalogue tracer luminosity modeling: Ly$\alpha$**

We used luminosity functions to determine the Ly$\alpha$ luminosity of the DM halos, as these tend to be better studied than models directly connecting luminosity of galaxies to their host DM halo masses (*i.e.*, $L(M_{\rm h})$ models) for optical line tracers

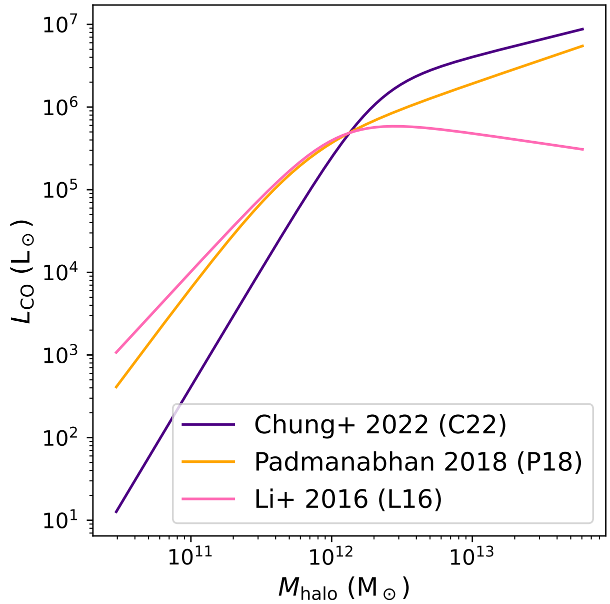


Figure 3.3: CO luminosity as a function of halo DM mass for the different CO models tested. The models are listed in §3.3.

such as Ly$\alpha$. We abundance-matched these luminosity functions onto the simulated DM mass function in order to assign to each DM halo a catalog luminosity. For the purposes of this work, we assumed Schechter luminosity functions (Schechter, 1976),

$$\frac{dn}{dL} = \frac{\phi^*}{L^*}\left(\frac{L}{L^*}\right)^{\alpha} e^{-L/L^*}, \tag{3.7}$$

where $L^*$ is a characteristic luminosity, $\phi^*$ is the normalisation density, and $\alpha$ is the faint-end power law slope. We varied only the input parameters here, although any functional form could be used for the catalog luminosity function in the `joint_limlam_mocker` framework.

We tested three different models for the Ly$\alpha$ luminosity function:

1. The Lyman-$\alpha$ Emitter (LAE) luminosity function at $z = 3.1$ from Ouchi et al., 2020 based on observations from Ouchi et al., 2008 and Konno et al., 2016. These should be a fairly accurate estimate of the HETDEX luminosity function, as HETDEX probes LAEs in the same redshift range. This is the model we treated as the default.

2. The quasar Ly$\alpha$ luminosity function calculated at $z \sim 3$ using the Physics of the Accelerating Universe Survey (PAUS; Torralba-Torregrosa et al., 2024).

Table 3.1: Schechter parameters for the catalog luminosity function models. 'Default' and 'faint' values are taken from Ouchi et al., 2020, and 'bright' values from Torralba-Torregrosa et al., 2024.

| Model | $L^*$ (erg s$^{-1}$) | $\phi^*$ (Mpc$^{-3}$) | $\alpha$ |
|---|---|---|---|
| Default | $8.49 \times 10^{42}$ | $3.9 \times 10^{-4}$ | -1.8 |
| Bright | $9.77 \times 10^{44}$ | $1.9 \times 10^{-7}$ | -1.34 |
| Faint | $7.10 \times 10^{42}$ | $1.12 \times 10^{-4}$ | -1.8 |

This model is much more bright-end heavy than typical LAE luminosity functions at this redshift range. Several large quasar catalogs (such as eBOSS, Ahumada et al. 2020 or DESI, Chaussidon et al. 2023) may make interesting potential targets for stacking, so this is an interesting regime to probe. These catalogs use different spectral lines but trace similar populations of DM halos. We refer to this as our 'bright' model throughout.

3. A lower-redshift ($z = 0.3$) Ly$\alpha$ luminosity function from Ouchi et al., 2008. This is a luminosity function for a galaxy population dominated by many fainter galaxies, and we used it as an example case for these types of populations. We refer to it as the 'faint' model.

The Schechter parameters for each model are listed in Table 3.1, and the luminosity functions themselves are compared in Figure 3.4. For each choice of Ly$\alpha$ function, we used a scatter $\sigma_{\mathrm{Ly}\alpha} = 0.41$ dex, determined from the uncertainty in the Schechter parameters of the default model.

**Correlated scatter in luminosities**

While the above formalism simulates how the two tracer luminosities vary with halo properties such as DM mass (thus roughly recreating the astrophysical bias for each tracer), it does not account for relationships between the two tracers in individual galaxies, which will set the level of the (cross-)shot noise. To account for this effect, we calculated the final luminosity values for each halo by introducing a correlated log-normal scatter $d$ between the luminosity determined from the $M_{\mathrm{h}}$ relations for the LIM tracer (here $L_{\mathrm{CO}}$) and the tracer of the resolved galaxy catalog (here $L_{\mathrm{Ly}\alpha}$). This was done by scaling the luminosity values output from the previous steps by

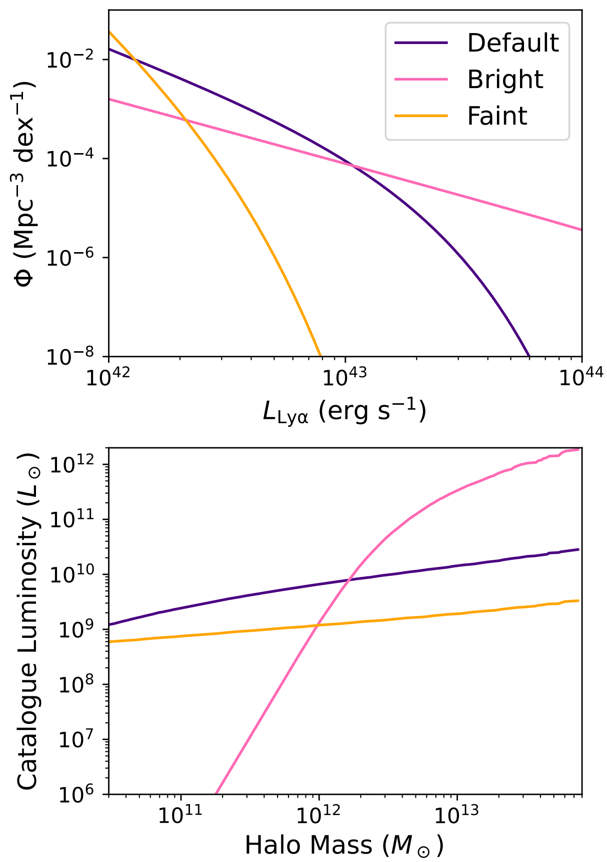


Figure 3.4: The Ly$\alpha$ models used in this work. *Top:* Various luminosity functions used to model the $L_{\text{Ly}\alpha}$ assigned to each DM halo to generate the galaxy catalog being stacked. Each model is a Schechter function, with parameters described in §3.3. *Bottom:* Resulting luminosity as a function of DM halo mass.

an exponential factor,

$$\begin{pmatrix} L_{\text{CO}} \\ L_{\text{Ly}\alpha} \end{pmatrix} = \begin{pmatrix} e^{\left[d_{\text{CO}}-\sigma_{\text{CO}}^2/2\right]} L_{\text{CO}}(M_{\text{h}}) \\ e^{\left[d_{\text{Ly}\alpha}-\sigma_{\text{Ly}\alpha}^2/2\right]} L_{\text{Ly}\alpha}(M_{\text{h}}) \end{pmatrix}, \tag{3.8}$$

where $\sigma_{\text{CO}}$ and $\sigma_{\text{Ly}\alpha}$ are the base-$e$ standard deviations calculated from base ten values in decimal exponents (dex) as

$$\sigma_{\text{CO}} = \sigma_{\text{CO, dex}} \ln 10, \ \sigma_{\text{Ly}\alpha} = \sigma_{\text{Ly}\alpha\text{, dex}} \ln 10 \tag{3.9}$$

($\sigma_{\text{CO, dex}}$ and $\sigma_{\text{Ly}\alpha\text{, dex}}$ are the amounts of logarithmic scatter in the CO and Ly$\alpha$ values, respectively, and are choices of the models used). The scaling factors $d_{\text{CO}}$ and $d_{\text{Ly}\alpha}$ were pulled from a two-dimensional normal distribution with zero mean and covariance given by

$$\begin{pmatrix} d_{\rm CO} \\ d_{{\rm Ly}\alpha} \end{pmatrix} \leftarrow \mathcal{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\rm CO}^2 & \rho\sigma_{\rm CO}\sigma_{{\rm Ly}\alpha} \\ \rho\sigma_{\rm CO}\sigma_{{\rm Ly}\alpha} & \sigma_{{\rm Ly}\alpha}^2 \end{pmatrix} \right\}. \tag{3.10}$$

The parameter $\rho$ (where $-1 < \rho < 1$) is the correlation coefficient, and it scales how the scatter in the two tracers relates for a given halo.

For the example case we are exploring (a stack on a CO LIM map using a Ly$\alpha$ catalog) an empirically motivated choice of $\rho$ is difficult due to the lack of available simultaneous observations of CO and Ly$\alpha$ in galaxies (especially at higher redshifts). For the purposes of this work, we looked at the relationship between CO and Ly$\alpha$ via galaxy metallicity. Davis et al., 2023a found the LAEs cataloged by HETDEX to be primarily composed of galaxies with young metal-poor stellar populations ($Z \sim 0.2 - 0.3\ Z_\odot$). This agrees with other surveys of (especially lower-luminosity) LAEs (e.g., Sobral et al., 2018; Matthee et al., 2021). CO emission, however, is typically suppressed in actively star-forming low-metallicity galaxies (Bolatto et al., 2013). We can thus expect at least a slight anti-correlation between emission in the two lines on an individual galaxy level, no matter how their overall populations correlate with DM halo mass. With this as justification, we defaulted to a value of $\rho = -0.5$ for the purposes of this paper. We note that, particularly for our default CO and Ly$\alpha$ configuration, this choice does not significantly affect our results – we explore the effects of varying $\rho$ in §3.5.

**Simulated intensity map**

We then generated a synthetic line intensity map, basing the simulated experimental parameters for this map on COMAP. We first simulated astrophysical sources of spectral line broadening by calculating a velocity to be associated with each halo. We used one of three methods:

1. The relationship between peak halo mass (essentially identical to the halo mass at these redshifts) and $v_{M_{\rm peak}}$ provided by Behroozi et al., 2019 in their UM framework,

$$v_{M_{\rm peak}}(M_{\rm h}) = (200\ {\rm km\ s^{-1}}) \left[ \frac{M_{\rm h}}{M_{200\ {\rm km\ s^{-1}}}(a)} \right]^{0.3}, \tag{3.11}$$

where $a$ is the scale factor at redshift $z$ and

$$M_{200\ {\rm km\ s^{-1}}} = \frac{1.64 \times 10^{12} M_\odot}{(a/0.378)^{-0.142} + (a/0.378)^{-1.79}}. \tag{3.12}$$

We added an additional log-normal scatter of 0.1 dex, and we scaled the velocity of each halo with a randomly generated galaxy inclination, $i$:

$$v_{\rm max} = v_{M_{peak}} \sin i/0.866. \tag{3.13}$$

This is the prescription that aligns most closely with observations (see Appendix 3.9), and thus the one we treated as the default.

2. The virial velocity ($v_{\rm vir}$) associated with the DM mass of each halo:

$$\begin{aligned} v_{\rm vir} &= \sqrt{\frac{GM_{\rm h}}{r_{\rm vir}}} \\ &= \left(\frac{\Delta_c}{2}\right)^{1/6} [GM_{\rm h}H(z)]^{1/3} \\ &\approx 35 \text{ km s}^{-1} \left(\frac{\Delta_c}{2}\right)^{1/6} \\ &\times \left(\frac{M_{\rm h}}{10^{10}\ M_\odot} \frac{H(z)}{100 \text{ km s}^{-1} \text{ Mpc}^{-1}}\right)^{1/3}. \end{aligned} \tag{3.14}$$

This was again scaled by a randomly generated galaxy inclination $\sin i$.

3. The virial velocity, calculated as above, but with a flat cut-off implemented at $v_{\rm max} = 1000\text{km s}^{-1}$. This primarily affects the largest DM halos ($M_{\rm h} \gtrsim 10^{12}\ M_\odot$), which can have virial velocities much larger than those seen in CO linewidths in observations (see Appendix 3.9, or Carilli and Walter 2013). Because the halos included in the stack are filtered through their catalog luminosity, which correlates positively with mass for all catalog models, these are the halos most likely to be included in the stack, and thus this cut-off may make a significant difference in the stacked linewidth.

We compare these different strategies for calculating halo linewidths, as well as our choice of $v_{M_{\rm peak}}$ as a default prescription, in Appendix 3.9. The effects of each on the stack output are shown in §3.5.

Following Chung et al. (2021), we binned halos into $N_{\rm bin}$ linearly spaced bins by their $v_{\rm max}$. We then generated a CO intensity cube separately for each velocity bin, summing the halo luminosities into voxels (three-dimensional pixels) of sizes $\delta_x \times \delta_x \times \delta_\nu = 2$ arcmin $\times$ 2 arcmin $\times$ 31.25 MHz ($\sim 3.7 \times 3.7 \times 4.1$ Mpc, or $k_\perp \sim 1.69\,\text{Mpc}^{-1}$ and $k_\parallel \sim 1.53\,\text{Mpc}^{-1}$). The intensity cubes cover $\Delta_x \times \Delta_x = 4^\circ \times 4^\circ$ total in their angular coordinates, and $\Delta_\nu = 8$ GHz spectrally. We smoothed the intensity cube corresponding to each velocity bin along the line of sight by a Gaussian kernel with width given by the median $v_{\rm max}$ in the bin, and summed together each smoothed cube. We approximated beam smoothing by convolving each spectral channel with a 2D Gaussian kernel with a full width at half maximum (FWHM) of $\theta_{\rm beam}$ defaulting to $\theta_{\rm beam} = 4.5$ arcmin ($\sim 8.3$ Mpc, or $k_\perp \sim 0.75$ Mpc$^{-1}$). In order to differentiate the LIM signal from the CO luminosity of a given halo (as the LIM signal comes from many halos summed together), we refer to the luminosity in the resulting map as $L_{\rm LIM}$. We show a frequency slice of a mock LIM cube at this stage (*i.e.*, noiseless) in Figure 3.2.

We also inserted varying amounts of simulated white noise into the map. We based the noise model on radiometer noise for the COMAP experiment, which is a focal-plane array with 19 feeds. Thus, assuming purely Gaussian noise (and even coverage over the simulated field), the COMAP noise response in each voxel is

$$\sigma_{\rm vox} = \frac{T_{\rm sys}}{\sqrt{2\delta_\nu N_{\rm f}\tau/(\Delta_x/\delta_x)^2}}, \tag{3.15}$$

where $T_{\rm sys}$ is the system temperature of the instrument, $\delta_\nu$ is the frequency width of each spectral channel, $N_{\rm f}$ is the number of feeds, $\tau$ is the total integration time of the experiment, and $(\Delta_x/\delta_x)^2$ is the number of spaxels (spatial pixels) that coverage of the field is split across, assuming a scanning strategy with uniform coverage across the map. The values we used for each parameter are listed in Table 3.2, and were based on the COMAP Pathfinder telescope (e.g., Cleary et al., 2022). However, to ensure we have high-significance stacks to analyse, we used a predicted per-field integration time $\tau$ that corresponds to a proposed future evolution of COMAP in which the currently operating Pathfinder instrument and two duplicates each observe for a total of ten years. This results in 29 000 hours spent integrating on each of the three COMAP fields (c.f. Breysse et al., 2022). We added a random value pulled from a distribution with this standard deviation and a zero mean to each voxel.

Additionally, we simulated line fluctuations from foreground or background populations of galaxies, which may be redshifted to the same observed frequency as the LIM data. For COMAP, this interloping emission will likely be CO(2-1) from galaxies at $6 < z < 8$. For [CII]-based surveys such as CONCERTO, TIME, or FYST, interloper emission will come from other $\sim 1$ mm emission lines, including [CI] and the CO rotational ladder (e.g., Béthermin et al., 2022). To mimic this emission, we added simulated interloping spectral line emission to the map. For efficiency, we generated the interloper map by applying one of several linear transformations to the existing simulated map, rather than simulating the interloper emission from a simulated population of galaxies at the interloper redshift. This is an inaccurate representation of the large-scale structure of the interloper emission, but the stack is not very sensitive to large-scale structure, so this should not affect our results. The transformation was done using one of 11 rotations or reflections:

1. One of three rotations (by 90°, 180°, or 270°) in the spatial axes.
2. One of two reflections (in either the RA or Dec axis).
3. A reflection across the frequency axis.
4. A reflection across the frequency axis, combined with any of the spatial rotations or reflections.

We then scaled the brightness of this transformed map to match the expected luminosity of foreground lines compared to CO. A factor of ten roughly corresponds to the difference between CO(1–0) at cosmic noon and the next most significant source of emission in our frequency range: CO(2-1) at $6 < z < 8$, which is expected to be roughly an order of magnitude fainter (Chung et al., 2024a). We then added the interloper map to the map of simulated CO emission.

Finally, we subtracted from the map a mean across the spatial dimensions in each spectral channel and across the spectral axis in each spaxel. This emulates the high-pass spatial and spectral filters in the actual COMAP data pipeline, which reject continuum emission almost entirely (Lunde et al., 2024). For this reason, we ignored continuum emission when inserting an interloper signal.

**Simulated galaxy catalog**

We generated a galaxy catalog from the set of simulated halos simply by logging the three-dimensional positions (and, optionally, luminosities) of some subset of the DM halos. This subset was generated by first cutting on luminosity, including only halos above a certain tracer luminosity $L_{\mathrm{Ly}\alpha,\mathrm{cut}}$, in order to simulate observational completeness. We note that this should more properly be a cut in flux, as flux (rather than luminosity) is the quantity observed by the telescope. The conversion between flux and luminosity is redshift-dependent, so the completeness limit should also be redshift dependent. However, the effect of this change on the returned number counts of cataloged galaxies even at the edges of the COMAP redshift range is small (roughly 10%). Additionally, most spectrographs have sensitivities that vary with wavelength, so the limiting catalog luminosity will in practice be a function of redshift due to instrumental effects. Here, however, we treated it as constant.

We then selected $N_{\text{obj}}$ halos with $L_{\text{Ly}\alpha} > L_{\text{Ly}\alpha,\text{cut}}$ randomly, weighting linearly by $L_{\text{Ly}\alpha}$ as brighter objects are more likely to be detected. $N_{\text{obj}}$ was set by the observational parameters of the galaxy survey being approximated. In practice, this selection is subject to the observing pattern of the galaxy survey used, but the stack is not sensitive to large scales, and thus large-scale spatial correlations should not be important here. Figure 3.2 shows an example simulation of the Ly$\alpha$ population as well as the halos that are selected by the detectability cuts to be included in the mock catalog.

To explore their effects, we also applied some other modifications to the galaxy catalog:

*Tracer velocity uncertainties*: To simulate astrophysical velocity offsets that can occur between different spectral lines emitted by a galaxy, due to, for example, bulk inflows and outflows as well as redshift uncertainties due to varying spectral resolution of galaxy surveys, we optionally applied a scatter $\sigma_v$ to the cataloged redshift. We note that bulk velocity offsets may also be present between the two tracers, but assume that these will be well characterized and thus easily accounted for.

*False positives:* Especially at low signal-to-noise (S/N) ratios, galaxy surveys can occasionally misidentify either noise peaks or other emission lines originating from foreground or background galaxies as the spectral line of interest. When this happens, false positive detections occur, creating catalog entries that do not correspond to actual galaxies in the target redshift range. We simulated this effect by generating catalog entries with random spatial positions and redshifts and inserting these spurious entries into the catalog. We quantified how many false positives are present in the simulated catalog using a fraction, $f_{\text{FP}}$, of total entries.

## 3.4 Three-dimensional stacking

### Summary of stacking methodology

We primarily followed the methodology for three-dimensional stacking established in Dunne et al., 2024, and we refer the reader to this work for the full details of the stacking pipeline. In essence, the stack is a simple co-addition – three-dimensional cutouts are taken from the LIM data cube at the angular and line-of-sight positions of objects included in a galaxy survey, and the cutouts are averaged together using

Table 3.2: Default values used in the generation of simulated LIM maps. Values are based on the COMAP Pathfinder experiment and the HETDEX catalog of LAEs. Where applicable, distances are comoving distances calculated at the average LAE redshift of 2.8.

| | |
|---|---|
| *LIM instrument parameters* | |
| $T_{\rm sys}$ | 44 K |
| $N_{\rm f}$ | 19 |
| $\Delta_\nu$ | 8 GHz (1028 Mpc) |
| $\Delta_x$ | 4 deg (218 Mpc) |
| $\delta_\nu$ | 31.25 MHz (3 Mpc) |
| $\delta_x$ | 2 arcmin (3.6 Mpc) |
| $\theta_{\rm beam}$ | 4.5 arcmin (8.2 Mpc) |
| $\tau$ | 29 000 h |
| *LIM modeling parameters* | |
| CO $L(M_{\rm h})$ | Chung et al. (2022) |
| $A$ | −2.85 |
| $B$ | −0.42 |
| $\log C$ | 10.63 |
| $\log \frac{M}{M_\odot}$ | 12.3 |
| $\sigma_{\rm CO}$ | 0.42 dex |
| Line broadening method | $v_{M_{\rm peak}}$ (Behroozi et al. 2019) |
| *Catalogue modeling parameters* | |
| Ly$\alpha$ luminosity function | Ouchi et al. (2020) |
| $L^*_{{\rm Ly}\alpha}$ | $8.49 \times 10^{42}$ erg s$^{-1}$ |
| $\phi^*_{{\rm Ly}\alpha}$ | $3.90 \times 10^{-4}$ Mpc$^{-3}$ |
| $\alpha$ | −1.8 |
| $\sigma_{{\rm Ly}\alpha}$ | 0.41 dex |
| $N_{\rm obj}$ | 1000 |
| $L_{{\rm Ly}\alpha,{\rm cut}}$ | $8.26 \times 10^9 L_\odot$ |
| $f_{\rm FP}$ | 0% |
| *Joint parameters* | |
| $\rho$ | −0.5 |

inverse-variance weighting. The resulting stacked ‘cubelet’ contains the average distribution of CO flux around the halos included in the Ly$\alpha$-based galaxy catalog. Finally, the cubelet is summed over a square central aperture of $N_{\rm spax}^2$ spaxels (spatial pixels) in the spatial axes to create a one-dimensional spectrum. This spectrum is integrated over $N_{\rm chan}$ spectral channels to yield an average luminosity measurement of the regions that correspond to a catalog object and are included in the stack. A stack performed on simulation realizations generated using the default parameters (listed in Table 3.2) is shown in Figure 3.5.

Unlike the $N_{\rm spax} = 3$ value used in Dunne et al., 2024, in this work we took $N_{\rm spax} = 7$ ($\sim 25.9$ Mpc, or $k_\perp \sim 0.25$ Mpc$^{-1}$) as the default. In Dunne et al., 2024, we defaulted to the $3 \times 3$ aperture, as this is the nearest odd number to 1.5× the FWHM of the beam using the COMAP beam width and spatial pixelisation. With a $3 \times 3$ aperture, a catalog object can be located anywhere within the central spaxel and still have its entire beam FWHM fall into the $N_{\rm spax} \times N_{\rm spax}$ aperture. However, as we discuss in detail below, we find that larger-scale clustering contributes extensively to the stacked signal, and a $7 \times 7$ aperture actually maximizes the stack S/N. We also differed in our choice of $N_{\rm chan}$ – here, we chose $N_{\rm chan} = 3$ ($\sim 12.3$ Mpc, or $k_\perp \sim 0.51$ Mpc$^{-1}$). We chose this after inspecting the default stacked spectrum (e.g., Figure 3.5), which on average has a stack FWHM of 835 km/s (83.6 MHz). Three spectral channels (936 km/s, or 93.75 MHz) is the nearest odd number of channels to this width, and encompasses nearly all the flux from the stacked spectral line. This is narrower than the expected value in Dunne et al., 2024, because that stack was performed on a catalog of quasars with a known large velocity uncertainty. In the case of a galaxy catalog with no significant uncertainty in redshift, the linewidth is much narrower (we investigate these effects further in §3.5).

**Formalism for stack sensitivity**

As the goal of this work is to explore what experimental design and analysis parameters are optimal for the purposes of the stack, we attempt here to establish an analytical formalism for the stack sensitivity. We present this in terms of the stack detection S/N ratio, as we explore factors that impact both the signal and the uncertainty in the final stack. We write this ratio as

$$\frac{S}{N} = \frac{\langle L_{\rm LIM} \rangle}{\sigma_{\rm stack}}, \tag{3.16}$$

where $\langle L_{\rm LIM} \rangle$ is the average luminosity in the LIM tracer across the stacked regions (the target signal), and $\sigma_{\rm stack}$ is the uncertainty in that average.

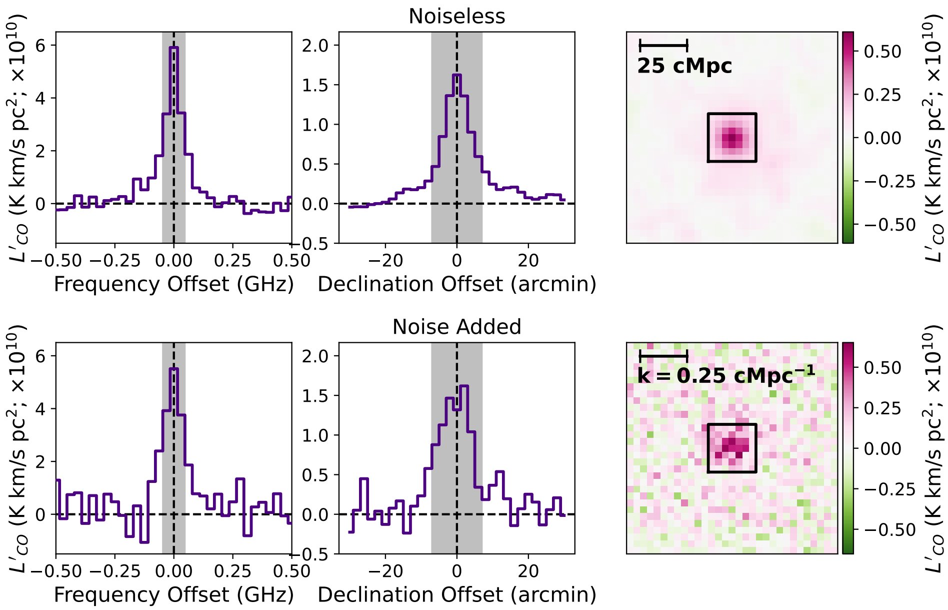


Figure 3.5: One- and two-dimensional projections of a simulated stack cubelet generated using the default simulation parameters. The top row shows a simulation realization with no added noise (variations are due to individual simulated halos), and the bottom row shows the same simulation realization with added radiometer noise (following §3.3). *Left:*Frequency spectrum of the central stack aperture. The $N_{\text{chan}}$ frequency channels that are integrated over to generate the final stack luminosity are highlighted in grey. *Center:* Spatial profile of the stack determined by summing over the three central frequency channels (those highlighted in the spectrum) into an image (shown in the right panel) and then collapsing the RA axis by summing over the three central spaxels. The spatial profile plots this quantity as a function of the angular offset in the declination direction from the stack center. As in the spectrum, the width of the spatial aperture over which the emission is integrated to generate the final stack luminosity is highlighted in grey. *Right:* Two-dimensional image with the $N_{\text{spax}} \times N_{\text{spax}}$ spatial aperture boxed in black.

The nature of the stack means that the signal $\langle L_{\rm LIM}\rangle$ is a biased tracer of the overall CO signal at the target redshifts. Because the average in the stack is only over cutouts of the LIM cube centered around DM halos included in the galaxy catalog, a region of space is only included in the average if it contains a halo emitting a detectable amount of (in this case) Ly$\alpha$ emission. We represent this effect with a 'detectability' factor, which we represent as $S\left(L_{\rm Ly\alpha}\left(M_{\rm h},\rho\right)\right)$. This is a delta function (either a region of space is included or it is not) set by factors including the minimum luminosity detectable by the galaxy survey ($L_{\rm Ly\alpha,cut}$), the mass of the halo, the size of the region of space cut out for each catalog object, and the correlation (or anti-correlation) between the scatter in the catalog tracer luminosities and the LIM tracer luminosities of each halo, which we parametrize with $\rho$ (§3.3).

Additionally, because LIM experiments are of relatively low spatial/spectral resolution (by design) compared to the galaxy survey from which the catalog is derived, the region of space surrounding the cataloged object included in the stack is quite large. This means it is almost certain there will be neighboring galaxies included in the $N_{\rm spax}\times N_{\rm spax}\times N_{\rm chan}$ central aperture, so $\langle L_{\rm LIM}\rangle$ is not a direct average over the cataloged galaxies themselves. Instead, there is another term contributing to $\langle L_{\rm LIM}\rangle$, which is the sum over the luminosities $L'_{\rm LIM}(M_{\rm h})$ of the $N(M_{{\rm h},i})$ neighboring galaxies for each $i^{\rm th}$ cataloged galaxy. Because we are only probing the regions surrounding objects in a galaxy catalog, the neighboring galaxies are also not truly representative of the global population of DM halos (they are only included in the stack if they are adjacent to a cataloged halo, and thus are biased via proximity to the catalog-selected objects).

Overall, this results in a model for the stack luminosity that can be written as

$$\begin{aligned}\langle L_{\rm LIM}\rangle = \frac{1}{N_h}\sum_i S(L_{\rm Ly\alpha}(M_{h,i},\rho))\times \\ \left[L'_{\rm LIM}(M_{h,i}) + \sum_{n=0}^{N(M_{h,i})} L'_{\rm LIM}(M_{h,n})\right].\end{aligned} \tag{3.17}$$

In other words, the stack luminosity is sensitive to both the LIM tracer luminosity of the objects included in the galaxy catalog, and the degree to which those objects trace larger-scale overdensities of emitters in the LIM tracers. Both a catalog of bright emitters or a highly biased catalog will thus in theory return higher stack luminosities. Each of these parameters is tunable using the simulation setup described in §3.3, and we explore the extent to which these different factors affect the overall significance of the stack in the following sections.

The uncertainty $\sigma_{\text{stack}}$ is somewhat more straightforward. The stack is an inverse-variance weighted average, so assuming constant noise across the LIM map (which we do assume for the simulations presented in this work, although this is not true in general), the sensitivity in the stack relates to the sensitivity in an individual cutout's stack aperture as

$$\sigma_{\text{stack}} = \frac{\sigma_{\text{aper}}}{\sqrt{N_{\text{obj}}}}, \tag{3.18}$$

where $N_{\text{obj}}$ is the number of catalog objects included in the stack. The signal in each stack aperture is from a sum across voxels, so using error propagation and again assuming a constant root mean square (RMS) noise value across the LIM cube, this can be written in terms of the noise level in each individual voxel (3D pixel), $\sigma_{\text{vox}}$, as

$$\sigma_{\text{stack}} = \frac{1}{\sqrt{N_{\text{obj}}}}\sqrt{\sum_{N_{\text{chan}}}\sum_{N_{\text{spax,x,y}}}\left(\sigma_{\text{vox}}^2\right)} \tag{3.19}$$

$$= \sqrt{\frac{N_{\text{chan}}N_{\text{spax}}^2}{N_{\text{obj}}}}\sigma_{\text{vox}}. \tag{3.20}$$

In this case (using COMAP as an example), $\sigma_{\text{vox}}$ is the radiometer noise described in Equation 3.15. The uncertainty is thus written as

$$\sigma_{\text{stack}} = T_{\text{sys}}\sqrt{\frac{N_{\text{chan}}N_{\text{spax}}^2\Delta_x^2}{N_{\text{obj}}2\delta_\nu\tau\delta_x^2 N_{\text{f}}}}. \tag{3.21}$$

This means that, in addition to the usual parameters that affect the sensitivity of a LIM experiment (such as the system temperature, the integration time, the number of feeds, etc.), which we do not explore here, the stack uncertainty is specifically dependent on the number of voxels included in the stack aperture in all three axes, as well as the voxel size in all three axes, and the number of objects included in the stack.

Finally, we note that this uncertainty calculation only accounts for thermal noise in the COMAP data, and leaves out sample variance. Chen et al., 2025 have shown that the sample variance is an important consideration when calculating cosmological parameters from a stack. However, as this work is framed primarily in terms of detecting signal from a stack, we ignore the contribution to the uncertainty from the sample variance, as thermal noise should be the main driver of what is detectable with a given instrument and spectral line. The S/N ratio values listed throughout this work can be thought of as 'detection significances.'

### 3.5 Simulated stack results

We explore the effects of varying different facets of both the LIM map and the galaxy catalog on the stack sensitivity. We divide these effects into choices made while generating the stack itself (§3.5), factors in the experimental design of either the galaxy catalog or the LIM experiment (§3.5 and §3.5), astrophysical factors affecting the LIM or the galaxy catalog tracer spectral line (§3.5), and interloper contamination in either experiment (§3.5).

#### Stack parameters

Some of the most obvious things that can be varied are in the choices made for the stack itself. Here, we test the size of the central aperture—the number of pixels in the spatial ($N_{\rm spax}$) and spectral ($N_{\rm chan}$) directions that are summed over to determine the total luminosity of the stacked cubelet. Under the stacking methodology we use here, these must be an odd integer, and they should be as small as possible (as $\sigma_{\rm stack} \propto \sqrt{N_{\rm chan}N_{\rm spax}^2}$) while still capturing as much as possible of the flux from the stack.

#### Spatial aperture size

Here, we tested the various choices of spatial aperture size to see which returns the best stack S/N. We performed stacks on 99 different simulation realizations using the default configuration. For each, we determined the stack luminosity and S/N using square apertures with side lengths varying from one spaxel to nine spaxels, in steps of two (holding the aperture width constant in the spectral axis) as well as larger apertures of side lengths 15, 21, and 27. We plot the results in Figure 3.6.

We find that the choice of central aperture size that maximizes the stack S/N is $N_{\rm spax} = 7$, although $N_{\rm spax} = 5$ and $N_{\rm spax} = 9$ provide very comparable results. This is contrary to the logic laid out in Dunne et al., 2024, where we had chosen $N_{\rm spax} = 3$ so that a single object located anywhere inside the central spaxel will have the majority of its beam FWHM fall into the central aperture, while simultaneously integrating in as little noise as possible. While the noise in the stack does increase as $\sigma_{\rm stack} \propto N_{\rm spax}$, increasing the aperture size to $7 \times 7$ integrates in so much more luminosity that the S/N actually increases. The stack luminosity continues to increase considerably with increasing aperture size, although it flattens out in the larger apertures. This reflects the spatial profile of the stack luminosity, which has large wings extending out to $\sim 20$ arcmin (Figure 3.5). As the aperture gets larger

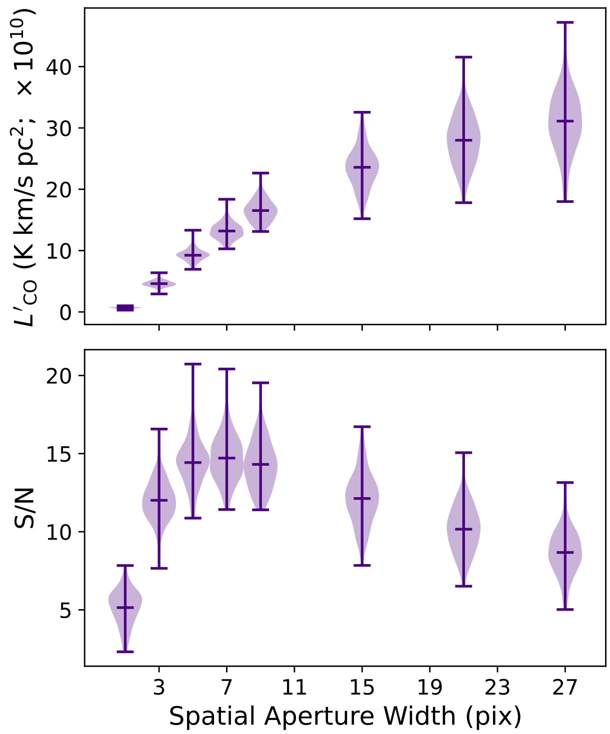


Figure 3.6: Violin plot showing how the luminosity (top) and the S/N (bottom) of the stack changes as the central aperture being summed over to determine the final stack luminosity is widened in the spatial axes. The default configuration uses a $3 \times 3$-spaxel aperture.

than $\sim 7 \times 7$ the effects of the increased $N_{\rm spax}$ dominate the S/N, which begins to decrease again even as the luminosity continues to increase.

The fact that there is so much more luminosity at wider radii in the stack suggests that the stack is not well-explained by point sources in the central spaxel (*i.e.*, the objects actually included in the galaxy catalog alone cannot account for all of the stack signal). We plot the average spatial profile of 99 noiseless simulations in Figure 3.7. We also show a two-component Gaussian fit to this profile, including both a brighter, narrower component and a broader, dimmer component. This is motivated by the logic that the profile would be dominated by luminosity from the cataloged object and its immediate surroundings, and larger-scale cosmological clustering would also contribute. The profile is fit well with two Gaussian profiles (reduced $\chi^2 = 0.09$). The standard deviations of the two Gaussian components are $3.2 \pm 0.1'$ and $9.6 \pm 0.5'$, respectively.

To confirm that the stack signal is the product of larger-scale clustering, and not extended signal from the cataloged objects themselves, we also generate the profile of a stack on only the cataloged DM halos. We generate noiseless simulation realizations with the $L_{CO}$ of all halos but those included in the mock galaxy catalog artificially set to zero, and perform the stack on these simulation realizations. The resulting stack is also shown in Figure 3.7. We find that not only is the spatial extent of this stack considerably reduced, but the luminosity of the stack itself is nearly negligible compared to the luminosity of the stack on the full mock LIM data. This suggests that the stack signal is so extended because, at least when using the default models for both CO and Ly$\alpha$, the signal is coming almost entirely from halos surrounding the catalog object, rather than from the catalog object itself. We explore the extent of this clustering in Section 3.6, and its dependence on model choices in Section 3.6. Although somewhat outside of the scope of this paper, we explore how model choices affect the spatial profile of the stack specifically in Appendix 3.10.

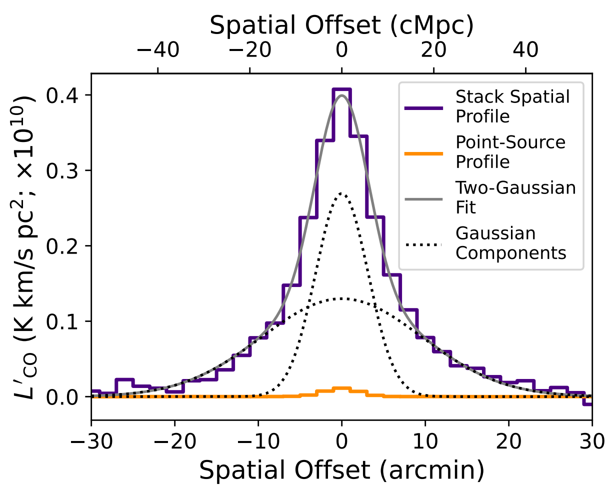


Figure 3.7: Average one-dimensional spatial profile of 99 noiseless simulated stack realizations. A double-Gaussian fit to the profile is shown in grey, with each of the Gaussian components shown as black dotted lines. The orange profile is a stack performed on simulations that have CO luminosity painted only onto the halos actually included in the mock galaxy catalog, to demonstrate the extent to which the stacked signal is the result of halo clustering.

#### Spectral aperture size

We default to an aperture width of three channels in the spectral axis. Here, in addition to cosmological clustering, the stacked signal is broadened by astrophysical line broadening (rather than instrumental resolution). The default aperture width of three channels is chosen from visual inspection of the output stack spectrum, as astrophysical line broadening is not well characterized for CO at $z \sim 3$. As above, we test the choice of aperture width by performing stacks on 99 different simulation realizations, and extracting the central aperture luminosity using varying aperture widths. Here, we hold the spatial axes constant at $3 \times 3$ spaxels and vary only the spectral axis, stepping from one channel to nine channels in steps of two. The results are shown in Figure 3.8.

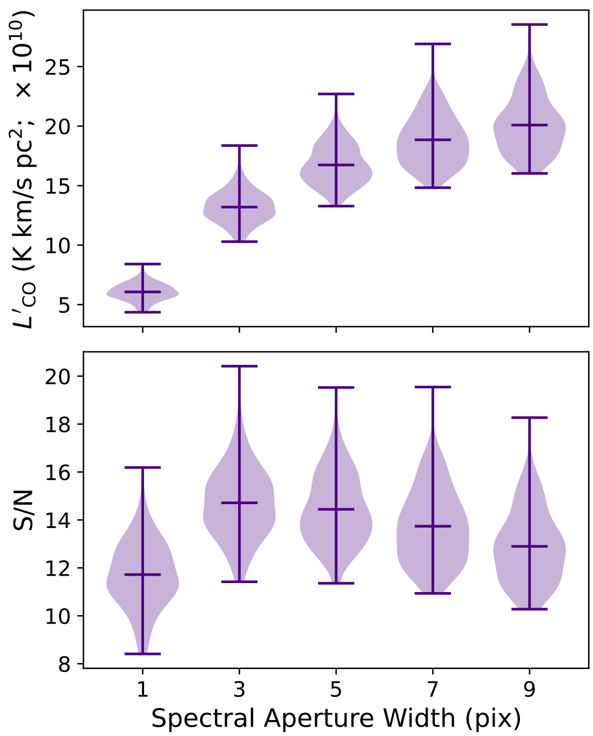


Figure 3.8: Violin plot showing how the luminosity (top) and the S/N (bottom) of the stack changes as the central aperture being summed over to determine the final stack luminosity is widened in the spectral axis. The default configuration uses a three-channel aperture.

In this case, we find that the three-channel aperture does indeed maximize the S/N of the stack. As with the spatial axes, the output stack luminosity continues to increase as the aperture gets wider. The noise increases as $\sigma_{\text{stack}} \propto \sqrt{N_{\text{chan}}}$, and begins to dominate after the three-channel aperture.

As above, we tested the extent of the contribution from the objects actually included in the mock catalog by comparing the spectrum of the stack to the spectrum of a stack performed on a map where only the sources included in the mock catalog are assigned CO luminosity. This is shown in Figure 3.9. We fit a two-component Gaussian to the spectral profile of the stack. We find that there is also considerable clustering contributing to the width of the signal in the spectral axis. However, the spectral channels correspond to larger physical sizes than the spaxels do, so the three-channel aperture has roughly the same comoving size as the 7-spaxel aperture in the spatial directions.

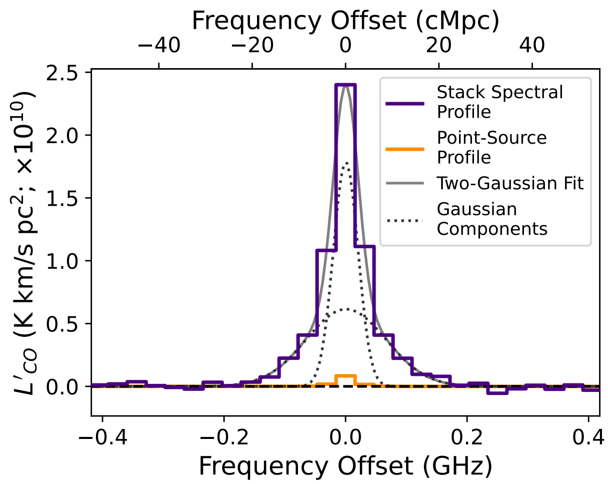


Figure 3.9: Average spectrum of 99 noiseless simulated stack realizations. A double-Gaussian fit to the profile is shown in grey, with each of the Gaussian components shown as black dotted lines. The orange profile is a stack performed on a simulated map with only the halos included in the Ly$\alpha$ catalog assigned CO luminosities; all the other halos were given zero luminosity.

**Galaxy catalog experimental factors**

As a resolved survey of galaxies is a much more established technique than LIM, there may potentially be several existing galaxy surveys to choose from when deciding to perform a stacking analysis with a LIM experiment. In this section, we explore how the makeup of the galaxy survey affects the stack result.

**Number of catalog objects**

The main factor driving the sensitivity of the stack should be the number of catalog objects being stacked, $N_{\text{obj}}$, with the sensitivity improving as $\sqrt{N_{\text{obj}}}$ for a case where the noise response is constant across the LIM map. To verify this hypothesis, we perform stacks of 100, 1000, and 10 000 catalog objects on each of 99 different simulation realizations. For all other parameters, we use the 'default' values listed in Table 3.2.

We plot the distribution of stack luminosity and S/N values for each $N_{\rm obj}$ in Figure 3.10. We also include for reference the theoretical $S/N \propto \sqrt{N_{\rm obj}}$ expectation for a stack with only white noise. Using the catalog selection strategy we adopt (where catalog objects can be selected above a cut-off luminosity, and the selection probability is linearly dependent on the luminosity of the catalog object), the stacks with varying $N_{\rm obj}$ values all return extremely consistent luminosities. The variation of S/N with $N_{\rm obj}$ agrees well with theory.

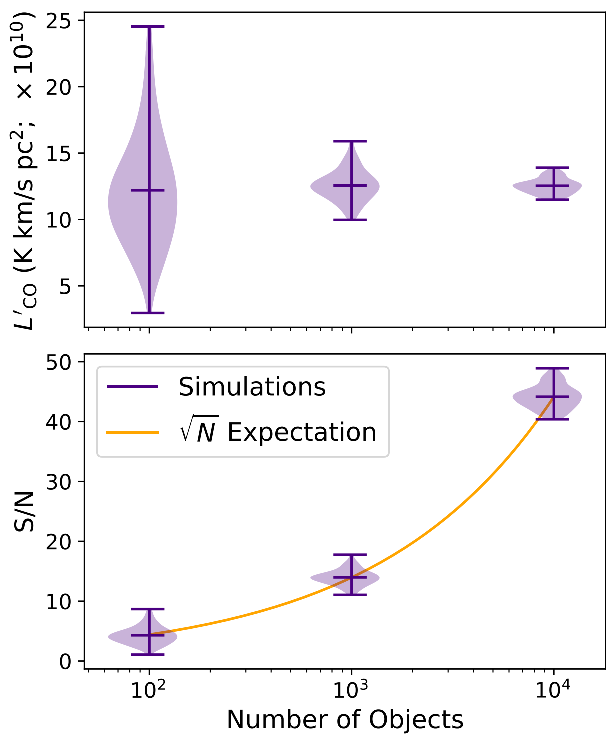


Figure 3.10: Violin plot showing how the luminosity (top) and the S/N (bottom) of the stack changes as more objects are included in the stacked catalog. The orange line shows the $\sqrt{N_{\rm obj}}$ improvement, which is to be expected for a LIM map with only white noise.

**Redshift uncertainty**

Thus far (including in our derivation in §3.4) we have assumed a galaxy catalog with perfect redshift accuracy. However, this is unlikely to be the case in practice. Galaxy catalogs range from high-precision spectral surveys (such as DESI; DESI

Collaboration et al. 2024b) to photometric redshift surveys with large uncertainties (such as DES; DES Collaboration et al. 2022). Additionally, while galaxy catalogs may have precise redshifts themselves, certain objects have inherent uncertain offsets between different redshift tracers (for example, quasars can have large offsets between the optical lines used for eBOSS and molecular lines, due to inflows or outflows in the quasar host galaxy, Dunne et al., 2024). Often, spectral resolution is sacrificed in order to gain larger numbers of objects in a galaxy catalog.

Because scatter in the redshift axis of a galaxy catalog can typically be characterized, although not removed, it is possible to account for the redshift uncertainty when determining the spectral width of the stack aperture. We thus chose three uncertainty values corresponding to steps in the spectral width of the stack aperture. We measure the frequency width of a stack with no redshift uncertainty (this still contains some inherent width in its spectrum due to astrophysical line broadening), which we treat with a three-channel spectral aperture, and then scale this frequency width up to widths that match a 7-, 11-, and 15-channel aperture. We then convert this frequency width to a velocity width at the mean redshift of the catalog objects ($z_{\rm mean} \sim 2.8$) and perform 99 different stacking realizations at each velocity uncertainty, matching the spectral aperture width to the line width for each stack. We plot the results in Figure 3.11. We repeat this analysis for a much larger redshift uncertainty approximating that of the Dark Energy Survey photometric catalog at $z = 1.2$, which is roughly $\Delta_z = 0.045$ (Sevilla-Noarbe et al., 2021; DES Collaboration et al., 2022).

We find that collecting the signal scattered out of the original stack aperture by redshift uncertainty works very well—the returned luminosity value for each of these stacks with added uncertainty is the same and is actually larger than the returned luminosity for the stack on the unscattered catalog. This is because, as shown above, the linewidth of the stack is not actually a perfect Gaussian—galaxy clustering and astrophysical line broadening, in addition to the redshift uncertainty, all contribute to widening the line. While the luminosity stays constant, the S/N ratio decreases considerably with increasing uncertainty, even for the smaller, spectroscopic levels of uncertainty. This is because this strategy adds additional spectral channels to the stack (*i.e.*, increases $N_{\rm chan}$), which adds to the noise as $\sqrt{N_{\rm chan}}$ (Equation 3.21). In simpler terms, the signal in the stack is spread over a larger region of the noise floor. At the photometric redshift uncertainty $\Delta_z = 0.045$, the decrease in sensitivity is extremely significant—the average S/N in the $\Delta_z = 0.045$ case is 5.3, only a third of

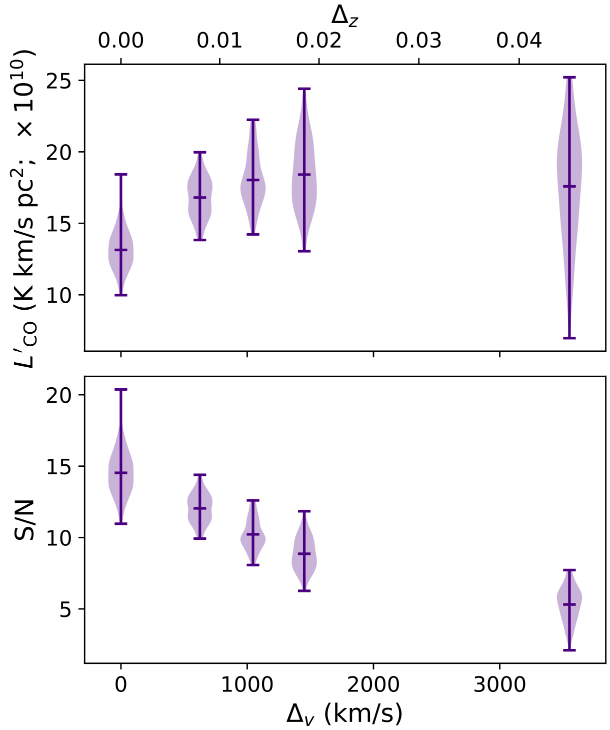


Figure 3.11: Violin plots showing the behavior of the stack in returned luminosity (*top*) and S/N (*bottom*) with uncertainty, $\Delta_v$, in the redshifts of the galaxy catalog being stacked on, assuming that uncertainty can be measured and the spectral aperture of the stack can be widened to account for it. Each of these stacks is on 1000 catalog objects. This analysis does not account for the increased numbers of objects available in some photometric catalogs.

the sensitivity of the unscattered case.

However, photometric galaxy catalogs also benefit from considerably increased source densities. We treat the DES Y6 Gold catalog and the DESI ELG sample as examples of photometric and spectroscopic galaxy catalogs, respectively. This is merely an illustrative example, as neither catalog has coverage in the redshift range we are simulating in the rest of this work. The DES Y6 Gold catalog has approximately 28.9 confident galaxy detections per arcmin$^2$, or 104,040 deg$^{-2}$. In comparison, the DESI ELG sample has an on-sky density of roughly 2400 deg$^{-2}$ (DESI Collaboration et al., 2025b)—the DES photometric catalog has a factor of 43.5 greater object density. This is more than enough to offset the attenuation due to redshift uncertainty, at least in this case. We find that a stack on a catalog with a DES-like redshift uncertainty ($\Delta_z$ = 0.045), but also a DES-like increase in object density (43,500 stack objects) returns a stack S/N ratio of 37.5. Photometric catalogs may then work very well for a stacking analysis, provided they are rich in available objects. To yield a stack detection at the same significance, a catalog with the DES redshift uncertainty would need to have 7.5× as many objects as a catalog with no redshift uncertainty.

**LIM data experimental factors**

The resolution of various LIM experiments varies considerably in both the spatial and spectral axes, which affects the stack outcome. In this section, we test the effects of varying the spectral resolution, the pixelisation, and the beam width of the LIM maps themselves.

**LIM spectral resolution**

The spectral resolution of the map is particularly interesting, as there is a large variation in spectral resolution across existing and proposed LIM experiments, based on the spectral lines being pursued and the technologies being employed at the receiver front end. For example, CONCERTO (CONCERTO Collaboration et al., 2020), a Kinetic Inductance Detector (KID) experiment with an absolute spectral resolution of $\sim$ 1.5 GHZ (Fasano et al., 2022), FYST (CCAT-Prime Collaboration et al., 2023, KID-based with $R \sim 100$), or TIME (Crites et al., 2014; Sun et al., 2021, using transition edge sensor bolometers with $R \sim 100$) all have velocity resolutions ($\Delta_v \sim 1000 - 3500$ km s$^{-1}$). This is up to an order of magnitude broader than the default value used here ($\Delta_v \sim 300$ km s$^{-1}$), for COMAP. SPHEREx (Doré et al., 2014; Doré et al., 2016; Doré et al., 2018), using optical linear variable filters at $R = 35 - 130$, will have a resolution ranging from 2300 to 8500 km s$^{-1}$.

We vary the spectral resolution into which the maps are being binned while preserving the size of the central frequency aperture at 91.75 MHz. Because the current stacking methodology requires there to be an odd number of frequency channels in the central aperture, we place one, three, and five channels across the aperture, corresponding to channel widths of 91.75 MHz, 31.25 MHz (the COMAP science resolution), and 18.75 MHz. We create 99 different simulation realizations at each spectral resolution, and perform stacks on each realization to measure the effects of changing spectral resolution on the sensitivity of the stack. In each case, we extract the central 91.75 MHz to calculate line luminosity values. Theoretically, broadening the spectral resolution while maintaining the same stack aperture width in frequency should not affect the stack sensitivity (from Equation 3.21, $\sigma_{\text{stack}} \propto \sqrt{N_{\text{chan}}/\delta_v}$, so the two parameters balance each other out).

Additionally, we perform stacks on each of 99 different simulation realizations at resolutions of 187.5 and 375 MHz (corresponding to roughly 1900 km/s and 3700 km/s, respectively). We do this to explore the effects of a LIM map with spectral resolution broad enough that a single frequency channel is much larger than the FWHM of the spectral line. These resolutions are wider than the chosen aperture size, so in each case we take the aperture size equal to the spectral resolution, and include only the central frequency channel in the stack aperture. Because the aperture size is now increasing, the assumption that the stack sensitivity should remain the same no longer holds. The results of these stacks are shown in Figure 3.12.

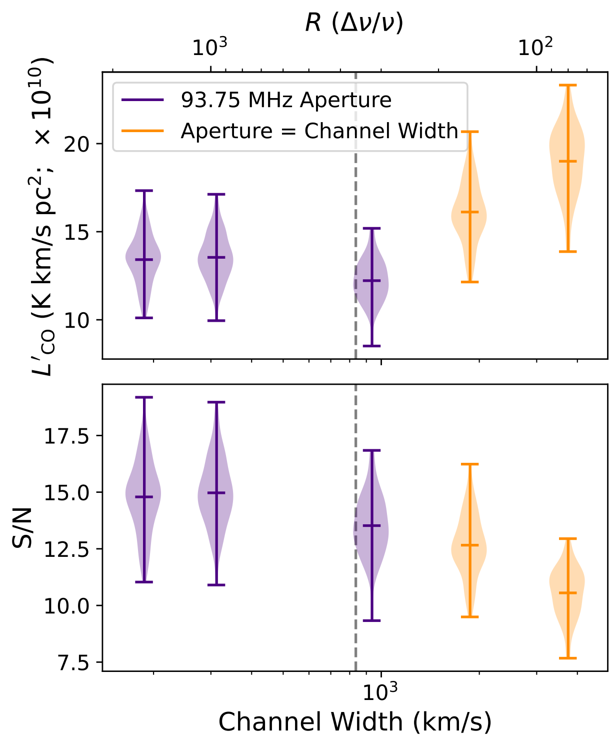


Figure 3.12: Violin plot showing the distributions of luminosity values (*top*) and S/N ratios (*bottom*) output from the stack for different simulation realizations at various LIM map spectral resolutions. Channel widths with indigo violins are those where the stack aperture width stays constant at 93.75 MHz as the channelisation changes. Orange violins have individual channels larger than 93.75 MHz – their stack apertures are the size of a single spectral channel (the smallest possible size). The grey dashed line indicates the FWHM of the stack linewidth.

We find that when all other parameters (including the size of the stack aperture) are kept constant, improving the spectral resolution (decreasing the channel width) actually marginally improves the S/N ratio of the output stack, albeit with a large scatter between different simulation realizations. This trend is likely an artefact of the current aperture extraction strategy: currently, the cutouts of the LIM data corresponding to each catalog object are only centered by lining up the central spectral channel. Thus, catalog objects located anywhere within the central channel are treated identically. This introduces a velocity uncertainty the width of the central channel, in practice equivalent to convolving the average spectral profile of the objects being stacked with a boxcar function the size of the central channel. Shrinking the size of the channel shrinks the boxcar being convolved in, and retains more flux in the central stack aperture.

We note that this effect is mitigated by improving the spectral resolution of the LIM data, but it could equally be mitigated by exploring more optimal aperture extraction techniques, such as matched filtering (e.g., Zubeldia et al., 2021) or an interpolation-based centring. Still, even when more optimal techniques are used, a better LIM spectral resolution is likely to be advantageous for the stack, as it will put more data points across the stacked spectrum. This will enable a more accurate measurement of the average line profile of the objects included in the stack or make matched filtering or interpolation extraction techniques more effective.

In the case where the channel width is considerably larger than the stack FWHM (so the aperture width of the stack in the spectral axis is forced to increase), the overall measured stack luminosity increases, as more of the wings of the line emission are included in the central aperture. However, as the aperture size increases, the balance between $N_{\mathrm{chan}}$ and $\delta_\nu$ is no longer maintained in $\sigma_{\mathrm{stack}}$, so as expected the S/N ratio degrades. As the stack aperture gets wider, more regions of the spectrum that contain only noise are being integrated into the final stack, and the line luminosity increasing is not sufficient to counteract the effects of this increased noise. Experiments with larger spectral channels will have to account for this effect if they choose to perform stacking experiments.

**LIM spatial pixelisation**

While the angular resolution (beam) of a single-dish LIM experiment is set by the size of the antenna being used and the illumination of that antenna by the feed(s), the pixelisation of the angular axes is decided in the map-making steps (e.g., Lunde et al., 2024). As in the spectral axis, where the ability of the aperture extraction to concentrate signal under the current stacking methodology is limited by the width of the central frequency channel, it may be possible to improve the sensitivity of a stacking analysis by oversampling the beam and concentrating signal more effectively.

We explore this possibility by varying the spatial resolution of the simulated maps, following a similar methodology to §3.5. Here, we hold the spectral resolution constant and instead vary the pixel size in the angular axes. Similar to the spectral analysis, we maintain the central aperture at a constant area and vary the number of spaxels placed across it. Again, the stack methodology requires an odd number of spaxels included in the central aperture in each angular dimension, so we place $1^2$, $3^2$, $5^2$, and $7^2$ spaxels across the area of the central aperture. These correspond to pixel scales of $6'$, $2'$, $1.2'$, and $0.857'$ per side. As above, $\sigma_{\rm stack} \propto \sqrt{N_{\rm spax}^2/\delta_x^2}$, so increasing the spectral resolution while maintaining the same aperture area should not affect the sensitivity of the stack.

We create 99 different simulation realizations at each pixel size, performing stacks on each realization. In each case, we extract a central $6' \times 6'$ aperture in the spatial axes. We show the S/N ratio of each stack as a function of its resolution in Figure 3.13. As with the spectral axis, improving the spatial pixelisation does marginally improve the stack S/N ratio, likely for the same reasons—the catalog object could be anywhere inside the central spaxel, so the average stack spatial profile is convolved with a 2D boxcar the width of the central spaxel. This effect is especially apparent in the case with $14'$ spaxels, where only the central spaxel is included in the stack—objects located anywhere near the outskirts of the central spaxel will have almost half of their signal left out of the stack. Again, improved aperture extraction techniques could improve the concentration of signal as effectively as decreasing the pixel size.

#### Beam size

Stacking analyses are likely to be far from the primary consideration setting the actual spatial resolution (in the case of COMAP, the beam size) of a LIM experiment. Nevertheless, for completeness, we explore the effects of the beam size on the stack sensitivity. We hold both the spatial and spectral pixel sizes constant, and vary the size of the beam being convolved into the synthetic LIM map.

We explore four steps in beam area—we choose beam FWHM values of $2.25'$, $6.75'$, and $9.0'$ as well as the default value of $4.5'$. Because of the extended spatial

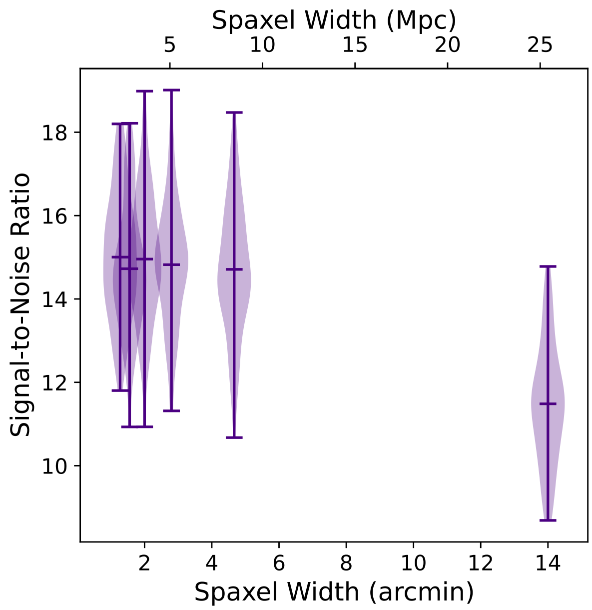


Figure 3.13: Violin plot showing the distributions of S/N ratios output from the stack as a function of the pixel size of the LIM data while holding the beam size constant.

distribution of the signal, it is difficult to scale the aperture size to the beam such that the same amount of signal falls into the central aperture at each beam size—we have already shown that the stack cannot be treated as a point source. Instead, we compare the output S/N ratios of the stacks on different beam sizes in two separate ways. First we hold the aperture size constant at $N_{\mathrm{spax}} = 7$ in the spatial axes, and measure the output S/N ratio at each beam size. Second we determine the aperture size that maximizes the S/N ratio for each beam FWHM, and compare the maximum-S/N stacks against each other. For the 2.25′ case the maximum S/N ratio is at $N_{\mathrm{spax}} = 5$, for the 4.5′ case the maximum S/N ratio is at $N_{\mathrm{spax}} = 7$, and both the 6.75′ and 9.0′ cases have their S/N ratio maximized at $N_{\mathrm{spax}} = 9$. The results of these tests performed on stacks using 99 different simulation realizations at each beam FWHM are compared in Figure 3.14.

We find that, in the case where the aperture size is held constant as the beam FWHM changes, the stack luminosity (and thus S/N ratio) does indeed decrease with increasing beam FWHM. Increasing the beam FWHM smears the signal more, pushing more of the luminosity from the stack outside the central aperture, so this is perhaps trivial. More interestingly, the stack S/N ratio falls off with beam FWHM at a nearly identical rate when the aperture size is chosen to maximize the S/N ratio.

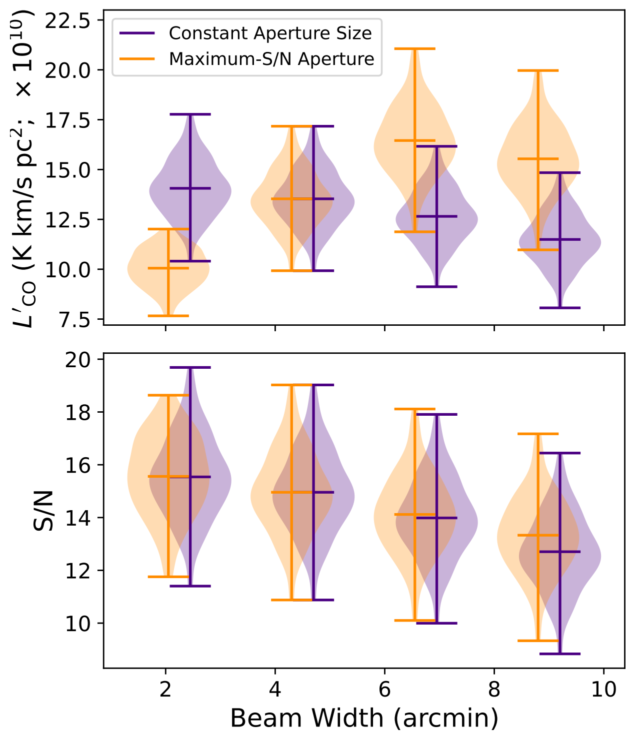


Figure 3.14: Violin plot showing the distribution of S/N ratios output from simulated stacks as a function of the beam FWHM (*i.e.*, angular resolution) of the LIM experiment while holding the pixelisation constant. We show a version where we hold the size of the spatial aperture constant as well as a version where we set the spatial aperture to the size that maximizes the S/N ratio of the stack. Violins are offset slightly in the x-axis for legibility.

As the beam FWHM gets larger, the signal is being spread over a larger region of the noise floor. The aperture that maximizes the S/N ratio is balancing integrating in more luminosity from the outskirts of the stacked signal with the additional contribution of more noise as more spaxels are included in the central aperture. In any case, a more concentrated signal leads to higher S/N ratio.

We note that, while this analysis may seem very similar to §3.5, there is a key difference in the noise properties of the maps. In §3.5, where the pixelisation is being changed, the fraction of the observing time being spent in each spaxel and thus the per-spaxel noise properties vary across the different resolutions (*i.e.*, $\delta_x$ and $N_{\rm spax}$ are both changing). Thus, while smaller spaxel sizes sampled the beam better, they simultaneously had more noise per spaxel, so the two effects were working in concert. Here, $\delta_x$ is staying the same. The only variation is in the amount of signal spread outside the central aperture by the beam and potentially $N_{\rm spax}$.

**Astrophysical factors affecting the stack**

In addition to the experimental factors affecting the LIM data or the galaxy catalog (which mostly come into the stack S/N ratio through the uncertainty in the stack, Equation 3.21), many basic astrophysical factors will also affect the outcome of the stack. These will directly affect the measured stack luminosity and affect the S/N ratio as shown by Equation 3.17. These factors tend to be considerably more nebulous than the experimental factors, as they have not yet been reliably measured at $z = 3$, and, in the case of the CO and Ly$\alpha$ lines targeted here, are extremely difficult to model even in hydrodynamic simulations. Indeed, measuring these quantities is part of the goal of LIM experiments.

**Choice of LIM model**

A large source of uncertainty in the modeled stacks is the CO luminosity of DM halos at $z \sim 3$, $L_{\rm CO}(M_{\rm h})$. Although considerable effort has been made to accurately model these emitters on cosmological scales (e.g., Lidz et al., 2011; Li et al., 2016; Padmanabhan, 2018; Keating et al., 2020; Chung et al., 2022), this is a phase space that has largely been unconstrained by observations. The best current constraints at these redshifts come from COMAP (Chung et al., 2024b) but the models that have not yet been excluded still predict luminosities that differ by an order of magnitude. We test three different models of CO emission, each described in §3.3, to see how this affects the stacked signal. These models are C22 (Chung et al., 2022, the default model), P18 (Padmanabhan, 2018), and L16 (Li et al., 2016). We chose not to explore any bulk changes to the CO luminosity, as these would increase the luminosity of all DM halos by the same factor and thus simply raise the stack luminosity by that same factor.

We ran 99 different simulation realizations for each model of CO emission, and compare the output S/N ratios for each in Figure 3.15. The C22 and L16 $L(M_{\rm h})$ functions yield very similar S/N ratios, while the P18 $L(M_{\rm h})$ function with $f_{\rm duty} = 0.1$ has considerably lower significance. This is an interesting result, especially when compared against the $L(M_{\rm h})$ functions plotted in Figure 3.3. At most values of $M_{\rm h}$, the returned luminosity under the P18 function falls between the $L_{\rm CO}$ values from C22 and L16. L16 provides higher $L_{\rm CO}$ values at low $M_{\rm h}$, and C22 provides higher $L_{\rm CO}$ values at high $M_{\rm h}$. The cross-over is at roughly $1.5 \times 10^{12}$ $M_\odot$, at which point all three $L(M_{\rm h})$ functions yield the same $L_{\rm CO}$. Based on this, one would naively expect either the C22 or L16 model to dominate, based on the halo

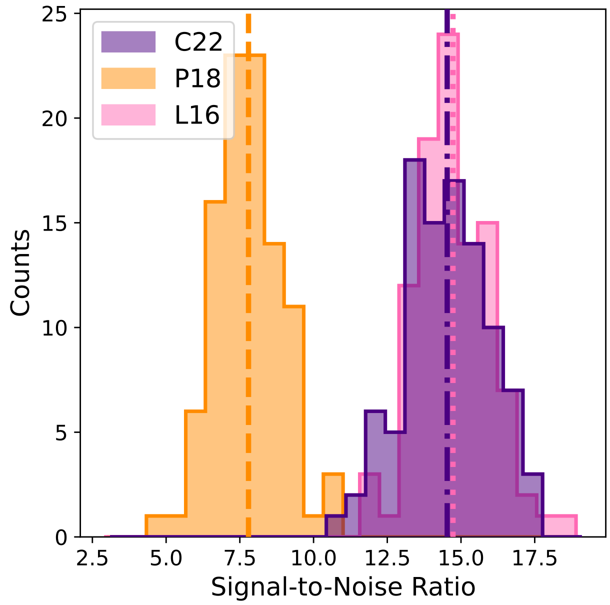


Figure 3.15: Histograms of the output S/N ratio for each model of the CO luminosity of a DM halo as a function of its mass, $L(M_{\rm h})$. Mean values for each model are shown as vertical lines. The three models are described in §3.3.

mass selected by the catalog luminosity function, and P18 to fall somewhere in the middle. We explore the implications of this more thoroughly in §3.6.

#### Choice of catalog luminosity function

The returned luminosity of the stack does not depend directly on the Ly$\alpha$ luminosity of the halos. However, the $L_{\rm Ly\alpha}$ of the halos does determine which DM halos are actually included in the stack (via the selection function $S(L_{\rm Ly\alpha}(M_h, \rho))$ in Equation 3.17), which has the potential to significantly impact the stack outcome.

To test how the galaxy catalog affects the behavior of the stack we run 99 different simulation realizations for each of the Schechter function parameters described in Section 3.3. These include parameters fit to observations of LAEs at $z = 3$ (the 'default' model), parameters fit to quasars (which trace a more stochastic population of higher-mass halos; the 'bright' model), and parameters fit to LAEs at $z = 0.3$ (a luminosity function more dominated by the many low-mass halos with faint luminosities; the 'faint' model). The output S/N ratios of stacks performed using each model are shown in Figure 3.16. Because we hold all other parameters constant, stacks on each model have the same noise levels—differences in S/N ratio are driven solely by the output stack luminosity.

We find that the 'faint' galaxy catalog has properties that are very similar to the default LAE catalog, while the 'bright' catalog is considerably brighter (by a factor of more than three). For all catalog models, larger halo masses correspond to brighter luminosities in the catalog tracer, so the bright catalog model should be selecting only the largest DM halos. These will both be the brightest halos themselves and be the most biased (they will have more, and brighter, neighboring halos contributing to the second term in Equation 3.17). The combination of these effects should explain the extremely bright stack. We test this further in §3.6.

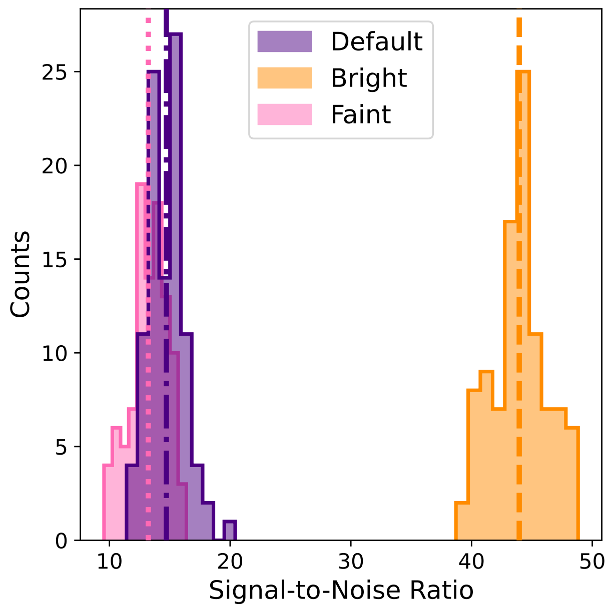


Figure 3.16: Histograms of the output S/N ratio for each model of the luminosity function of the emission line used in the galaxy catalog being stacked on (in this case, Ly$\alpha$). The different catalog models are described in §3.3.

#### Correlation in scatter between tracer luminosities

Finally, because the stack is an average of the LIM tracer luminosities of the DM halos included in the galaxy catalog, it will be affected by the halo-to-halo correlation between the LIM tracer luminosity and the catalog tracer luminosity. As in §3.5, this is only a selection effect (coming in through the $S(L_{\mathrm{Ly}\alpha}(M_h, \rho))$ factor in Equation 3.17), and not a direct scaling of the luminosity values being averaged.

We stack 99 simulation realizations at five different values of $\rho$ spanning from completely anti-correlated to completely correlated scatter in halo luminosities ($\rho = [-1, -0.5, 0, 0.5, 1]$). The output luminosity distributions of the resulting stacks are shown in Figure 3.17, both as absolute values and relative to the output luminosity for completely uncorrelated scatter ($\rho = 0$). We repeat this exercise for the bright, AGN-like catalog luminosity function, to explore interactions between the catalog luminosity function and $\rho$ in determining how halos are selected and the resulting stack luminosity.

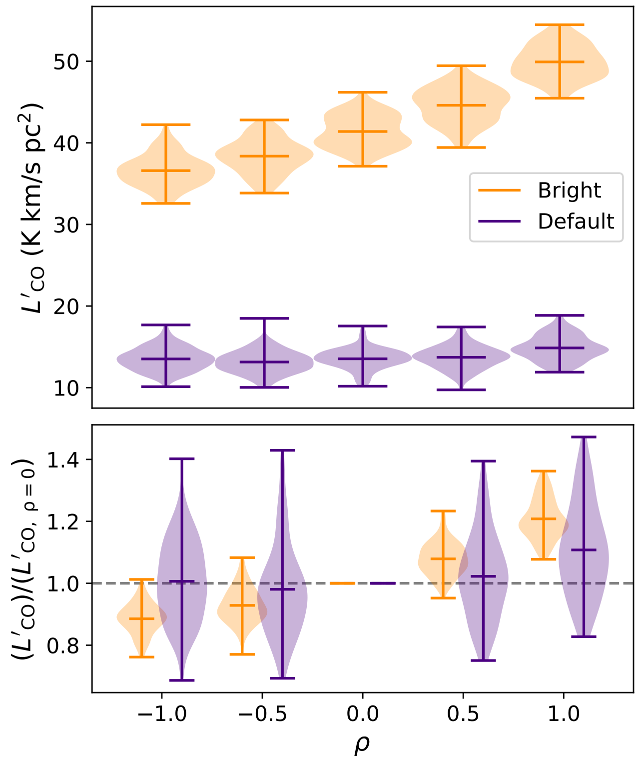


Figure 3.17: Effect of varying the correlation in the two luminosity values' scatter on the output stack luminosity. *Top:* Violin plots showing the stack luminosity for each $\rho$ value for both the default choice of the Ly$\alpha$ luminosity function $z \sim 3$ and the bright AGN-like version. *Bottom:* Output luminosity values normalised to the value at $\rho = 0$ to show relative changes between values of $\rho$. Violins are offset slightly in the x-axis for clarity.

We find that $\rho$ does have an effect on the returned luminosity, but this effect is very subdominant compared to the luminosity function of the catalog tracer. Additionally, we find that varying $\rho$ has proportionally more of an effect on the output stack luminosity for the more stochastic, top-heavy 'bright' catalog luminosity function than for the flatter default one. We explore this behavior in more detail in §3.6.

**Astrophysical line broadening**

Astrophysical line broadening of the LIM tracer emission is likely to have an effect on the stack degenerate with redshift uncertainty in the galaxy catalog—it widens the signal in the spectral axis, thus possibly moving more signal outside the stack aperture. Unlike redshift uncertainty, this is inherent to the LIM tracer and not instrumental in origin.

Because CO linewidths are not well characterized at $z = 3$, we use the three theoretical prescriptions described in §3.3 to test the effects of line broadening on the stack. For each prescription, we perform stacks on 99 different simulation realizations. Unlike in the redshift uncertainty case, we do not attempt to correct for the CO linewidth prescription when determining the spectral aperture width of the stack—we use three spectral channels, or 91.75 MHz in all cases. This is because we do not know a priori how the CO linewidths will behave at $z \sim 3$, (whereas the redshift uncertainty of a galaxy catalog can typically be measured). It is possible that this effect could be measured using more advanced signal extraction techniques (than the straight sum we are currently using) in the spectral axis. The output S/N ratios for the different simulation realizations are compared in Figure 3.18. Also shown are the average stacked spectra (across simulation realizations) for each line broadening prescription.

We find that using the virial velocity of the DM halo as the rotational velocity predicts the worst S/N ratio for the stack, where either using $v_{M_{\rm peak}}$ or the virial velocity with a cut-off predict higher S/N ratios. The difference in S/N ratio is due to the stacked linewidth being smaller for the $v_{M_{\rm peak}}$ method than for either of the $v_{\rm vir}$-based methods, as can be seen in their average stacked spectra. The average FWHM of the stacked spectral line is 596 km/s using the $v_{M_{\rm peak}}$ method. The $v_{\rm vir}$ method without a 1000 km/s cut-off yields a line FWHM of 1212 km/s, and adding a cut-off reduces the FWHM to 813 km/s.

This is consistent with the virial velocity method assigning high-mass DM halos much higher linewidths than have been observed (see Appendix 3.9). These high-mass halos are also likely the brightest halos and thus contribute more strongly to the stack. If the high-mass halos have larger linewidths, the stack as a whole is likely to be broader and thus have more signal falling outside of the spectral aperture used to calculate the final S/N ratio. Both the cut-off virial velocity and the $v_{M_{\rm peak}}$ prescription allow for narrower spectral lines at high halo masses.

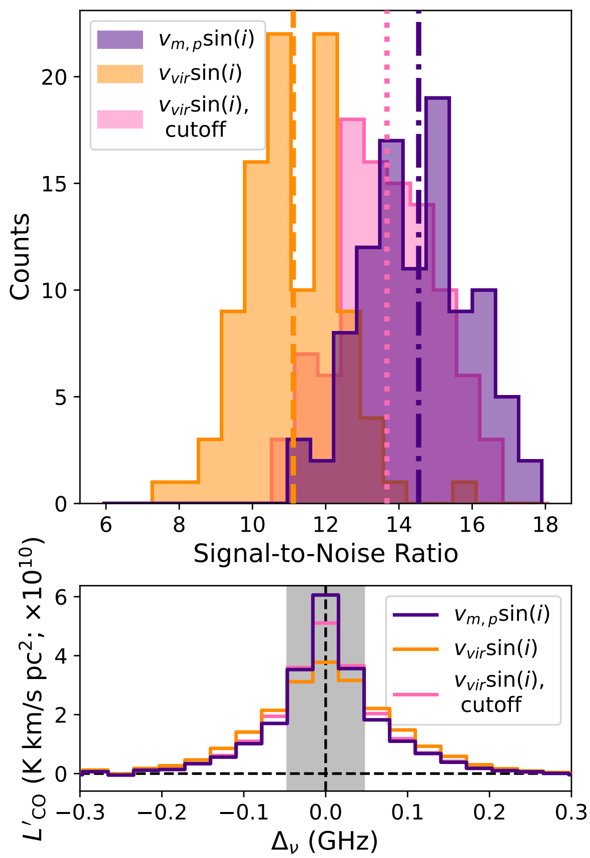


Figure 3.18: Effects of different line broadening prescriptions on the stack. *Top:* Histogram of the output S/N ratio for each prescription for astrophysical line broadening of CO. The different prescriptions are described in §3.3. Mean values are indicated with vertical lines. *Bottom:* Average stacked spectrum across simulation realizations for each prescription. The central 93.75 MHz, which is integrated across to determine the final stack luminosity, is shown in shaded grey.

**Interloper contamination**

While we have thus far been assuming a perfect LIM map and galaxy catalog, with only uncorrelated noise being added to the map, this is unlikely to be the case in a real intensity mapping situation. Potential sources of contamination to the stack include foreground or background interloper emission being present in the LIM map (for example, the COMAP data cubes will contain CO(2-1) emission from galaxies at $6 < z < 8$, which is expected to be roughly an order of magnitude fainter than the target CO(1-0) signal; Breysse et al. 2022, Chung et al. 2024a), and interloper galaxies being present in the galaxy catalog (due to either noise peaks or [OII]-emitting foreground galaxies). We explore both cases here.

**LIM interlopers**

In order to test the robustness of the stacking pipeline against background and foreground interloper line emission being included in the LIM maps, we add in simulated background line emission following the methodology outlined in §3.3. For each of nine different LIM map realizations, we generate 11 different interloper iterations. We then perform stacks on the maps with interloper emission present, at 1000%, 100%, 10%, and 1% of the target emission, and compare the output luminosity to the luminosity without interlopers. We repeat this analysis for stacks with 100 and 10 000 objects, to see if a larger galaxy catalog helps to mitigate the effects of spectral line contamination. The results are shown in Figure 3.19.

We find that the stack should be robust against the presence of faint spectral line interloper emission in the LIM map, although it is more affected by extremely bright contamination. Interloper emission has two effects on the maps: it biases the stacked luminosity towards higher values (as it is functionally adding more line signal into the maps, although this signal is not correlated with the actual population of galaxies the stack is aiming to probe), and it introduces scatter, as the catalog objects will randomly be associated with regions of stronger or fainter interloper emission. On average, the maps with interlopers output the same values as the maps without: at the level of contamination expected in COMAP (roughly 10%), the average deviation from the uncontaminated stack values is less than 0.5% for all catalog sizes. Even in the most extreme case, where the strength of the interlopers is $10\times$ that of the target line, the average deviation is 37% in the 100-object stack, but only 4% in the 10 000-object stack, showing that larger galaxy catalogs are more robust against interloper emission.

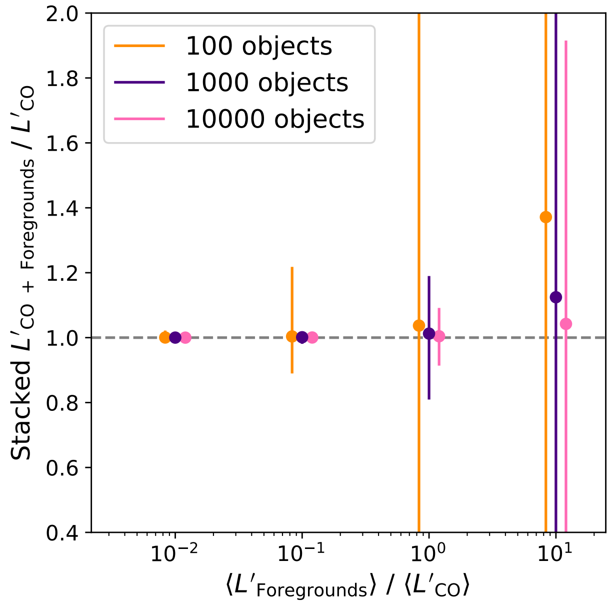


Figure 3.19: Average ratio of stack luminosity with added spectral line interlopers to the stack luminosity without interlopers for various scalings of the interloper emission. For each scaling, a CDF was created for the ratios from different simulation realizations, and the 95% confidence interval is shown.

The amount of scatter added to the maps increases as the interlopers become brighter. At the expected COMAP interloper strength, the (95%) scatter is 20% (2%, 0.9%) for the stack with 100 (1000, 10 000) objects. This increases to 120% (19%, 9%) for the same stacks when the interloper is of equal brightness to that of the target line, and more than 100% for the stacks on all catalogs except the 10 000-object catalog when the interloper is 10× brighter than the target emission. The scatter is not solely to higher luminosity values because the maps are mean-subtracted – adding extra emission to the maps will bring the overall mean up, so regions with few interloper galaxies will have their CO luminosity brought down by the mean subtraction. If the stacked catalog is more associated with voids in the interloper map, the stacked luminosity will be brought down. We note that this analysis involves no attempt whatsoever to mitigate interloper contamination—using a map-based strategy to handle contamination in LIM maps (e.g., Sun et al., 2018; Karoumpis et al., 2024) would likely reduce this scatter.

### Catalogue false positives

As with the intensity map, galaxy catalogs (especially those based on a single spectral line) may include some objects that are not galaxies at the target redshift emitting the target spectral line, but come from other spectral lines being emitted at other redshifts, or more rarely noise peaks being picked up by the signal-detection algorithm. HETDEX, for example, is expected to contain some population of $z < 0.5$ [O II] emitters misidentified as LAEs. These false positive detections will serve to bring down the overall signal in the stack, as each false positive object will add only noise to the stack. From this, one would expect the signal in a stack with a fraction $f_{\rm FP}$ of false positive objects to be attenuated by $f_{\rm FP}$.

In order to test this prediction, we stacked on catalogs where 0, 10, 20, 30, and 40% of the catalog was replaced with false positive (random) values (see §3.3). The results of these stacks are shown in Figure 3.20. We find that the attenuation of the stack due to the inclusion of false positive detections in the catalog agrees well with the theoretical expectation, with some scatter. The scatter is slightly higher above the theoretical expectation, suggesting that this scatter may come from the random 'false positive' catalog entries falling onto brighter regions (regions of higher signal) of the map.

## 3.6 Discussion

### Ideal experimental setup

We list the results of the various tests of the LIM and galaxy catalog experimental parameters in Table 3.3. We find that the main consideration when selecting a galaxy catalog for stacking is that the catalog should have as many objects as possible. This drives down the noise in the stack by $\sqrt{N_{\rm obj}}$, but also helps to protect the stack against interloper emission in the LIM map. The uncertainty in the catalog redshifts is also important, as any uncertainty spreads the signal over more noise and reduces the stack S/N ratio. Redshift uncertainties up to $\Delta_{\rm v} \sim 1000$ km s$^{-1}$ may be tractable for smaller catalogs. High-precision photometric catalogs may also be useable, provided the large catalog redshift uncertainty is offset by a significant increase in the number of catalog objects across the LIM experiment's footprint.

The masses of the DM halos selected by the galaxy catalog are also very important—a galaxy catalog that selects high-mass halos can improve the S/N ratio of the stack (when compared to a less-biased, Ly$\alpha$-like catalog) more than any other single factor. We also note that galaxy surveys targeting different galaxy types could also

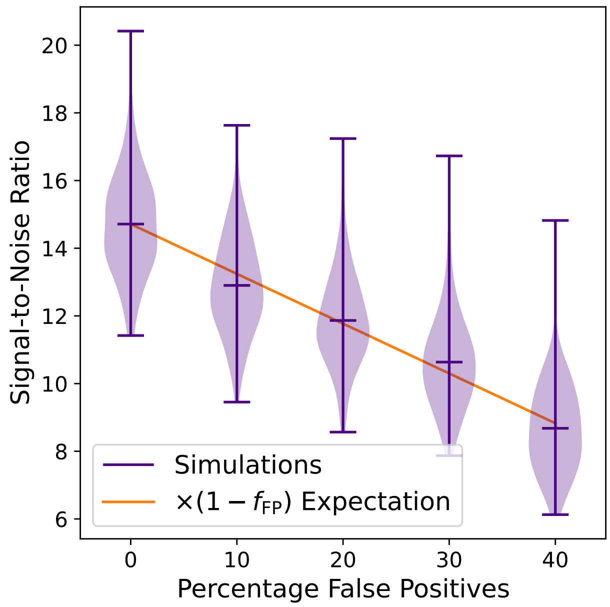


Figure 3.20: Signal-to-noise ratio of the stack as a function of the percentage of 'false positive' interlopers included in the galaxy catalog. Plotted in orange is the theoretical expectation, assuming the value of the map is identically zero at the (random) location of each false positive.

Table 3.3: Summary of optimal parameters for the stack S/N ratio.

| Parameter | Best value |
|---|---|
| *Stack parameters* | |
| $N_{\text{spax}}$ | 7 |
| $N_{\text{chan}}$ | 3 |
| $N_{\text{obj}}$ | As high as possible |
| $\Delta_z$ | 0 km/s |
| *LIM experimental parameters* | |
| $\delta_\nu$ | Sampling stack spatial extent |
| $\delta_x$ | Sampling stack spectral extent |
| $\theta_{\text{beam}}$ | As small as possible |
| *Astrophysical factors* | |
| CO model | L16 or C22 |
| Ly$\alpha$ model | bright |
| $\rho$ | 1.0 |
| Line broadening | $v_{\text{m,peak}}$ |
| *Astrophysical factors* | |
| Interloper emission | Tolerable to 10% |
| False positives | Tolerable |

be combined to maximize the collective $N_{\mathrm{obj}}$. The expected output signal would then become an average of the expected signal from each survey tracer, weighted by the number of objects included in each. In this way it may be possible to combine the benefits of, for example, a LAE survey with high $N_{\mathrm{obj}}$ and a more biased quasar survey with fewer objects. This may be especially valuable in the current LIM regime, where experiments are primarily searching for a detection, and thus the S/N is the primary concern.

Similarly to increasing the number of objects in the catalog, the simplest way to adjust the LIM data to improve the sensitivity of the stack is to decrease the RMS noise in the data, through increased integration time or decreased $T_{\mathrm{sys}}$. This is not something we pursue in this work. Instead we explore the resolution and pixelisation of the map, given a fixed integration time. We find that the stack becomes less viable when the channel width of the LIM data is wider than the FWHM of the stacked spectral line—LIM experiments with poorer spectral resolution may find stacking more difficult. Marginal improvements are also possible through having smaller pixel sizes in each axis of the LIM data, primarily because the stacked objects are able to be lined up more precisely when the central pixel in the stack is smaller.

As an additional point of consideration, most LIM experiments are designed to optimize for the auto-correlation power spectrum. None of the stack-driven experimental preferences that we list above will negatively impact power spectrum analysis, particularly at the high-$k$ end also probed by the stack. Improving the resolution, both spatially and spectrally, will increase the number of accessible $k$-modes without changing the noise power spectrum (the increased noise per voxel will be offset by the increased number of voxels, Li et al. 2016), thus improving the overall power spectrum signal-to-noise. In both cases, the primary constraints preventing this improved resolution are instrumental.

Finally, we explore the effects of various $L(M_{\mathrm{h}})$ models for the LIM tracer spectral line. While we framed this in the context of the largely unconstrained CO model space, this could equally be used as an exploration of which LIM tracers are most ideal for a stacking analysis. While a brighter tracer will automatically improve the stack luminosity—any bulk increase to the tracer luminosity will result in a commensurate increase in the stack luminosity—we also test models that have varying luminosity distributions across mass. We find that a tracer that is bright in the high-mass end (e.g., C22) or a tracer that is bright in the low-mass end (e.g., L16) each give greater S/N ratios than one that is a more moderate luminosity across all masses. If selecting a LIM tracer specifically for a stacking analysis, this should be the primary consideration—we find that the correlation on an individual halo level between the LIM tracer luminosity and the galaxy catalog luminosity is not nearly as important. In order to explain these two effects, we must look in more detail at the halos contributing to the stack.

**Origin of the stack luminosity**

In order to understand the behavior of the stack with parameters such as the CO $L(M_{\rm h})$ function or the correlation between the two tracers' scatter in luminosity, we explore the underlying distribution of halos, and how they drive the returned stack luminosity. Firstly, we explore where the DM halos that are contributing to the stack are located. We do this by generating a 3D histogram of the positions of all DM halos (without performing any cuts on $M_{\rm h}$, $L_{\rm CO}$, or $L_{\rm Ly\alpha}$, or weighting by any of the above) in a simulated stack run performed using default parameters. We then cut out the regions surrounding the DM halos included in the mock galaxy catalog. We sum over the central three frequency channels (as is done when generating the actual 2D images and spatial profiles), yielding the DM halo density per spaxel, $n_{\rm halo}$, in the 2D image. We then average these halo density profiles across the cutouts associated with all 1000 catalog objects included in the stack. The resulting distribution is shown in Figure 3.21.

We find that on average 217 halos are contributing to the measured $\langle L_{\rm LIM}\rangle$ across the stack aperture. In the central voxel alone, there are, on average, ten halos contributing. There is also a fairly high baseline number of halos, with $\sim 20$ halos present even at the outskirts of the spatial profiles. This is consistent with the fact that the spatial profiles show luminosity extending nearly the full width of the stacked cubelet, even in the mean-subtracted simulated maps (we note that the halo number counts in Figure 3.21 are not mean-subtracted in any way, unlike the actual mock data cubes). At least with the COMAP experimental parameters, we are operating on a broad enough resolution that DM halos can almost never be resolved individually.

We also extract a 1D spatial profile from the $\langle n_{\rm halo}\rangle$ map, and fit this profile with a two-Gaussian fit as is done for the actual simulated data in Figure 3.7. The narrow and broad components of the profile fit have standard deviations of $1.7' \pm 0.1'$ and $6.6' \pm 0.4'$ respectively. We note that the values of $3.2'$ and $9.6'$ reported for the full simulations are convolved with the $4.5'$ beam, where this $\langle n_{\rm halo}\rangle$ profile is not—

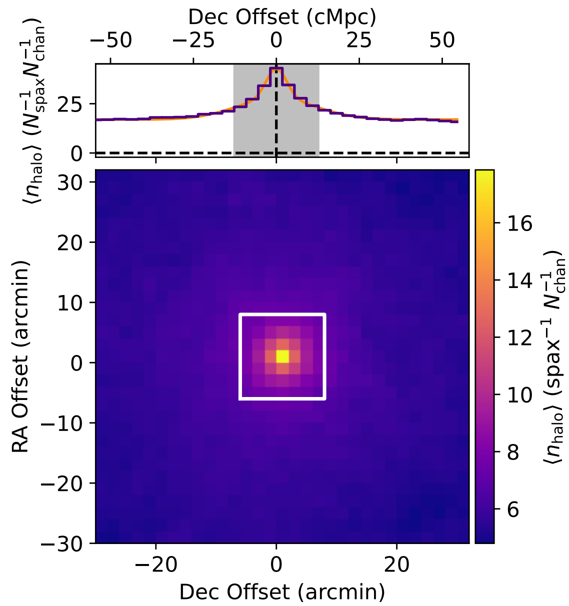


Figure 3.21: The spatial distribution of DM halos in the stack. *Bottom:* Number of DM halos included in each spaxel of the 2D spatial profile of the stack (*i.e.*, summed across the $N_{\rm chan} = 3$ frequency channels of the stack aperture), averaged across cutouts. The actual cataloged DM halo is included in this count, but is outnumbered by more than an order of magnitude in the central spaxel alone. There are a significant number of DM halos in spaxels extending beyond the central spatial aperture. *Top:* Spatial profile generated by integrating the $\langle n_{\rm halo}\rangle$ map across the three channels making up the stack aperture in the RA direction. Unlike the regular simulations, this is not mean-subtracted, and thus the baseline number of halos is high. A double-Gaussian fit to this profile is shown in orange.

after deconvolving from the beam, the widths of the two Gaussian components in the full simulated stacks are $2.5'$ and $9.4'$. The broad component agrees within error between the two profiles, but the narrow component is narrower in the $\langle n_{\rm halo}\rangle$ stack profile than the full $L_{\rm LIM}$ version. This could be due to halos at radii slightly offset from the central objects being less frequent but brighter in CO.

**What the stack reveals about the galaxies contributing to the signal**

Though we have found that the stack signal originates from the large-scale clustering of galaxies, the extent of that clustering (and the centering of the stack objects in relation to it) will depend on the CO and Ly$\alpha$ luminosity properties of the galaxies being stacked. As we have found, the output stack luminosity depends heavily on the mass (and thus the bias) of the halos being targeted, and also depends on the CO luminosity at both the low- and high-mass ends of the DM mass function.

As such, we look into the luminosity contribution to the stack from both the cataloged objects themselves and their neighbors. We do this using the simulated catalogs of halo masses and CO luminosities, before generating any mock observations. We isolate two populations of halos that contribute to the stack. Firstly, we take the DM halos actually selected by the galaxy survey (via their Ly$\alpha$ luminosity) to be stacked on. The CO luminosity of each of these halos corresponds to the first term in Equation 3.17. Secondly, we took all of the halos that are not specifically selected by the galaxy survey but will still be counted in the stack due to their proximity to a selected halo. Practically, this amounts to all halos within 3.5 pixels along any axis of the LIM map. The contribution from these halos (the second term in Equation 3.17) will be the sum over the $L'_{\rm CO}$ of all neighbors of a given selected halo.

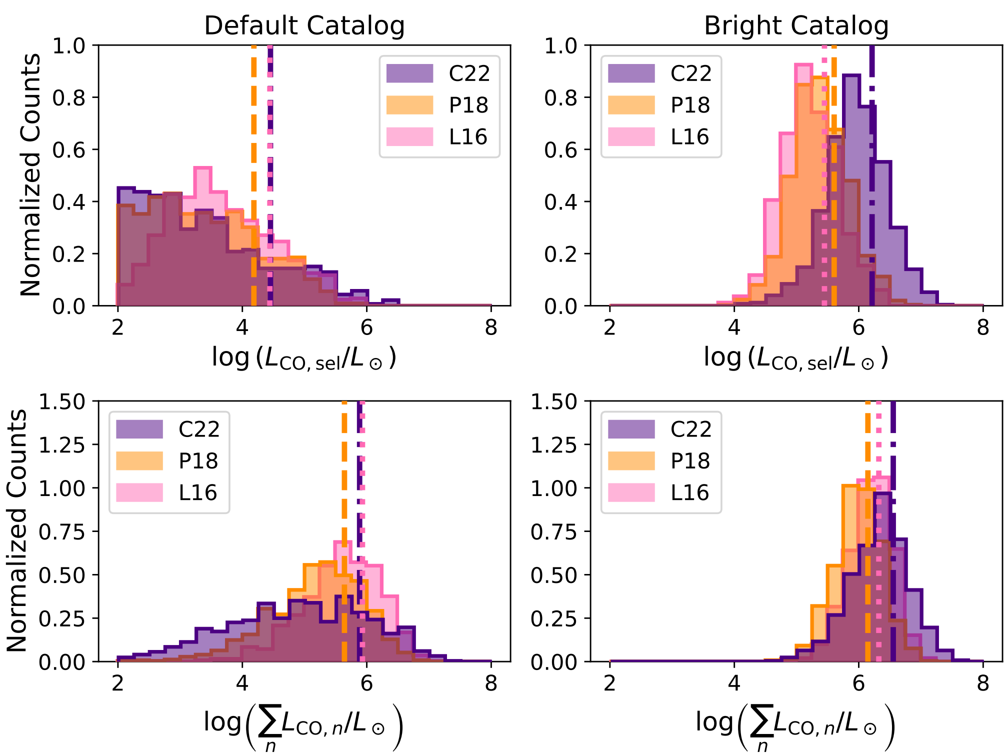


Figure 3.22: Comparison of the terms contributing to the overall stack luminosity for a single simulation realization using the three different CO prescriptions. The top panel shows histograms of the $L_{\rm CO}$ of each DM halo included in the galaxy catalog and thus directly selected to be included in the stack. The bottom panel shows the summed $L_{\rm CO}$ of all the halos neighboring a given selected halo – these halos also fall into the stack aperture and thus contribute to the stack as well. In both panels, the (linear) averages are shown as vertical lines.

For a single stack realization, we plot a histogram of the $L_{\mathrm{CO}}$ values of each halo in the first population, and a histogram of the sum over the neighboring halos of a given selected halo in Figure 3.22. We repeat this analysis for each of the CO luminosity prescriptions. We also plot the (linear) averages over all apertures included in the stack as vertical lines, keeping to the default Ly$\alpha$ catalog model. The first, and most important, thing that we note from this analysis is that the stack luminosity is almost entirely driven by the neighbors of the halo selected by the galaxy catalog, rather than the selected halo itself. For the default catalog model, the integrated luminosity of the neighboring halos for a given cutout is on average nearly two orders of magnitude greater than the luminosity of the selected halo. This agrees with the finding that there are on average more than 200 other halos contributing to the stack luminosity (Figure 3.21), and explains the fact that the stack signal is more extended in all three dimensions than a point-source contribution (§3.5).

We find that for both the neighboring halos and the selected halo itself, the value predicted for P18 is lower than either of the other two prescriptions. In the case of the first term, this seems to be because the C22 value is driven by contributions from extremely bright high-mass values, and the L16 value, though its high-mass halos are not as bright, has many brighter low-mass halos. The P18 model has neither particularly bright high-mass halos nor low-mass halos. This also seems to be true of the second term, which is primarily a sum over lower-mass halos: the L16 model is brighter on average ($\sum_n L_{\mathrm{CO},n}$ is peaked high, with narrow spread), and the C22 model has sufficient bright contributors to drive its luminosity up (its spread in $\sum_n L_{\mathrm{CO}.n}$ is much larger, but it contains the brightest outliers).

As a comparison point, we repeat this exercise for the bright, AGN-like galaxy catalog, that selects more massive halos. This leads to two separate qualitative changes—the selected halos are considerably more massive and thus brighter (the average luminosity increases by almost two orders of magnitude for each CO model), and they are more biased, so they have more (and brighter) neighbors. In this case, the contribution from neighboring halos is still greater than the contribution from the cataloged halos themselves for all models, but the contributions are much more comparable. We do not show an equivalent test for the faint catalog model here, as its behavior is extremely similar to that of the default model.

The variation between the CO models in average luminosity of the selected halos follows from the fact that more massive halos are being selected. The C22 model is brightest at high halo mass, followed by P18 and then L16 (Figure 3.3). This is the same distribution seen in the selected halos. The average contribution from the neighboring halos to the overall luminosity does not change as dramatically as the selected halo contribution from the default Ly$\alpha$ catalog to the brighter one, but they are still more luminous for each CO model. Additionally, the distribution between neighbors of selected halos for a given model is much tighter. This is likely because more of the neighboring halos are present due to actual clustering and not random chance. The C22 model also dominates considerably more in this case, probably because the more massive, more biased selected halos are more likely to have massive neighbors, which will be very bright in the C22 model.

The origin of the stack signal also explains the behavior of the stack with varying $\rho$, seen in Figure 3.17 (§3.5). Variations in $\rho$ do not strongly impact the output stack luminosity, whereas variations in the catalog luminosity function do, suggesting $S(L_{\mathrm{Ly}\alpha}(M_h, \rho))$ is mostly driven by $L_{\mathrm{Ly}\alpha}$ rather than $\rho$. This follows from the fact that the neighboring halos in the stack have a greater contribution to the total stack luminosity than the selected halos themselves for the default catalog model – a direct anti-correlation between the scatter in the CO and Ly$\alpha$ luminosities would mean the brightest Ly$\alpha$ halos have lower CO luminosities, so the halos directly selected by the galaxy catalog would be fainter overall, but the neighboring halos have a variety of luminosities in both tracers. It is possible that the small variations come from the selected halos being slightly less massive (and thus less biased) on average when $\rho$ is negative.

Additionally, we found that varying $\rho$ has proportionally more of an effect on the output stack luminosity for the more stochastic, top-heavy 'bright' catalog luminosity function than for the flatter default one. Figure 3.22 shows that, when the bright catalog luminosity function is used, the actual selected halos are contributing to the overall stack luminosity almost equally to the neighboring halos. Correlating or anti-correlating the scatter will thus have a larger effect on the overall stack, as these selected halos are where the correlation matters.

**The stack in comparison to other summary statistics**

Using the survey parameters that we default to here (which were intended to simulate realistic small-scale CO and Ly$\alpha$ distributions while still achieving a solid stack detection with as few catalog objects as possible), the mean stack detection S/N is 14.5. For the same parameters, the expected detection S/N ratio (*i.e.*, ignoring sample variance, enabling a direct comparison) of the cross-spectrum over all $k$ is

18.7. This is roughly 1.3× the stack S/N ratio. We note for completeness that the auto-spectrum S/N ratio is 53.4 across all $k$-modes, although the parameters we use here have far too few catalog objects spread across too wide an area to allow for a fair comparison between the auto- and cross-spectrum analyses.

It is expected that the cross-spectrum be more significant than the stack—the stack is probing the same signal as the smaller-scale modes ($k \lesssim 0.25$) of the cross-spectrum, but in real space rather than Fourier space. The cross-spectrum is sensitive to both these small-scale modes and the larger modes that the stack is not and thus has access to information that the stack does not. Additionally, this implies that any of the optimizations for stack signal we have discussed in this work (whether astrophysical or experimental) will also yield the strongest cross-correlation signal possible at small scales. Thus, in terms of raw statistics, one might expect to see a cross-spectrum detection before a stack detection. However, as we discuss below, practical analysis considerations mean the ordering may not be so clear-cut.

In addition to increased significance, there are analysis benefits to having information over multiple spatial scales—many of the astrophysical and cosmological parameters that may be extricable from a cross-correlation (e.g., Pullen et al., 2013b; Breysse and Alexandroff, 2019) are extremely degenerate in the stack. More detailed comparisons between the stack and the cross-power spectrum are outside of the scope of this work, but we imagine it would be very difficult to extract more information than this overall amplitude from a stacking analysis.

While the cross-spectrum is likely a more valuable analysis tool in an idealized case, in practice LIM experiments are subject to many real-world complications that will likely be more difficult to disentangle from the cross-spectrum than the real-space stack. These include effects such as on-sky survey asymmetries in either the galaxy catalog or LIM survey (HETDEX, for example, has an extremely complicated window function, Gebhardt et al. 2021), systematic errors (ground pickup, standing waves, etc., all of which preferentially appear on the larger scales that the cross-spectrum is more sensitive to), and, particularly in the case of 21 cm experiments, continuum foregrounds. None of these complications are insurmountable in the cross-power spectrum, and there exists a wealth of literature already on addressing them (e.g., Keenan et al., 2022; Lunde et al., 2024; MeerKLASS Collaboration et al., 2025). Still, they are likely easier to address in the stack. This has recently been shown in Chen et al., 2025, where the authors were able to directly isolate and identify a specific source of systematic error in the MeerKLASS 21 cm data using their stacking analysis. In a field where we are still searching for first detections, the simplicity of the stack is a major advantage, and the stack and the cross-spectrum have the potential to work very well together. We will explore the complementarity of these two analyses in more detail in future work.

### 3.7 Summary and conclusions

In this work, we have developed a simulation framework for multi-tracer LIM analyses and applied it to three-dimensional stacking of LIM data on the positions of objects from a resolved galaxy catalog. We find that although stacking is conceptually simple, many degenerate parameters may affect both the stack signal and its noise. These parameters could come from the galaxy catalog or the LIM data itself and could be experimental or astrophysical in origin. We posed three framing questions in the introduction to this work, and we aim to answer them here.

1. What experimental design is optimal for a stacking analysis? We find that the optimal galaxy catalog for stacking is a spectroscopic one that (a) has a maximum number of objects included and (b) targets high-mass DM halos. The ideal LIM experiment has a spectral resolution that is as high as possible (it must at least sample the stacked linewidth) and uses a spectral line tracer that is bright across the entire halo mass range, particularly at high halo masses. Other factors such as the spatial resolution of the LIM experiment or the correlation between the two spectral line tracers on an individual halo level may also affect the stack S/N ratio, although only marginally. The stack is quite resilient to faint ($\sim 10\%$) interloper emission in the LIM data.

2. Where does the signal in a stacking analysis originate? For all choices of the galaxy catalog luminosity function, the majority of the signal in the stack is not from the galaxies actually included in the catalog itself. Instead, the signal is dominated by the integrated luminosity of the DM halos neighboring the cataloged galaxies. This is true by more than an order of magnitude for our 'default' galaxy catalog luminosity function and by a factor of a few for a 'bright' AGN-like luminosity function. This makes the stack very useful as a tool for LIM, where the goal is to measure galaxy properties and cosmology via aggregated galaxy emission.

3. What can the stack tell us about the properties of the galaxies contributing to the signal? A galaxy catalog targeting more massive biased halos leads to an

increase in the stack signal. The stack signal is also improved by increased correlation between the luminosities of the two tracers being used (on a single-halo level) and brighter CO emission on either the high- or low-mass end of the $L(M)$ function. However, all of these parameters are degenerate, and all other parameters would need to be known in order to constrain any single parameter.

As much of the field is still in a detection phase, we have framed this work only in terms of maximising the stack S/N ratio. We believe that the stack will be a useful tool for detection of the LIM signal, as the stack is easy to compute and has all the benefits of a joint analysis. It will be particularly valuable as an accompaniment to an auto-correlation detection based on the same LIM dataset, validating the origin of this signal in high-redshift galaxies. The stack will also be very useful in concert with a cross-correlation of LIM data with galaxy catalogs. Once a detection is made, characterising each of the astrophysical factors explored here will become a major goal of LIM surveys. The output of the stack is primarily a single data point, and we have shown that many of these astrophysical factors will have degenerate effects on that data point. Cross-correlations are sensitive to more spatial scales than the stack and thus allow for a more detailed comparison with models. However, cross-correlation analyses are more sensitive to systematic errors and may in turn rely on stacking analyses to constrain those errors. We will explore the interplay between these other analyses in future work. Ultimately, stacking is simple and robust, and it has the potential to be a useful tool in the LIM toolbox.

### 3.8 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’. HP’s research is supported by the Swiss National Science Foundation via Ambizione Grant PZ00P2_179934. JK was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00340759). JG acknowledges support from the Keck Institute for Space Science, NSF AST-1517108 and University of Miami.

## 3.9 Appendix: Modelling of astrophysical line broadening

In Section 3.3, we selected from several different theoretical models for CO line broadening at $z = 3$ without much justification. Here, we attempt to provide a basis for our choice of the $v_{M_{\rm peak}}$ method from Behroozi et al., 2019 through comparison with observations of CO-emitting galaxies. Ideally this comparison would be made directly with observations of CO(1–0) at $z = 3$, but these observations are so limited that we include other CO transitions and lower redshifts in order to collect a larger sample. This yields an extremely heterogeneous collection of galaxies. We will aim to explore, and correct for, the effects of redshift and $J$ in future works.

In total, we collected 119 high-z CO-detected galaxies from various sources:

- ASPECS: We include 16 galaxies from the ALMA Spectroscopic Survey in the HUDF (ASPECS, Aravena et al., 2019; González-López et al., 2019; Boogaard et al., 2019; Boogaard et al., 2020). This was a blind search for CO emission in ALMA's Band 3, so there is some confusion between CO transitions. The bulk of these objects are likely CO(2–1) emission from $z \sim 1$ and CO(3–2) emission from $z \sim 2.5$. As we are only interested in the line FWHM for this application, we make no cuts on the confidence of the line identification. Line widths are reported directly as FWHM values in km s$^{-1}$.

- VLAASPECS: This is an extension of the original ASPECS program using the Very Large Array (Riechers et al., 2020). We include the three confident CO(1–0) detections at $z \sim 2.5$ from this survey in our sample. Again, linewidths are reported directly as FWHM values in km s$^{-1}$.

- PHIBSS: We include 53 galaxies from the IRAM Plateau de Bure high-z blue sequence CO(3-2) survey (PHIBSS, Tacconi et al., 2013). These are all CO(3–2) detections, at $z \sim 1$, and are primarily massive, main-sequence star-forming galaxies. The linewidths for these galaxies have been converted to circular velocities ($v_{\rm c}$), so we undo this conversion to return the observed line FWHM values. This is done in two different ways for different galaxy types – galaxies that have measured disc-like morphologies from HST and significant CO velocity gradients have been converted into $v_{\rm c}$ using the equation $v_{\rm c} = 1.3(\Delta v_{\rm blue-red}/2\sin i)$, and galaxies that have measured disc-like morphologies but no strong CO velocity gradient, or those with irregular morphologies, have been converted using $v_{\rm c} = \sqrt{3/8\ln 2}\Delta v_{\rm FWHM}$.

- PHIBSS2: We include an additional 43 galaxies from the second iteration of PHIBSS (PHIBSS2, Freundlich et al., 2019; Lenkić et al., 2020), including both targeted and serendipitous detections. These galaxies include CO(1–0) emitters $0.5 < z < 0.8$, CO(2–1) emitters $z \sim 1$, and some (16) higher-redshift, higher-$J$ CO emitters. The best-fit line FWHM is reported directly here.

- COLDz: We include the four CO(1–0) detections from the CO Luminosity Density at high-z (COLDz, Pavesi et al., 2018; Riechers et al., 2019) survey, all at $z \sim 2.5$. The best-fit line FWHM is reported directly.

The FWHM for each of these galaxies is plotted in Figure 3.23. They have an average linewidth of 385 km s$^{-1}$, with a long tail towards broader linewidths.

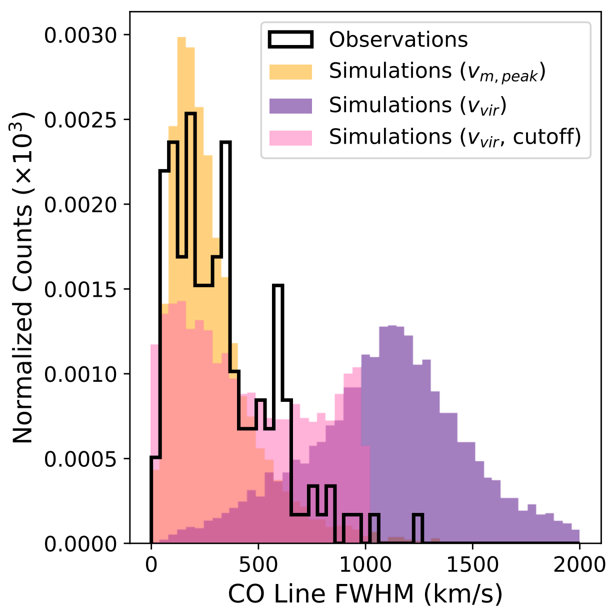


Figure 3.23: Relative occurrence of different CO line FWHM values across the simulated DM halo population for each of the line broadening prescriptions outlined in §3.3. These are compared to observations (black line) of galaxies emitting in various CO lines up to $z = 3.5$.

Also in Figure 3.23, we plot the calculated line FWHM for each of the simulated galaxies included in the stack (*i.e.*, all galaxies with $L_{\mathrm{Ly}\alpha}$ above the cut-off value for observation, $L_{\mathrm{Ly}\alpha,\mathrm{cut}} = 8.26 \times 10^9\ L_\odot$) for each of the three line broadening prescriptions described in §3.3. We note that these are not the only galaxies included in the simulated stack, as the neighbors of these galaxies that are near enough to fall into the stack aperture will also be averaged in. However, the selected halos are likely to be more massive than average and thus will have greater linewidths under any of the broadening prescriptions. They will thus dominate the stack linewidth, and are the relevant halos for this analysis.

We find that the $v_{M_{\rm peak}}$ prescription yields a distribution of line widths that matches the observations quite well. The $v_{\rm vir}$ prescription provides a distribution that peaks significantly higher than observations, and extends to much higher linewidths than are observed. Adding a 1000 km s$^{-1}$ cut-off to the $v_{\rm vir}$ prescription does shift the average down to match more closely with observations, but the shape of the distribution is much more uniform, especially at high FWHM, than is actually observed.

## 3.10 Appendix: Model-dependence of the stack spatial extent

As outlined in §3.5 and §3.5, the extent of the signal in the stack cannot be explained solely by point sources at the location of the objects in the galaxy catalog. This is particularly interesting in the angular axes, where the only broadening factors are beam smoothing and astrophysical clustering. If the large-scale signal visible in the stacks is indeed coming from clustering effects, the extent of this signal is dependent on both the CO and Ly$\alpha$ model chosen. While it is unlikely that the sensitivity of current and upcoming LIM experiments will allow for measurement of this spatial extent in detail in the near future (thus making a full exploration of its model-dependence outside the scope of this work), there will still be cosmological information in this distribution that may be exploitable in the longer-term.

Table 3.4: Standard deviation values from double-Gaussian fits to spatial profiles of stacks generated using different model configurations. The value $\sigma_{\rm narrow}$ always refers to the narrower component and $\sigma_{\rm broad}$ to the broader component.

| CO model | Ly$\alpha$ model | $\sigma_{\rm narrow}$ (arcmin) | $\sigma_{\rm narrow}$ (Mpc) | $\sigma_{\rm broad}$ (arcmin) | $\sigma_{\rm broad}$ (Mpc) |
|---|---|---|---|---|---|
| C22 | Default | $2.47 \pm 0.10$ | $4.5 \pm 0.2$ | $9.2 \pm 0.4$ | $16.7 \pm 0.8$ |
| P18 | Default | $2.59 \pm 0.07$ | $4.5 \pm 0.2$ | $10.2 \pm 0.4$ | $18.4 \pm 0.7$ |
| L16 | Default | $2.53 \pm 0.07$ | $4.7 \pm 0.1$ | $9.9 \pm 0.4$ | $18.0 \pm 0.6$ |
| C22 | bright | $1.77 \pm 0.04$ | $3.23 \pm 0.07$ | $8.4 \pm 0.3$ | $15.3 \pm 0.6$ |
| C22 | faint | $2.44 \pm 0.07$ | $4.4 \pm 0.1$ | $11.5 \pm 0.5$ | $20.8 \pm 1.0$ |

Thus, we performed an initial characterisation by fitting two-Gaussian profiles to the spatial profiles of simulated stacks following §3.5. For each stack exploring a different CO model, we adopt the default Ly$\alpha$ model. In those stacks testing different

Ly$\alpha$ models, we adopt the fiducial C22 CO model. We also plot the average spatial profile generated from a set of nine catalog-only simulation runs for each model. This is done following §3.5—we set the $L_{\mathrm{CO}}$ of every halo not included in the galaxy catalog to zero, to show the contribution from the cataloged halos themselves. The results of this are shown in Figure 3.24. For most models, the contribution from the cataloged halos themselves is nearly negligible compared to the rest of the field.

The standard deviation values returned from the two-Gaussian fits are tabulated in Table 3.4. Unlike in §3.5 above, we deconvolve out the FWHM = 4.5′ beam from the standard deviation values before reporting them. We find that changing the CO model does not significantly change the spatial extent of the stack signal, while changing the Ly$\alpha$ model does result in some significant changes. This is because altering the CO model does not alter the DM halo population underlying the stack signal, as this is determined by which halos are included in the catalog, which is set only by the Ly$\alpha$ luminosity (through $S(L_{\mathrm{Ly}\alpha}(M_{h,i}, \rho))$, of the halos following Equation 3.17). Thus changing the Ly$\alpha$ prescription used to generate the simulated catalog alters the underlying population of DM halos behind the stack signal.

The standard deviation of the narrower Gaussian profile, $\sigma_{\mathrm{narrow}}$, corresponds to roughly 4.5 Mpc for most models. The only notable outlier is the AGN-like ('bright') Ly$\alpha$ model, which has a significantly smaller $\sigma_{\mathrm{narrow}}$. This is likely because the bright Ly$\alpha$ model has the most contribution from the actual cataloged objects (Figure 3.24), so its narrow Gaussian component is the closest to a point-source fit. The broad component's standard deviation $\sigma_{\mathrm{broad}}$ is significantly offset from the default for both the bright and the narrow Ly$\alpha$ models. The bright model again has a narrower broad component than the default model, possibly because it is more heavily clustered. In contrast, the faint Ly$\alpha$ model has a much wider spatial extent than the default configuration. The stacks on catalogs generated using the faint model are centered around lower-mass halos than either of the other catalog models' stacks. These halos are less biased, and thus stacks of their surroundings show less clustering than other models.

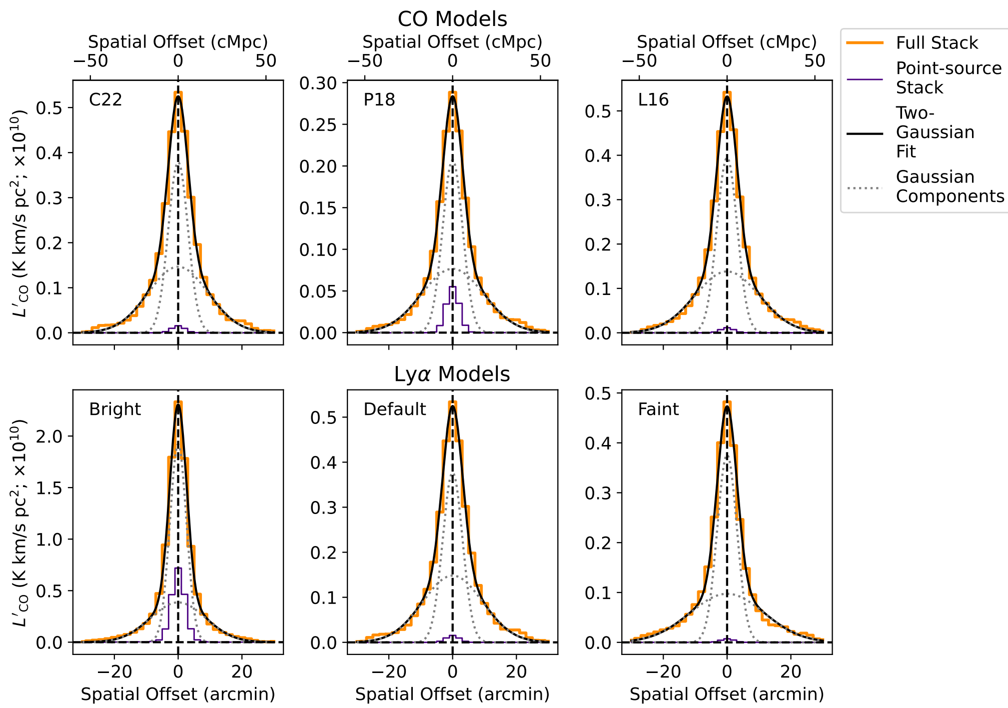


Figure 3.24: One-dimensional spatial profiles of stacks performed on simulations generated using various CO (left) and Ly$\alpha$ (right) models. For each profile, we plot the stack on the full simulated map (the orange lines) as well as a stack performed on a simulated map that contains only the DM halos included in the galaxy catalog being stacked on (so there is no contribution at all from neighboring halos; the indigo lines). We also plot a two-Gaussian fit (black lines) as well as the two Gaussian components individually (grey dotted lines).

*Chapter 4*

# COMAP PATHFINDER – SEASON 2 RESULTS IV. A STACK ON EBOSS/DESI QUASARS

Dunne, Delaney A. et al. (Mar. 2026). “COMAP Pathfinder – Season 2 results: IV. A stack on eBOSS/DESI quasars.” In: *Astronomy & Astrophysics* 707, A256, A256. DOI: 10.1051/0004-6361/202557864. arXiv: 2510.23568 [astro-ph.CO].

## 4.1 Chapter 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 $10.0 \times 10^{10}$ K km s$^{-1}$ pc$^2$ at 95% confidence within an $\sim$ 18 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).

## 4.2 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 and 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., 1988a; 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 cataloged 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; Keating et al., 2020; CONCERTO Collaboration et al., 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 et al., 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 catalogs to improve sensitivity and to mitigate systematic error; as major systematic errors are primarily unique to one survey (e.g., Fronenberg and 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 et al., 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 catalogs 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 (Breysse and 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., 2024b), 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 et al., 2025a). 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 §4.3, and describe the simulated LIM and galaxy catalog data to which we compare the stack in §4.4. The stack methodology is described in §4.5. In §4.6 we present the stack results, to which we compare models in §4.7. Throughout this work, we take the nine-year WMAP cosmology (Hinshaw et al., 2013) for consistency with the peak-patch simulations used in §4.4 and other COMAP works. This is a ΛCDM model with $\Omega_\mathrm{m} = 0.286$, $\Omega_\Lambda = 0.714$, $\Omega_\mathrm{b} = 0.047$, and $H_0 = 100h$ km s$^{-1}$ Mpc$^{-1}$ with $h = 0.7$. We assume base-10 logarithms unless otherwise stated.

## 4.3 Data

### 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 continuing to observe. Thus far, two COMAP observing seasons have taken place. Season 1 encompassed 5,200 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., 2024b), 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 $\mu$K (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-neighbor 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 §4.5. 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 neighbors 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%.

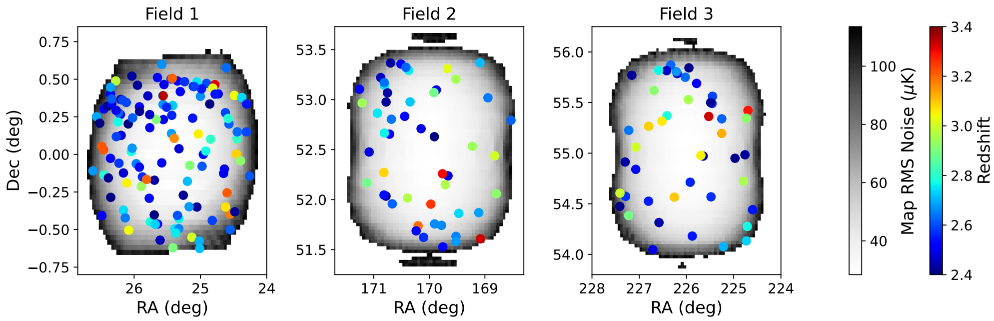


Figure 4.1: Spatial distribution of the eBOSS/DESI quasars compared to the footprint of the COMAP data. The COMAP footprint is colored by the RMS in each voxel, averaged across all spectral channels, and the quasars are colored by their redshift.

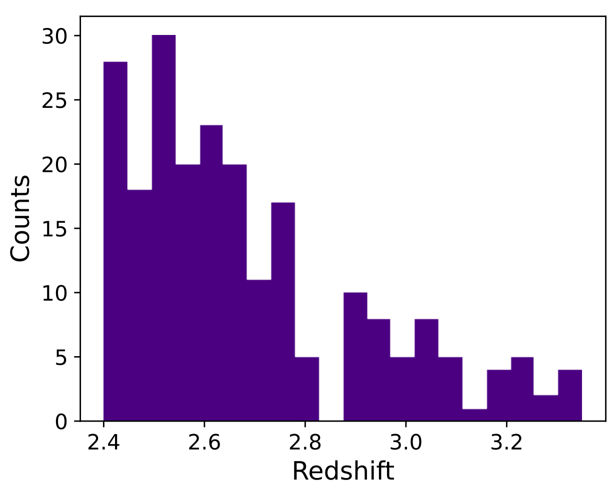


Figure 4.2: Redshift distribution of the eBOSS and DESI quasars included in the stack catalog.

### eBOSS and DESI

We performed the stack on a combined catalog 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 et al., 2025a). 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 catalog reduction. We took DESI data from their first data release (DR1; DESI Collaboration et al., 2025a). DESI is still ongoing; there will be more objects in the COMAP fields 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 $1185\,\mathrm{km\,s^{-1}}$ 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 (DESI Collaboration et al., 2024a). Where duplicates were present, we used the DESI position and redshift.

Additionally, we applied a systematic velocity offset of 1185 km $\mathrm{s}^{-1}$ to each QSO in the combined eBOSS-DESI catalog. 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 center 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 km $\mathrm{s}^{-1}$, or several COMAP spectral channels) scatter in redshift offsets, which we address in §4.7.

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 4.1, and their redshift distribution in Figure 4.2. 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.

## 4.4 Simulations

We generated a suite of simulated LIM data cubes and associated galaxy catalogs for comparison with the real COMAP and quasar data. These were generated using the `joint_limlam_mocker`[1] code described in Dunne et al., 2025, based on peak-patch DM halo simulations (Bond and 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-$\alpha$ emitters from the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX; Hill et al., 2024; Hill et al., 2021). In the modeling for this work, we primarily adopted the same parameters for the CO LIM map, but altered the parameters to match the quasar catalog. 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.

### CO modeling

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

$$\frac{L'_{\mathrm{CO}}(M_{\mathrm{h}})}{\mathrm{K\ km\ s^{-1}\ pc^2}} = \frac{C}{(M_{\mathrm{h}}/M)^A + (M_{\mathrm{h}}/M)^B}. \tag{4.1}$$

$A$, $B$, $C$, and $M$ are free parameters. Here, we adopted values determined using priors from UniverseMachine (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 $\sigma_{\mathrm{CO}} = 0.42$ 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 4.3. 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 $L_{\mathrm{IR}}$,

$$\log L'_{\mathrm{CO}} = \frac{1}{\alpha}\left[\log L_{\mathrm{IR}} - \beta\right], \tag{4.2}$$

$$L_{\mathrm{IR}} = \frac{\mathrm{SFR}}{\delta_{\mathrm{MF}}} \times 10^{10}. \tag{4.3}$$

[1] `https://github.com/delaneydunne/joint_limlam_mocker`

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 and Walter, 2013. These values have the additional advantage of connecting the CO luminosities directly to $L_{\mathrm{IR}}$, without the SFR intermediary. The Kamenetzky et al., 2016 parameter values are $\delta_{\mathrm{MF}} = 1.0$, $\alpha = 1.27$, $\beta = -1.0$. We used a scatter of 0.3 dex, and took

$$\frac{L_{\mathrm{CO}}}{L_\odot} = 4.9 \times 10^{-5} \frac{L'_{\mathrm{CO}}}{\mathrm{K\ km\ s^{-1}\ pc^2}}. \tag{4.4}$$

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,

$$L = 2N(z)M \left[ \left( \frac{M}{M_1(z)} \right)^{-\alpha(z)} + \left( \frac{M}{M_1(z)} \right)^{\beta(z)} \right]^{-1}, \tag{4.5}$$

with a duty cycle

$$f_{\mathrm{duty}} = \frac{1}{1 + \left( \frac{M}{M_2(z)} \right)^{\gamma(z)}}. \tag{4.6}$$

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,

$$L_{\mathrm{CO}}(M_{\mathrm{h}}, z) = \frac{2N(z)M_{\mathrm{h}}}{\left( \frac{M_{\mathrm{h}}}{M_1(z)} \right)^{-b(z)} + \left( \frac{M_{\mathrm{h}}}{M_1(z)} \right)^{y(z)}}, \tag{4.7}$$

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 $f_{\mathrm{duty}} = 0.1$. 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 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.

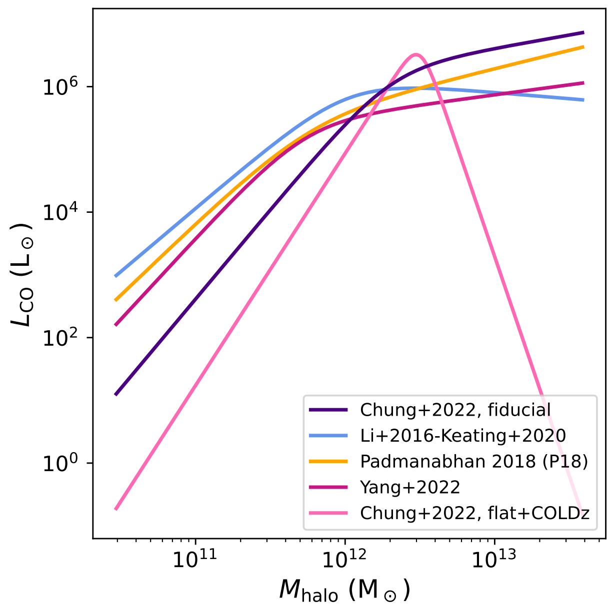


Figure 4.3: Models for CO luminosity as a function of halo mass used to generate simulated stacks. We defaulted to the C22 fiducial model.

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

#### Quasar modeling

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 catalog 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 catalog, 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

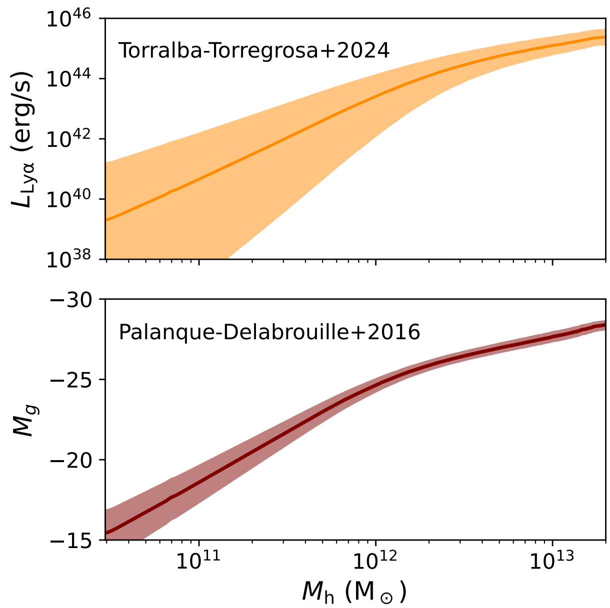


Figure 4.4: Models for quasar luminosity used to generate simulated catalogs, created from abundance-matching onto observed luminosity functions. The top panel shows the PAU model from (Torralba-Torregrosa et al., 2024) and the bottom panel shows the Palanque-Delabrouille et al., 2016. We defaulted to the Torralba-Torregrosa et al., 2024 model.

modeled quasar luminosities by abundance-matching onto observed QLFs. We show the parametrizations of the QLF we use in this work in Figure 4.4. The default model we used is based on a Ly$\alpha$ 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,

$$\frac{dn}{dL} = \frac{\phi^*}{L^*}\left(\frac{L}{L^*}\right)^{\alpha} e^{-L/L^*}. \tag{4.8}$$

We took the lowest-redshift bin, which is centered slightly above the $z = 2.62$ median redshift of our combined eBOSS-DESI catalog, yielding $\log\left(L^*/\mathrm{erg\ s^{-1}}\right) = 44.89$, $\log\left(\phi^*/\mathrm{Mpc^{-3}}\right) = -6.48$, and $\alpha = -1.44$. Through propagating forward the uncertainty on these best-fit values, we found an average scatter at the high-mass end ($M_{\mathrm{halo}} > 5 \times 10^{12}\ \mathrm{M}_\odot$) of $\sigma_{\mathrm{QSO}} = 0.39$ dex. We took the limiting luminosity $L_{\mathrm{cut}} = 10^{44}$ 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,

$$\Phi(M_g, z) = \frac{\Phi^*}{10^{0.4(\alpha+1)(M_g - M_g^*)} + 10^{0.4(\beta+1)(M_g - M_g^*)}} \tag{4.9}$$

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

$$\log\left[\Phi^*(z)\right] = \log\left[\Phi^*(z_\mathrm{p})\right] + c_{1a}(z - z_\mathrm{p}) + c_{1b}(z - z_\mathrm{p})^2, \tag{4.10}$$

$$M_g^*(z) = M_g^*(z_\mathrm{p}) + c_2(z - z_\mathrm{p}), \tag{4.11}$$

and

$$\alpha(z) = \alpha(z_\mathrm{p}) + c_3(z - z_\mathrm{p}). \tag{4.12}$$

In this high-redshift limit, there are eight fitted parameters ($M_g^*(z_\mathrm{p})$, $\Phi^*(z_\mathrm{p})$, $\alpha$, $\beta$, $c_{1a}$, $c_{1b}$, $c_2$, 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 catalog), 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 ($M_\mathrm{halo} > 5 \times 10^{12}\ \mathrm{M}_\odot$). This yields a scatter of 0.11 dex.

The actual magnitude limits of the DESI quasar catalog 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 catalog in two steps. First, we removed from the mock catalog 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 ($N_\mathrm{obj}$) 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 catalog varies between models, and thus the degree of clustering of other DM halos around the cataloged objects also varies.

## 4.5 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.

### Coaddition

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

$$L'_{\rm CO} = \frac{c^2}{2k_{\rm B}} (S\Delta v) \frac{D_{\rm L}^2}{\nu_{\rm obs}^2 (1+z)^3}, \tag{4.13}$$

where

$$(S\Delta v) = \frac{2\nu^2 k_{\rm B} T_{\rm b}}{c^2} \Omega_{\rm vox} \Delta v. \tag{4.14}$$

Here $(S\Delta v)$ is the integrated CO line flux in the voxel, $D_{\rm L}$ is the luminosity distance of the catalog object, $\nu_{\rm obs}$ is the observed frequency and $z$ is the redshift. The parameter $\Omega_{\rm vox}$ refers to the solid angle of the sky subtended by the voxel. We then extracted from the COMAP data cubes a 3D 'cubelet' centered around the spatial position and observed CO(1-0) frequency of each catalog object. This centering is performed only at the level of the COMAP voxel. We explore more advanced sub-voxel centering in Mansfield et al., 2025. The cutout cubelets were $31\times31\times81$ voxels ($62'\times62'\times2.53$ GHz, or 113 cMpc$\times$113 cMpc$\times$337 cMpc), with the catalog object itself located within the central voxel of the cubelet. For context, the full COMAP cubes are $120\times120\times256$ 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 catalog 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 cataloged objects, for each field. We refer to this average cubelet as $\mathbf{s}$. For presentation purposes, we summed the averaged cubelet across its central $5\times5$ spatial voxels or 3 spectral voxels to generate 1D spectra or 2D images, respectively. These are shown in Figure 4.6.

**Extraction of line luminosity**

The goal output value of the stack is the average CO luminosity of the quasars in the eBOSS/DESI catalog and their surroundings over some given cosmic volume, $\langle L'_{\mathrm{CO}}\rangle$. In Dunne et al., 2024, we obtained this value by integrating the stacked cubelet data $\mathbf{s}$ over a central aperture $3\times3$ 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 $\langle L'_{\mathrm{CO}}\rangle$ from the stacked datacube. We determined $\langle L'_{\mathrm{CO}}\rangle$ 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 $\langle L'_{\mathrm{CO}}\rangle$. 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 behavior in RA and Dec for each field.

**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_mocker`. 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 §4.5, and shown in Figure 4.5). 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 $\mathbf{s}$,

$$a = (\mathbf{t}^{\mathrm{T}}\mathbf{t})^{-1}\mathbf{t}^{\mathrm{T}}\mathbf{s}. \tag{4.15}$$

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 $\langle L'_{\rm CO}\rangle$ by integrating over the scaled template signal $a\mathbf{r}$ in three dimensions,

$$\langle L'_{\rm CO}\rangle = a\sum_{k=-1}^{+1}\sum_{i,j=-2}^{+2} r_{i,j,k}, \tag{4.16}$$

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 $5\times5$ spaxels (in contrast to the $3\times3$ 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 $f_{\rm VI}$ (see §4.7).

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 $\langle L'_{\rm CO}\rangle$ 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 $\langle L'_{\rm CO}\rangle$. We describe how we accounted for this in §4.7.

**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 4.11). 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 catalog using the `joint_limlam_mocker` code and default parameters described in §4.4. 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 §4.7). We then injected each simulation realization into an 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 4.11. 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-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 km s$^{-1}$ (§4.3). This resulted in a 3D stacked cubelet for each realization. The 131-object Field 1 stacks with a 50× 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 4.5. 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.

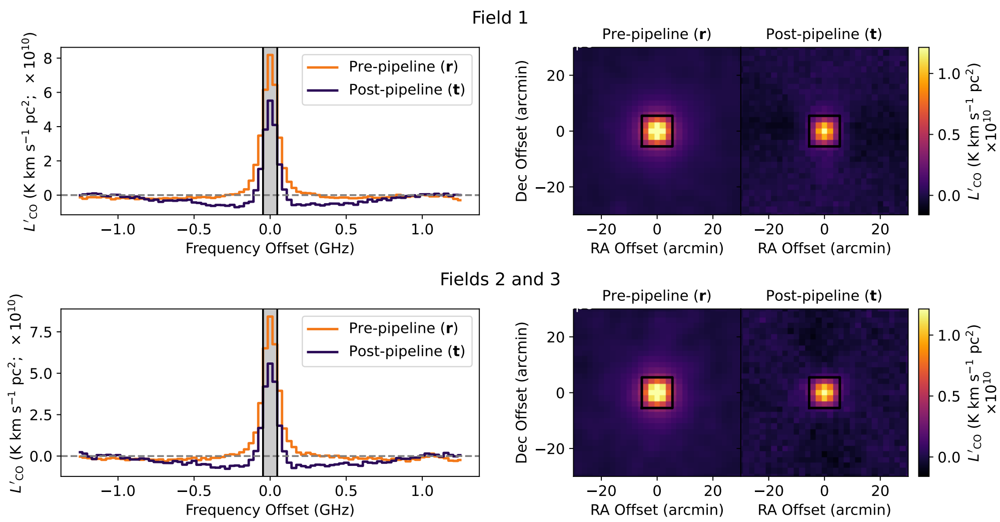


Figure 4.5: Averaged simulated stacks used to account for pipeline signal loss in the stack, shown here for COMAP Field 1 above and Field 2 (which is also used for Field 3) below. We show the stacked spectra on the left (averaging across the central $5 \times 5$ spaxels for presentation purposes) and 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.

### Variance

While it is possible to determine the uncertainties in each stack from propagating through the RMS noise from the TOD, this methodology does not 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 catalogs and measured the distribution of their measured line luminosity values. We randomized the catalogs using the same method as was used in Dunne et al., 2024, offsetting the coordinates of each object in the stack catalog 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 catalogs, as both are fairly spatially homogeneous. We ran the stack 10,000 times, each time using a different random catalog realization generated using this method. We calculated $\langle L'_{\rm CO} \rangle$ from the resulting stacked cubelets using the methodology outlined in §4.5. We then fit a Gaussian distribution to the output $\langle L'_{\rm CO} \rangle$ 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 modeling; we discuss this in §4.7.

## 4.6 Results

We show the final COMAP-eBOSS/DESI stacks for each field in Figure 4.6, 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 $a$t are also shown in Figure 4.6. We then determined $\langle L'_{\rm CO}\rangle$ for each field using its respective model $\mathbf{r}$, finding $\langle L'_{\rm CO}\rangle = [8.44, 0.56, 0.59] \times 10^{10}$ K km s$^{-1}$ pc$^2$ for Fields [1,2,3].

We calculated the variance for each field using bootstrapping, as described in §4.5. We combined the values extracted for each field using inverse-variance weighting, using the outputs from the bootstraps in each field to determine the variances. We found an overall output stack luminosity for the COMAP-eBOSS/DESI stack of $\langle L'_{\rm CO}\rangle = 4.15 \times 10^{10}$ K km s$^{-1}$ pc$^2$, with an uncertainty $\sigma_{\langle L'_{\rm CO}\rangle} = 2.94 \times 10^{10}$ K km s$^{-1}$ pc$^2$. The stack therefore has a significance of $1.41\sigma$, and is a non-detection. The $2\sigma$ (95%) upper limit is $\langle L'_{\rm CO}\rangle \leq 10.0 \times 10^{10}$ K km s$^{-1}$ pc$^2$. A histogram showing the bootstrapped overall output luminosity values is shown in Figure 4.7, 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.

## 4.7 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

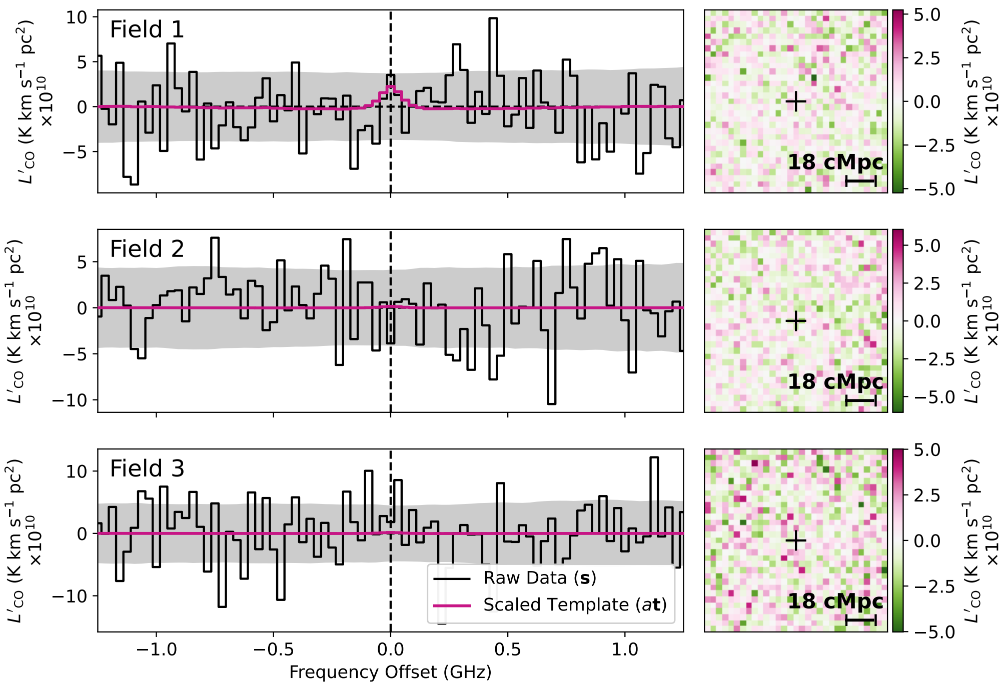


Figure 4.6: Stacked spectra (left) and 2D images (right) for COMAP data in each of the three COMAP fields at the positions of the eBOSS/DESI QSOs. The per channel RMS values are shown as grey shaded regions in the stacked spectrum, and the scaled template $a\mathbf{t}$ is overlaid. The stack is not corrected for any pipeline signal-loss effects, and the templates have the same signal loss present.

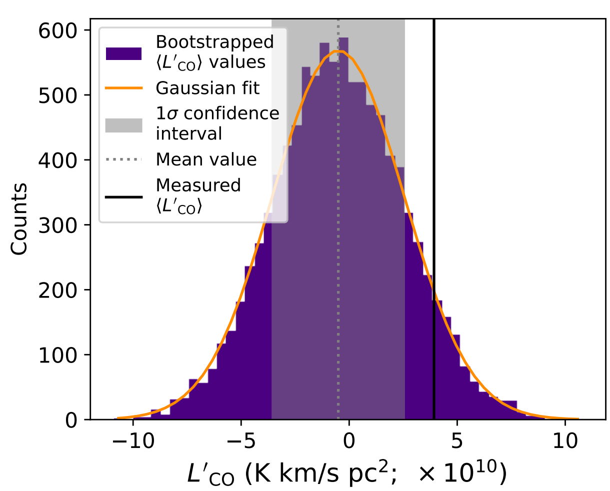


Figure 4.7: Bootstrapped $\langle L'_{\mathrm{CO}} \rangle$ values determined from stacks on randomized eBOSS/DESI catalogs, with a Gaussian fit. The mean and $1\sigma$ confidence interval of the Gaussian are shown in grey.

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 modeling 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 modeled versions. These are in addition to signal loss from the COMAP analysis pipeline, which is accounted for when extracting $\langle L'_{\rm CO} \rangle$. We discuss these factors below. The effects of each on the simulated signal are shown in Figure 4.9.

**Modelled luminosity values**

We tested the five CO models and two QLF models listed in §4.4. In each case, we determined the modeled $\langle L'_{\rm CO} \rangle$ 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 §4.5, 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 $\mathbf{r}_{\rm model}$ cubes). In each field, we generated a mock catalog cut to the same number of objects as that field in the real stack ($N_{\rm obj}$ = 131, 53, and 47 for Fields 1, 2, and 3), and stacked over that catalog. 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 and Chatterjee, 2025). If either the cosmic volume covered by the LIM experiment or the number of cataloged 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 §4.4. The $\langle L'_{\rm CO} \rangle$ values for each model are listed in Table 4.1, both before any corrections for experimental sources of signal attenuation and with uncertainties from raw sample variance (§4.7), and after the corrections are made. The uncorrected spectra are shown in Figure 4.8.

**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 4.9.

Table 4.1: Stack $\langle L'_{\rm CO} \rangle$ values for each of the CO and galaxy catalog models tested. 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.

| Model | TT24[a] $\langle L'_{\rm CO} \rangle$ (K km s$^{-1}$ pc$^2$) $\times 10^{10}$ | PD16[b] $\langle L'_{\rm CO} \rangle$ (K km s$^{-1}$ pc$^2$) $\times 10^{10}$ |
|---|---|---|
| *Before corrections* | | |
| 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 |
| *After corrections* | | |
| 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 |

[a]The quasar model from Torralba-Torregrosa et al., 2024.
[b]The quasar model from Palanque-Delabrouille et al., 2016.
[c]From C22.

Table 4.2: Experimental sources of signal attenuation and variance, and their contribution to the stack.

| Description | Factor | Value |
|---|---|---|
| *Signal attenuation* | | |
| Pointing uncertainties | $f_{\rm PO}$ | 1.05 |
| Redshift uncertainties | $f_{\rm VI}$ | 1.21 |
| *Variance* | | |
| Model-dependence | $d_{\rm MD}$ | 1.07 |
| LIM interloper emission | $d_{\rm LI}$ | 1.04 |

### 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 4.12), finding no systematic offset in the continuum sources and very little additional spread in the signal.

This effect can be ignored in the COMAP autocorrelation, but the stack is predicated on the ability to center the LIM signal on regions of sky containing catalog objects, and thus will be attenuated by any imprecision in the instrumental pointing. We determined the magnitude of this attenuation by convolving kernels modeling 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.

### Catalogue redshift uncertainty

Redshift uncertainty in the galaxy catalog 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 (§4.7), offsetting the velocity of each catalog 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 $f_{\rm VI} = 1.21$. Instead of simply dividing the $\langle L'_{\rm CO}\rangle$ output from each model by this value when accounting for redshift uncertainty, we used the actual $\langle L'_{\rm CO}\rangle$ value from the stack with the introduced redshift scatter.

**Error budget**

Certain experimental and modeling 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 §4.6) or model values. Again, we treated each of these as multiplicative factors, labelling them as $d$.

**Model dependence**

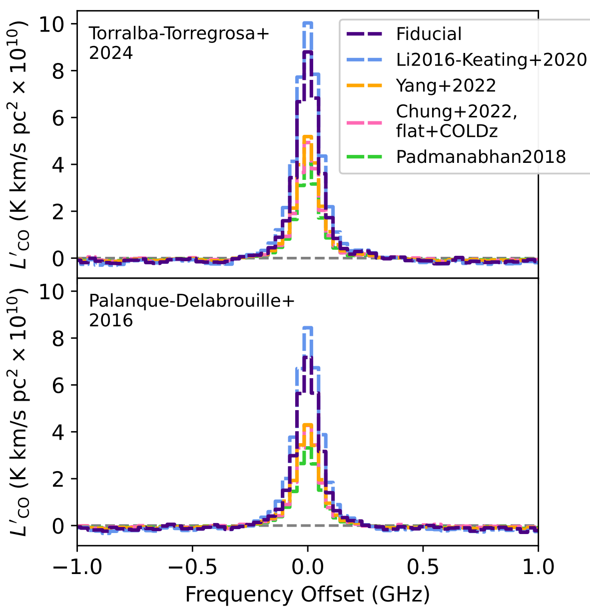


Figure 4.8: Spectra for each CO model choice, shown for a quasar catalog generated using the Ly$\alpha$-based Torralba-Torregrosa et al., 2024 model (top) and the eBOSS g-band luminosity function Palanque-Delabrouille et al., 2016 (bottom). Spectra are shown before correcting for any of the attenuation factors discussed in §4.7.

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 catalogs 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 $\mathbf{r}_{\rm fid}$ to the pre-pipeline model stack for each other model $\mathbf{r}_{\rm model}$, and then extracting the $\langle L'_{\rm CO}\rangle$ from the scaled fiducial model $a_{\rm model}\mathbf{r}_{\rm fid}$. The output model $\langle L'_{\rm CO}\rangle$ 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 $d_{\rm MD} = 1.069$ applied directly to the uncertainty on the data value.

### 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 §4.6, there is in 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 catalogs. 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 modeling, 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 $\langle L'_{\rm CO}\rangle$ 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 4.1. In this particular stacking configuration, these uncertainty values are of similar magnitude to those from the random noise in the COMAP stack data.

**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., 2024a). 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 catalog) 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 catalog 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 $d_{\rm LI} = 1.038$.

## 4.8 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 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 modeled 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 4.9), which brings the modeled luminosity values closer to the limit placed by the COMAP-eBOSS/DESI stack, but still leaves a $3.3\sigma$ disagreement between the two using our fiducial modeling 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

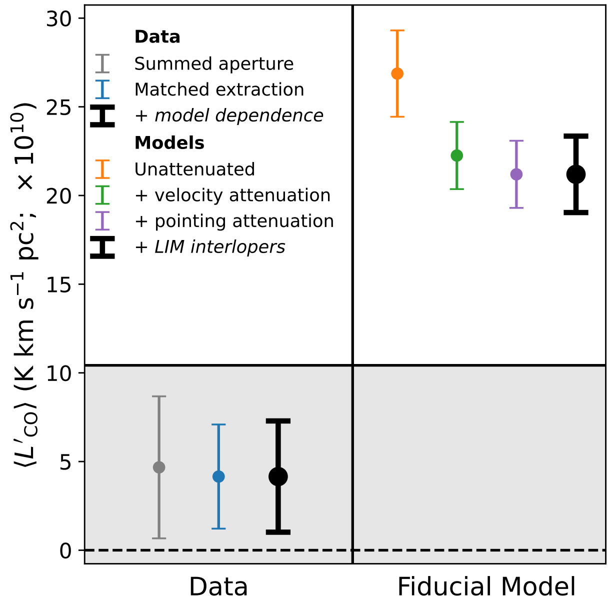


Figure 4.9: Factors attenuating signal or increasing variance, shown for the upper limit from the COMAP-eBOSS/DESI stack and the fiducial model from C22. The corrections applied to the other models are qualitatively similar. The shaded region indicates the $2\sigma$ upper limit from the data alone. The various corrections are outlined in §4.7.

other models in Figure 4.10. Stacks on the L16-K20 CO model are discrepant with our upper limit to $> 3\sigma$ for both choices of quasar luminosity function.

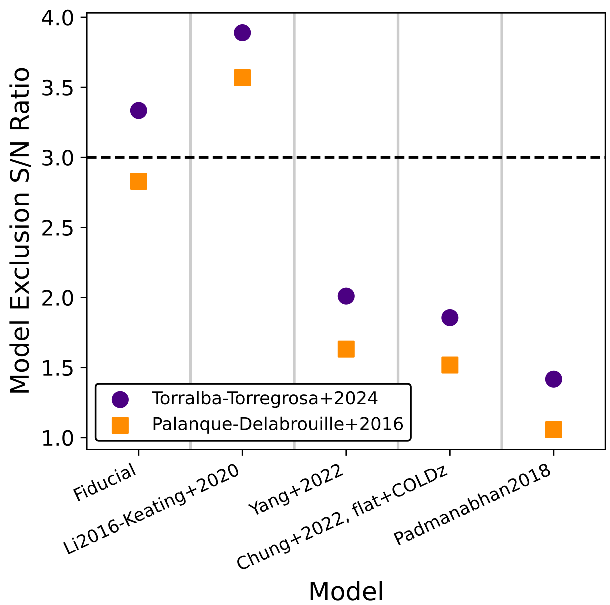


Figure 4.10: S/N at which models are discrepant with the COMAP-eBOSS/DESI upper limit for various models of CO emission, including all potential sources of uncertainty and signal attenuation listed in §4.7.

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 cataloged 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 modeling), but the host galaxies are still a minor factor. For the purposes of discussion, we also assume that all experimental considerations affecting $\langle L'_{\rm CO}\rangle$ have been accounted for in §4.7, 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 modeling we use, and fainter average CO emission.

**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 neighboring those containing the quasars. In our modeling, 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 $L'_{\rm cut}$) used to generate the actual mock catalog 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., Shen 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 modeling is drastically incorrect. Still, small differences in the QLF can significantly affect the modeled 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 modeling.

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 (§4.7), 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 modeled luminosities down and reduce the tension between the models and the data.

**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 possibly 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 modeled stack luminosities down. Additionally, quasar-specific feedback effects such as jets (which could extend up to several megaparcsecs and have been observed at redshifts as high as $z \sim 5$; Oei et al. 2024; Gloudemans et al. 2025) could quench galaxies and disrupt CO emission around the eBOSS/DESI objects.

This sensitivity to large-scale environmental effects is an 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.

**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., 2024b).

While considerable effort has been dedicated to realistically modeling 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 $L'_{\mathrm{CO}} \sim$

$10^{10}$ K km s$^{-1}$ 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 $L'_{\mathrm{CO}(1-0)} \sim 10^8$ K km s$^{-1}$ 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.

## 4.9 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 $2\sigma$ upper limit on the CO emission surrounding 231 quasars from eBOSS and DESI of $\langle L'_{\mathrm{CO}} \rangle = 10.0 \times 10^{10}$ K km s$^{-1}$ 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 $\langle L'_{\rm CO} \rangle$ extraction, and interloper emission in the LIM data. The COMAP fiducial/Torralba-Torregrosa et al., 2024 model is also strongly disfavoured at $3.3\sigma$.

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 modeling, 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 modeling 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.

## 4.10 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, Observatário 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.

## 4.11 Appendix: Signal injection boost strength

We determine pipeline signal loss and extract the stack $\langle L'_{\rm CO} \rangle$ 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 §4.5, 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, $\langle {\rm SNR_{vox}} \rangle$. They find that the map-level PCA filter behaves linearly with respect to the CO signal up to $\langle {\rm SNR_{vox}} \rangle = 0.02$. The boosted, signal-injected data subsets we use in this work have $\langle {\rm SNR_{vox}} \rangle \sim 0.005$, 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 $\langle {\rm SNR_{vox}} \rangle = 2 \times 10^{-4}$ in each subset of the full COMAP S2 dataset, and this rises to $\langle {\rm SNR_{vox}} \rangle = 0.002$ 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 catalog. 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.

## 4.12 Appendix: 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 during 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 4.11).

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 4.12. We find a source profile that is somewhat more spread-out than the 4.5′ Gaussian to be expected from stacking on sources exactly centered 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.

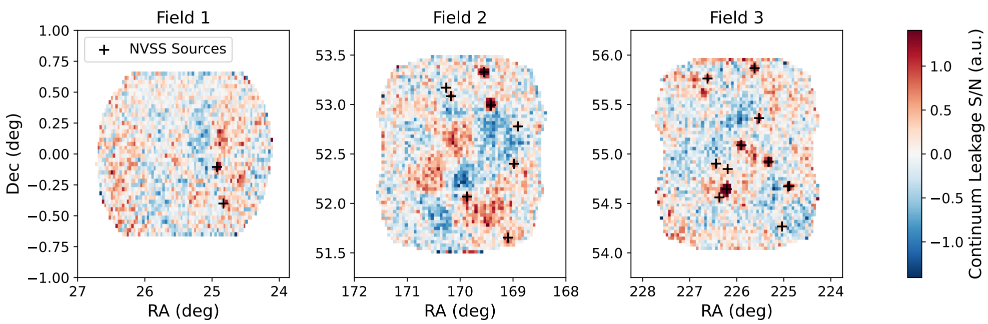


Figure 4.11: Continuum-leakage maps of the three COMAP fields, with radio continuum sources from NVSS (Condon et al., 1998) overplotted.

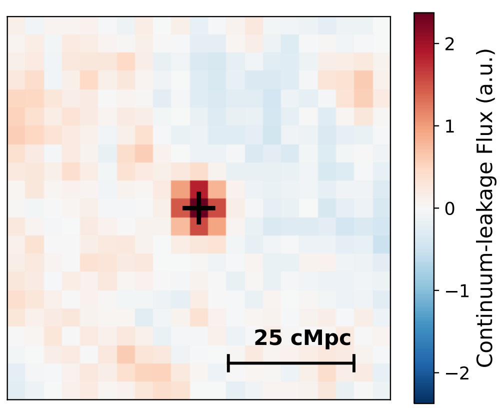


Figure 4.12: Stack of 22 NVSS radio sources on the COMAP continuum-leakage maps. The resulting image is roughly beam-sized, indicating good pointing accuracy. The scale bar is shown for $z = 2.62$ (the median redshift of the eBOSS/DESI sources). It corresponds to $\sim 14'$.

*Chapter 5*

# COMAP PATHFINDER – SEASON 2 RESULTS V. A STACK ON HETDEX LAES

## 5.1 Chapter abstract

We stack data from the line intensity mapping (LIM) experiment COMAP on the positions of 5879 Lyman-$\alpha$ emitting galaxies (LAEs) from the Hobby-Eberly Telescope Dark Energy Survey (HETDEX). COMAP is targeting the global evolution of the molecular gas fuel for star formation in galaxies $2.4 < z < 3.4$ through measurements of the CO(1–0) rotational line. We find a non-detection of CO emission in the LAE stack to the level of $1.5\sigma$, yielding an upper limit of $\langle L'_{\rm CO} \rangle \leq 2.5\times10^{10}$ K km s$^{-1}$ pc$^2$ over an aperture of $18 \times 18 \times 9$ cMpc centered on the HETDEX LAEs. We find that this limit is in mild to moderate tension with models of CO emission stacked on $z \sim 3$ LAEs. In particular, we find a $> 3\sigma$ discrepancy between our fiducial CO model and the COMAP-HETDEX upper limit irrespective of the luminosity function chosen to represent the modeled LAEs. While the average CO luminosity of the Universe is degenerate with the clustering properties of the LAEs in the stack, these clustering properties are well-measured $z \sim 3$, and we propose that the tension between models and observations is tentative evidence against the brighter CO models tested.

## 5.2 Introduction

It has been well-established that the Universe's global star formation rate (SFR) peaks at $z \sim 2$ (Madau and Dickinson, 2014). What remains a mystery, however, are the mechanisms by which the star-forming molecular gas assembles in galaxies in order to enable this peak. Molecular gas (typically traced by rest-frame millimetre or sub-millimetre transitions such as those of carbon monoxide, CO) is not well-observed above $z \sim 2$. Existing galaxy surveys (Pavesi et al., 2018; Aravena et al., 2019; Lenkić et al., 2020, e.g.) are limited at the faint end by completeness and at the bright end by observing area.

Complementing these galaxy surveys is a growing observational technique known as line intensity mapping (LIM). In this technique, the galaxy population is measured as a 3D fluctuation field traced by a specific spectral line rather than a catalog of discrete objects (see Chang and Lidz 2026 for a review). In LIM datacubes, all sources of luminosity within a given resolution element contribute to its measured luminosity, making it ideal for applications where the average cosmic luminosity of some galaxy formation tracer (typically CO, CII, H$\alpha$ Ly$\alpha$, or 21 cm neutral hydrogen emission) needs to be measured, such as the build-up to the peak of cosmic star formation. The CO Mapping Array Project (COMAP; Cleary et al., 2022) is the leading LIM experiment targeting cosmic CO emission via its 1–0 transition, probing the redshift range $2.4 < z < 3.4$, immediately preceding the peak of cosmic star formation. Although no LIM experiment (including COMAP) has yet detected cosmic CO, COMAP has provided the tightest constraints yet available on CO emission in this redshift range with its Season 2 results, ruling out the brightest proposed models with its autocorrelation analysis (Chung et al., 2024b).

The COMP sensitivity is limited by white noise as of Season 2. In general, raw sensitivity, as well as instrumental and astrophysical sources of systematic error, are expected to be what limit the ability of LIM experiments to constrain bulk line emission. One family of analysis techniques that has been found to be effective in addressing these issues (e.g., Keenan et al., 2022; Dunne et al., 2024; MeerKLASS Collaboration et al., 2025; Chen et al., 2025; Lujan Niemeyer et al., 2026) folds in information from external catalogs of galaxies through cross-correlation or stacking. These techniques add information, using the external galaxy catalogs as priors on where signal in the LIM datasets can be expected, and mitigate systematic errors that are not correlated with the distribution of the galaxy catalog objects.

In the previous paper in this series (Dunne et al., 2026c, hereafter referred to as D26), we used quasars from the extended Baryon Oscillation Spectroscopic Survey (eBOSS; Ahumada et al., 2020) and the first data release of the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al., 2025a) as a reference catalog on which to stack. We did not detect CO emission in this stack, but found a mild (1-4$\sigma$) tension between the upper limit we derived and the predicted luminosity from several different models of CO emission stacked on quasar positions. Quasars are heavily biased galaxies, leading in theory to a bright stack signal (D25). However, the eBOSS/DESI quasars also have some limitations in terms of stack sensitivity: quasars are rare, so there are relatively few objects in this catalog, and there are large velocity uncertainties between the optical lines used to measure the quasar redshift and the bulk velocity of the host galaxy. Additionally, quasars occupy unique (dense) positions in the large-scale structure (LSS) that could affect the CO properties of their surroundings in extremely nonlinear ways. In the stack, these nonlinear effects cannot be distinguished from variations in the bulk CO luminosity of the Universe.

In this work, we perform the stack instead on a catalog of Lyman-$\alpha$ emitting galaxies (LAEs) from the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX; Gebhardt et al., 2021). LAEs have lower bias than quasars, meaning their stack signal will not be as bright. However, there are more than an order of magnitude more HETDEX LAEs in the COMAP volume than there are eBOSS/DESI quasars, which more than makes up for the decreased expected signal, and the decreased bias means that the regions probed by LAEs are much more representative of the Universe as a whole.

We introduce the COMAP LIM survey and the HETDEX LAE catalog in §5.3 of this work. In §5.4, we describe a set of mock COMAP-HETDEX stacks that we use to put our stacked signal into context. We present the stack methodology in §5.5 and compare the observed stack to models (which requires correcting for a variety of experimental effects) in §5.7. Finally, we provide a discussion, summary, and conclusions in §5.8 and §5.9. We use the nine-year WMAP cosmology ($\Omega_{\rm m} = 0.286$, $\Omega_\Lambda = 0.714$, $\Omega_{\rm b} = 0.047$, and $H_0 = 100h$ km s$^{-1}$ Mpc$^{-1}$ with $h = 0.7$; Hinshaw et al. 2013) throughout this work for consistency with other COMAP works as well as the peak-patch DM simulations on which our modeling is based, and we assume base-10 logarithms unless stated otherwise.

## 5.3 Data

### COMAP

We use data from the first two seasons (May 2019 – November 2023) of the COMAP Pathfinder instrument, a 19-feed focal plane array receiver fielded on a 10.4 m Leighton (Leighton, 1977) dish at the Owens Valley Radio Observatory. In total, this comprises 17,500 hours on-sky, split between three cosmological fields. The COMAP experiment and the Pathfinder instrument are described in detail in Cleary et al., 2022 and Lamb et al., 2022, respectively. The full data reduction process for the Season 2 dataset is described in Lunde et al., 2024 and Stutzer et al., 2024. In short, the raw detector readout (in the form of time-ordered data, TOD) for a given field is filtered to reduce correlated noise and systematic errors and calibrated into brightness temperature units feed-by-feed. These data are then binned into three-dimensional data cubes (maps) with dimensions $2^\circ \times 2^\circ \times 8$ GHz (centered at 30 GHz) and voxel sizes of $2' \times 2' \times 31.25$ MHz. The COMAP beam is $4.5'$

at 30 GHz. For the stacking analysis in this work, we average the maps for each feed into a single, co-added, map for a given field. At the current sensitivity of COMAP—a per-voxel root-mean-square (RMS) noise of $25 - 50\,\mu$K—the noise has been shown to be Gaussian and each of the filters involved in this reduction have been shown to behave linearly (Lunde et al., 2024).

Only two of the three COMAP fields (Fields 1 and 2) overlap with the HETDEX coverage. We show the RMS distribution, averaged across all spectral channels, for each of these two fields in Figure 5.1. We note that the noise response across each field is inhomogeneous due to the COMAP scanning strategy. We follow the masking strategy established in D26 to remove low-sensitivity regions of the map before stacking for efficiency.

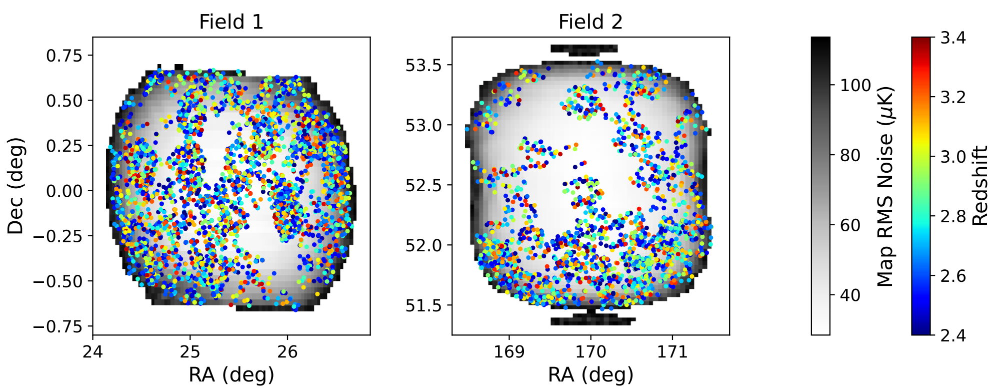


Figure 5.1: Spatial distribution of the HETDEX LAEs compared to the footprint of the COMAP data, for the two COMAP fields in which there is HETDEX overlap. The COMAP footprint is colored by the RMS in each voxel, averaged across all spectral channels, and the LAEs are colored by their redshift.

**HETDEX**

The Hobby-Eberly Telescope Dark Energy eXperiment (HETDEX; Gebhardt et al., 2021) is a large (540 deg$^2$), blind survey of Lyman-$\alpha$ Emitters (LAEs) $1.9 < z < 3.5$, designed to measure potential evolution of the dark energy density. The HETDEX survey ran from January 2017 to January 2025. HETDEX uses the integral field unit VIRUS (Visible Integral-field Replicable Unit Spectrograph; Hill et al., 2021; Hill et al., 2024) on the Hobby-Eberly Telescope (HET; Ramsey et al., 1998; Hill et al., 2021), targeting the Ly$\alpha$ 1215 Å line over a bandpass of 3500−5500 Å with resolving power $R \sim 800$. HETDEX surveys the sky sparsely, with a fill factor of 1/4.6. The HETDEX footprint overlaps significantly with the COMAP fields: Fields 1 and 2 both have complete (if sparse) HETDEX coverage (see Figure 5.1). In redshift, HETDEX overlaps with the COMAP coverage over the full extent $2.4 < z < 3.4$, with a mean redshift of 2.81 in the catalog overlap.

The HETDEX data reduction and selection process is outlined in full in Davis et al., 2023a, Mentuch Cooper et al., 2023, and Mentuch Cooper et al., 2026, and we will not reproduce it here. We use the catalog selection choices outlined in Mentuch Cooper et al., 2026 – in particular, we take the same cut on overall catalog S/N ratio ($\geq 4.6$). There are also two machine learning classification schemes applied to the HETDEX PDR1 (Mukae et al., 2026), and we take the same values as for PDR1 (`p_cnn` > 0.5 and `p_conf` > 0.5). However, we only apply the `p_conf` cut to objects with HETDEX detection S/N ratios above 5.5. There are 4577 HETDEX objects in the overlap with COMAP after applying these cuts.

In addition to the mainline HETDEX dataset, we include data from an additional 77 individual HET observations in COMAP Field 1, observed between December 2024 and January 2026. In total, these comprise an additional 28 hours of data. We process these observations through the same data reduction and line-identification process as the main HETDEX catalog, described above. However, the machine learning classification schemes are not available for these additional observations. Instead, we visually inspect every potential line detection in these additional observations (see §5.3 below).

Of the HETDEX objects in the catalog, 58 objects have been identified as AGN (following the procedure outlined in Liu et al., 2025). As there are relatively few AGN (and as several of the LAE models we test include a high-luminosity component for Ly$\alpha$ emission from AGN, which commonly are included in LAE catalogs), we treat them with the rest of the LAEs. We find that excluding the known AGN does not affect our results.

**Visual inspection**

The majority of the cataloged HETDEX objects are relatively low-significance, with signal-to-noise (S/N) ratios < 6 (Figure 5.2). At this significance, the catalog is contaminated by two different sources of false positive detections: (1) noise peaks or instrumental artifacts and (2) a small percentage of low-redshift galaxies misidentified as $z \sim 3$ LAEs. This latter source of contamination is because roughly 95% (Davis et al., 2023b) of detections are based on a single spectral line, and there are other spectral lines (primarily [O II] at $z < 0.5$) which fall into this wavelength range. The presence of false positive detections in the stack is not catastrophic, but false positives linearly attenuate the stack signal (e.g., D25). It is thus necessary to remove as many as possible to maximize S/N, and to accurately characterize the number of remaining false positives in order to correct for this attenuation. We thus attempt to increase the fidelity of our catalog by performing a visual inspection.

We visually inspect each HETDEX object with a S/N ratio $\geq$ 5.0, as well as 500 objects in each field 4.8 $\leq$ S/N ratio $<$ 5.0 and the entire additional dataset in COMAP Field 1, using the `EliXerWidget` tool developed by the HETDEX collaboration (Davis et al., 2023b). We remove any confident false positives from the data (any detection clearly associated with an extended source in auxiliary $gri$ imaging, with significant continuum, or that is likely noise or an artifact based on the 2D spectra). We remove 5% of the HETDEX catalog in this way. For the remaining objects, we assign them a confidence level—either ‘maybe’ or ‘likely’ an LAE (roughly 50% and 75% confident that the object is an LAE, as complete confidence is reserved for any object with more than one spectral line visible, and these objects are typically AGN). Of the remaining catalog objects, we classified 47.7% as ‘maybe’ distant and 45.5% as ‘likely’ distant.

#### False positive ratio

Because false positives attenuate the stacked signal (by averaging noise in to the overall stack), we must characterize the remaining rate of false positive detections in the catalog. This was done by Mentuch Cooper et al., 2026 using repeat HETDEX observations in a deep Legacy field in COSMOS. Using the same selection criteria used here, they found an 80% confirmed rate, and thus a false positive ratio of 0.2. However, this is based on a sample of only 414 LAEs in a single field.

As a secondary strategy, we use the visual inspection to estimate the false positive fraction. We assume that, as the ‘maybe’ distant objects are assigned roughly 50% confidence, half of them are real LAEs and the other half are false positive detections. Assuming the subset that we visually inspected is representative of the entire catalog, this gives a false positive fraction of 0.238. Visual inspection is inherently subjective, relying on human pattern recognition and incorporating human biases. It is therefore difficult to assess the reliability of the false positive fraction based on visual inspection alone. However, it is encouraging that this value is reasonably consistent with the value determined from repeat observations.

As neither of our strategies for calculating the false positive fraction are completely reliable, we add a confidence range of ±0.1 to this quantity. This is one of the dominant sources of uncertainty in the COMAP-HETDEX stack.

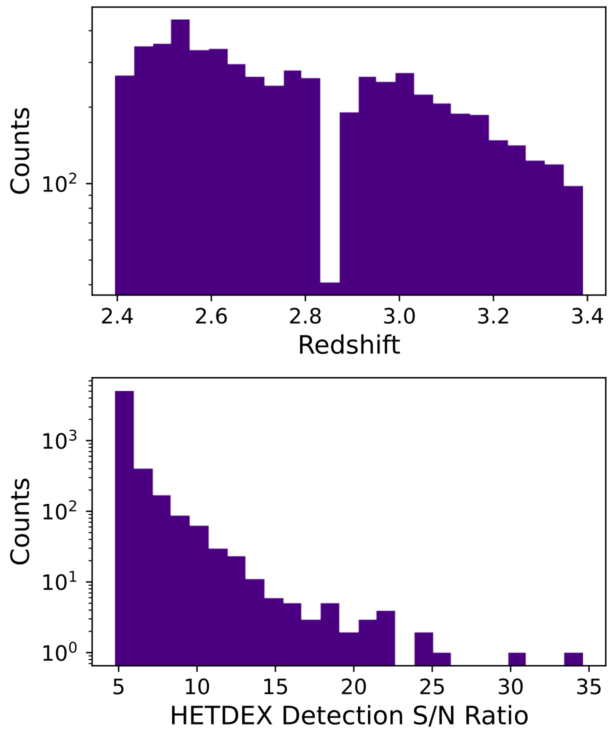


Figure 5.2: Distribution in redshift (top) and S/N ratio (bottom) of the HETDEX objects included in the stack catalog.

### CO-Ly$\alpha$ velocity offset

Because HETDEX is a survey of LAEs, the galaxies on which we are stacking have redshifts determined by their Ly$\alpha$ wavelengths. There is a known radiative transfer effect where flux is suppressed in galaxies near and below the rest-frame Ly$\alpha$ wavelength of 1215.6 Å, meaning that redshift values from Ly$\alpha$ emission are often from fitting only to the red peak of the spectrum, and are therefore systematically offset with respect to other emission lines (e.g., Byrohl and Gronke, 2020). The CO emission in the stacks will thus be offset from the central channels if this is not accounted for. Based on stacks on a subset of the HETDEX catalog itself, Davis et al., 2023a find that this offset is on average $235 \pm 18$ km s$^{-1}$ for several known ISM and IGM lines. This agrees well with other determinations of the Ly$\alpha$ velocity offset (e.g., Shibuya et al., 2014; Steidel et al., 2018; Verhamme et al., 2018; Muzahid et al., 2020), and corresponds to slightly less than one COMAP spectral channel. We therefore determine the systemic redshift of each object in the HETDEX catalog by offsetting by $\Delta v_{\mathrm{Ly}\alpha} = 235$ km s$^{-1}$,

$$z_{\mathrm{sys}} = z_{\mathrm{Ly}\alpha} - \frac{\Delta v_{\mathrm{Ly}\alpha}}{c}(1 + z_{\mathrm{Ly}\alpha}). \tag{5.1}$$

We address the uncertainty in this relation in §5.7.

## 5.4 Simulations

We generate a suite of simulated LIM data cubes and associated galaxy catalogs for comparison with the real COMAP and HETDEX data. These are generated using the `joint_limlam_mocker`[1] code described in D25, based on peak-patch DM halo simulations (Bond and Myers, 1996; Stein et al., 2019). We refer the reader to D25 for the full details of the simulation code. In short, the catalog of DM halos generated by the peak-patch simulations are painted with CO luminosities (based on data-driven $L(M)$ models calculating the CO luminosity directly as a function of DM halo mass) and Ly$\alpha$ luminosities (through abundance-matching to measured Ly$\alpha$ luminosity functions, LFs). These models account for the large-scale distribution of the CO and Ly$\alpha$ luminosity in the universe, and the per-halo interaction between the two tracers is modeled using a correlated scatter with correlation coefficient $\rho$. The CO luminosities are then binned into a simulated LIM data cube, accounting for effects such as astrophysical line broadening and smoothing by the main beam, and a HETDEX-like mock LAE catalog is generated including the halos which pass observability cuts based on their Ly$\alpha$ luminosity. The suite of simulation parameters in D25 was designed to emulate a COMAP-HETDEX stack, so in the modeling for this work we primarily adopt the same parameters, adding additional models for the CO $L(M)$ function and the Ly$\alpha$ LF. We also alter two additional simulation choices. First, we assume that the scatter in the CO and Ly$\alpha$ luminosity values for a given halo are not correlated or anti-correlated (*i.e.*, we assume $\rho = 0$). This simplifies the modeling, and we showed in D25 that $\rho$ has negligible effect on the output stack luminosity for an LAE-like catalog. Additionally, we alter the observational parameters used to select the mock LAE catalog from the full distribution of DM halos. This is described in more detail in §5.4.

[1] `https://github.com/delaneydunne/joint_limlam_mocker`

**CO modeling**

As in D26, we attempt to account for the large uncertainty in the CO model space $z \sim 3$ by testing a variety of $L(M)$ models from the literature. We default to the 'COMAP fiducial' ("UM+COLDz+COPSS") CO model from Chung et al., 2022 (hereafter C22), which is a double-power law model, taking the form

$$\frac{L'_{\rm CO}(M_{\rm h})}{\rm K\ km\ s^{-1}\ pc^2} = \frac{C}{(M_{\rm h}/M)^A + (M_{\rm h}/M)^B}, \tag{5.2}$$

where $A$, $B$, $C$, and $M$ are free (fitted) parameters. We take $A = -2.85$, $B = -0.42$, $\log C = 10.63$, and $\log \frac{M}{M_\odot} = 12.3$, which are the C22 parameters from a prior from UniverseMachine (Behroozi et al., 2019) conditioned on the luminosity function constraints from the COLDz blind CO galaxy survey (Riechers et al., 2019) and power spectrum constraints from the CO Power Spectrum Survey (COPSS; Keating et al., 2015; Keating et al., 2016). We apply a log-normal scatter with $\sigma_{\rm CO} = 0.42$ dex. We also test the suite of other models: the Keating et al., 2020 extension to the Li et al., 2016 model (L16-K20), the Yang et al., 2022 semi-analytic model based model (Y22), the double-power-law model from Padmanabhan, 2018 (P18), and the flat-prior model from C22, constrained on the COLDz blind CO galaxy survey (Riechers et al., 2019) alone (C22 flat+COLDz). A full description of each of these models is available in D26.

**LAE modeling**

We model the Ly$\alpha$ luminosity of our simulated DM halos by abundance-matching to empirical LAE luminosity functions (LFs). We default to the $z = 3.1$ model from (Ouchi et al., 2020), which is based on LAE observations from Ouchi et al., 2008 and Konno et al., 2016. This model takes a Schechter, 1976 form,

$$\frac{dn}{dL} = \frac{\phi^*}{L^*}\left(\frac{L}{L^*}\right)^{\alpha} e^{-L/L^*}, \tag{5.3}$$

with $L^* = 8.49 \times 10^{42}\,{\rm erg\,s^{-1}}$, $\phi^* = 3.9 \times 10^{-4}\,{\rm Mpc^{-3}}$ and $\alpha = -1.8$. We add a lognormal scatter on the resulting Ly$\alpha$ luminosities of $\sigma_{{\rm Ly}\alpha} = 0.49$ dex, determined by averaging the uncertainty on the Schechter luminosity function across luminosity values, weighted by the luminosity distribution of the abundance-matched simulation. This is a large scatter, but it is broadly consistent with other simulations of Ly$\alpha$ radiative transfer (e.g., Byrohl and Nelson, 2023), and is consistent with the expectation that the luminosity of an LAE is not well-correlated with its host halo's clustering properties owing to the behavior of the Ly$\alpha$ line (Orsi et al., 2008). Hereafter, we refer to this model as O20.

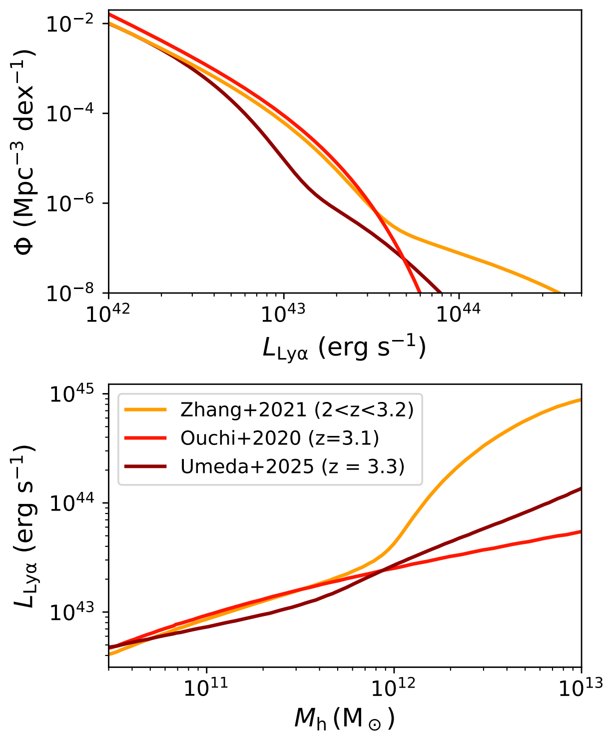


Figure 5.3: Empirical $z \sim 3$ LAE luminosity functions used for our LAE modeling (*top*) and the resulting $L(M)$ functions after abundance-matching (*bottom*). We default to the Ouchi et al., 2020 model.

While the $z \sim 3$ LAE LF is considerably better-constrained by data than the CO $L(M)$ functions, we still test two other LFs, for completeness and to better span the COMAP redshift range ($2.4 < z < 3.4$). Both of these models add an additional high-luminosity component to the Schechter functional form to account for the population of AGN within the catalog. In both cases, this additional component is treated as a double-power law (DPL) model,

$$\frac{dn}{dL} = \frac{\phi^*_{\mathrm{DPL}}\left(\frac{L}{L^*_{\mathrm{DPL}}}\right)}{\left[\left(\frac{L}{L^*_{\mathrm{DPL}}}\right)^{-\alpha_{\mathrm{DPL}}-1} + \left(\frac{L}{L^*_{\mathrm{DPL}}}\right)^{-\beta_{\mathrm{DPL}}-1}\right]} dL, \tag{5.4}$$

where $L^*_{\mathrm{DPL}}$, $\phi^*_{\mathrm{DPL}}$, $\alpha_{\mathrm{DPL}}$, and $\beta_{\mathrm{DPL}}$ are the fitted components. We take fit parameters from:

- The HETDEX-based luminosity function from Zhang et al., 2021. The bright end is fit to an intermediate stage of the HETDEX survey (internal HETDEX Data Release 2), which comprises observations from January 2017 to

Table 5.1: Luminosity function fit values used to model Ly$\alpha$ emission. Parameters, in order, are from Ouchi et al., 2020, Zhang et al., 2021 and Umeda et al., 2025.

| Model | Schechter Parameters | | | DPL Parameters | | | | Uncertainty |
|---|---|---|---|---|---|---|---|---|
| | $L^*$ (erg s$^{-1}$) | $\phi^*$ (Mpc$^{-3}$ dex$^{-1}$) | $\alpha$ | $L^*_{\rm DPL}$ (erg s$^{-1}$) | $\phi^*_{\rm DPL}$ (Mpc$^{-3}$ dex$^{-1}$) | $\alpha_{\rm DPL}$ | $\beta_{\rm DPL}$ | $\sigma_{\rm Ly\alpha}$ dex |
| O20 | $8.49 \times 10^{42}$ | $3.9 \times 10^{-4}$ | -1.8 | – | – | – | – | 0.49 |
| Z21 | $7.41 \times 10^{42}$ | $3.9 \times 10^{-4}$ | -1.7 | $4.79 \times 10^{44}$ | $1.2 \times 10^{-8}$ | -1.23 | -3.06 | 0.41 |
| U25 | $1.95 \times 10^{42}$ | $7.4 \times 10^{-3}$ | -1.19 | $3.72 \times 10^{43}$ | $2.7 \times 10^{-7}$ | -1.88 | -4.19 | 0.38 |

June 2020 combined with $r$-band Subaru/HSC imaging, and the faint end ($\log(L_{\rm Ly\alpha}/{\rm erg\,s^{-1}}) < 43.0$) to LAEs from the Vimos-VLT Deep Survey (Cassata et al., 2011). Both of these datasets are selected to cover redshifts $2 < z < 3.2$, slightly lower on average than the HETDEX catalog we use here. For this model, we measure a (weighted) average scatter of $\sigma_{\rm Ly\alpha} = 0.41$ dex. Hereafter, we refer to this model as Z21.

- The SILVERRUSH luminosity function from Umeda et al., 2025, based on narrowband Ly$\alpha$ data. We use the luminosity function at $z = 3.3$ to provide a higher redshift comparison to the Z21 LF above. We take an average scatter of $\sigma_{\rm Ly\alpha} = 0.38$ dex. We refer to this model as U25.

We tabulate the Schechter and DPL parameters for these three LFs in Table 5.1, and show the LFs in Figure 5.3. We assign Ly$\alpha$ luminosities to the simulated DM halos by abundance-matching the LF to the peak-patch simulations' mass function. The resulting $L(M)$ functions are also shown in Figure 5.3.

Once each DM halo is assigned a Ly$\alpha$ luminosity, we generate a mock catalog by cutting all halos below a HETDEX survey sensitivity limit of $L^{\rm cut}_{\rm Ly\alpha} = 3.16 \times 10^{42}\,{\rm erg\,s^{-1}}$ (this corresponds to the fifth percentile luminosity in the HETDEX catalog) and then randomly selecting $N_{\rm obj}$ simulating LAEs from the resulting cut catalog. Unlike in D25, we apply a logarithmic (rather than linear) weighting by Ly$\alpha$ luminosity when selecting halos. This is because, above $L^{\rm cut}_{\rm Ly\alpha}$, the probability of observing a HETDEX halo is fairly flat with luminosity. The number of objects in the HETDEX catalog is mostly limited by the sparse on-sky coverage. Finally, we offset each mock catalog object in the spectral axis by a value pulled randomly from a Gaussian distribution with standard deviation $104\,{\rm km\,s^{-1}}$, to simulate the velocity uncertainty in the HETDEX catalog (§5.3).

Because clustering completely dominates the stack signal, particularly for an LAE-like catalog (D25), each of these steps are designed to recover mock catalogs with similar clustering properties to the real HETDEX LAEs. We test each mock catalog for consistency with HETDEX in two ways: by measuring its (linear) bias, and by testing its average DM halo mass. We find a $b_{\rm LAE} = 1.72 \pm 0.03$ ($2.31 \pm 0.04$, $1.78 \pm 0.03$) for the mock catalogs generated using the O20 (Z21, U25) LF. These in general agree well with literature values of $b \sim 1-2$ at $z \sim 2-3$ (Ouchi et al., 2020). The bias of the Z21 model is driven upwards by the high-luminosity/high-mass AGN component. We also find that the median DM halo masses of the mock catalogs are $6.7 \times 10^{10}\,{\rm M}_\odot$ ($6.8 \times 10^{10}\,{\rm M}_\odot$, $6.5 \times 10^{10}\,{\rm M}_\odot$) for the O20 (Z21, U25) LFs, which each agree well with observational findings of $M_{\rm DM,\,LAE} \sim 10^{10} - 10^{11}\,{\rm M}_\odot$ (Ouchi et al., 2020; Umeda et al., 2025).

## 5.5 Stacking methodology

We stacked the COMAP data on the positions of the HETDEX objects in order to probe the average CO luminosity in the regions surrounding LAEs. We followed the procedure established in D26 nearly exactly, only modifying certain steps to match the HETDEX LAEs rather than the eBOSS/DESI QSOs. We summarize this procedure and describe the modifications here, and refer to D26 for a full description of the methodology.

### Summary of stacking methodology

Firstly, we converted the COMAP data from brightness temperature ($T_{\rm b}$) units into CO line luminosity values (Solomon et al., 1997), performing the same conversion for both the measured $T_{\rm b}$ and its RMS uncertainty $\sigma_{\rm vox}$. We then extracted from this line luminosity datacube large $31 \times 31 \times 81$ voxel ($62' \times 62' \times 2.53$ GHz, or 112.6 cMpc $\times$ 112.6 cMpc $\times$ 337.4 cMpc) 'cubelets' centered around the spatial position and observed CO(1-0) frequency of each object in the HETDEX catalog. We discarded cubelets which have more than half of their central $5 \times 5 \times 3$ voxels (where the majority of the signal is expected to lie) masked. We combined the cubelets associated with each catalog object in voxel-by-voxel, accounting for the inhomogeneous noise response across the COMAP data cubes by inverse-variance weighting by $\sigma_{\rm vox}$. We did this separately for each field. We will refer to the resulting averaged cubelets as \.

Next, we generated a template of pipeline-processed stack emission by creating simulated realizations of CO fluctuation cubes and corresponding HETDEX-like catalogs, boosting the brightness temperature of these realizations by a factor of 50 and injecting them into subsets of the COMAP TOD (to serve as noise models), processing these through the COMAP data reduction pipeline and mapmaker, and performing stacks on the resulting maps. We discuss our choice of boost factor in detail in D26. We refer to this processed template as **t**. We also performed a stack on the unprocessed simulation to create a raw, reference cubelet, which we refer to as **r**. The processed stack **t** is therefore the raw stack **r** passed through each of the filters in the COMAP pipeline.

We then used the processed template **t** as a matched filter. We scaled this to the real stack data **s** by a single-parameter amplitude $a$,

$$a = (\mathbf{t}^{\mathrm{T}}\mathbf{t})^{-1}\mathbf{t}^{\mathrm{T}}\mathbf{s}. \tag{5.5}$$

We undid the effects of the COMAP pipeline (equivalent to deconvolving out the data-reduction transfer function) by scaling the reference template **r** by the amplitude $a$. This is the input signal best-explained by the actual data. Finally, we summed this signal profile across a central aperture to return a single, average CO luminosity $\langle L'_{\mathrm{CO}} \rangle$ for the regions surrounding the LAEs,

$$\langle L'_{\mathrm{CO}} \rangle = a \sum_{k=-1}^{+1} \sum_{i,j=-2}^{+2} r_{i,j,k}, \tag{5.6}$$

where $k$ indexes the spectral axis and $i, j$ the spatial (plane-of-sky) axes. We used an aperture size of 3 spectral channels and 5 spatial channels, as these are the parameters which we found to optimize the stack S/N for LAEs in D25. The COMAP pipeline attenuates the signal by a factor of 1.58 in Field 1 and by a factor of 1.55 in Field 2, when summed across the central aperture.

**LAE-specific modifications**

The only modifications from the procedure described in D26 are in generating the processed and unprocessed template stacks, **t** and **r**. We used the HETDEX-like LAE default parameters described in 5.4 rather than the QSO parameters adopted in D26 – the LF from O20, the observable luminosity limit of $L_{\mathrm{Ly}\alpha}^{\mathrm{cut}} = 3.16 \times 10^{42}\,\mathrm{erg\,s^{-1}}$ and the logarithmic weighting in the random catalog generation, and the velocity uncertainty of $104\,\mathrm{km\,s^{-1}}$. We used the LAE-correct $N_{\mathrm{obj}}$ in each field (3533 for Field 1 and 2346 for Field 2), and disregarded Field 3 as there are no HETDEX LAEs in its footprint. We show the resulting templates pre- and post-processing in Figure 5.4. With the boost factor of 50, each individual simulation realization had a S/N of 6.5 (8.0) in the central voxel for Field 1 (2). After averaging the realizations together, the template S/N was 65 (80) for the two fields.

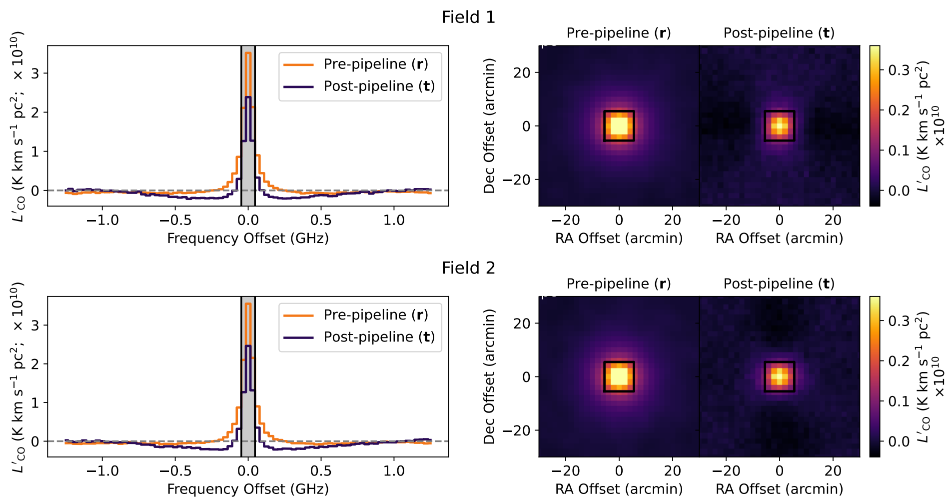


Figure 5.4: Matched-filter templates used for extraction of the signal from the stack for Field 1 (*top*) and Field 2 (*bottom*). The left panels show the spectral extent of the signal before and after processing by the COMAP pipeline, summed across the central $5 \times 5$ spaxels in each spectral channel. The right panels show the spatial extent of the signal, summed across the central 3 spectral channels. In both cases, the central aperture is indicated with the black lines. Differences in the post-processed template **t** are primarily in the spatial axes, and are driven by differences in scanning strategy between the two fields, which lie at different declination.

### Variance

While the Season 2 data have been tested extensively and found to be broadly consistent with white noise (Lunde et al., 2024), the stack is averaging the uncertainty down and may uncover persistent systematic errors below the level of the COMAP RMS noise. In order to return robust measurements of the stack variance (and to account for pixel-to-pixel correlations in the stacking strategy and check for underlying systematic uncertainties), we calculated the variance by bootstrapping, again following the procedure established in D26. We generated many random catalog realizations, performed a stack on each following the steps in §5.5 and §5.5, and measured the distribution of the output $\langle L'_{\rm CO} \rangle$ from each random stack.

Because both the HETDEX and COMAP on-sky footprints have irregular shapes, there is also a strong potential for systematic errors to be amplified by interactions between the two. It is thus extremely important that the random catalogs used for bootstrapping mimic the on-sky footprint of the HETDEX catalog. As such, we generated random catalogs by pulling spatial positions randomly from the distribution of on-sky coverage in the HETDEX catalog (including the additional observations in the SHELA field added for this work), weighted inversely by the RMS noise in each

HETDEX fibre and shot. We assigned each spatial position a redshift pulled randomly from the distribution of redshifts in the real HETDEX catalog $2.4 < z < 3.4$. We generated 10 000 random catalogs using this method.

## 5.6 Results

We show the final COMAP/HETDEX stacks in Figure 5.5, noting that the extended 2D image and 1D spectrum representations of the stack are uncorrected for any pipeline signal loss effects. We find scale factors of $a = 0.283, 0.027$ for Fields 1 and 2, respectively, corresponding to line luminosity values of $\langle L'_{\mathrm{CO}} \rangle = 2.2 \times 10^{10}\,\mathrm{K\,km\,s^{-1}\,pc^2}$ and $\langle L'_{\mathrm{CO}} \rangle = 2.1 \times 10^{9}\,\mathrm{K\,km\,s^{-1}\,pc^2}$.

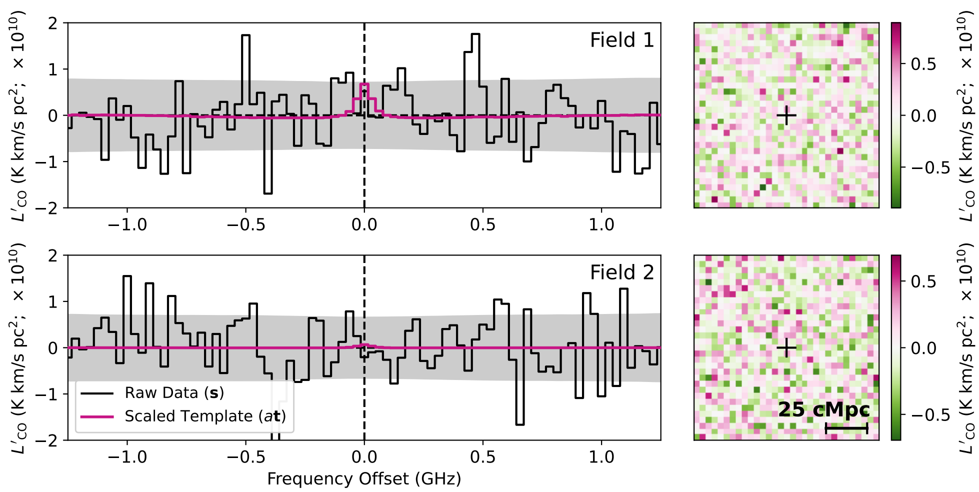


Figure 5.5: Stacked spectra (left) and 2D images (right) for COMAP data at the positions of the HETDEX LAEs in the two COMAP fields. In the stacked spectra, we show the per-channel RMS uncertainty as grey shaded regions, as well as the processed template **t** scaled to the stack data. In the 2D images, we show the central voxel (the location of the galaxies included in the stack) with the black crosses.

The distribution of output $\langle L'_{\mathrm{CO}} \rangle$ values from bootstrapping for each field is shown in Figure 5.6. We also show a fit to a Gaussian distribution. We find that the bootstrap output is consistent with zero-centered Gaussian noise in each field (mean values of $0.11 \times 10^{10}\,\mathrm{K\,km\,s^{-1}\,pc^2}$ and $-0.10 \times 10^{10}\,\mathrm{K\,km\,s^{-1}\,pc^2}$ and reduced-$\chi^2$ values of 1.3 and 1.5 in Fields 1 and 2, respectively. The standard deviations $\sigma_{\langle L'_{\mathrm{CO}} \rangle}$ of the distributions in each field are $1.1 \times 10^{10}\,\mathrm{K\,km\,s^{-1}\,pc^2}$ and $9.5 \times 10^{9}\,\mathrm{K\,km\,s^{-1}\,pc^2}$ for Fields 1 and 2, respectively. While there are more HETDEX objects in Field 1, this field also has the worse COMAP sensitivity, which is driving the difference in uncertainty.

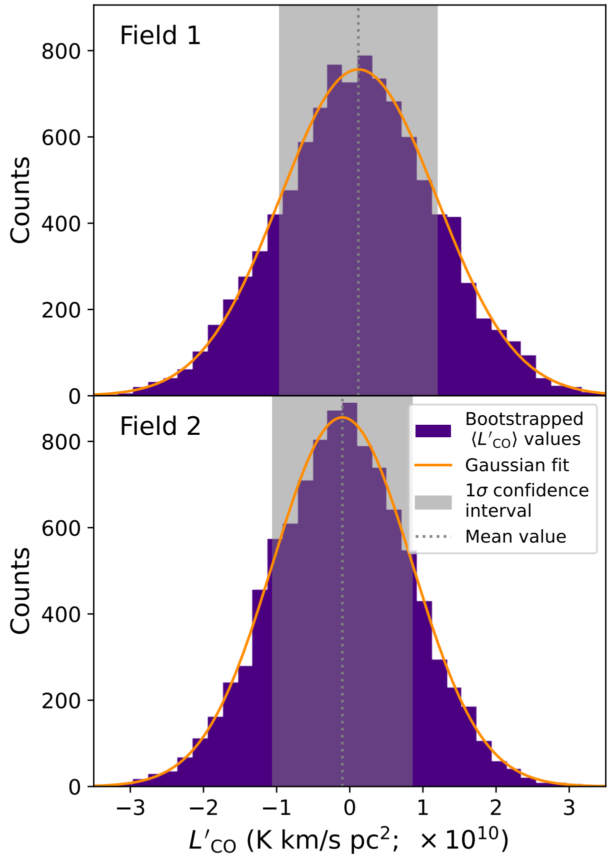


Figure 5.6: Bootstrapped $\langle L'_{\mathrm{CO}} \rangle$ values determined from stacks on randomized HETDEX in COMAP Field 1 (top) and Field 2 (bottom), with a Gaussian fit. The mean and $1\sigma$ confidence interval from the Gaussian fit in each field are shown in grey.

When averaged together (weighted by the uncertainty in each field, derived from bootstrapping) this yields an output line luminosity of $\langle L'_{\mathrm{CO}} \rangle = (1.1 \pm 0.7) \times 10^{10}\,\mathrm{K\,km\,s^{-1}\,pc^2}$. The stack is therefore $1.5\sigma$ significant, and is a non-detection. The $2\sigma$ (95%) upper limit is $\langle L'_{\mathrm{CO}} \rangle \leq 2.5 \times 10^{10}\,\mathrm{K\,km\,s^{-1}\,pc^2}$.

## 5.7 Comparison to models

As with the quasar stack, this is an unusual observational constraint (it is a constraint on CO emission biased to the regions surrounding LAEs, but not necessarily originating from the LAEs themselves), so we compare this upper limit directly to models of the COMAP/HETDEX stacks in order to put it into context. We describe the stack luminosities resulting from our modeling in §5.7. In order to correctly perform this comparison, we must account for each of the sources of experimental attenuation and uncertainty in the real stack when generating model stacks. We account for these factors in different ways: by multiplying them into the measured stack variance (in the case of the model-dependence of the stack extraction), or by treating them as additional sources of attenuation or variance on our modeling. We note that these factors are nearly identical to those considered for the quasar stack in D26. However, they must be adjusted for a HETDEX LAE stack rather than an eBOSS/DESI QSO stack. Additionally, in this work we consider both attenuation and uncertainty resulting from several of the factors. We show the cumulative effects of each correction in Figure 5.7. We note that the dominant source of uncertainty in the stack is the native stack variance, even after accounting for each of these observational effects.

Finally, there are several factors that could affect the stack luminosity that are accounted for inherently in either the modeling or the luminosity extraction in the stack. We provide a discussion of experimental factors which could affect the stack luminosity, but do not in this specific case, in §5.7.

**Modelled luminosities and sample variance**

We test the five CO models listed in §5.4, as well as the three Ly$\alpha$ models listed in §5.4. For each model, we aim to accurately reproduce the number of objects in the real HETDEX catalog and the COMAP on-sky sensitivity pattern in order to sample the distribution of halo luminosities in the same way as the real data. As such, we generate 50 unique simulation realizations for each model. For each realization, we simulate both COMAP fields with HETDEX data: we apply the real COMAP RMS distribution (and mask) to the simulated LIM data cube, and cut the mock HETDEX catalog to the same number of sources as are in the real catalog in that field (3533 and 2346 for Fields 1 and 2, respectively). Unlike the simulations used to generate the template **t** in §5.5, we do not pass the simulated maps through the COMAP data reduction pipeline. Instead, we work directly with the raw, unaltered signal for each realization. We stack the noiseless mock COMAP data on the positions of the mock catalog objects, and extract the signal by summing directly across the central aperture, as is done to extract signal from **r** (Equation 5.6.

We take the mean across the 50 realizations to be the output stack luminosity for a given model. These values are listed for each model in Table 5.2. To characterize sample variance for each of the models we test, we calculate the mean absolute deviation (MAD) between the stack $\langle L'_{\rm CO}\rangle$ value returned from each realization.

Table 5.2: Stack $\langle L'_{\rm CO}\rangle$ values for each of the CO and galaxy catalog models we test. Values are shown before correcting for any attenuation due to experimental parameters. Uncertainties are raw sample variances.

| Model | Stack $\langle L'_{\rm CO}\rangle$ (K km s$^{-1}$ pc$^2$; $\times 10^{10}$) |
|---|---|
| Default (C22+O20) | 7.99 ± 0.49 |
| *Other CO Models* | |
| LK20 | 11.70 ± 0.58 |
| Y22 | 5.79 ± 0.30 |
| C22fc | 4.95 ± 0.24 |
| P18 | 4.29 ± 0.23 |
| *Other LAE Models* | |
| Z21 | 10.54 ± 0.58 |
| U25 | 7.89 ± 0.45 |

When we test different CO models, we use the default (O20) Ly$\alpha$ LF, and when we are testing different LAE models we use the default (C22) CO $L(M)$ function.

We find that the stack $\langle L'_{\rm CO}\rangle$ values for the LAE catalog are roughly 3× smaller than the QSO numbers listed in D26 for a given CO model. This is because the QSO halos are both more biased (and thus have more massive neighboring DM halos) and are more CO-bright themselves (D25). In relation to each other, the CO models behave as expected: the L16-K20 model provides the brightest expected stack signal, and the P18 the faintest. The Z21 LAE model, which includes a significant bright AGN component, provides a brighter stack signal than the O20 model, and the higher-redshift U25 model is slightly fainter overall. Finally, the sample variance is smaller for each of the LAE models presented here than it is for the QSO models, even after accounting for the decreased absolute luminosity. This is for two reasons – there are more LAEs in the HETDEX catalogs, and the LAEs themselves are less biased overall, meaning they probe more average regions of the Universe and there is less stochasticity going into the average.

**Extraction model-dependence**

The extraction methodology we describe in §5.5 effectively corrects for the transfer function imprinted on the data by the COMAP data reduction pipeline. However, it assumes that the three-dimensional distribution of the actual stack signal matches that of a modeled template. The signal is not a point source, and its distribution is affected by nuanced factors such as clustering, astrophysical line broadening, and the fingers-of-god effect (Mansfield et al., 2025) along the various axes. The

Table 5.3: Experimental sources of signal attenuation and variance, and their contribution to the stack.

| Description | Attenuation factor ($f$) | Uncertainty factor ($d$) |
|---|---|---|
| *Explicit corrections to measured luminosity* | | |
| Model-dependence (MD) | – | 1.033 |
| *Explicit corrections to modeled luminosity* | | |
| Pointing offset (PO) | 1.054 | 1.027 |
| LIM interloper emission (LI) | 1.003 | 1.006 |
| False positives (FP) | 1.207 | 1.107 |

extracted stack luminosity therefore depends on the choice of CO and LAE models, introducing an additional source of variance to the output luminosity that must be accounted for.

We determine the stack uncertainty due to model-dependence following D26. We treat each of the (pipeline-uncorrected) CO and LAE models as the template, and scale these template cubes ($\mathbf{r}_{\text{model}}$) to the fiducial model's raw model cube ($\mathbf{r}_{\text{fiducial}}$), yielding a scaling amplitude $a_{\text{model}}$ as in Equation 5.5. We then measure the output luminosity $a_{\text{model}}\mathbf{r}_{\text{fid}}$. The model $\langle L'_{\text{CO}}\rangle$ values calculated using this method differ from the fiducial $\langle L'_{\text{CO}}\rangle$ value by at most 3.3% (for the Z21 LAE model). We thus assume a 3.3% uncertainty due to model dependence, $d_{\text{MD}} = 1.033$, and increase the variance on the measured stack by $d_{\text{MD}}$ to account for this effect. We assume that the model-dependence does not introduce any systematic offset to the output luminosity.

**Corrections to the modeled luminosities**

While we account for many observational effects directly in our modeling or in our signal extraction, some observational effects which directly attenuate signal cannot be applied to the real data as we have no actual signal level to compare them to. As such, we apply these effects to the modeled luminosities for a fair comparison. We represent each correction in two ways – as an attenuation factor, $f$, and as a factor by which the correction increases the variance, $d$. The modeled stack luminosities become

$$\langle L'_{\text{CO}}\rangle_{\text{corr}} = \frac{\langle L'_{\text{CO}}\rangle_{\text{model}}}{f_{\text{PO}} f_{\text{LI}} f_{\text{FP}}}, \tag{5.7}$$

and the modeled variances become

$$\sigma_{\text{corr}} = d_{\text{PO}} d_{\text{LI}} d_{\text{FP}} \sigma_{\text{model}}. \tag{5.8}$$

We list the correction factors in Table 5.3, and plot the modeled luminosities after each correction in Figure 5.7

### COMAP pointing uncertainty

There is a known issue with the COMAP Pathfinder telescope's pointing, leading to both systematic and random offsets between real and assumed source locations. These are described in detail in D26. In each field, both the systematic and random offsets are smaller than a single COMAP spaxel. While the magnitude and direction of these offsets has not changed from that work, the three-dimensional distribution of the stack signal is different for the HETDEX LAE stack, thus affecting how much this effect attenuates (and adds uncertainty to) the stack $\langle L'_{\rm CO}\rangle$. In order to quantify this effect for the LAEs, we performed 1000-object stacks on the fiducial LAE simulations in each field after passing the simulated data cubes through the pipeline. We then convolved the simulated data cubes with kernels modeling the RA-Dec pointing scatter described in D26 and repeated the 1000-object stacks on the same catalogs. We found that the pointing uncertainty attenuates the LAE stacks by a factor of 1.066 in Field 1 and 1.044 in Field 2. Weighted by the uncertainty in each field, this is an overall attenuation factor of $f_{\rm PO} = 1.054$. It additionally increases the variance in the two fields by factors of 1.008 and 1.013 for Fields 1 and 2, respectively, for an overall factor of $d_{\rm PO} = 1.011$.

### LIM interloper emission

Although the $z \sim 3$ CO(1–0) emission is relatively isolated in its observed frequency band, there will be a small amount of contamination in the LIM data from other spectral line emission. In particular, CO(2–1) emission from $z \sim 8$ will be observed at around 30 GHz. While the exact luminosity of this emission is extremely uncertain, it is estimated to be roughly 30% of the $z \sim 3$ CO(1–0) emission at the COMAP map level (Chung et al., 2024a). To quantify the effect of this interloper emission on the HETDEX LAE stack, we generate 99 simulation realizations using the fiducial parameters, adding in an additional galaxy fluctuation field that is uncorrelated with the mock galaxy catalog and has 30% of the mean luminosity of the simulated CO(1–0) emission before mean subtraction. We then performed a stack with 5879 catalog objects (the total number of objects in the real COMAP/HETDEX stack) on each simulation realization, as the severity of the interloper contamination in the stack is dependent on the number of objects being stacked (D25).

We find that, for a 5879-object stack on the LAEs, interloper contamination has a negligible effect on the overall stack $\langle L'_{\rm CO} \rangle$: it attenuates the signal by a factor of $f_{\rm LI} = 1.003$, and increases the variance by a factor of $d_{\rm LI} = 1.006$. We incorporate this correction into our model comparison for completeness, but note that in the future it could easily be ignored for a HETDEX-sized stack. This is consistent with our findings in D25 that the effect of interloper emission on the stack is very small when the stack catalog is large. As in previous works, we ignore contamination by continuum sources such as the CMB or radio point sources entirely, as the COMAP data reduction pipeline is extremely effective at removing continuum emission (Lunde et al., 2024).

**False positive detections**

As discussed in §5.3, it is known that the HETDEX survey contains a fraction $f_{\rm FP}$ of false positive detections, which will linearly attenuate the signal by the false positive fraction. The actual fraction of false positives in the HETDEX data is difficult to determine, and depends considerably on the exact choice of catalog. Based on the HETDEX parameters we adopt, which are consistent with the choices of Mentuch Cooper et al., 2026, we expect a 20% false positive rate in our stack reference catalog, which should correspond to a 20% attenuation in the signal level (D25).

To verify this, and to measure any additional variance resulting from the addition of false positives into the catalog, we compare 5879-object stacks on simulations to stacks with the same number of objects, but with a percentage of the mock catalog entries replaced by randomly-generated 3D positions. As expected, adding 20% simulated false positives to the catalog attenuates the stack signal by a factor of $f_{\rm FP} = 1.207$. The false positives also add an additional scatter to the stack, increasing the variance by a factor of $d_{\rm FP} = 1.107$. Due to the uncertainty on the HETDEX false positive rate discussed in §5.3, we also test a 10% false positive rate and a 30% false positive rate, and find that at 10% false positives, $f_{\rm FP,10} = 1.109$ and $d_{\rm FP,10} = 1.079$ and at 30% false positives, $f_{\rm FP,30} = 1.338$ and $d_{\rm FP,30} = 1.260$.

**Experimental effects that are implicitly corrected for**

Several other experimental factors which might affect the stack luminosity are either negligible or are implicitly corrected for in our modeling or stacking methodology. Aliasing between COMAP spectral channels is negligible (Lamb et al., 2022). Our modeling accounts for beam smearing, astrophysical spectral-line broadening, and pixel-window effects, as well as the relationship between CO and Ly$\alpha$ luminosity and the LSS and the interaction between the two tracers at the level of an individual DM halo (see D25, for more details). The matched-filter technique we adopt here corrects for the effects of the COMAP pipeline transfer function, and we additionally correct for the COMAP main beam efficiency while stacking. In this section, we discuss the effects of the LAE velocity uncertainty and CMB heating for posterity, although we find both effects to be negligible.

**LAE velocity uncertainty**

Due to the intricacies of Ly$\alpha$ radiative transfer, we expect the LAE redshifts to have a small ($235\,\mathrm{km\,s^{-1}}$) systematic offset from the CO redshifts. We correct for the systematic offset directly by simply shifting the HETDEX objects along the LOS. However, there is some scatter in this relation. We take the scatter to be $86\,\mathrm{km\,s^{-1}}$, which is the MAD between the $\Delta v_{\mathrm{Ly}\alpha}$ values reported by the Shibuya et al., 2014 LAE census. There is also on average $18\,\mathrm{km\,s^{-1}}$ uncertainty on the measured HETDEX redshifts. We take the total uncertainty on the HETDEX redshift to be the sum of these two quantities, $104\,\mathrm{km\,s^{-1}}$. These uncertainties will serve to move the line-of-sight position of the LAEs randomly away from the overdensity within which they are contained, thus spreading out the stack signal in the spectral axis. This will attenuate $\langle L'_{\mathrm{CO}} \rangle$ as well as introduce some additional variance.

To measure the magnitude of these effects, we compare the simulations generated in §5.7 to a set of the same simulation realizations which do not include the random velocity offsets. We find that the velocity uncertainty attenuates the signal by a $f_{\mathrm{VI}} = 1.018$, and increases the variance by a factor of $d_{\mathrm{VI}} = 1.013$. This effect is fairly negligible.

**CMB heating**

At higher redshifts, many of the conditions of the Universe change. In particular, the temperature of the cosmic microwave background (CMB) is increased from its local temperature, meaning the observed CO brightness temperature (and thus luminosity) is attenuated by a factor $\sim 25\%$ depending on the exact conditions of the galaxy (da Cunha et al., 2013). Were this factor not accounted for in the CO models shown in this work, it would be another direct attenuation factor that must be accounted for when comparing them to our data. However, we find that this is not necessary. Several of the models are based directly on empirical observations, which also suffer from this heating-based attenuation and in general have not been corrected upwards to account for this. This is true of both Chung et al., 2022 models, L16-K20, and P18. The Yang et al., 2022 model is based on semi-analytic models rather than empirical data, but the CO radiative transfer follows Popping et al., 2019 in using Despotic (Krumholz, 2014), which accounts for the effects of the CMB. As such, the modeled luminosity values we present here do not need to be further corrected for CMB heating.

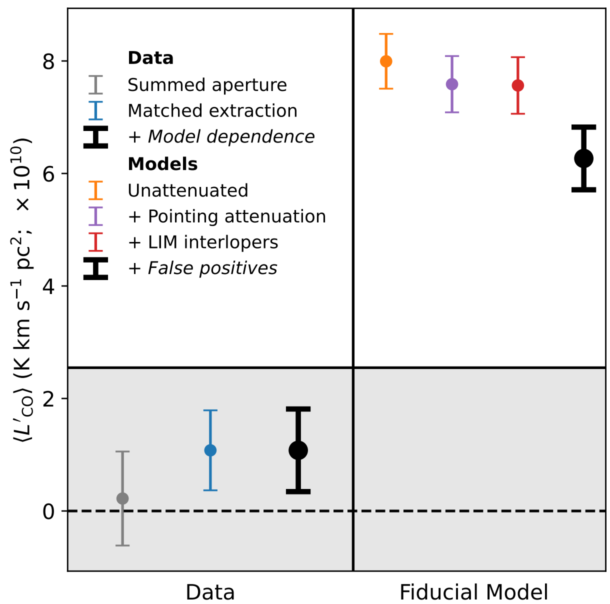


Figure 5.7: Factors attenuating signal or increasing variance, shown for the upper limit from the COMAP-HETDEX stack and the fiducial model from Chung et al., 2022. Shaded region indicates the $2\sigma$ upper limit from the data alone. The various corrections are outlined in §5.7.

## 5.8 Discussion

In this work, we present stacks of CO emission from COMAP on the positions of galaxies from the HETDEX LAE survey. We do not detect any cosmic CO, and the upper limit we place on the stacked signal falls below the expected luminosity of the stack based on simulations that incorporate several different CO and LAE model choices. We find that this tension cannot be explained by any of the several experimental sources of signal attenuation or variance that we test. We quantify the tension between models and observations, after accounting for these experimental factors, in Figure 5.8. In particular, our fiducial model for CO emission (C22) stacked on the positions of LAEs is discrepant with the COMAP-HETDEX upper limit at $> 3\sigma$ for each of the three LAE LFs and for the full range of HETDEX false positive ratios that we test. The L16-K20 CO model is also consistently $> 3\sigma$ discrepant with the real stack.

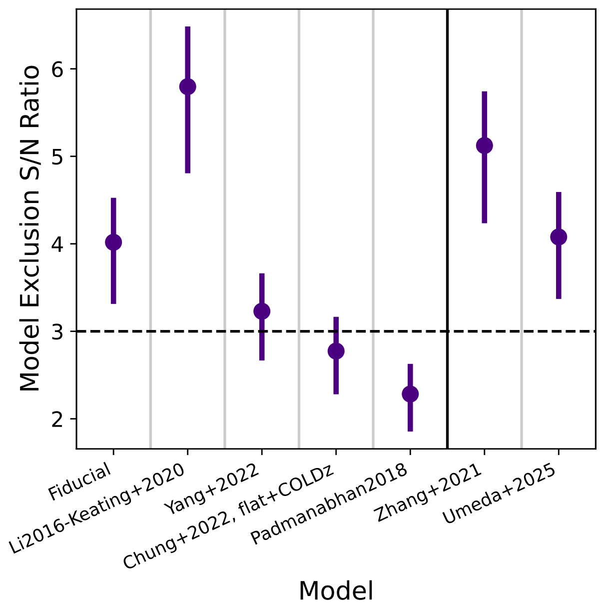


Figure 5.8: S/N ratio at which models are excluded by the COMAP-HETDEX upper limit for various models of CO emission (to the left of the vertical black line) and LAE LF (to the right). When varying the CO model we hold the LAE model constant to O20, and when varying the LAE model we hold the CO model constant to C22. Error bars show the range in false positive ratios expected for the HETDEX catalog.

**Comparison to the COMAP-eBOSS/DESI stack**

This work follows directly from the stack of COMAP data on the positions of eBOSS and DESI quasars presented in D26. In that work, we found an upper limit on the CO luminosity stacked on quasar positions of $\langle L'_{\rm CO}\rangle_{\rm QSO} \leq 10.0 \times 10^{10}\,{\rm K\,km\,s^{-1}\,pc^2}$. This is four times less sensitive than the LAE-based limit of $\langle L'_{\rm CO}\rangle_{\rm LAE} \leq 2.5 \times 10^{10}\,{\rm K\,km\,s^{-1}\,pc^2}$. However, quasars are much more biased tracers of the matter density field than LAEs, and we expect the signal from a quasar stack to be roughly three times brighter than an equivalent LAE stack for a given CO model (D25). Combined with the increased sample variance (and increased uncertainty/attenuation from spectral line broadening and LIM interlopers) caused by the smaller sample size in the quasar stack, the HETDEX LAE stack is roughly

$20 - 50\%$ more sensitive to the expected stack signal than the eBOSS/DESI quasar stack.

### Implications for cosmic CO emission

The $\langle L'_{\rm CO} \rangle$ returned from the stack is the average luminosity summed over the $18 \times 18 \times 9$ cMpc region surrounding the LAEs detected by HETDEX. This is not a direct measurement of the CO luminosity from the HETDEX LAEs themselves (in fact, these provide a very small contribution to the overall stack luminosity, D25), but it is also not an unbiased tracer of the average CO luminosity of the Universe. As such, any effect which could cause a decrease in the clustering around the observed HETDEX LAEs in a way that is not captured by our modeling could drive our measured discrepancy between observations and models. Equivalently, fainter average CO emission could also drive this discrepancy. Because the stack returns just a single value, these two explanations are unfortunately degenerate. We discuss each below. For the purposes of this discussion, we ignore the possibility of an unaccounted-for systematic error.

### Clustering around the HETDEX LAEs

The Ly$\alpha$ luminosity function is fairly well-studied even at $z \sim 3$, thanks to several large surveys including SILVERRUSH (Umeda et al., 2025), MUSE (Herenz et al., 2019), and HETDEX itself. However, there is still significant evolution in the global LAE properties with redshift, particularly in the period $z \lesssim 3$ (see the review in Ouchi et al., 2020, although we note that this transition is to higher bias at higher redshifts and would serve to increase the stack luminosity relative to what is modeled here). We have attempted to bracket this evolution with the range of empirical LAE models we test, but we do not apply any direct redshift evolution of the LAE luminosity modeling across the COMAP redshift range within a given simulation, which could affect the clustering properties when integrated over redshift.

Additionally, we transform the measured LF into modeled galaxy Ly$\alpha$ luminosities using abundance matching. Because of the unique properties of the Ly$\alpha$ line, which is a resonant transition of neutral hydrogen, the escape of Ly$\alpha$ emission from galaxies is stochastic and depends sensitively on galaxy properties such as dust extinction, ISM kinematics, neutral gas content, and gas geometry (e.g., Kunth et al., 1998; Rivera-Thorsen et al., 2015; Weiss et al., 2021). This means host halo masses and galaxy clustering properties correlate weakly with Ly$\alpha$ luminosity, particularly at lower luminosities (Orsi et al., 2008; Chiang et al., 2015), making abundance-matching difficult. We are also somewhat limited by our peak-patch DM simulations, which have a minimum mass of $2.9 \times 10^{10}\,\mathrm{M}_\odot$.

The resonance of the Ly$\alpha$ transition can also cause photons to be absorbed even after escaping the LAEs themselves. This can bias the properties of the LAE catalog, especially when this absorption is happening due to large-scale structures intervening between the LAE and the observer. The ODIN narrowband survey has recently found that the number of faint LAEs observed in protoclusters is suppressed compared to the field (Nagaraj et al., 2025), which would make LAE local environments less dense and attenuate stack signal. Additionally, studies of HETDEX LAEs themselves suggest that they may be preferentially located towards the observer side of galaxy overdensities (Weiss et al., 2024; Weiss et al., 2025). This would bias the location of the stacked CO signal towards higher redshift along the spectral axis, and thus would also attenuate the measured $\langle L'_{\mathrm{CO}} \rangle$. The magnitude of this bias is not yet well-quantified.

However, we note that we have checked the clustering properties of our modeling against the same surveys used to measure many of the above effects, and found them to be broadly in agreement (§5.4). In particular, we found that our simulations have an LAE bias $\sim 1.8$ for the fiducial O20 model, as well as an average DM halo mass $< 10^{11}\,\mathrm{M}_\odot$. Both of these quantities are consistent with expectations for this redshift range (Gawiser et al., 2007; Guaita et al., 2010; Kusakabe et al., 2018; Umeda et al., 2025). We also find that the $L(M)$ function we determine agrees well with other Ly$\alpha$ simulations (Garel et al., 2015; Byrohl and Nelson, 2023). While our simulations may not capture the exact nuances of the Ly$\alpha$ radiative transfer described above, these global clustering properties are likely what is driving the stack luminosity. In combination with the lower bias of the LAE catalog (implying that the environments being probed by the LAEs are much more representative of the Universe as a whole than, for example, the D26 quasar stack), it seems less likely that the discrepancy between observations and models could be attributable to LAE clustering alone.

**Brightness of the CO models**

The other potential explanation for the discrepancy between the modeled luminosities and our observed upper limit is that the average CO luminosity of our models is brighter than the actual cosmic CO luminosity. Cosmic CO is not well measured, particularly at $z \sim 3$ and at the faint end of the luminosity function, and each of the models we test includes a degree of extrapolation. This is also what is driving the large discrepancy between the $\langle L'_{\rm CO}\rangle$ values for different CO models. As with the quasar stack in D26, it is possible for the COMAP-HETDEX upper limit to be in tension with the models without being in tension with the observations underpinning the models, due to these extrapolations.

This is particularly true for the two models with which our observed upper limit is in $> 3\sigma$ tension – the C22 COMAP fiducial model and the L16-K20 model. The L16-K20 model is based on observations at $z \sim 0$ extrapolated to higher-redshift by assuming a relationship with FIR luminosity. If this relationship changes with redshift (*i.e.*, if there is less CO luminosity for a given amount of FIR luminosity in a galaxy at higher-$z$), that could explain the discrepancy between the COMAP-HETDEX upper limit and the L16-K20 model.

The C22 fiducial model is conditioned on luminosity function constraints from the COLDz CO survey, which is complete only down to $L'_{\rm CO} \sim 10^{10}\,{\rm K\,km\,s^{-1}\,pc^2}$ at $z = 2.5$ (Pavesi et al., 2018), at or above the knee of the CO luminosity function. The faint-end slope is not meaningfully constrained. If the discrepancy with C22 is driven by the average CO luminosity, therefore, the likely cause is that faint CO emitters are less common than expected, and are thus contributing less to the overall CO luminosity density than our modeling predicts. This explanation is consistent with the fact that our upper limit is not significantly in tension with the C22 flat+COLDz model, which uses the same COLDz LF conditioning but a flat prior and thus predicts almost no faint-end contribution to the CO luminosity density.

## 5.9 Summary and conclusions

In this work we perform a stack of the CO LIM measurements from COMAP's Season 2 dataset on the positions of 5879 LAEs surveyed by HETDEX, following the stack on quasars from eBOSS/DESI presented in D26. We find a $1.5\sigma$ nondetection of the stacked CO emission, leading to a 95% upper limit of $\langle L'_{\rm CO}\rangle \leq 2.5 \times 10^{10}\,{\rm K\,km\,s^{-1}\,pc^2}$. After correcting for experimental attenuation factors, we find that this result is in mild to moderate ($3 - 6\sigma$) tension with certain models of COMAP LIM data stacked on the positions of HETDEX LAEs. In particular, we are in tension with the C22 and L16-K20 CO models for every choice of the LAE LF that we adopt.

The stack returns only a single averaged luminosity value, so the average CO luminosity of the Universe is degenerate with more complicated Ly$\alpha$ radiative transfer effects that could decouple the detectability of a galaxy by HETDEX from its clustering properties. These include the stochasticity of the Ly$\alpha$ escape fraction, as well as absorption of Ly$\alpha$ photons by intervening large-scale structure. However, there exist several large LAE surveys, even at $z \sim 3$, that are able to measure these effects directly, and our modeling is largely consistent with these observations. In combination with the fact that the D26 quasar stack (which probes a very different galaxy population with a very different relation to the large-scale structure) is also returning an upper limit that tentatively rules out the brighter CO stack models, we suggest that the most likely explanation for our results is that cosmic CO is slightly fainter than is predicted by the C22 model. This is possible without being in tension with other $z \sim 3$ CO observations. Ultimately, this result can only be confirmed with a direct measurement of the luminosity of the faint end of the CO luminosity function, which will primarily be possible with a CO LIM autocorrelation detection.

*Chapter 6*

# NOVEL CONSTRAINTS ON DARK-MATTER DECAYS USING COMAP

## 6.1 Chapter abstract

One potential solution to the mystery of dark matter's nature is the axion, an ultra-light, long-lived particle which can decay directly into two photons via the axion-photon coupling. In this work, we use Line Intensity Mapping (LIM) to test for radiative decay from axionic dark matter for the first time. We search for radiation with the LIM experiment COMAP, using the HETDEX and DESI galaxy catalogs as tracers of extragalactic dark matter overdensities within the COMAP volume and stacking on the positions of these tracer catalogs. We probe rest frequencies from 88–150 GHz at $2.4 < z < 3.4$ (the COMAP redshift range of interest), corresponding to axion masses $0.7 - 1.0$ meV. We find no evidence for dark matter decay. In our most sensitive frequency range, we place an upper limit on the axion decay rate $\Gamma < 2.9 \times 10^{-31}\,\mathrm{s}^{-1}$. In our mass range, this is the only constraint on radiative dark matter decay directly measured from dark matter itself, as opposed to from other potential axion interactions such as with magnetic fields. This corresponds to a constraint on the axion-photon coupling constant $|g_{a\gamma\gamma}| < 6.7 \times 10^{-9}\,\mathrm{GeV}^{-1}$. We find that stacking of LIM data on the positions of galaxy surveys will be an effective technique for searching for dark-matter decay, especially at lower redshifts and higher frequencies.

## 6.2 Introduction

Despite the overwhelming evidence for its existence, the dark matter (DM) that makes up nearly 30% (e.g., Planck Collaboration et al., 2020) of our universe remains a mystery, with a large family of potential explanations all equally viable. Axions, a potential solution arising out of a solution to the strong charged-parity problem in quantum chromodynamics (QCD) (Peccei and Quinn, 1977a; Peccei and Quinn, 1977b; Weinberg, 1978; Wilczek, 1978), are ultra-light, long-lived particles which can interact with standard-model particles. In particular, axions decay directly to two photons (e.g., Cadamuro and Redondo, 2012), each with frequency $\nu_a = m_a c^2/(4\pi\hbar)$. A broader family of axion-like particle models allows for more variable coupling between axions and photons, but still results in photons produced from axion decay. Axions are thus well-motivated candidates for DM, with clear mechanisms for how to generate the correct abundance today (Abbott and Sikivie, 1983; Preskill et al., 1983; Dine and Fischler, 1983), and obvious astrophysical signatures.

Indeed, many astrophysical searches for axionic DM have already taken place (see, e.g., Di Luzio et al. 2017 or Boddy et al. 2022 for reviews). Although these can take different forms, several are searches for otherwise-unexplained spectral line emission, whether arising from the Milky Way's DM halo (e.g., Strigari, 2013; Pinetti, 2025), the cosmic microwave background (e.g., Bolliet et al., 2021; Capozzi et al., 2023; Goldstein et al., 2025) or other diffuse cosmic backgrounds (e.g., Kohri and Kodama, 2017; Nakayama and Yin, 2022; Carenza et al., 2023; Janish and Pinetti, 2025), or from known DM overdensities such as in the halos of galaxies or galaxy clusters (e.g., Wang et al., 2024; Todarello and Regis, 2025). The full extent of the potential axion parameter space is functionally unconstrained in mass (although with preferred 'windows' set by the phenomenology of the particle, e.g., Di Luzio et al., 2017), so searches for photon signatures of axion decay must span a broad range of wavelengths.

One emerging technique which could be extremely well-suited to searches for photon signatures of axion decay is line intensity mapping (LIM), which is designed to map large cosmological volumes efficiently (see Chang and Lidz, 2026, for a review). Unlike traditional, resolved surveys of galaxies, LIM surveys treat the distribution of galaxies as a three-dimensional fluctuation field, and map this field both spatially and spectrally. This means that LIM experiments are sensitive to all emission from a region without being subject to completeness issues. This makes LIM experiments ideal for searches for faint spectral-line emission, particularly from diffuse fields, and LIM has been put forward as an excellent technique for axion searches (Creque-Sarbinowski and Kamionkowski, 2018; Bernal et al., 2021; Bernal et al., 2022b).

Although typically designed for a known spectral-line signature of galaxy formation or evolution which is expected to dominate its observing band, such as carbon monoxide (CO, e.g., Keating et al., 2015; Cleary et al., 2022) or singly-ionized carbon, ([C II], e.g., Staniszewski et al., 2014; Vieira et al., 2020; CCAT-Prime Collaboration et al., 2023; Karkare et al., 2022b), LIM experiments can be adapted to be spectral-line searches at any frequency by introducing a prior on the regions of the universe from which the signal is originating. This can be done through stacking or cross-correlations with existing surveys of galaxies, both techniques which are established within the LIM field (e.g., Keenan et al., 2022; Dunne et al., 2024; Dunne et al., 2026c; Chen et al., 2025; Amiri et al., 2024). By varying the target rest frequency, one can search for spectral lines. This has the potential to be particularly effective at millimetre and sub-millimetre wavelengths, where it is otherwise difficult to probe large regions at high sensitivity.

In this work, we search for photon signatures of axion decay by stacking cm-wave data from the CO Mapping Array Project (COMAP, Cleary et al., 2022) LIM experiment on the positions of galaxies from the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX, Gebhardt et al., 2021) and the Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration et al., 2025a). We describe the datasets in §6.3 and §6.4, the methodology used to obtain constraints on the axion decays in §6.5, and then present and discuss the results in §6.6-6.8. We use the nine-year WMAP cosmology (a ΛCDM model with $\Omega_\mathrm{m} = 0.286$, $\Omega_\Lambda = 0.714$, $\Omega_\mathrm{b} = 0.047$, and $H_0 = 100h$ km s$^{-1}$ Mpc$^{-1}$ with $h = 0.7$; Hinshaw et al. 2013) for consistency with other COMAP works, as well as the peak-patch simulations used to determine DM masses. All distances are presented as comoving quantities.

## 6.3 COMAP

COMAP is a LIM experiment designed to measure CO(1–0) emission from galaxies in the lead-up to the peak of cosmic star formation at $z \sim 2$ (Cleary et al., 2022). In practice, this means that COMAP provides a highly-sensitive, low spatial-resolution (main beam 4.5′) map of three cosmological fields (covering roughly 12 deg$^2$ total) at frequencies ranging from 26 to 34 GHz and a spectral resolution of 31.25 MHz (thus mapping CO(1-0) emission from redshift 2.4 to 3.4). The coordinates of the three COMAP fields are given in Table 6.1. These were chosen to avoid bright sources of Galactic or extragalactic emission in the observed frequency range.

Table 6.1: The locations of the three fields used in the COMAP Pathfinder survey.

| Field | RA (J2000) | Dec (J2000) |
|---|---|---|
| Field 1 | $01^\mathrm{h}\ 41^\mathrm{m}\ 44$ | 00° 00′ 00″ |
| Field 2 | $11^\mathrm{h}\ 20^\mathrm{m}\ 00$ | 52° 30′ 00″ |
| Field 3 | $15^\mathrm{h}\ 04^\mathrm{m}\ 00$ | 55° 00′ 00″ |

The COMAP Pathfinder telescope is a 19-feed focal plane array that maps each field by scanning across it with constant-elevation scans (Lamb et al., 2022). The resulting time-ordered data undergoes several filtering and calibration steps before being made into three-dimensional data-cubes using a simple binned mapmaker

(Lunde et al., 2024). For this work, we coadd the maps from each feed, and perform the analysis on the resulting data cubes. We make use of the COMAP Season 2 dataset in this work. These data are white-noise limited in each field (Lunde et al., 2024), with a per-voxel (three-dimensional pixel) root-mean-square (RMS) noise of $25 - 50\,\mu$K (Lunde et al., 2024). The sensitivity varies considerably across each field due to the on-sky scanning strategy. For a full description of the COMAP data reduction and analysis process, see Lunde et al., 2024 and Stutzer et al., 2024, as well as the Early Science references within Cleary et al., 2022.

## 6.4 Galaxy catalogues

In order to target DM overdensities at known redshifts within the LIM data from COMAP, we use two different galaxy surveys as reference catalogs. We coadd the COMAP data surrounding the three-dimensional positions of the objects in each of these catalogs.

### HETDEX

HETDEX is a large, blind survey of Lyman $\alpha$ Emitting galaxies (LAEs), intended for constraining dark energy (Gebhardt et al., 2021). HETDEX covers a redshift range of $1.9 < z < 3.5$, and a sky area of 540 $\mathrm{deg}^2$, of which roughly 8 $\mathrm{deg}^2$ overlaps with the COMAP sky coverage. Observations are performed with the VIRUS integral field unit (IFU) on the Hobby-Eberly Telescope (Hill et al., 2021; Hill et al., 2024), which covers the sky sparsely, at a fill factor of roughly $1/4.5$. Spectrally, VIRUS has resolving power $750 < R < 950$, depending on the wavelength covered.

The bulk of the HETDEX data reduction involves isolating sources of spectral line emission from the IFU data, and then identifying those sources as either an LAE or a contaminating line source, such as a low-redshift [OII] emitter. This reduction process is described in detail in Mentuch Cooper et al., 2023 and Davis et al., 2023b. The exact catalog we use follows the choices established in Mentuch Cooper et al., 2026, although we perform additional cuts to limit the catalog to the COMAP fields and to the redshift range $2.4 < z < 3.4$ (the redshift range covered by our modeling, which is necessary for estimating the DM masses we probe). The mean redshift of the HETDEX objects in this range is 2.81, and we show the full redshift distribution in Figure 6.1. As we are using the LAEs only as tracers of the underlying DM, we use only the three-dimensional positions of the detected LAEs, and do not incorporate any information about their flux, etc.

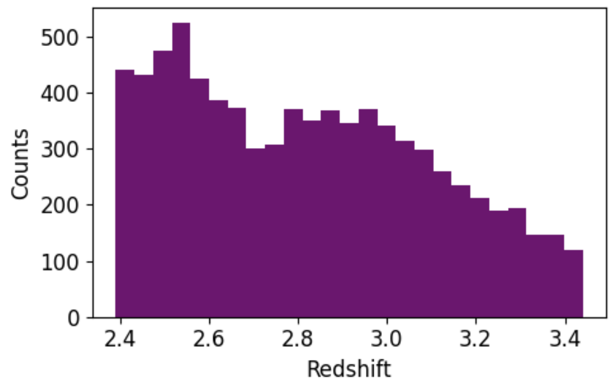


Figure 6.1: Redshift distribution of the HETDEX reference catalog in the COMAP volume.

The primary sources of uncertainty in the three-dimensional positions of the LAEs are both in the spectral axis—the uncertainty in the HETDEX redshifts, and the known systemic offset in the Ly$\alpha$ line due to its resonant radiative transfer. Uncertainties in the spatial position of the HETDEX galaxies are negligible (on the order of a few arcseconds Mentuch Cooper et al., 2023) compared to the 4.5$'$ COMAP beam. We address the systemic Ly$\alpha$ offset by offsetting each object in the HETDEX catalog by 235 km s$^{-1}$ (an offset value measured from HETDEX galaxies themselves Davis et al., 2023a) along the line of sight before carrying out any stacking. The uncertainty in the spectral position of the LAEs arises both from native HETDEX redshift uncertainty (18 km s$^{-1}$, Davis et al., 2023b) and from scatter in the systemic Ly$\alpha$ offset (which we take to be 86 km s$^{-1}$, the standard deviation across galaxies used to measure the Ly$\alpha$ systemic offset in LAEs by Shibuya et al., 2014). We sum these two quantities to estimate a total uncertainty on the LAE redshifts of 104 km s$^{-1}$. This will attenuate the potential spectral-line signal in the stack (Dunne et al., 2025), and we address this effect in our DM modeling (§6.5).

Finally, there is a significant fraction of false positive detections (primarily due to noise peaks being misidentified as LAEs) in the HETDEX catalog. These false positives will attenuate potential signal by averaging regions of the universe that do not correspond to DM overdensities into the stack. With our choice of parameters for HETDEX catalog generation, which matches the choices made in the HETDEX Public Data Release, we expect a false positive fraction $f_{\mathrm{FP}}$ of 0.2 (Mentuch Cooper et al., 2026). As with the velocity uncertainty in the catalog, we account for this effect in our modeling of the DM mass being probed (§6.5).

### DESI

DESI is a multi-object spectrograph mounted on the 4 m Mayall Telescope at Kitt Peak National Observatory DESI Collaboration et al., 2025a. Also intended for finding precise constraints on the equation of state for dark energy, DESI has obtained over 18 million optical spectra of extragalactic objects in its first data release (DR1). DR1 is subdivided into four primary target surveys: luminous red galaxies (LRGs), emission line galaxies (ELGs), quasars (QSOs), and bright galaxies (BGS). In this work, we limit ourselves to only the high-redshift QSOs, but in principle any of the DESI surveys could be used for a stacking-based blind spectral line search and we will do so in future work.

We assemble a QSO catalog from the DESI DR1 by first cutting to the three COMAP fields, and then to the redshift range $2.4 \leq z \leq 3.4$, to remain consistent with our modeling. The redshift distribution of the resulting reference catalog is shown in Figure 6.2. We additionally remove any potential duplicates. The DESI catalog does not suffer from false positive contamination to the same degree as the HETDEX catalog, but the quasar redshifts are considerably more uncertain than the LAE redshifts. This is due to the large in- or outflows common in quasar systems, which can bias the optical lines with respect to the systemic velocity of the host galaxy (and thus the center of mass of the DM halo, e.g., Banerji et al., 2017). As with the LAE offset, we address two components of this uncertainty – a systemic blueshift of $1100\,\mathrm{km\,s^{-1}}$, and an additional scatter in the spectral position of $1600\,\mathrm{km\,s^{-1}}$. Both of these quantities are derived from comparing optically-derived redshifts to CO redshifts for quasars at this redshift range (Dunne et al., 2024), and agree well with other estimates of the same quantity (Tytler and Fan, 1992; Orellana et al., 2011). We address these redshift uncertainties in the same way as the corresponding LAE components, although we also widen the region of the spectral axis over which we integrate in order to account for the much larger QSO redshift uncertainty (§6.5).

## 6.5 Methodology

### Stacking

At each rest frequency, we calculate line luminosity values using LIM stacking, which involves cross referencing galaxy catalogs with a three dimensional LIM data cube, and coadding emissions at the galaxy catalog positions. We source galaxy catalogs from DESI and HETDEX (§6.4), which are used as tracers for emission within the COMAP fields (§6.3). A three dimensional cutout is made around each catalog object in the LIM cube, yielding a ‘cubelet’ encompassing the object’s

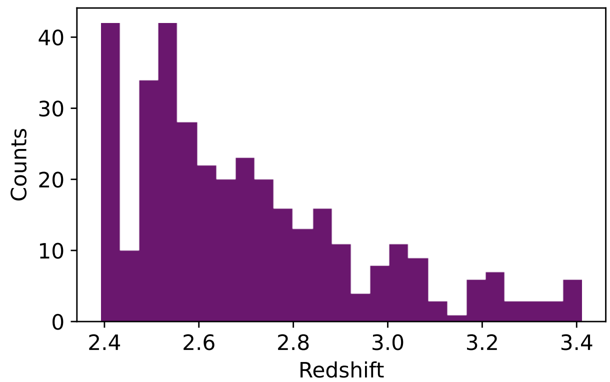


Figure 6.2: Redshift histogram of the DESI QSO reference catalog in the COMAP volume.

surrounding emission, centered on its sky position and redshift. Each of these cubelets is averaged together using inverse-variance weighting, yielding a single stacked cubelet for a corresponding rest frequency. To obtain a one-dimensional stacked spectrum, the emission within a central spatial aperture of $N_{\rm spax} \times N_{\rm spax}$ spatial pixels (spaxels) is summed over, along with a set of $N_{\rm chan}$ spectral channels. For LAEs, we use $N_{\rm spax} = 5$, and $N_{\rm chan} = 3$, while for QSOs, we use $N_{\rm spax} = 5$, and $N_{\rm chan} = 7$, corresponding to $18 \times 18 \times 9$ cMpc and $18 \times 18 \times 21$ cMpc, respectively. The resulting spectrum is integrated over to yield an average line luminosity value associated with the catalog objects. The relative uncertainty of this value is estimated from voxel-level noise in the stack aperture, which is propagated through the stacking pipeline. This stack pipeline is described in more detail in Dunne et al., 2024 and Dunne et al., 2025.

We then repeat the stack process across a spectrum of rest frequencies: 88.4 – 149.6 GHz, chosen to correspond to the COMAP observed frequencies from $2.4 < z < 3.4$. This encompasses the redshift range which we can accurately simulate the mock catalogs for (see §6.5). Galaxies outside this range are affected by significant redshift evolution, meaning that the size and extent of the DM fluctuations will be inaccurately described with our current simulation parameters. The stack parameter choices $N_{\rm spax}$, $N_{\rm chan}$, etc., are also optimized for galaxies in the redshift range $2.4 < z < 3.4$. This frequency range corresponds to a range of axion masses from 0.7 – 1.0 meV, allowing constraints across a large parameter space. We adopt a step size of 0.1 GHz, allowing us to oversample the stack aperture in the spectral axis. Our sensitivity to a given spectral line is improved if the rest frequency we probe is directly on top of the spectral line, so oversampling is beneficial in this case. We note that, due to this oversampling, there is a degree of correlation visible in the line search spectra.

### Conversion into the axion-photon coupling constant

We assume that all DM in the Universe is made of axions of a single mass $m_a$ which decay radiatively into two photons of equal energy $E_a = 2h\nu_a$ at a decay rate $\Gamma$. Thus, a mass $M_{\rm DM}$ of DM releases energy from radiative decay at a rate $\frac{dE}{dt} = \frac{M_{\rm DM}}{m_a} E_a \Gamma$. At a given frequency the measured line flux $F_\nu = \frac{dE}{dt\, dA\, d\Omega\, d\nu}$ becomes

$$F_\nu = \frac{\Omega_{\rm ap} E_a \Sigma_{\rm DM} \Gamma}{4\pi m_a \Delta\nu (1+z)^4}, \tag{6.1}$$

where $\Sigma_{\rm DM} = \int M_{\rm DM} V_{\rm ap}^{-1}\, ds$ is the surface density of DM ($M_{\rm DM}$, divided by the aperture volume $V_{\rm ap}$ and integrated along the line of sight $ds$), $\Omega_{\rm ap}$ is the solid angle subtended by the stack aperture on the sky, and $\Delta\nu$ is the bandwidth of the stack aperture. We describe how we determine $\Sigma_{\rm DM}$ in §6.5 below. The factor of $(1+z)^{-4}$ comes from cosmological surface brightness dimming. This equation is equivalent to that used by Wang et al., 2024 for a similar analysis using DESI stacks. We then obtain a limit on the axion-photon coupling constant $|g_{a\gamma\gamma}|$, from the limit on $\Gamma$ at each rest frequency,

$$\Gamma = \frac{g_{a\gamma\gamma}^2 m_a^3}{64\pi}, \tag{6.2}$$

where $|g_{a\gamma\gamma}|$ is expressed in units of 1/GeV (e.g., Balázs et al., 2022).

### Dark matter mass

The stack aperture is much larger than an individual DM halo, and is also biased by being centered around objects from a galaxy catalog. As such, we obtain the average DM mass in the regions being stacked over using simulations designed for LIM. We stack dark-matter only simulations on the positions of simulated mock catalogs for both LAEs and QSOs. We base these simulations on the simulation pipeline described in Dunne et al., 2025, which is designed for LIM-galaxy catalog joint analyses and is derived from peak-patch N-body simulations (Bond and Myers, 1996; Stein et al., 2019). We alter this pipeline to output maps in terms of DM density rather than line luminosity or brightness temperature. From this underlying dark-matter simulation, two components are generated—a three-dimensional DM fluctuation map with the same dimensions as the LIM data cube, and a mock galaxy catalog.

To generate the DM map, we begin by binning the DM mass of each peak-patch halo into a three-dimensional voxel grid matching the voxel dimensions of the actual COMAP data cubes. We then convolve this binned grid with a $4.5'$-FWHM Gaussian profile in each channel to simulate the COMAP instrumental beam. This corresponds to a FWHM of roughly 8 cMpc at these redshifts. As the peak-patch simulation only captures DM which has been virialized into DM halos, we recapture the missing diffuse DM component by converting the convolved DM mass map into a fluctuation map, and then scaling this to the cosmic critical DM density at the mean redshift of the HETDEX catalog, $\bar{z} = 2.81$,

$$\rho_{\mathrm{DM,\,total}}(\mathbf{x}) = \frac{M_{\mathrm{DM,\,vir}}(\mathbf{x})}{\langle M_{\mathrm{DM,\,vir}} \rangle_{\mathbf{x}}} \rho_{\mathrm{crit}}(\bar{z})\left(\Omega_{\mathrm{m}}(\bar{z}) - \Omega_{\mathrm{b}}(\bar{z})\right). \tag{6.3}$$

As the voxel- and beam-level resolution of the COMAP data cube is so poor, this is functionally similar to extending the simulated DM halos past their virial radii to capture the diffuse DM component. Even the most massive halo in the map has an $R_{\mathrm{vir}} < 400$ kpc, nearly $10\times$ less than the side length of a voxel, so accounting for the increased spatial extent of the DM halos does not meaningfully change the distribution of DM in the map. Finally, we subtract the mean from the density map across each spectral channel and in each spaxel (spatial pixel), to match a high-pass filtering step performed in the actual COMAP data reduction and bring the mass determination more in line with the COMAP transfer function. We generate three mock COMAP fields for each simulation realization. In each field, we apply the actual COMAP mask and RMS distribution to the DM cube. This ensures cosmological sample variance (which depends on the cosmic volume being surveyed) is accurately captured.

The mock galaxy catalog is created by first assigning each simulated DM halo a luminosity (for the LAE catalog, this is a Ly$\alpha$ luminosity, and for the quasars, this is a either a Ly$\alpha$ luminosity or a $g$-band magnitude). In each case, luminosities are assigned by abundance-matching an empirical luminosity function to the DM halo mass function. For our default models we assume a Schechter, 1976 luminosity function,

$$\frac{dn}{dL} = \frac{\phi^*}{L^*}\left(\frac{L}{L^*}\right)^{\alpha} e^{-L/L^*}, \tag{6.4}$$

where $L^*$ is the characteristic ('knee') luminosity, $\phi^*$ is the normalization density, and $\alpha$ is the faint end power law slope, in each case. For LAEs, we use the Ouchi et al., 2020 Ly$\alpha$ luminosity function at $z \approx 3.1$ as our default Schechter parameters. For QSOs, we use the quasar Ly$\alpha$ luminosity function from the Physics of the

Accelerating Universe (PAUS; Torralba-Torregrosa et al., 2024) at $z \approx 3$, as default. The Schechter parameters for each are given in Table 6.2. We then generate a mock galaxy catalog using these luminosity values by applying an observational cut (removing from the catalog all halos below a certain observable cutoff luminosity, $L^{\text{cut}}$, which would be too faint to be observed by the instrument we are simulating), and selecting $N_{\text{obj}}$ halos from the resulting cut catalog. We determine $N_{\text{obj}}$ using the number of catalog objects included in the real stack for each catalog at 110 GHz.

Table 6.2: Schechter parameters used in generating mock galaxy catalogs. LAE parameters are from Ouchi et al., 2020 and QSO parameters are from Torralba-Torregrosa et al., 2024.

| | LAEs | QSOs |
|---|---|---|
| $L^*$ | $8.49 \times 10^{42}$ | $9.77 \times 10^{44}$ |
| $\phi^*$ | $3.9 \times 10^{-4}$ | $1.9 \times 10^{-7}$ |
| $\alpha$ | $-1.8$ | $-1.34$ |

Finally, we account for two potential factors in the catalogs which could attenuate potential signal: the velocity uncertainty in the redshift determinations for both catalogs, and the false positive detections in the HETDEX catalogs. In both cases, the attenuation of potential signal is caused by displacement of the measured 3D positions of the catalog objects away from their real locations within matter overdensities (see Dunne et al., 2025). This is functionally the same as decreasing the DM mass contained within the stack aperture, so we account for sources of attenuation by applying these same displacement effects to the DM mass calculation. The HETDEX and DESI catalogs have random redshift uncertainties of $104\,\text{km}\,\text{s}^{-1}$ and $1600\,\text{km}\,\text{s}^{-1}$, respectively (§6.4), which we account for by randomly scattering the positions of the simulated catalog objects by values pulled from a Gaussian distribution with standard deviation equal to the redshift uncertainty. This yields a 1.1% (55%) suppression of the probed DM mass for the HETDEX (DESI) catalogs respectively, and increases the sample variance by 2.1% (17%), respectively.

Additionally, the HETDEX catalog contains an estimated 20% false positive detections (see §6.4), which will average random locations within the COMAP volume into the stack (rather than the expected overdensities within which LAEs are contained). We account for false positives by replacing a fraction $f_{\text{FP}}$ of the simulated catalog with randomly-generated 3D positions. As expected, this suppresses the measured DM mass by 19.5%. It also adds a small amount (roughly 7%) of additional variance, due to the increased sample variance introduced by the random (*i.e.*, unbiased) catalog sample.

For each of the two catalogs we target, we generate 50 simulation realizations, each encompassing the three COMAP fields. We then stack the DM cube on the positions of the objects in the mock galaxy catalog, following the procedure described in §6.5. We stack on only a single rest frequency value (at 110 GHz), rather than performing the full spectral line search, and extrapolate this value to each rest frequency. We average across the 50 realizations to obtain the average DM mass being probed by the stacks, and take the standard deviation across realizations to be the sample variance. We find an average DM mass of $M_{\rm DM} = (1.6 \pm 0.1) \times 10^{15}\,{\rm M}_\odot$ for the HETDEX-like catalog and $M_{\rm DM} = (4.4 \pm 1.2) \times 10^{15}\,{\rm M}_\odot$ for the DESI-like QSO catalog. After dividing by the area of the stack aperture in the two perpendicular axes (13.2 cMpc$^2$ at 110 GHz), this yields a DM surface density $\Sigma_{\rm DM}$ of $(2.5 \pm 0.2) \times 10^{-2}\,{\rm g\,cm^{-2}}$ for the HETDEX-like catalog and $(6.9 \pm 1.9) \times 10^{-2}\,{\rm g\,cm^{-2}}$ for the DESI QSO-like catalog.

The use of simulations to determine the average DM mass being probed does introduce a model-dependence to our determined limits on $\Gamma$ and $|g_{a\gamma\gamma}|$, as different prescriptions for assigning luminosities to halo masses may result in differently-biased catalogs, and thus different values for the average DM surface density. We test two other LAE models, based on the luminosity functions determined in Zhang et al., 2021 and Umeda et al., 2025, and find that the $\Sigma_{\rm DM}$ value varies by 7-11% depending on the model chosen. Similarly, we find a 20% offset using the quasar luminosity function from Palanque-Delabrouille et al., 2016. We take the maximum model-based offset for each galaxy type, and treat these as additional sources of uncertainty on the DM mass probed (*i.e.*, we scale $\sigma_M$ by these values).

## 6.6 Results

We show the line search spectra for each galaxy catalog in Figure 6.3, as well as their associated S/N values. In both cases, the spectra are consistent with white noise—we do not detect any unexplained spectral line emission (or indeed, any of the known molecular lines in this rest frequency range). As described in §6.5, the step size used for our search is less than the size of the frequency aperture we integrate across to obtain the stack values, so adjacent points are not statistically independent. This aperture is 93.35 MHz in observed frequency space, corresponding to a rest-frame frequency width of roughly 0.35 GHz. There are therefore 157 independent

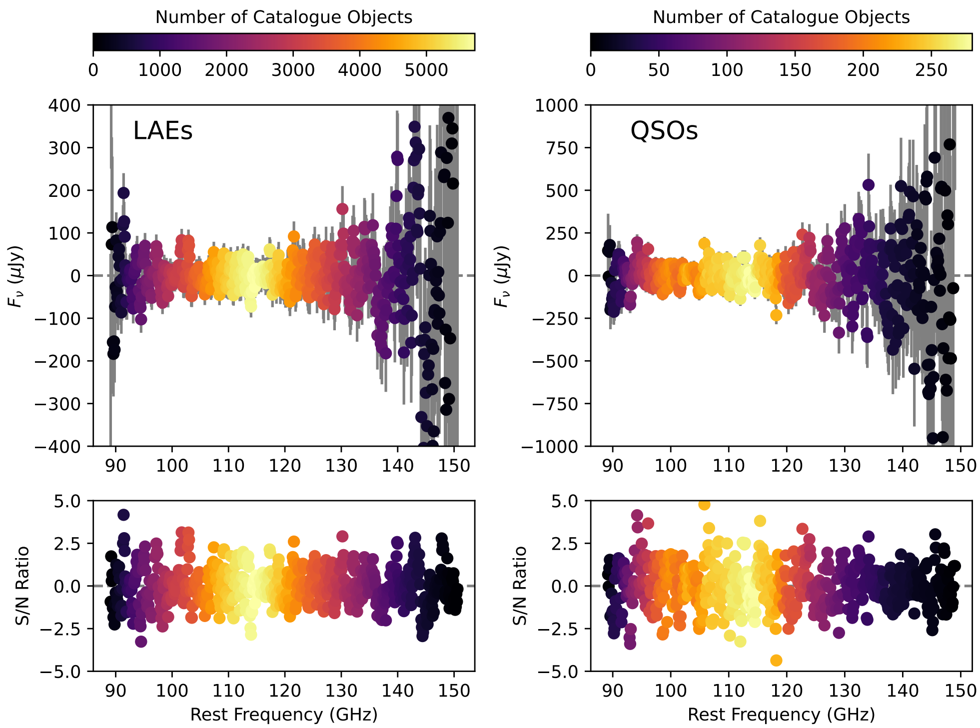


Figure 6.3: Line search spectra for the HETDEX LAE catalog (left) and the DESI QSO catalog (right). Points are colored by the number of catalog objects available. The S/N ratio for each point is shown in the bottom panel. We find no evidence for spectral line emission in either catalog.

measurements possible within our frequency range. After correcting for this, the highest $p$-value we find for a spectral line in either stack is $p \leq 0.006$. At the most sensitive rest frequency values (between $110 - 120$ GHz for both catalogs), we obtain $2\sigma$ (95%) flux density limits $< 50\,\mu$Jy for the LAE catalog and $< 75\,\mu$Jy for the QSO catalog.

We then convert these flux limits into limits on the two-photon axion decay rate from DM ($\Gamma$) at each rest frequency by propagating through both the uncertainty on the measured value of $F_\nu$ from the COMAP data and the uncertainty on $\Sigma_{\rm DM}$ (§6.5), which comes from cosmological sample variance and the various sources of signal attenuation described in §6.4 and §6.5. These limits are shown in Figure 6.4. In the same highest-sensitivity range ($m_a \sim 0.95$ meV), we obtain 95% upper limits of $\Gamma < 3.5 \times 10^{-31}$ s$^{-1}$ for the LAE catalog and $\Gamma < 5.9 \times 10^{-31}$ s$^{-1}$ for the quasars. While the limit from the quasar catalog is probing a higher DM surface density (due to the higher bias of the quasars), the LAE catalog provides the more constraining limit. This is because there is more than an order of magnitude more objects in

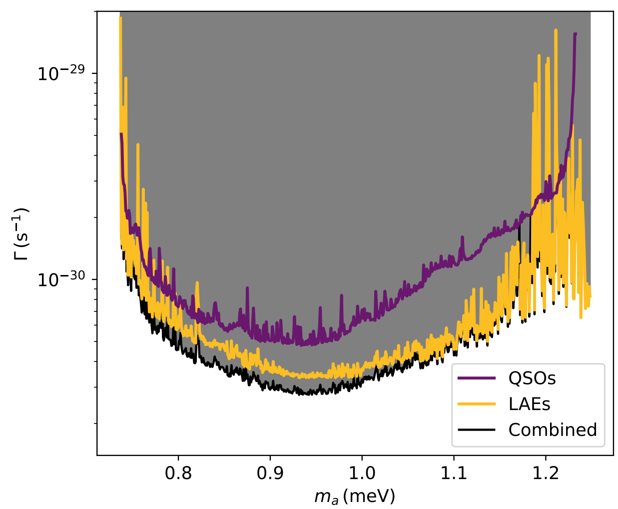


Figure 6.4: Limits on the decay rate of DM axions from COMAP stacking. Limits from the DESI QSO catalog alone are shown in purple, the HETDEX LAE catalog alone are shown in orange, and the combined constraints from both catalogs are shown in black.

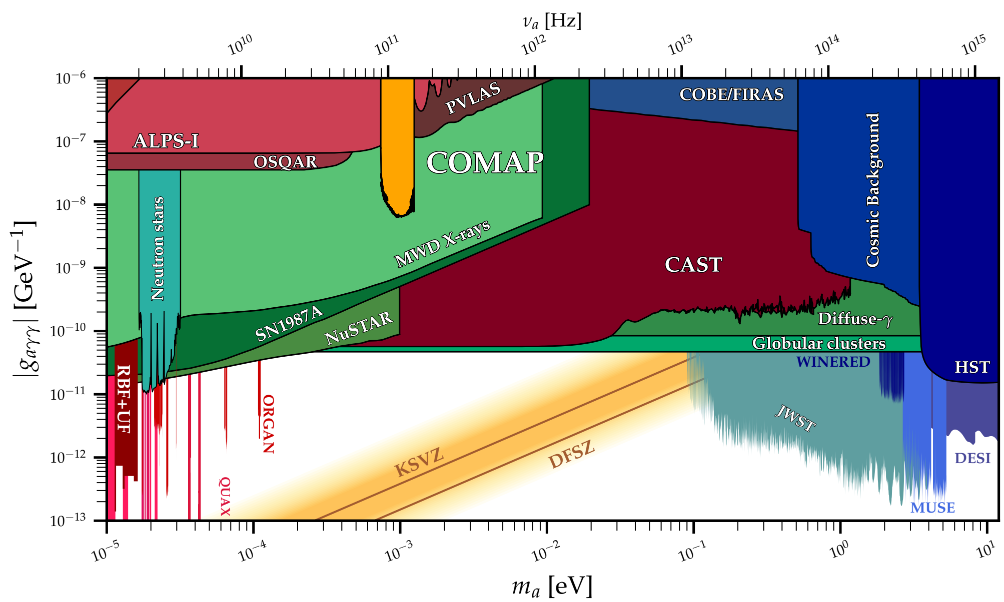


Figure 6.5: The COMAP constraint (in orange) on the axion-photon coupling constant $|g_{a\gamma\gamma}|$ in comparison with other existing constraints. The orange lines correspond to the two model predictions from Kim, 1979; Shifman et al., 1980 and Dine et al., 1981; Zhitnitskii, 1980. We use the `AxionLimits` package (O'Hare, 2020) to compile and plot constraints. The full set of references are listed therein.

the HETDEX catalog than the DESI QSO catalog, which decreases the uncertainty on the input line flux measurements considerably. Our most constraining limit is from combining the two catalogs' constraints. There, we obtain an upper limit on the axion decay rate of $2.9 \times 10^{-31}\,\mathrm{s}^{-1}$, or a particle half-life of $\tau > 3.4 \times 10^{30}\,\mathrm{s}$. A particle with a decay timescale equivalent to the age of the Universe would have a decay rate $\sim 2.3 \times 10^{-18}\,\mathrm{s}^{-1}$, so this implies that only one particle in $10^{13}$ would have decayed in the lifetime of the observable Universe.

Finally, we convert these limits into direct limits on the axion-photon coupling constant, $|g_{a\gamma\gamma}|$. At the most sensitive frequencies, we obtain 95% upper limits of $|g_{a\gamma\gamma}| < 6.7 \times 10^{-9}\,\mathrm{GeV}^{-1}$ using the combined limits from the two catalogs. The full COMAP limit as a function of axion mass is shown in Figure 6.5.

## 6.7 Validation

Previously, when presenting uncertainties associated with a stack, we have determined those uncertainties through bootstrapping, *i.e.*, performing the same stack repeatedly on randomized catalogs and capturing the distribution of output line luminosity values. This method captures not only the inherent uncertainty due to noise in the map, but also naturally incorporates pixel-pixel covariance, and any additional spread attributable to possible systematic errors in the LIM data. However, bootstrapping is computationally expensive (it requires performing $O(10^3)$ stacks) for a given rest frequency, and would be much more expensive for this application, which requires scanning over rest frequencies. For the line-search spectra in this work we instead propagate map-level uncertainties through the stack pipeline and report these propagated uncertainties at each rest frequency.

The COMAP Season 2 data have been tested several times and found to be close to white noise in each case (e.g., Lunde et al., 2024; Dunne et al., 2026c). Still, in order to ensure that these propagated uncertainties are correctly reflecting the full potential variation at each rest frequency, we cross-check the propagated uncertainties against bootstrapped uncertainties for a 1 GHz chunk of the LAE line-search spectrum, from 100 to 101 GHz. We follow the bootstrapping method outlined in Dunne et al., 2026c, where we use the actual COMAP data, but generate a random catalog following the spatial distribution of the real galaxy catalog for each bootstrapping realization. For the QSO stack, we simply add a random offset in all three dimensions (between 1 and 5 spaxels/channels) to the position of each QSO object. For the LAE stack, we use HETDEX-generated random catalogs (Davis et al., 2023b). We perform 1000 bootstrap realizations at each of the 11 rest frequency values tested in this 1 GHz chunk.

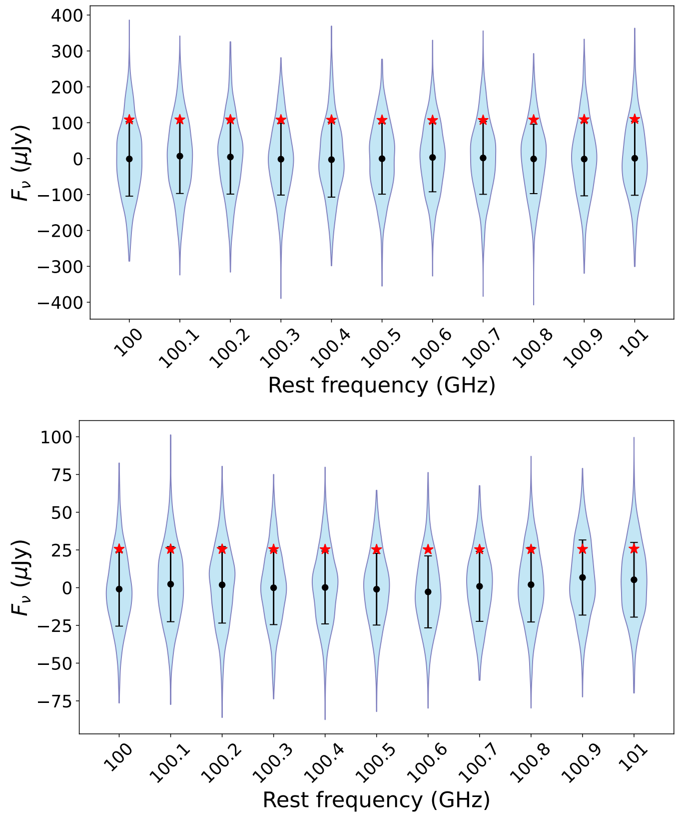


Figure 6.6: Comparison between bootstrapped and propagated uncertainties for the QSO stack (top) and for the LAE stack (bottom), in a 1 GHz region. The blue shaded regions show the full distribution of output $\langle L'_{\rm CO}\rangle$ values from the bootstraps in each channel, with the $1\sigma$ (68.2%) region shown in black. The propagated uncertainty for the actual stack performed in each channel is shown with the red stars.

In both cases, we find excellent agreement between uncertainties derived from bootstrapping and those derived from propagating the map-level RMS noise through the stack pipeline. We show the propagated errors against the bootstrapped values in the 1 GHz range in Figure 6.6. On average, the two uncertainty values agree within 4.3% for the LAEs and 6.9% for the QSOs. This suggests that the propagated uncertainties used to derive the $|g_{a\gamma\gamma}|$ limit are accurately reflecting the total uncertainty in the stacks, including pixel-to-pixel covariance and residual systematic error. This is also confirmation that the covariance between spectral channels driven by the size of the stack aperture (§6.5) is not affecting the noise properties of our axion constraint.

## 6.8 Discussion

We present a novel technique for testing axions as a DM candidate at high redshift, searching for two-photon decay from DM overdensities by stacking LIM data on the positions of known galaxies. We find no evidence for spectral line emission at the level of the COMAP sensitivity, placing an upper limit on the axion decay rate $\Gamma$ of $2.9 \times 10^{-31}\,\mathrm{s}^{-1}$ at axion masses $\sim 6\,\mathrm{meV}$. This technique is most similar to the searches for dark-matter decay performed by stacking galaxy spectra using HST, JWST, DESI, or the VLT (e.g., Todarello et al., 2024; Todarello and Regis, 2025; Saha et al., 2025; Wang et al., 2024; Grin et al., 2007; Regis et al., 2021), although these searches tend to be limited to the infrared through ultraviolet part of the electromagnetic spectrum. In this mass range, this is the only constraint available on radiative DM decay.

The COMAP limit on $\Gamma$ corresponds to a constraint on the axion-photon coupling constant of $|g_{a\gamma\gamma}| < 6.7 \times 10^{-9}\,\mathrm{GeV}^{-1}$, assuming all DM is made of axions which undergo two-photon decay. We show this constraint in Figure 6.5 in comparison with other limits obtained using a wide variety of different techniques (see O'Hare, 2020, and the included references). In the axion mass range that we probe, the state-of-the-art limits are from analysing stellar populations in globular clusters (Dolan et al., 2022) and from the CERN Axion Solar Telescope (CAST, Andriamonje et al., 2007; Anastassopoulos et al., 2017; Melcón et al., 2020; Altenmüller et al., 2024). Other strong limits in this mass range include the limits from Manzari et al., 2024 who search for a gamma-ray signature due to axions formed in the progenitor star of SN1987a, and Dessert et al., 2022 who use *Chandra* to search for ALPs formed in the core of a magnetic white dwarf and converted to x-rays in its magnetosphere, as well as Ruz et al., 2025, who use NuSTAR to search for axions converted similarly in the magnetosphere of the Sun. None of these constraints, however, are directly searching for radiative decay from DM itself. Our COMAP-based limit is the only constraint in this mass range on axions as a DM candidate.

As a technique for constraining $|g_{a\gamma\gamma}|$, stacking LIM data on high-redshift galaxy catalogs is not particularly sensitive. It is unlikely that COMAP will be able to produce a unique constraint with the experimental parameters in this work—following the radiometer equation, the line flux sensitivity $\sigma_{F_\nu}$ scales with integration time $t_{\mathrm{obs}}$ as $\sigma_{F_\nu} \propto (t_{\mathrm{obs}})^{-1/2}$, so $|g_{a\gamma\gamma}|$ scales with integration time as $\sigma_{g_{a\gamma\gamma}} \propto t_{\mathrm{obs}}^{-1/4}$. This weak $\sigma^{-1/4}$ scaling also applies to the other obvious way to improve the sensitivity of a stack, which is increasing the number of objects in the galaxy catalog, $N_{\mathrm{obj}}$. To produce a limit on $|g_{a\gamma\gamma}|$ which improves on that from the globular clusters, this stack would require the product $(t_{\mathrm{obs}}N_{\mathrm{obj}})$ to improve by a factor of roughly $10^{10}$.

An obvious direction for future work is to perform the same analysis on lower-redshift galaxy catalogs. These will provide a sensitivity improvement in two immediate ways: the effect of cosmological surface brightness dimming (the factor of $(1+z)^{-4}$ in Equation 6.1) will be considerably lessened (although this is somewhat offset by the lower rest frequencies being probed resulting in smaller axion masses—Equation 6.1 is also proportional to $\nu_{\mathrm{a}}^3$ through $m_{\mathrm{a}}$ and $E_{\mathrm{a}}$), and galaxy catalogs are more common and more complete at these redshifts than at $z \sim 3$, so $N_{\mathrm{obj}}$ will increase. Using the lower-redshift DESI catalogs and the [O II] emitters from HETDEX as reference catalogs for a COMAP stack, these two factors will likely improve the $|g_{a\gamma\gamma}|$ limit by more than a factor of 4. The value of $\Sigma_{\mathrm{DM}}$ will also likely increase (also improving the $|g_{a\gamma\gamma}|$ limit) in the lower-redshift Universe, which is more highly clustered. The degree to which $\Sigma_{\mathrm{DM}}$ will increase is beyond the scope of this work. Transitioning to For the same reasons, LIM experiments targeting higher-frequency spectral lines and lower redshifts, such as TIM (Vieira et al., 2020), SPT-SLIM (Karkare et al., 2022b), SPHEREx (Doré et al., 2014), or HETDEX itself (Gebhardt et al., 2021) will likely provide the best stacking-based LIM constraints on $|g_{a\gamma\gamma}|$.

Finally, other strategies have been proposed for providing constraints on radiative axion decay by looking for interloper emission in LIM data. Bernal et al., 2021 suggest searching for (extragalactic) radiative decay from axions using power spectrum projection effects or the shape of the voxel intensity distribution (VID). In this work, they forecast 95% upper limits on $\Gamma$ from the COMAP power spectrum and VID that are $\sim 10^{-29} - 10^{-30}\,\mathrm{s}^{-1}$, broadly consistent with the limits we obtain in this work. Galactic axion searches can also be performed by observing a region of sky that is (near-)blank extragalactically, and searching for line emission from the Milky Way's own DM halo. Because the COMAP data is high-pass filtered both spatially and spectrally to mitigate atmospheric contamination, such searches are not possible with COMAP, but could be viable with LIM experiments that do not high-pass filter spatially.

## 6.9 Summary and conclusions

In this work, we explore a novel strategy for studying DM decay through LIM stacking, and apply this strategy to data from the COMAP experiment. Our technique utilizes existing galaxy catalogs as tracers for DM overdensities within the COMAP LIM data. We coadd the LIM data at the positions of the galaxies to perform a blind spectral-line search at a range of rest frequencies $88.4 - 149.6$GHz, corresponding to axion masses $0.7 - 1.0$ meV. We find no evidence for spectral-line emission in this frequency range. Assuming all DM is made of axion particles which undergo two-photon decay, we find an upper limit on the axion decay rate of $\Gamma < 2.9 \times 10^{-31}\,\mathrm{s}^{-1}$. We use these values to obtain constraints on the axion-photon coupling constant $|g_{a\gamma\gamma}| < 6.7 \times 10^{-9}\,\mathrm{GeV}^{-1}$. These are the first observational constraints to be placed on axionic DM using LIM.

As a constraint on axionic DM, this is the deepest limit currently available for this axion mass range. With the inclusion of lower-redshift galaxies (for which observational data are readily available), we expect to improve this limit further in future work, after adjusting both our determination of $\Sigma_{\mathrm{DM}}$ and of the optimal stack parameters to the lower redshifts. While COMAP itself is unlikely to obtain a unique limit on $|g_{a\gamma\gamma}|$, stacking-based LIM analyses in general are likely to be highly competitive. Our stack-based limit is comparable to those forecast for the power spectrum and VID by Bernal et al., 2021 for COMAP. We expect stacking to be a competitive strategy for axion searches using other LIM experiments, particularly those targeting higher-frequency, lower-redshift spectral lines.

*Chapter 7*

# SUMMARY AND PATH FORWARD

In this thesis, I have presented the development, testing, and application of a 3D stacking pipeline for COMAP. This includes robust treatments for the uncertainty on the stack statistic (Chapter 2), the transfer function of the LIM experiment (Chapter 4), cosmological sample variance (Chapter 4), and a variety of other experimental attenuation factors (Chapters 4 and 5). In service of this effort, I have also built a simulation pipeline designed for LIM-galaxy joint analyses, and used it to develop a theoretical understanding of this novel statistic. While these pipelines were developed for COMAP, they could each be used for any of the rich landscape of upcoming LIM experiments. With these tools, I have tested models of the molecular gas fueling star formation at a key epoch in the Universe's history. The main results of the thesis are as follows:

1. I have shown that the signal in a LIM stack is completely dominated by the clustering of neighboring objects around the galaxies being stacked upon (Chapter 3; Dunne et al., 2025), with more than 200 other DM halos falling into the stack aperture for a stack on an LAE-like catalog with COMAP resolution. This conclusion has wide-reaching applications for stacking analyses. In particular, the stack signal is extended beyond a point-source profile (important for selecting a stack aperture size to maximize S/N), and the magnitude of the stack signal is driven in large part by the bias of the reference catalog (important for selecting a galaxy catalog to use in a stack).
2. I have found a mild to moderate ($2-5\,\sigma$) tension between simulations and observations for the upper limits derived from COMAP stacks on eBOSS/DESI quasars (Chapter 4; Dunne et al., 2026c) and HETDEX LAEs (Chapter 5). This tension holds even after accounting for a variety of experimental factors that could serve to systematically attenuate the stacked signal or increase the variance. In the stack, the overall cosmic CO luminosity is degenerate with any factor (linear or non-linear) that could affect the clustering properties of the reference catalog. This tension, thus, could be attributable either to incorrect modeling of the clustering (or the galaxy-halo connection) in the simple mock datasets, or to the average CO luminosity of the Universe being fainter

than models suggest. If the latter, this is likely due to the faint end of the CO luminosity function (which is nearly entirely unconstrained by observations $z \sim 3$) being extrapolated to be too bright by models.

3. I have found a novel application for the stack in the form of a search for two-photon emission from decaying axionic dark matter (Chapter 6). By stacking COMAP data on DESI and HETDEX, I have placed constraints on the axion decay rate $\Gamma$ and the axion-photon coupling constant $|g_{a\gamma\gamma}|$. In the axion mass range to which COMAP is sensitive, these are the strictest available constraints on an axion model of dark matter.

This work has aimed to build a foundation for a future of stacking analyses for the entire LIM field. There are many directions that stacking could take in the future, both within COMAP and across the LIM field as a whole. Below, I discuss some promising future directions for stacking and other joint analyses with COMAP.

## 7.1 Future datasets

The datasets I use for stacking applications in this thesis are by no means exhaustive. Additionally, near future releases of new datasets (or improvements to existing ones) have the potential to greatly improve the stack sensitivity. In this section, I discuss some promising possibilities, both for reference galaxy catalogs and for COMAP itself.

### Galaxy catalogs

One of the easiest ways to obtain new stack-based constraints with a given LIM experiment is simply to fold in more galaxies, as the stack sensitivity increases $\propto \sqrt{N_{\mathrm{obj}}}$. Especially in this era where large, public galaxy surveys are being released, there is a wealth of data available. Unfortunately, for the purposes of COMAP, there are two primary factors limiting the utility of a dataset for the stack. Firstly, spectroscopic galaxy catalogs are vastly preferable (while it is possible to stack on a photometric catalog, the sensitivity is greatly decreased per catalog object because the signal is spread over a larger noise floor, and the COMAP pipeline filters out any signal correlated over more than $\sim$ a GHz, Chapter 3). Secondly, the catalog must be coincident with the COMAP on-sky coverage and redshift range. The number densities of most galaxy catalogs drop off after $z \sim 2$, so this can also be a challenge. These criteria rule out some surveys that might otherwise be prime reference partners, such as the Roman Space Telescope HLSS survey (Wang et al.,

2022) and the *Euclid* WFS (Euclid Collaboration et al., 2022). Still, there are some future prospects for galaxy catalogs.

1. The DESI survey has not yet released its full catalog. The QSO catalog used in Chapter 4 is from DESI's first data release (DESI Collaboration et al., 2025a). DESI's DR2, which is not yet public, will have nearly twice as many quasars available (DESI Collaboration et al., 2025b; DESI Collaboration et al., 2025c). If coverage in the COMAP fields doubles, this would improve the sensitivity of the quasar stack to models by more than a factor of 1.4.

2. HETDEX, which is by far the largest spectroscopic survey available in the COMAP volume, still covers the sky very sparsely. There are also no HETDEX observations in COMAP's Field 3. We have already begun to fill in the sparse HETDEX coverage in Field 1 (the additional observations included in the stack in Chapter 5). We will continue to propose to fill in the sparse HETDEX coverage in Field 2, and to begin to observe Field 3. Optimistically, more observations could improve the HETDEX $N_{\rm obj}$ by 30 – 50%.

**COMAP datasets optimized for stacking**

The COMAP data reduction pipeline is optimized for the auto-correlation power spectrum, and is thus designed to remove as much contamination as possible across all spatial scales at the cost of data retention. Because the stack traces only a defined range of spatial scales, it will likely be possible to loosen some of the cuts made in the filtering and mapmaking steps, thus including additional data without introducing instrumental systematic errors into the final statistic. In particular, these could include cuts on weather, distance to the Sun or Moon, or other cuts based on TOD or map-level summary statistics. Currently, these remove roughly 75% of the raw COMAP data (Lunde et al., 2024), so there is great potential to increase the data volume contributing to the COMAP datacubes.

**COMAP-*wide***

The next iteration of COMAP (and the biggest imminent step forward for COMAP stacking) will be COMAP-*wide*. This is a near-duplicate of the Pathfinder telescope designed to study the EoR through cross-correlation of the CO(2–1) transition (which is redshifted to $\nu_{\rm obs} \sim 30\,$GHz at $z \sim 7$) with 21 cm data from the LOw Frequency Array (van Haarlem et al., 2013). COMAP-*wide* is currently under construction and should be on-sky by the end of 2026. To better match coverage in Fourier space with the 21 cm experiment, COMAP-*wide* will cover a wider field than any of the Pathfinder fields (up to 100 deg$^2$) at the North Ecliptic Pole (NEP). The proposed coverage is shown in Figure 7.1. In addition to the COMAP-LOFAR cross-correlation, which I discuss in §7.3, this new, wider, field will provide several opportunities for additional joint analyses with galaxy catalogs $z \sim 3$.

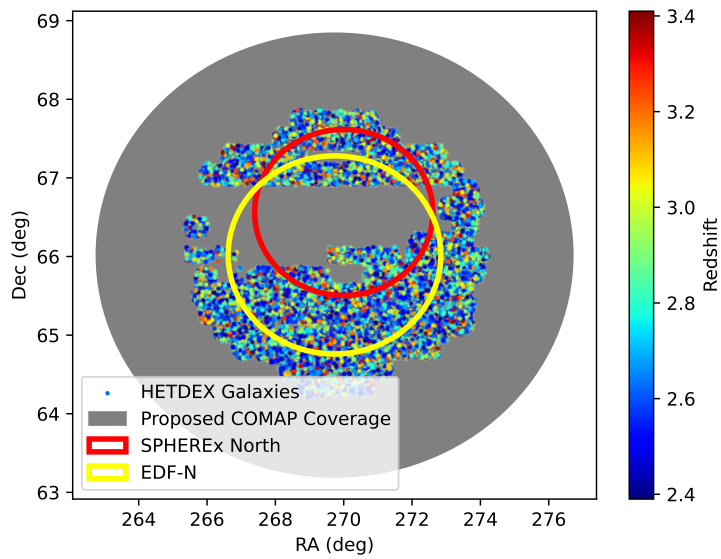


Figure 7.1: Proposed COMAP-*wide* survey area at the NEP, together with potential cross-correlation partners. HETDEX has $\sim 45,000$ additional LAEs within this area, and the SPHEREx Northern deep field also covers much of the same volume.

When discussing prospects for improving stack sensitivity with a wider-field survey, it is important to note that the stack sensitivity is constant with survey area for a fixed integration time $t_{\text{obs}}$ and on-sky reference catalog number density. This is because the number of objects in the stack increases $\propto \sqrt{\Omega_{\text{survey}}}$, but at fixed $t_{\text{obs}}$ the observing time spent in any given resolution element ($\tau$) decreases by the same factor. The stack RMS uncertainty $\sigma$ is $\propto (\tau N_{\text{obj}})^{-1/2}$. Thus, COMAP-*wide* will enable stacks on many more catalog objects, but will not improve the stack sensitivity over the COMAP Pathfinder by $\sqrt{N_{\text{obj}}}$ directly.

The NEP is a popular field for extragalactic galaxy surveys, some of which are shown in Figure 7.1. The eBOSS and DESI quasar surveys both overlap with this field. There is considerable HETDEX coverage (more than 40,000 LAEs) in a 10 deg$^2$ field centered on the NEP through the TESLA (Chávez Ortiz et al., 2023) survey. This subset of the HETDEX catalog also has denser coverage than the typical 1/4.5 HETDEX fill factor, and a wide variety of ancillary data, making it ideal as a stack partner. The *Euclid* Deep Field North (EDF-N; Euclid Collaboration et al., 2025) is also contained within this field, and will have both a large catalog of high-precision photometric redshifts that overlaps with the lower end of the COMAP redshift range ($z \lesssim 2.5$) and spectroscopic observations via *Euclid*'s NISP instrument. The SPHEREx Northern deep field is also in this patch. SPHEREx is spectrophotometric, and its spectral resolution approaches the lower limit of the COMAP resolution. Still, SPHEREx will provide a large catalog of H$\alpha$ emission from galaxies $z < 3$ (Doré et al., 2018), and with some adjustment to the low-level COMAP data analysis there is high potential for either a stacking or cross-correlation based joint analysis with these data.

## 7.2 Stacking methodology improvements

The stacking pipeline I use throughout this thesis is extremely simple, and there are also opportunities to improve the sensitivity of stacking analyses through more sophisticated methodologies. In this section, I discuss potential future directions for improving the stack pipeline.

### Matched filtering

In a recently published work (Mansfield et al., 2025), I co-supervised a student through exploring how 2D and 3D matched-filtering techniques could improve the stack sensitivity. These techniques (similar to forced photometry in optical/IR parlance) weight voxels that contain more signal more highly, by a assuming a 2- or 3D distribution of signal and fitting the stack cubelet to this distribution. This is in contrast to early stacking efforts, where I simply summed over a rectangular aperture when determining the final $\langle L'_{\rm CO} \rangle$ from a cubelet. In Mansfield et al. (2025), we tested the efficacy of this technique on the stack by matched filtering simulations from `joint_limlam_mocker` to known functional forms (Gaussian or Lorentzian functions). We found that the S/N of a stack can be improved by more than 20% when fitting a function that is well-matched to the stack signal(Figure 7.2).

The COMAP Season 2 stacks (Chapters 4 and 5) include a version of this matched-filtering technique that is used to simultaneously correct for the pipeline transfer function. What has not yet been implemented is the sub-pixel centering that we developed in Mansfield et al. (2025)—instead of coadding the cutouts from every galaxy included in the stack and then fitting a 3D profile, we fit the profile to

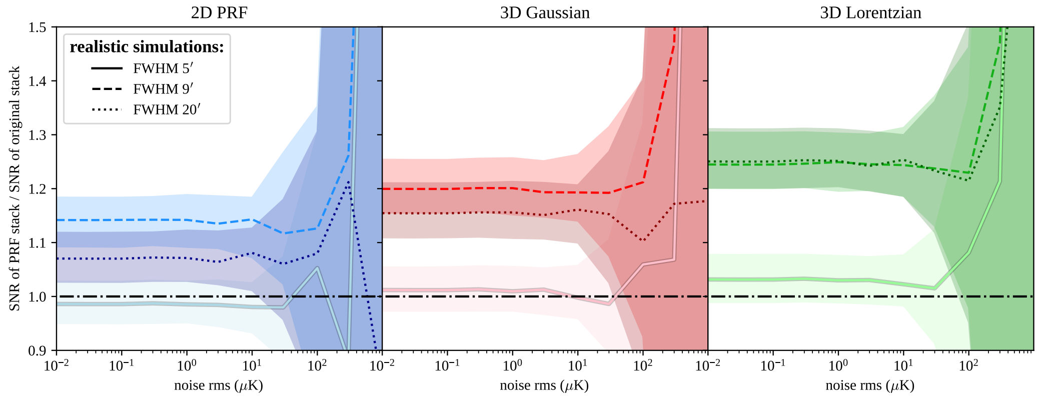


Figure 7.2: Relative S/N of a simulated stack $\langle L'_{\rm CO} \rangle$ from three matched-filtering techniques in comparison to the simple summed-aperture version we use in Chapters 2 and 3, shown as a function of map-level noise RMS. From left to right, stack cutouts are fit to a Gaussian profile in the spatial axes only, a spatial-spectral Gaussian profile, and a Gaussian profile in the spatial axes and a Lorentzian profile in the spectral axis (to better for the fingers-of-God effect). Matched filtering can improve the S/N by more than 20%. See details in Mansfield et al. (2025).

each galaxy individually, centered on the precise location of the galaxy within the COMAP pixel grid, and then average together the resulting fit amplitudes. Because the COMAP voxels are considerably larger than the photometric precision of any reference catalog we use, this technique is able to prevent attenuation due to imprecisions in the profile centroid (see Chapter 3). In Mansfield et al. (2025), we found a $\sim$ 6% improvement to the stack S/N when centering profiles at the sub-pixel level with the COMAP resolution. I will continue to work to apply this sub-pixel centering to stacks on actual data.

### Simultaneous stacking

'Simultaneous stacking' is a method that has been used successfully with 2D fluctuation tracers such as the cosmic infrared background (Viero et al., 2013; Viero et al., 2015; Viero et al., 2022; Agrawal et al., 2026). With this technique, each of the sources in a reference galaxy catalog are fit simultaneously to the fluctuation map. If the catalog is highly complete, then the average properties of the sources in the galaxy catalog can be determined in a way that accounts for the effects of source confusion. Catalogued galaxies can also be binned by properties such as redshift (allowing redshift evolution to be extracted from a projected 2D field) or mass (allowing for the change in average luminosity with mass to be measured).

This is similar to the cross-correlation-based tomographic analyses of Chiang et al. (2025) and Chiang (2026).

LIM stacks have inherent access to line-of-sight information, but there are still potential benefits to a simultaneous stacking methodology: it is much more efficient than fitting each cutout individually, and with a highly complete galaxy survey it could undo the clustering/confusion effects that are driving the stacks in this thesis. Binning by galaxy properties such as stellar mass would allow for measurements of the relative contribution from different galaxy types to the overall CO luminosity of the Universe. I am working with a collaborator (Prof. Dongwoo Chung) and a PhD student (Megan Shultze) at Cornell University to develop a simultaneous stacking pipeline for COMAP LIM data.

#### Weighting

For each of the stack analyses in this thesis, cutouts of the LIM data are coadded together weighted only by the RMS noise in the LIM map, using a simple inverse-variance scheme (following, for example, the HI stacks of Fabello et al., 2011). Under the assumption of inhomogeneous but Gaussian noise across the LIM dataset this optimizes the noise properties of the stack, but does nothing to maximize the stacked signal. Other 21 cm stacks have explored an additional weighting by some power-law function of the luminosity distance, and found that this positively impacts the S/N (Delhaize et al., 2013; Hu et al., 2019; Sinigaglia et al., 2022). Other ancillary information available in the galaxy catalog, such as the luminosity of a spectral line or even fitted quantities such as the galaxy's stellar mass, could be used to weight galaxies that we expect to be more clustered more highly in a stack.

### 7.3 Cross-correlation

While the stack has been practical as an analysis technique in the early stages of COMAP, the ultimate target for LIM-galaxy joint analysis is a full cross-correlation. The cross-correlation folds in information from all spatial scales, so the S/N is higher than the stack S/N (when integrated across the full cross-spectrum), and measurements at different spatial scales make it easier to fit models. This helps to break the degeneracies that are present in the stack alone. In this section, I discuss the most immediate prospects for cross-correlations with COMAP.

**COMAP × the Ly$\alpha$ forest**

An intermediate step between the stack and a full 3D cross-correlation is a cross-correlation with the Ly$\alpha$ forest, In this scheme, spectra along lines of sight to eBOSS background quasars can be extracted from the COMAP LIM data and cross-correlated with the flux transmission spectra of the Ly$\alpha$ forest. Ly$\alpha$ forest fluctuations result from absorption of the quasar continuum by HI regions along the line of sight. The CO signal is in emission, so the cross-correlation signal should be negative at small scales. This strategy was used by Amiri et al. (2024) to obtain a high-redshift detection of the 21 cm line with CHIME. There are roughly 300 quasar spectra available in the eBOSS Ly$\alpha$ forest catalog (du Mas des Bourboux et al., 2020) as cross-correlation partners. Together with a collaborator (Prof. Junhan Kim) and two PhD students (MinSu Ko and HeeBeom Yang) at KAIST, I am working to carry out this cross-correlation. We have produced cross-power spectra and are currently in the process of characterizing the transfer function and the uncertainty.

**COMAP × HETDEX LAEs**

HETDEX is the largest survey available in our fields and redshift range, and the COMAP fields were originally chosen to overlap with HETDEX for the purposes of cross-correlation. This is the obvious next step for COMAP joint analyses, and a HETDEX cross-correlation detection seems extremely plausible in the lifetime of the COMAP survey. I have generated forecasts for the cross-correlation using `joint_limlam_mocker`, following the CO and LAE modeling in Chapter 5 and found forecast S/N ratios of 8 for the C22 fiducial model, 13 for the L16-K20 model, and 5 for the P18. Although the two brighter models (C22 and L16-K20) models are somewhat in tension with the stack result I present in that chapter, the P18 model is not, and I still predict a $5\sigma$ detection for that model by COMAP Year 8. The P18 model also can be thought of as a lower limit to the possible CO LIM signal, as it assumes contributions to the average CO only from parts of the CO LF where CO has been observed directly (*i.e.*, there is no extrapolation to the faint end).

There are several factors which will make this analysis more accessible in the near future. Firstly, in preparation for the first HETDEX public data release (PDR1), the HETDEX catalog selection has been finalized (Mentuch Cooper et al., 2026). This will minimize false positive detections while maximizing the number of objects that are retained. Additionally, HETDEX has internally developed robust random catalogs that match the on-sky coverage and catalog selection. These will be necessary for addressing the HETDEX window function, which has correlated

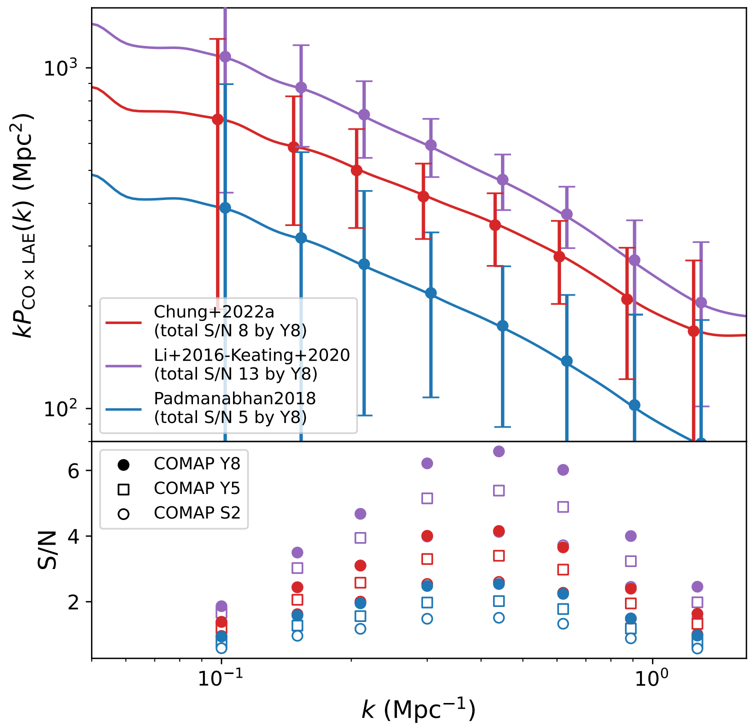


Figure 7.3: Forecasts for COMAP cross-correlations with LAEs. (*top*) Forecast power spectra $kP(k)$ generated for a COMAP×HETDEX cross-correlation. Three CO models, C22, L16-K20, and P18, are shown. (*bottom*) Forecast S/N in each $k$-bin for the two models, at the current COMAP noise level (S2) and projected forward to five and eight years of COMAP Pathfinder-only observations (Y5/Y8).

repeating structures at several spatial scales relevant to the COMAP-HETDEX cross-correlation (see Chapter 5). Finally, COMAP-*wide* will dramatically increase the available large-scale modes, moving the cross-correlation away from scales where the HETDEX window function is highly structured.

**COMAP × 21 cm intensity maps**

In theory, the most sensitive COMAP joint analyses will combine the COMAP LIM dataset with other LIM surveys, as both experiments will have access to all of the luminosity available in their given tracer and cosmic volume. While it will eventually be possible to cross-correlate COMAP with other CO and [CII] experiments (in particular, TIME; Staniszewski et al., 2014 and EXCLAIM; Oxholm et al., 2020), the most immediate prospects for COMAP-LIM cross-correlations are with 21 cm experiments across cosmic time.

During the Epoch of Reionization (e.g., Barkana and Loeb, 2001; Loeb and Barkana, 2001), 21 cm emission traces the neutral hydrogen in the intergalactic medium (IGM) that is not yet ionized by the heating influence of galaxies. It is anti-correlated with CO emission (which traces the ionizing sources) inside the ionized 'bubbles' and transitions to being correlated outside of the bubbles, so the average bubble size can be measured from the two lines' cross-correlation coefficient (e.g., Libanore et al., 2025; Thélie et al., 2026). It is for this purpose that COMAP-*wide* was proposed. By measuring the CO-21 cm cross-correlation coefficient from the COMAP-*wide*×LOFAR power spectrum, the amount of ionizing radiation escaping galaxies could be constrained. Conversely, by making assumptions about the cross-correlation coefficient, the cross-correlation could be used to extract limits on the auto-spectra for both the CO and 21 cm lines (e.g., Beane et al., 2019).

During the EoGA, the COMAP CO(1–0) coverage overlaps very well in Fourier space with the coverage of the 21 cm line by the Canadian Hydrogen Observatory and Radio Transient Detector (CHORD; Vanderlinde et al., 2019). In this epoch, the IGM is fully ionized, and neutral gas is located within galaxies. The 21 cm emission will thus be correlated with the CO emission, as it traces the same structures. A cross-correlation between the two lines would yield information about the interaction between the molecular and neutral medium in galaxies in the buildup to the peak of cosmic star formation. This analysis will be one of the primary goals of my postdoctoral work. Looking forward, a similar analysis will also be possible with the Deep Synoptic Array (DSA; Hallinan et al., 2019) in the coming years.

## 7.4 The future of LIM

As I have aimed to demonstrate throughout this thesis, LIM is a very rich field. Even when the focus is narrowed to just CO and COMAP, there is a wide variety of possible analyses available that my research has only begun to address. COMAP is already beginning to place constraints on galaxy formation (through CO at $z \sim 3$) and particle physics (through constraints on axion properties) and soon will extend its range to the Epoch of Reionization. Over the next few years many other LIM datasets will become available, covering different observed frequencies, spectral lines, redshift ranges and spatial scales. Joint analysis techniques do not simply sum together the possibilities from these experiments—they multiply them. LIM is about to become a powerful tool for charting the history of the Universe.

# BIBLIOGRAPHY

Abbott, Larry F. and Pierre Sikivie (Jan. 1983). “A cosmological bound on the invisible axion.” In: *Physics Letters B* 120.1-3, pp. 133–136. DOI: 10.1016/0370-2693(83)90638-X.

Abdurashidova, Zara et al. (Feb. 2022). “First results from HERA Phase I: Upper limits on the epoch of reionization 21 cm power spectrum.” In: *The Astrophysical Journal* 925.2, 221, p. 221. DOI: 10.3847/1538-4357/ac1c78. arXiv: 2108.02263 [astro-ph.CO].

Agrawal, Shubh, James Aguirre, and Ryan Keenan (Jan. 2026). “Far-infrared lines hidden in archival deep multi-wavelength surveys: Limits on [CII]-158 $\mu$m at z $\sim$ 0.3–2.9.” In: *Astronomy & Astrophysics* 705, A246, A246. DOI: 10.1051/0004-6361/202556503. arXiv: 2509.24211 [astro-ph.GA].

Ahumada, Romina et al. (2020). “The 16th data release of the sloan digital sky surveys: first release from the APOGEE-2 southern survey and full release of eBOSS spectra.” In: *The Astrophysical Journal Supplement Series* 249.1, p. 3.

Altenmüller, Konrad et al. (Nov. 2024). “New upper limit on the axion-photon coupling with an extended CAST run with a Xe-based micromegas detector.” In: *Physical Review Letters* 133.22, 221005, p. 221005. DOI: 10.1103/PhysRevLett.133.221005. arXiv: 2406.16840 [hep-ex].

Amiri, Mandana et al. (Apr. 2023). “Detection of cosmological 21 cm emission with the Canadian Hydrogen Intensity Mapping Experiment.” In: *The Astrophysical Journal* 947.1, 16, p. 16. DOI: 10.3847/1538-4357/acb13f. arXiv: 2202.01242 [astro-ph.CO].

Amiri, Mandana et al. (Mar. 2024). “A detection of cosmological 21 cm emission from CHIME in cross-correlation with eBOSS measurements of the Ly$\alpha$ forest.” In: *The Astrophysical Journal* 963.1, 23, p. 23. DOI: 10.3847/1538-4357/ad0f1d. arXiv: 2309.04404 [astro-ph.CO].

Anastassopoulos, Vassilis et al. (2017). “New CAST limit on the axion-photon interaction.” In: *Nature Physics* 13, pp. 584–590. DOI: 10.1038/nphys4109. arXiv: 1705.02290 [hep-ex].

Andriamonje, Samuel et al. (2007). “An Improved limit on the axion-photon coupling from the CAST experiment.” In: *JCAP* 04, p. 010. DOI: 10.1088/1475-7516/2007/04/010. arXiv: hep-ex/0702006.

Aravena, Manuel et al. (Sept. 2019). “The ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: Evolution of the molecular gas in CO-selected galaxies.” In: *The Astrophysical Journal* 882.2, 136, p. 136. DOI: 10.3847/1538-4357/ab30df. arXiv: 1903.09162 [astro-ph.GA].

Balázs, Csaba et al. (Dec. 2022). “Cosmological constraints on decaying axion-like particles: A global analysis.” In: *Journal of Cosmology and Astroparticle Physics* 2022.12, 027, p. 027. doi: 10.1088/1475-7516/2022/12/027. arXiv: 2205.13549 [astro-ph.CO].

Banerji, Manda et al. (Mar. 2017). “The discovery of gas-rich, dusty starbursts in luminous reddened quasars at z ∼ 2.5 with ALMA.” In: *Monthly Notices of the Royal Astronomical Society* 465.4, pp. 4390–4405. doi: 10.1093/mnras/stw3019. arXiv: 1610.04907 [astro-ph.GA].

Banerji, Manda et al. (Sept. 2018). “The interstellar medium properties of heavily reddened quasarsand companions at z∼ 2.5 with ALMA and JVLA.” In: *Monthly Notices of the Royal Astronomical Society* 479.1, pp. 1154–1169. doi: 10.1093/mnras/sty1443. arXiv: 1805.09825 [astro-ph.GA].

Barkana, Rennan and Abraham Loeb (2001). “In the beginning: the first sources of light and the reionization of the universe.” In: *Physics Reports* 349.2, pp. 125–238. issn: 0370-1573. doi: 10.1016/S0370-1573(01)00019-9. url: https://www.sciencedirect.com/science/article/pii/S0370157301000199.

Bautista, Julian E. et al. (June 2017). “Measurement of baryon acoustic oscillation correlations at z = 2.3 with SDSS DR12 Ly$\alpha$-Forests.” In: *Astronomy & Astrophysics* 603, A12, A12. doi: 10.1051/0004-6361/201730533. arXiv: 1702.00176 [astro-ph.CO].

Beane, Angus, Francisco Villaescusa-Navarro, and Adam Lidz (Apr. 2019). “Measuring the EoR power spectrum without measuring the EoR power spectrum.” In: *The Astrophysical Journal* 874.2, 133, p. 133. doi: 10.3847/1538-4357/ab0a08. arXiv: 1811.10609 [astro-ph.CO].

Behroozi, Peter et al. (Sept. 2019). “UNIVERSEMACHINE: The correlation between galaxy growth and dark matter halo assembly from z = 0-10.” In: *Monthly Notices of the Royal Astronomical Society* 488.3, pp. 3143–3194. doi: 10.1093/mnras/stz1182. arXiv: 1806.07893 [astro-ph.GA].

Behroozi, Peter S., Risa H. Wechsler, and Charlie Conroy (Jan. 2013a). “On the lack of evolution in galaxy star formation efficiency.” In: *The Astrophysical Journal Letters* 762.2, L31, p. L31. doi: 10.1088/2041-8205/762/2/L31. arXiv: 1209.3013 [astro-ph.CO].

– (June 2013b). “The average star formation histories of galaxies in dark matter halos from z = 0-8.” In: *The Astrophysical Journal* 770.1, 57, p. 57. doi: 10.1088/0004-637X/770/1/57. arXiv: 1207.6105 [astro-ph.CO].

Belitsky, Victor et al. (Apr. 2018). “SEPIA - a new single pixel receiver at the APEX telescope.” In: *Astronomy & Astrophysics* 612, A23, A23. doi: 10.1051/0004-6361/20173145810.48550/arXiv.1712.07396. arXiv: 1712.07396 [astro-ph.IM].

Bernal, José Luis, Andrea Caputo, and Marc Kamionkowski (Mar. 2021). "Strategies to detect dark-matter decays with line-intensity mapping." In: *Physical Review D* 103.6, 063523, p. 063523. DOI: 10.1103/PhysRevD.103.063523. arXiv: 2012.00771 [astro-ph.CO].

Bernal, José Luis and Ely D. Kovetz (Dec. 2022a). "Line-intensity mapping: theory review with a focus on star-formation lines." In: *Astronomy & Astrophysics Review* 30.1, 5, p. 5. DOI: 10.1007/s00159-022-00143-0. arXiv: 2206.15377 [astro-ph.CO].

Bernal, José Luis, Gabriela Sato-Polito, and Marc Kamionkowski (Dec. 2022b). "Cosmic optical background excess, dark matter, and line-intensity mapping." In: *Physical Review Letters* 129.23, 231301, p. 231301. DOI: 10.1103/PhysRevLett.129.231301. arXiv: 2203.11236 [astro-ph.CO].

Béthermin, Matthieu et al. (Nov. 2022). "CONCERTO: High-fidelity simulation of millimeter line emissions of galaxies and [CII] intensity mapping." In: *Astronomy & Astrophysics* 667, A156, A156. DOI: 10.1051/0004-6361/202243888. arXiv: 2204.12827 [astro-ph.GA].

Bischetti, Manuela et al. (Jan. 2021). "The WISSH quasars project. IX. Cold gas content and environment of luminous QSOs at z ~ 2.4-4.7." In: *Astronomy & Astrophysics* 645, A33, A33. DOI: 10.1051/0004-6361/202039057. arXiv: 2009.01112 [astro-ph.GA].

Bock, James J. et al. (Mar. 2026). "The SPHEREx satellite mission." In: *The Astrophysical Journal* 999.1, 139, p. 139. DOI: 10.3847/1538-4357/ae2be2. arXiv: 2511.02985 [astro-ph.IM].

Boddy, Kimberly K. et al. (Aug. 2022). "Snowmass2021 theory frontier white paper: Astrophysical and cosmological probes of dark matter." In: *Journal of High Energy Astrophysics* 35, pp. 112–138. DOI: 10.1016/j.jheap.2022.06.005. arXiv: 2203.06380 [hep-ph].

Bolatto, Alberto D., Mark Wolfire, and Adam K. Leroy (2013). "The CO-to-H2 Conversion Factor." In: *Annual Review of Astronomy & Astrophysics* 51.1, pp. 207–268. DOI: 10.1146/annurev-astro-082812-140944. eprint: https://doi.org/10.1146/annurev-astro-082812-140944. URL: https://doi.org/10.1146/annurev-astro-082812-140944.

Bolliet, Boris, Jens Chluba, and Richard Battye (2021). "Spectral distortion constraints on photon injection from low-mass decaying particles." In: *Monthly Notices of the Royal Astronomical Society* 507.3, pp. 3148–3178. DOI: 10.1093/mnras/stab1997. arXiv: 2012.07292 [astro-ph.CO].

Bolton, Adam S. et al. (Nov. 2012). "Spectral classification and redshift measurement for the SDSS-III Baryon Oscillation Spectroscopic Survey." In: *Astronomical Journal* 144.5, 144, p. 144. DOI: 10.1088/0004-6256/144/5/144. arXiv: 1207.7326 [astro-ph.CO].

Bond, J. Richard and Steven T. Myers (Mar. 1996). "The peak-patch picture of cosmic catalogs. I. Algorithms." In: *The Astrophysical Journal Supplement Series* 103, p. 1. doi: 10.1086/192267.

Boogaard, Leindert A. et al. (Sept. 2019). "The ALMA Spectroscopic Survey in the HUDF: Nature and physical properties of gas-mass selected galaxies using MUSE spectroscopy." In: *The Astrophysical Journal* 882.2, 140, p. 140. doi: 10.3847/1538-4357/ab3102. arXiv: 1903.09167 [astro-ph.GA].

Boogaard, Leindert A. et al. (Oct. 2020). "The ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: CO excitation and atomic carbon in star-forming galaxies at z = 1-3." In: *The Astrophysical Journal* 902.2, 109, p. 109. doi: 10.3847/1538-4357/abb82f. arXiv: 2009.04348 [astro-ph.GA].

Bothwell, Matthew S. et al. (Mar. 2013). "A survey of molecular gas in luminous sub-millimetre galaxies." In: *Monthly Notices of the Royal Astronomical Society* 429.4, pp. 3047–3067. doi: 10.1093/mnras/sts562. arXiv: 1205.1511 [astro-ph.CO].

Bracks, Justin S. et al. (Feb. 2026). "Forecasting the cross correlation of Terahertz Intensity Mapper [CII] line intensity maps with Euclid galaxies." In: *arXiv e-prints*, arXiv:2602.24204, arXiv:2602.24204. doi: 10.48550/arXiv.2602.24204. arXiv: 2602.24204 [astro-ph.GA].

Breysse, Patrick C and Rachael M Alexandroff (Sept. 2019). "Observing AGN feedback with CO intensity mapping." In: *Monthly Notices of the Royal Astronomical Society* 490.1, pp. 260–273. issn: 1365-2966. doi: 10.1093/mnras/stz2534. url: http://dx.doi.org/10.1093/mnras/stz2534.

Breysse, Patrick C., Ely D. Kovetz, and Marc Kamionkowski (Oct. 2015). "Masking line foregrounds in intensity-mapping surveys." In: *Monthly Notices of the Royal Astronomical Society* 452.4, pp. 3408–3418. doi: 10.1093/mnras/stv1476. arXiv: 1503.05202 [astro-ph.CO].

Breysse, Patrick C. et al. (July 2022). "COMAP Early Science. VII. Prospects for CO intensity mapping at reionization." In: *The Astrophysical Journal* 933.2, 188, p. 188. doi: 10.3847/1538-4357/ac63c9. arXiv: 2111.05933 [astro-ph.CO].

Brusa, Marcella et al. (Apr. 2018). "Molecular outflow and feedback in the obscured quasar XID2028 revealed by ALMA." In: *Astronomy & Astrophysics* 612, A29, A29. doi: 10.1051/0004-6361/201731641. arXiv: 1712.04505 [astro-ph.GA].

Byrohl, Chris and Max Gronke (Oct. 2020). "Variations in shape among observed Lyman-$\alpha$ spectra due to intergalactic absorption." In: *Astronomy & Astrophysics* 642, L16, p. L16. doi: 10.1051/0004-6361/202038685. arXiv: 2006.10041 [astro-ph.GA].

Byrohl, Chris and Dylan Nelson (Aug. 2023). "The cosmic web in Lyman-alpha emission." In: *Monthly Notices of the Royal Astronomical Society* 523.4, pp. 5248–5273. doi: 10.1093/mnras/stad1779. arXiv: 2212.08666 [astro-ph.GA].

Cadamuro, Davide and Javier Redondo (Feb. 2012). "Cosmological bounds on pseudo Nambu-Goldstone bosons." In: *Journal of Cosmology and Astroparticle Physics* 2012.2, 032, p. 032. DOI: 10.1088/1475-7516/2012/02/032. arXiv: 1110.2895 [hep-ph].

Capozzi, Francesco et al. (June 2023). "CMB and Lyman-$\alpha$ constraints on dark matter decays to photons." In: *Journal of Cosmology and Astroparticle Physics* 2023.6, 060, p. 060. DOI: 10.1088/1475-7516/2023/06/060. arXiv: 2303.07426 [astro-ph.CO].

Carenza, Pierluca, Giuseppe Lucente, and Edoardo Vitagliano (Apr. 2023). "Probing the blue axion with cosmic optical background anisotropies." In: *Physical Review D* 107.8, 083032, p. 083032. DOI: 10.1103/PhysRevD.107.083032. arXiv: 2301.06560 [hep-ph].

Carilli, Chris L. and Fabian Walter (2013). "Cool has in high-redshift galaxies." In: *Annual Review of Astronomy and Astrophysics* 51.1, pp. 105–161. DOI: 10.1146/annurev-astro-082812-140953. eprint: https://doi.org/10.1146/annurev-astro-082812-140953. URL: https://doi.org/10.1146/annurev-astro-082812-140953.

Cassata, Paolo et al. (Jan. 2011). "The VIMOS VLT Deep Survey: Star formation rate density of Ly$\alpha$ emitters from a sample of 217 galaxies with spectroscopic redshifts $2 \leq z \leq 6.6$." In: *Astronomy & Astrophysics* 525, A143, A143. DOI: 10.1051/0004-6361/201014410. arXiv: 1003.3480 [astro-ph.CO].

CCAT-Prime Collaboration et al. (Jan. 2023). "CCAT-prime Collaboration: Science goals and forecasts with Prime-Cam on the Fred Young Submillimeter Telescope." In: *The Astrophysical Journal Supplement Series* 264.1, 7, p. 7. DOI: 10.3847/1538-4365/ac9838. arXiv: 2107.10364 [astro-ph.CO].

Chang, Tzu-Ching and Adam Lidz (Feb. 2026). "Line-intensity mapping." In: *arXiv e-prints*, arXiv:2602.03011, arXiv:2602.03011. DOI: 10.48550/arXiv.2602.03011. arXiv: 2602.03011 [astro-ph.CO].

Chang, Tzu-Ching et al. (July 2010). "An intensity map of hydrogen 21-cm emission at redshift z~0.8." In: *Nature* 466.7305, pp. 463–465. DOI: 10.1038/nature09187.

Chaussidon, Edmond et al. (Feb. 2023). "Target selection and validation of DESI quasars." In: *The Astrophysical Journal* 944.1, 107, p. 107. DOI: 10.3847/1538-4357/acb3c2. arXiv: 2208.08511 [astro-ph.CO].

Chávez Ortiz, Óscar A. et al. (Aug. 2023). "Introducing the Texas Euclid Survey for Ly$\alpha$ (TESLA) Survey: Initial study correlating galaxy properties to Ly$\alpha$ emission." In: *The Astrophysical Journal* 952.2, 110, p. 110. DOI: 10.3847/1538-4357/acc403. arXiv: 2304.03258 [astro-ph.GA].

Chen, Zhaoting et al. (July 2025). "Emission-line stacking of 21 cm intensity maps with MeerKLASS: Inference pipeline and application to the L-band deep-field data." In: *The Astrophysical Journal Supplement Series* 279.1, 19, p. 19. DOI: 10.3847/1538-4365/add897. arXiv: 2504.03908 [astro-ph.CO].

Cheng, Carina et al. (Nov. 2018). "Characterizing signal loss in the 21 cm reionization power spectrum: A revised study of PAPER-64." In: *The Astrophysical Journal* 868.1, 26, p. 26. DOI: 10.3847/1538-4357/aae833. arXiv: 1810.05175 [astro-ph.IM].

Cheng, Yun-Ting, Tzu-Ching Chang, and James J. Bock (Oct. 2020). "Phase-space spectral line deconfusion in intensity mapping." In: *The Astrophysical Journal* 901.2, 142, p. 142. DOI: 10.3847/1538-4357/abb023. arXiv: 2005.05341 [astro-ph.CO].

Cheng, Yun-Ting, Brandon S. Hensley, and Thomas S.-Y. Lai (Nov. 2025). "Feature intensity mapping: Polycyclic aromatic hydrocarbon emission from all galaxies across cosmic time." In: *The Astrophysical Journal* 993.1, 30, p. 30. DOI: 10.3847/1538-4357/ae044e. arXiv: 2506.13863 [astro-ph.CO].

Cheng, Yun-Ting et al. (Dec. 2016). "Spectral line de-confusion in an intensity mapping survey." In: *The Astrophysical Journal* 832.2, 165, p. 165. DOI: 10.3847/0004-637X/832/2/165. arXiv: 1604.07833 [astro-ph.CO].

Chenu, Jean-Yves et al. (2016). "The front-end of the NOEMA interferometer." In: *IEEE Transactions on Terahertz Science and Technology* 6.2, pp. 223–237. DOI: 10.1109/TTHZ.2016.2525762.

Chiang, Yi-Kuan (Mar. 2026). "Cosmic CO and [C II] backgrounds and the fuelling of star formation over 12 Gyr." In: *Nature Astronomy*. DOI: 10.1038/s41550-026-02798-6. arXiv: 2602.02658 [astro-ph.GA].

Chiang, Yi-Kuan, Ryu Makiya, and Brice Ménard (Oct. 2025). "Cosmic infrared background tomography and a census of cosmic dust and star formation." In: *The Astrophysical Journal* 992.1, 65, p. 65. DOI: 10.3847/1538-4357/adfb6a. arXiv: 2504.05384 [astro-ph.CO].

Chiang, Yi-Kuan et al. (July 2015). "Surveying galaxy proto-clusters in emission: A large-scale structure at z = 2.44 and the outlook for HETDEX." In: *The Astrophysical Journal* 808.1, 37, p. 37. DOI: 10.1088/0004-637X/808/1/37. arXiv: 1505.03877 [astro-ph.CO].

Chowdhary, Anirban and Suchetana Chatterjee (Oct. 2025). "Halo occupation distribution of quasars: Dependence on luminosity, redshift, black hole mass and feedback modes." In: *The Astrophysical Journal* 992.1, 21, p. 21. DOI: 10.3847/1538-4357/adfb6d. arXiv: 2508.08851 [astro-ph.CO].

Chung, Dongwoo T., Jens Chluba, and Patrick C. Breysse (July 2024a). "Carbon monoxide and ionized carbon line emission global signals: Foregrounds and targets for absolute microwave spectrometry." In: *Physical Review D* 110.2, 023513, p. 023513. DOI: 10.1103/PhysRevD.110.023513. arXiv: 2311.03297 [astro-ph.CO].

Chung, Dongwoo T. et al. (Feb. 2019). "Cross-correlating carbon monoxide line-intensity maps with spectroscopic and photometric galaxy surveys." In: *The Astrophysical Journal* 872.2, 186, p. 186. DOI: 10.3847/1538-4357/ab0027. arXiv: 1809.04550 [astro-ph.GA].

Chung, Dongwoo T. et al. (Dec. 2021). "A model of spectral line broadening in signal forecasts for line-intensity mapping experiments." In: *The Astrophysical Journal* 923.2, 188, p. 188. DOI: 10.3847/1538-4357/ac2a35. arXiv: 2104.11171 [astro-ph.CO].

Chung, Dongwoo T. et al. (July 2022). "COMAP Early Science. V. Constraints and forecasts at z 3." In: *The Astrophysical Journal* 933.2, 186, p. 186. DOI: 10.3847/1538-4357/ac63c7. arXiv: 2111.05931 [astro-ph.CO].

Chung, Dongwoo T. et al. (Nov. 2024b). "COMAP Pathfinder – Season 2 results: III. Implications for cosmic molecular gas content at z~3." In: *Astronomy & Astrophysics* 691, A337, A337. DOI: 10.1051/0004-6361/202451122. arXiv: 2406.07512 [astro-ph.CO].

Cleary, Kieran A. et al. (July 2022). "COMAP Early Science. I. Overview." In: *The Astrophysical Journal* 933.2, 182, p. 182. DOI: 10.3847/1538-4357/ac63cc. arXiv: 2111.05927 [astro-ph.CO].

CONCERTO Collaboration et al. (Oct. 2020). "A wide field-of-view low-resolution spectrometer at APEX: Instrument design and scientific forecast." In: *Astronomy & Astrophysics* 642, A60, A60. DOI: 10.1051/0004-6361/202038456. arXiv: 2007.14246 [astro-ph.IM].

Condon, James J. and Scott M. Ransom (2016). *Essential Radio Astronomy*. Princeton University Press.

Condon, James J. et al. (May 1998). "The NRAO VLA Sky Survey." In: *Astronomical Journal* 115.5, pp. 1693–1716. DOI: 10.1086/300337.

Creque-Sarbinowski, Cyril and Marc Kamionkowski (Sept. 2018). "Searching for decaying and annihilating dark matter with line intensity mapping." In: *Physical Review D* 98.6, 063524, p. 063524. DOI: 10.1103/PhysRevD.98.063524. arXiv: 1806.11119 [astro-ph.CO].

Crites, Abigail T. et al. (Aug. 2014). "The TIME-Pilot intensity mapping experiment." In: *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*. Ed. by Wayne S. Holland and Jonas Zmuidzinas. Vol. 9153. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 91531W, 91531W. DOI: 10.1117/12.2057207.

da Cunha, Elisabete et al. (Mar. 2013). "On the effect of the cosmic microwave background in high-redshift (sub-)millimeter observations." In: *The Astrophysical Journal* 766.1, 13, p. 13. DOI: 10.1088/0004-637X/766/1/13. arXiv: 1302.0844 [astro-ph.CO].

Daddi, Emanuele et al. (Apr. 2010). "Very high gas fractions and extended gas reservoirs in z = 1.5 disk galaxies." In: *The Astrophysical Journal* 713.1, pp. 686–707. DOI: 10.1088/0004-637X/713/1/686. arXiv: 0911.2776 [astro-ph.CO].

Davis, Dustin et al. (Sept. 2023a). "HETDEX Public Source Catalog 1: Stacking 50,000 Lyman alpha emitters." In: *The Astrophysical Journal* 954.2, 209, p. 209. DOI: 10.3847/1538-4357/ace4c2. arXiv: 2307.03096 [astro-ph.GA].

Davis, Dustin et al. (Apr. 2023b). "The HETDEX survey emission-line exploration and source classification." In: *The Astrophysical Journal* 946.2, 86, p. 86. DOI: 10.3847/1538-4357/acb0ca. arXiv: 2301.01799 [astro-ph.GA].

Dawson, Kyle S. et al. (Jan. 2013). "The Baryon Oscillation Spectroscopic Survey of SDSS-III." In: *Astronomical Journal* 145.1, 10, p. 10. DOI: 10.1088/0004-6256/145/1/10. arXiv: 1208.0022 [astro-ph.CO].

Dawson, Kyle S. et al. (Feb. 2016). "The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Overview and early data." In: *The Astrophysical Journal* 151.2, 44, p. 44. DOI: 10.3847/0004-6256/151/2/44. arXiv: 1508.04473 [astro-ph.CO].

Decarli, Roberto et al. (Jan. 2021). "A search for dust and molecular gas in enormous Ly$\alpha$ nebulae at z $\approx$ 2." In: *Astronomy & Astrophysics* 645, L3, p. L3. DOI: 10.1051/0004-6361/202039814. arXiv: 2012.04675 [astro-ph.GA].

Delhaize, Jacinta et al. (2013). "Detection of HI in distant galaxies using spectral stacking." In: *Monthly Notices of the Royal Astronomical Society* 433.2, pp. 1398–1410.

Delubac, Timothée et al. (Feb. 2017). "The SDSS-IV eBOSS: emission line galaxy catalogues at z $\approx$ 0.8 and study of systematic errors in the angular clustering." In: *Monthly Notices of the Royal Astronomical Society* 465.2, pp. 1831–1846. DOI: 10.1093/mnras/stw2741. arXiv: 1611.06934 [astro-ph.CO].

DES Collaboration et al. (Jan. 2022). "Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing." In: *Physical Review D* 105.2, 023520, p. 023520. DOI: 10.1103/PhysRevD.105.023520. arXiv: 2105.13549 [astro-ph.CO].

DESI Collaboration et al. (Nov. 2024a). "DESI 2024 II: Sample definitions, characteristics, and two-point clustering statistics." In: *arXiv e-prints*, arXiv:2411.12020, arXiv:2411.12020. DOI: 10.48550/arXiv.2411.12020. arXiv: 2411.12020 [astro-ph.CO].

DESI Collaboration et al. (Aug. 2024b). "The early data release of the Dark Energy Spectroscopic Instrument." In: *Astronomical Journal* 168.2, 58, p. 58. DOI: 10.3847/1538-3881/ad3217. arXiv: 2306.06308 [astro-ph.CO].

DESI Collaboration et al. (Mar. 2025a). "Data Release 1 of the Dark Energy Spectroscopic Instrument." In: *arXiv e-prints*, arXiv:2503.14745, arXiv:2503.14745. DOI: 10.48550/arXiv.2503.14745. arXiv: 2503.14745 [astro-ph.CO].

DESI Collaboration et al. (July 2025b). "DESI 2024 II: sample definitions, characteristics, and two-point clustering statistics." In: *Journal of Cosmology and Astroparticle Physics* 2025.7, 017, p. 017. DOI: 10.1088/1475-7516/2025/07/017. arXiv: 2411.12020 [astro-ph.CO].

DESI Collaboration et al. (Oct. 2025c). "DESI DR2 results. II. Measurements of baryon acoustic oscillations and cosmological constraints." In: *Physical Review D* 112.8, 083515, p. 083515. DOI: 10.1103/tr6y-kpc6. arXiv: 2503.14738 [astro-ph.CO].

Dessert, Christopher, Andrew J. Long, and Benjamin R. Safdi (2022). "No evidence for axions from Chandra observation of the magnetic white dwarf RE J0317-853." In: *Phys. Rev. Lett.* 128.7, p. 071102. DOI: 10.1103/PhysRevLett.128.071102. arXiv: 2104.12772 [hep-ph].

Di Luzio, Luca, Federico Mescia, and Enrico Nardi (Jan. 2017). "Redefining the axion window." In: *Physical Review Letters* 118.3, 031801, p. 031801. DOI: 10.1103/PhysRevLett.118.031801. arXiv: 1610.07593 [hep-ph].

Dine, Michael and Willy Fischler (Jan. 1983). "The not-so-harmless axion." In: *Physics Letters B* 120.1-3, pp. 137–141. DOI: 10.1016/0370-2693(83)90639-1.

Dine, Michael, Willy Fischler, and Mark Srednicki (Aug. 1981). "A simple solution to the strong CP problem with a harmless axion." In: *Physics Letters B* 104.3, pp. 199–202. DOI: 10.1016/0370-2693(81)90590-6.

Dolan, Matthew J., Frederick J. Hiskens, and Raymond R. Volkas (Oct. 2022). "Advancing globular cluster constraints on the axion-photon coupling." In: *Journal of Cosmology and Astroparticle Physics* 2022.10, 096, p. 096. DOI: 10.1088/1475-7516/2022/10/096. arXiv: 2207.03102 [hep-ph].

Doré, Olivier et al. (Dec. 2014). "Cosmology with the SPHEREX All-Sky Spectral Survey." In: *arXiv e-prints*, arXiv:1412.4872, arXiv:1412.4872. DOI: 10.48550/arXiv.1412.4872. arXiv: 1412.4872 [astro-ph.CO].

Doré, Olivier et al. (June 2016). "Science impacts of the SPHEREx All-Sky Optical to Near-Infrared Spectral Survey: Report of a community workshop examining extragalactic, galactic, stellar and planetary science." In: *arXiv e-prints*, arXiv:1606.07039, arXiv:1606.07039. DOI: 10.48550/arXiv.1606.07039. arXiv: 1606.07039 [astro-ph.CO].

Doré, Olivier et al. (May 2018). "Science impacts of the SPHEREx All-Sky Optical to Near-Infrared Spectral Survey II: Report of a community workshop on the scientific synergies between the SPHEREx survey and other astronomy observatories." In: *arXiv e-prints*, arXiv:1805.05489, arXiv:1805.05489. DOI: 10.48550/arXiv.1805.05489. arXiv: 1805.05489 [astro-ph.IM].

Draine, Bruce T. (2011). *Physics of the interstellar and intergalactic medium*. Vol. 19. Princeton University Press.

du Mas des Bourboux, Hélion et al. (Oct. 2020). "The completed SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Baryon acoustic oscillations with Ly$\alpha$ forests." In: *The Astrophysical Journal* 901.2, 153, p. 153. DOI: 10.3847/1538-4357/abb085. arXiv: 2007.08995 [astro-ph.CO].

Dunne, Delaney, Ella Mansfield, and Dongwoo Chung (Apr. 2026a). *limstacker*. DOI: 10.5281/zenodo.19688492. URL: https://doi.org/10.5281/zenodo.19688492.

Dunne, Delaney et al. (May 2026b). *delaneydunne/joint_limlam_mocker: Release 1*. Version v0.1.0. DOI: 10.5281/zenodo.20018743. URL: https://doi.org/10.5281/zenodo.20018743.

Dunne, Delaney A. et al. (Apr. 2024). "COMAP Early Science. VIII. A joint stacking analysis with eBOSS quasars." In: *The Astrophysical Journal* 965.1, 7, p. 7. DOI: 10.3847/1538-4357/ad2dfc. arXiv: 2304.09832 [astro-ph.GA].

Dunne, Delaney A. et al. (Oct. 2025). "Three-dimensional stacking as a line intensity mapping statistic." In: *Astronomy & Astrophysics* 702, A247, A247. DOI: 10.1051/0004-6361/202554780. arXiv: 2503.21743 [astro-ph.CO].

Dunne, Delaney A. et al. (Mar. 2026c). "COMAP Pathfinder – Season 2 results: IV. A stack on eBOSS/DESI quasars." In: *Astronomy & Astrophysics* 707, A256, A256. DOI: 10.1051/0004-6361/202557864. arXiv: 2510.23568 [astro-ph.CO].

Euclid Collaboration et al. (June 2022). "Euclid preparation. I. The Euclid Wide Survey." In: *Astronomy & Astrophysics* 662, A112, A112. DOI: 10.1051/0004-6361/202141938. arXiv: 2108.01201 [astro-ph.CO].

Euclid Collaboration et al. (May 2025). "Euclid: I. Overview of the Euclid mission." In: *Astronomy & Astrophysics* 697, A1, A1. DOI: 10.1051/0004-6361/202450810. arXiv: 2405.13491 [astro-ph.CO].

Fabello, Silvia et al. (2011). "ALFALFA HI data stacking–I. Does the bulge quench ongoing star formation in early-type galaxies?" In: *Monthly Notices of the Royal Astronomical Society* 411.2, pp. 993–1012.

Fasano, Alessandro et al. (Aug. 2022). "CONCERTO: a breakthrough in wide field-of-view spectroscopy at millimeter wavelengths." In: *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy XI*. Ed. by Jonas Zmuidzinas and Jian-Rong Gao. Vol. 12190. Society of Photo-Optical

Instrumentation Engineers (SPIE) Conference Series, 121900Q, 121900Q. DOI: 10.1117/12.2629228. arXiv: 2206.15146 [astro-ph.IM].

Feruglio, Chiara et al. (Nov. 2015). "The multi-phase winds of Markarian 231: from the hot, nuclear, ultra-fast wind to the galaxy-scale, molecular outflow." In: *Astronomy & Astrophysics* 583, A99, A99. DOI: 10.1051/0004-6361/201526020. arXiv: 1503.01481 [astro-ph.GA].

Foss, Marie K. et al. (July 2022). "COMAP Early Science. III. CO data processing." In: *The Astrophysical Journal* 933.2, 184, p. 184. DOI: 10.3847/1538-4357/ac63ca. arXiv: 2111.05929 [astro-ph.IM].

Freundlich, Jonathan et al. (Feb. 2019). "PHIBSS2: survey design and z = 0.5 - 0.8 results. Molecular gas reservoirs during the winding-down of star formation." In: *Astronomy & Astrophysics* 622, A105, A105. DOI: 10.1051/0004-6361/201732223. arXiv: 1812.08180 [astro-ph.GA].

Fronenberg, Hannah and Adrian Liu (Nov. 2024). "Forecasts and statistical insights for line intensity mapping cross-correlations: A case study with 21 cm × [C II]." In: *The Astrophysical Journal* 975.2, 222, p. 222. DOI: 10.3847/1538-4357/ad77cc. arXiv: 2407.14588 [astro-ph.CO].

Gao, Yu and Philip M. Solomon (May 2004). "The star formation rate and dense molecular gas in galaxies." In: *The Astrophysical Journal* 606.1, pp. 271–290. DOI: 10.1086/382999. arXiv: astro-ph/0310339 [astro-ph].

García-Burillo, Santiago et al. (July 2014). "Molecular line emission in NGC 1068 imaged with ALMA. I. An AGN-driven outflow in the dense molecular gas." In: *Astronomy & Astrophysics* 567, A125, A125. DOI: 10.1051/0004-6361/201423843. arXiv: 1405.7706 [astro-ph.GA].

Garel, Thibault et al. (June 2015). "The UV, Lyman $\alpha$, and dark matter halo properties of high-redshift galaxies." In: *Monthly Notices of the Royal Astronomical Society* 450.2, pp. 1279–1294. DOI: 10.1093/mnras/stv374. arXiv: 1503.06635 [astro-ph.GA].

Gawiser, Eric et al. (Dec. 2007). "Ly$\alpha$-emitting galaxies at z = 3.1: L* progenitors experiencing rapid star formation." In: *The Astrophysical Journal* 671.1, pp. 278–284. DOI: 10.1086/522955. arXiv: 0710.2697 [astro-ph].

Gebhardt, Karl et al. (Aug. 2000). "A relationship between nuclear black hole mass and galaxy velocity dispersion." In: *The Astrophysical Journal Letters* 539.1, pp. L13–L16. DOI: 10.1086/312840. arXiv: astro-ph/0006289 [astro-ph].

Gebhardt, Karl et al. (Dec. 2021). "The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX) survey design, reductions, and detections." In: *The Astrophysical Journal* 923.2, 217, p. 217. DOI: 10.3847/1538-4357/ac2e03. arXiv: 2110.04298 [astro-ph.IM].

Genzel, Reinhard et al. (Oct. 2010). “A study of the gas-star formation relation over cosmic time.” In: *Monthly Notices of the Royal Astronomical Society* 407.4, pp. 2091–2108. DOI: 10.1111/j.1365-2966.2010.16969.x. arXiv: 1003.5180 [astro-ph.CO].

Gloudemans, Anniek J. et al. (Feb. 2025). “Monster radio jet (>66 kpc) observed in quasar at z ∼ 5.” In: *The Astrophysical Journal Letters* 980.1, L8, p. L8. DOI: 10.3847/2041-8213/ad9609. arXiv: 2411.16838 [astro-ph.GA].

Goldstein, Samuel et al. (Feb. 2025). “Constraints on axions from patchy screening of the cosmic microwave background.” In: *Physical Review Letters* 134.8, 081001, p. 081001. DOI: 10.1103/PhysRevLett.134.081001. arXiv: 2409.10514 [astro-ph.CO].

Gong, Yan et al. (Feb. 2017). “Intensity mapping of H$\alpha$, H$\beta$, [OII], and [OIII] lines at z < 5.” In: *The Astrophysical Journal* 835.2, 273, p. 273. DOI: 10.3847/1538-4357/835/2/273. arXiv: 1610.09060 [astro-ph.GA].

González-López, Jorge et al. (Sept. 2019). “The Atacama Large Millimeter/submillimeter Array Spectroscopic Survey in the Hubble Ultra Deep Field: CO Emission Lines and 3 mm Continuum Sources.” In: *The Astrophysical Journal* 882.2, 139, p. 139. DOI: 10.3847/1538-4357/ab3105. arXiv: 1903.09161 [astro-ph.GA].

Greve, Thomas R. et al. (Oct. 2014). “Star formation relations and CO spectral line energy distributions across the J-ladder and redshift.” In: *The Astrophysical Journal* 794.2, 142, p. 142. DOI: 10.1088/0004-637X/794/2/142. arXiv: 1407.4400 [astro-ph.GA].

Grin, Daniel et al. (2007). “A telescope search for decaying relic axions.” In: *Physical Review D* 75, p. 105018. DOI: 10.1103/PhysRevD.75.105018. arXiv: astro-ph/0611502.

Guaita, Lucia et al. (May 2010). “Ly$\alpha$-emitting galaxies at z = 2.1 in ECDF-S: Building blocks of typical present-day galaxies?” In: *The Astrophysical Journal* 714.1, pp. 255–269. DOI: 10.1088/0004-637X/714/1/255. arXiv: 0910.2244 [astro-ph.CO].

Gunn, James E. et al. (Apr. 2006). “The 2.5 m telescope of the Sloan Digital Sky Survey.” In: *Astronomical Journal* 131.4, pp. 2332–2359. DOI: 10.1086/500975. arXiv: astro-ph/0602326 [astro-ph].

Hallinan, Gregg et al. (Sept. 2019). “The DSA-2000 — A radio survey camera.” In: *Bulletin of the American Astronomical Society*. Vol. 51, 255, p. 255. DOI: 10.48550/arXiv.1907.07648. arXiv: 1907.07648 [astro-ph.IM].

Harish, Santosh et al. (Oct. 2025). “COSMOS-Web: MIRI data reduction and number counts at 7.7 $\mu$m using JWST.” In: *The Astrophysical Journal* 992.1, 45, p. 45. DOI: 10.3847/1538-4357/adfa1e. arXiv: 2506.03306 [astro-ph.GA].

Harris, David W. et al. (June 2016). “The composite spectrum of BOSS quasars selected for studies of the Ly$\alpha$ forest.” In: *Astronomical Journal* 151.6, 155, p. 155. DOI: 10.3847/0004-6256/151/6/155. arXiv: 1603.08626 [astro-ph.GA].

Herenz, Edmund Christian et al. (Jan. 2019). “The MUSE-Wide Survey: A determination of the Lyman $\alpha$ emitter luminosity function at $3 < z < 6$.” In: *Astronomy & Astrophysics* 621, A107, A107. DOI: 10.1051/0004-6361/201834164. arXiv: 1810.05037 [astro-ph.GA].

Herrera-Camus, Rodrigo et al. (Mar. 2020). “AGN feedback in a galaxy merger: multi-phase, galaxy-scale outflows with a fast molecular gas blob ~6 kpc away from IRAS F08572+3915.” In: *Astronomy & Astrophysics* 635, A47, A47. DOI: 10.1051/0004-6361/201936434. arXiv: 1911.06326 [astro-ph.GA].

Hill, Gary J. et al. (Dec. 2021). “The HETDEX instrumentation: Hobby-Eberly Telescope wide-field upgrade and VIRUS.” In: *Astronomical Journal* 162.6, 298, p. 298. DOI: 10.3847/1538-3881/ac2c02. arXiv: 2110.03843 [astro-ph.IM].

Hill, Gary J. et al. (July 2024). “VIRUS2: a next generation replicated integral field spectrograph with wide field and broad wavelength coverage.” In: *Ground-based and Airborne Instrumentation for Astronomy X*. Ed. by Julia J. Bryant, Kentaro Motohara, and Joël. R. D. Vernet. Vol. 13096. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 130960F, 130960F. DOI: 10.1117/12.3017716.

Hill, Ryley et al. (May 2019). “The SCUBA-2 web survey: I. Observations of CO(3-2) in hyper-luminous QSO field.” In: *MNRAS* 485.1, pp. 753–769. DOI: 10.1093/mnras/stz429. arXiv: 1810.10655 [astro-ph.GA].

Hinshaw, Gary et al. (Sept. 2013). “Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Cosmological parameter results.” In: *The Astrophysical Journal Supplement Series* 208.2, p. 19. DOI: 10.1088/0067-0049/208/2/19. URL: https://dx.doi.org/10.1088/0067-0049/208/2/19.

Hopkins, Philip F. et al. (Mar. 2006). “A unified, merger-driven model of the origin of starbursts, quasars, the cosmic x-ray background, supermassive black holes, and galaxy spheroids.” In: *The Astrophysical Journal Supplement Series* 163.1, pp. 1–49. DOI: 10.1086/499298. arXiv: astro-ph/0506398 [astro-ph].

Hopkins, Philip F. et al. (Apr. 2008). “A cosmological framework for the co-evolution of quasars, supermassive black holes, and elliptical galaxies. I. Galaxy mergers and quasar activity.” In: *The Astrophysical Journal Supplement Series* 175.2, pp. 356–389. DOI: 10.1086/524362. arXiv: 0706.1243 [astro-ph].

Hu, Wenkai et al. (Oct. 2019). “An accurate low-redshift measurement of the cosmic neutral hydrogen density.” In: *Monthly Notices of the Royal Astronomical Society* 489.2, pp. 1619–1632. DOI: 10.1093/mnras/stz2038. arXiv: 1907.10375 [astro-ph.GA].

Ihle, Håvard T. et al. (July 2022). "COMAP Early Science. IV. Power spectrum methodology and results." In: *The Astrophysical Journal* 933.2, 185, p. 185. DOI: 10.3847/1538-4357/ac63c5. arXiv: 2111.05930 [astro-ph.CO].

Israel, Frank P. (Mar. 2020). "Central molecular zones in galaxies: $^{12}$CO-to-$^{13}$CO ratios, carbon budget, and X factors." In: *Astronomy & Astrophysics* 635, A131, A131. DOI: 10.1051/0004-6361/201834198. arXiv: 2001.04975 [astro-ph.GA].

Janish, Ryan and Elena Pinetti (Feb. 2025). "Hunting Dark Matter Lines in the Infrared Background with the James Webb Space Telescope." In: *Physical Review Letters* 134.7, 071002, p. 071002. DOI: 10.1103/PhysRevLett.134.071002. arXiv: 2310.15395 [hep-ph].

Jolly, Jean-Baptiste et al. (Aug. 2021). "ALMA Lensing Cluster Survey: A spectral stacking analysis of [C II] in lensed z ~ 6 galaxies." In: *Astronomy & Astrophysics* 652, A128, A128. DOI: 10.1051/0004-6361/202140878. arXiv: 2106.09085 [astro-ph.GA].

Kamenetzky, Julia et al. (Oct. 2016). "L'$_{CO}$/L$_{FIR}$ relations with CO rotational ladders of galaxies across the Herschel SPIRE archive." In: *The Astrophysical Journal* 829.2, 93, p. 93. DOI: 10.3847/0004-637X/829/2/93. arXiv: 1508.05102 [astro-ph.GA].

Karkare, Kirit S. and Simeon Bird (Aug. 2018). "Constraining the expansion history and early dark energy with line intensity mapping." In: *Physical Review D* 98.4, 043529, p. 043529. DOI: 10.1103/PhysRevD.98.043529. arXiv: 1806.09625 [astro-ph.CO].

Karkare, Kirit S. et al. (Mar. 2022a). "Snowmass 2021 Cosmic Frontier White Paper: Cosmology with millimeter-wave line intensity mapping." In: *arXiv e-prints*, arXiv:2203.07258, arXiv:2203.07258. DOI: 10.48550/arXiv.2203.07258. arXiv: 2203.07258 [astro-ph.CO].

Karkare, Kirit S. et al. (Dec. 2022b). "SPT-SLIM: A line intensity mapping rathfinder for the South Pole Telescope." In: *Journal of Low Temperature Physics* 209.5-6, pp. 758–765. DOI: 10.1007/s10909-022-02702-2. arXiv: 2111.04631 [astro-ph.IM].

Karoumpis, Christos et al. (Nov. 2024). "[CII] line intensity mapping the epoch of reionization with the Prime-Cam on FYST: II. CO foreground masking based on an external catalog." In: *Astronomy & Astrophysics* 691, A262, A262. DOI: 10.1051/0004-6361/202450304. arXiv: 2410.17330 [astro-ph.GA].

Keating, Garrett K et al. (Dec. 2015). "First results from COPSS: The CO power spectrum survey." In: *The Astrophysical Journal* 814.2, 140, p. 140. DOI: 10.1088/0004-637X/814/2/140. arXiv: 1510.06744 [astro-ph.GA].

Keating, Garrett K et al. (Oct. 2016). "COPSS II: the molecular gas content of ten million cubic megaparsecs at redshift z 3." In: *The Astrophysical Journal* 830.1, 34, p. 34. DOI: 10.3847/0004-637X/830/1/34. arXiv: 1605.03971 [astro-ph.GA].

Keating, Garrett K. et al. (Oct. 2020). "An intensity mapping detection of aggregate CO line emission at 3 mm." In: *The Astrophysical Journal* 901.2, 141, p. 141. DOI: 10.3847/1538-4357/abb08e. arXiv: 2008.08087 [astro-ph.GA].

Keenan, Ryan P., Garrett K. Keating, and Daniel P. Marrone (Mar. 2022). "An intensity mapping constraint on the CO-galaxy cross-power spectrum at redshift 3." In: *The Astrophysical Journal* 927.2, 161, p. 161. DOI: 10.3847/1538-4357/ac4888. arXiv: 2110.02239 [astro-ph.GA].

Kennicutt, Robert C. and Neal J. Evans (Sept. 2012). "Star formation in the Milky Way and nearby galaxies." In: *Annual Review of Astronomy & Astrophysics* 50, pp. 531–608. DOI: 10.1146/annurev-astro-081811-125610. arXiv: 1204.3552 [astro-ph.GA].

Kim, Jihn E. (July 1979). "Weak-interaction singlet and strong CP invariance." In: *Physical Review Letters* 43.2, pp. 103–107. DOI: 10.1103/PhysRevLett.43.103.

Klitsch, Anne et al. (Nov. 2019). "ALMACAL - VI. Molecular gas mass density across cosmic time via a blind search for intervening molecular absorbers." In: *Monthly Notices of the Royal Astronomical Society* 490.1, pp. 1220–1230. DOI: 10.1093/mnras/stz2660. arXiv: 1909.08624 [astro-ph.GA].

Kohno, Kotaro et al. (2024). "Sub/millimeter-Wave dual-band line intensity mapping using the terahertz integral field units with universal nanotechnology (TIFUUN) for the atacama submillimeter telescope experiment (ASTE)." In: *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy XII*. SPIE, PC1310209.

Kohri, Kazunori and Hideo Kodama (2017). "Axion-like particles and recent observations of the cosmic infrared background radiation." In: *Phys. Rev. D* 96.5, p. 051701. DOI: 10.1103/PhysRevD.96.051701. arXiv: 1704.05189 [hep-ph].

Konno, Akira et al. (May 2016). "Bright and faint ends of ly$\alpha$ luminosity functions at z= 2 determined by the subaru survey: Implications for agns, magnification bias, and ism hi evolution." In: *The Astrophysical Journal* 823.1, 20, p. 20. DOI: 10.3847/0004-637X/823/1/20. arXiv: 1512.01854 [astro-ph.GA].

Kovetz, Ely et al. (May 2019). "Astrophysics and cosmology with line-intensity mapping." In: *Bulletin of the American Astronomical Society* 51.3, 101, p. 101. arXiv: 1903.04496 [astro-ph.CO].

Kovetz, Ely D. et al. (Sept. 2017). "Line-intensity mapping: 2017 status report." In: *arXiv e-prints*, arXiv:1709.09066, arXiv:1709.09066. DOI: 10.48550/arXiv.1709.09066. arXiv: 1709.09066 [astro-ph.CO].

Krumholz, Mark R. (Jan. 2014). "DESPOTIC - a new software library to Derive the Energetics and SPectra of Optically Thick Interstellar Clouds." In: *Monthly Notices of the Royal Astronomical Society* 437.2, pp. 1662–1680. DOI: 10.1093/mnras/stt2000. arXiv: 1304.2404 [astro-ph.IM].

Kunth, Daniel et al. (June 1998). "HST study of Lyman-alpha emission in star-forming galaxies: the effect of neutral gas flows." In: *Astronomy & Astrophysics* 334, pp. 11–20. DOI: 10.48550/arXiv.astro-ph/9802253. arXiv: astro-ph/9802253 [astro-ph].

Kusakabe, Haruka et al. (Jan. 2018). "The stellar mass, star formation rate and dark matter halo properties of LAEs at z ∼ 2." In: *Publications of the Astronomical Society of Japan* 70.1, 4, p. 4. DOI: 10.1093/pasj/psx148. arXiv: 1707.09373 [astro-ph.GA].

Lamb, James W. et al. (July 2022). "COMAP Early Science. II. Pathfinder instrument." In: *The Astrophysical Journal* 933.2, 183, p. 183. DOI: 10.3847/1538-4357/ac63c6. arXiv: 2111.05928 [astro-ph.IM].

Leighton, Robert B. (1977). *Final Technical Report for National Science Foundation Project AST 73-04708*. Tech. rep. California Institute of Technology. URL: https://resolver.caltech.edu/CaltechAUTHORS:20121023-144727561.

Lenkić, Laura et al. (May 2020). "Plateau de Bure High-z Blue Sequence Survey 2 (PHIBSS2): Search for secondary sources, CO luminosity functions in the field, and the evolution of molecular gas density through cosmic time." In: *Astronomical Journal* 159.5, 190, p. 190. DOI: 10.3847/1538-3881/ab7458. arXiv: 1908.01791 [astro-ph.GA].

Li, Tony Y. et al. (Feb. 2016). "Connecting CO intensity mapping to molecular gas and star formation in the epoch of galaxy assembly." In: *The Astrophysical Journal* 817.2, 169, p. 169. DOI: 10.3847/0004-637X/817/2/169. arXiv: 1503.08833 [astro-ph.CO].

Libanore, Sarah et al. (Sept. 2025). "A new boundary condition on reionization." In: *arXiv e-prints*, arXiv:2509.08886, arXiv:2509.08886. DOI: 10.48550/arXiv.2509.08886. arXiv: 2509.08886 [astro-ph.CO].

Lidz, Adam and Jessie Taylor (July 2016). "On removing interloper contamination from intensity mapping power spectrum measurements." In: *The Astrophysical Journal* 825.2, 143, p. 143. DOI: 10.3847/0004-637X/825/2/143. arXiv: 1604.05737 [astro-ph.CO].

Lidz, Adam et al. (Nov. 2011). "Intensity mapping with carbon monoxide emission lines and the redshifted 21 cm line." In: *The Astrophysical Journal* 741.2, 70, p. 70. DOI: 10.1088/0004-637X/741/2/70. arXiv: 1104.4800 [astro-ph.CO].

Liu, Chenxu et al. (Feb. 2025). "The Hobby–Eberly Telescope Dark Energy Experiment Survey (HETDEX) active galactic nuclei catalog: The fourth data release." In: *The Astrophysical Journal Supplement Series* 276.2, 72, p. 72. DOI: 10.3847/1538-4365/ada4a5. arXiv: 2412.19414 [astro-ph.GA].

Loeb, Abraham and Rennan Barkana (Jan. 2001). "The reionization of the universe by the first stars and quasars." In: *Annual Review of Astronomy & Astrophysics* 39, pp. 19–66. DOI: 10.1146/annurev.astro.39.1.19. arXiv: astro-ph/0010467 [astro-ph].

Lujan Niemeyer, Maja et al. (Apr. 2022). "Surface brightness profile of Lyman-$\alpha$ halos out to 320 kpc in HETDEX." In: *The Astrophysical Journal* 929.1, 90, p. 90. DOI: 10.3847/1538-4357/ac5cb8. arXiv: 2203.04826 [astro-ph.GA].

Lujan Niemeyer, Maja et al. (Mar. 2026). "Ly$\alpha$ Intensity Mapping in HETDEX: Galaxy-Ly$\alpha$ Intensity Cross-power Spectrum." In: *The Astrophysical Journal* 999.2, 177, p. 177. DOI: 10.3847/1538-4357/ae3a98. arXiv: 2510.11427 [astro-ph.CO].

Lunde, Jonas Gahr Sturtzel et al. (Nov. 2024). "COMAP Pathfinder – Season 2 results: I. Improved data selection and processing." In: *Astronomy & Astrophysics* 691, A335, A335. DOI: 10.1051/0004-6361/202451121. arXiv: 2406.07510 [astro-ph.CO].

Lunde, Jonas Gahr Sturtzel et al. (Oct. 2025). "Separation of gain fluctuations and continuum signals in total power spectrometers with application to COMAP." In: *arXiv e-prints (Astronomy & Astrophysics Submitted)*, arXiv:2510.23502, arXiv:2510.23502. DOI: 10.48550/arXiv.2510.23502. arXiv: 2510.23502 [astro-ph.IM].

Lundgren, Andreas et al. (Sept. 2025). "CONCERTO: forward modeling of interferograms for calibration." In: *arXiv e-prints*, arXiv:2509.04929, arXiv:2509.04929. DOI: 10.48550/arXiv.2509.04929. arXiv: 2509.04929 [astro-ph.IM].

Lyke, Brad W. et al. (Aug. 2020). "The Sloan Digital Sky Survey Quasar Catalog: Sixteenth Data Release." In: *The Astrophysical Journal Supplement Series* 250.1, p. 8. ISSN: 1538-4365. DOI: 10.3847/1538-4365/aba623. URL: http://dx.doi.org/10.3847/1538-4365/aba623.

Madau, Piero and Mark Dickinson (Aug. 2014). "Cosmic star-formation history." In: *Annual Review of Astronomy & Astrophysics* 52, pp. 415–486. DOI: 10.1146/annurev-astro-081811-125615. arXiv: 1403.0007 [astro-ph.CO].

Madau, Piero, Avery Meiksin, and Martin J Rees (1997). "21 centimeter tomography of the intergalactic medium at high redshift." In: *The Astrophysical Journal* 475.2, pp. 429–444.

Mansfield, Ella M., Delaney A. Dunne, and Dongwoo T. Chung (Dec. 2025). "Clustering Confuses Spectro-photometry: An Investigation of 2D and 3D Forced Profile Matching for Stacking Line-intensity Mapping Data on Source Catalogues." In: *arXiv e-prints (The Astrophysical Journal Accepted)*, arXiv:2512.22356. DOI: 10.48550/arXiv.2512.22356. arXiv: 2512.22356 [astro-ph.CO].

Manzari, Claudio Andrea et al. (Nov. 2024). "Supernova axions convert to gamma rays in magnetic fields of progenitor stars." In: *Physical Review Letters* 133.21, 211002, p. 211002. DOI: 10.1103/PhysRevLett.133.211002. arXiv: 2405.19393 [hep-ph].

Masui, Kiyoshi W. et al. (Jan. 2013). “Measurement of 21 cm brightness fluctuations at z∼ 0.8 in cross-correlation.” In: *The Astrophysical Journal Letters* 763.1, L20, p. L20. DOI: 10.1088/2041-8205/763/1/L20. arXiv: 1208.0331 [astro-ph.CO].

Matthee, Jorryt et al. (July 2021). “The X-SHOOTER Lyman $\alpha$ survey at z = 2 (XLS-z2) I: what makes a galaxy a Lyman $\alpha$ emitter?” In: *Monthly Notices of the Royal Astronomical Society* 505.1, pp. 1382–1412. DOI: 10.1093/mnras/stab1304. arXiv: 2102.07779 [astro-ph.GA].

MeerKLASS Collaboration et al. (Mar. 2025). “MeerKLASS L-band deep-field intensity maps: entering the H I dominated regime.” In: *Monthly Notices of the Royal Astronomical Society* 537.4, pp. 3632–3661. DOI: 10.1093/mnras/staf195. arXiv: 2407.21626 [astro-ph.CO].

Melcón, Alejandro Álvarez et al. (2020). “First results of the CAST-RADES haloscope search for axions at 34.67 $\mu$eV.” In: *Journal of High Energy Physics* 21, p. 075. DOI: 10.1007/JHEP10(2021)075. arXiv: 2104.13798 [hep-ex].

Mentuch Cooper, Erin et al. (Feb. 2023). “HETDEX Public Source Catalog 1: 220 K sources including over 50 K Ly$\alpha$ emitters from an untargeted wide-area spectroscopic survey.” In: *The Astrophysical Journal* 943.2, 177, p. 177. DOI: 10.3847/1538-4357/aca962. arXiv: 2301.01826 [astro-ph.GA].

– (Apr. 2026). “HETDEX Public Data Release 1: Source catalog 2 and datacubes from ∼ 90 deg$^2$ of integral-field optical spectroscopy.” In: *The Astrophysical Journal (Accepted)*.

Moriwaki, Kana et al. (July 2020). “Deep learning for intensity mapping observations: component extraction.” In: *Monthly Notices of the Royal Astronomical Society: Letters* 496.1, pp. L54–L58. DOI: 10.1093/mnrasl/slaa088. arXiv: 2002.07991 [astro-ph.GA].

Mukae, Shiro et al. (Apr. 2026). “Enhancing Ly$\alpha$ emitter identification in HETDEX with a convolutional neural network.” In: *arXiv e-prints*, arXiv:2604.12414, arXiv:2604.12414. DOI: 10.48550/arXiv.2604.12414. arXiv: 2604.12414 [astro-ph.GA].

Muñoz-Elgueta, Nahir et al. (Mar. 2022). “APEX at the QSO MUSEUM: Molecular gas reservoirs associated with z 3 quasars and their link to the extended Ly $\alpha$ emission.” In: *Monthly Notices of the Royal Astronomical Society* 511.1, pp. 1462–1483. DOI: 10.1093/mnras/stac041. arXiv: 2201.11660 [astro-ph.GA].

Muzahid, Sowgat et al. (Aug. 2020). “MUSEQuBES: Calibrating the redshifts of Ly $\alpha$ emitters using stacked circumgalactic medium absorption profiles.” In: *Monthly Notices of the Royal Astronomical Society* 496.2, pp. 1013–1022. DOI: 10.1093/mnras/staa1347. arXiv: 1910.03593 [astro-ph.GA].

Nagaraj, Gautam et al. (Dec. 2025). “ODIN: Probing the LAE Ly$\alpha$ luminosity function across cosmic time and different environments.” In: *The Astrophysical Journal* 995.1, 126, p. 126. DOI: 10.3847/1538-4357/ae1969. arXiv: 2506.14510 [astro-ph.GA].

Nakayama, Kazunori and Wen Yin (2022). “Anisotropic cosmic optical background bound for decaying dark matter in light of the LORRI anomaly.” In: *Phys. Rev. D* 106.10, p. 103505. DOI: 10.1103/PhysRevD.106.103505. arXiv: 2205.01079 [hep-ph].

O’Hare, Ciaran (July 2020). *cajohare/AxionLimits: AxionLimits*. https://cajohare.github.io/AxionLimits/. Version v1.0. DOI: 10.5281/zenodo.3932430.

Oei, Martijn S. S. L. et al. (Sept. 2024). “Black hole jets on the scale of the cosmic web.” In: *Nature* 633.8030, pp. 537–541. DOI: 10.1038/s41586-024-07879-y. arXiv: 2411.08630 [astro-ph.HE].

Orellana, Gustavo et al. (July 2011). “Sub-millimeter detected z $\sim$ 2 radio-quiet QSOs. Accurate redshifts, black hole masses, and inflow/outflow velocities.” In: *Astronomy & Astrophysics* 531, A128, A128. DOI: 10.1051/0004-6361/201015807. arXiv: 1105.1149 [astro-ph.CO].

Orsi, Alvaro et al. (Dec. 2008). “The clustering of Ly$\alpha$ emitters in a $\Lambda$CDM Universe.” In: *Monthly Notices of the Royal Astronomical Society* 391.4, pp. 1589–1604. DOI: 10.1111/j.1365-2966.2008.14010.x. arXiv: 0807.3447 [astro-ph].

Ouchi, Masami, Yoshiaki Ono, and Takatoshi Shibuya (Aug. 2020). “Observations of the Lyman-$\alpha$ Universe.” In: *Annual Review of Astronomy & Astrophysics* 58, pp. 617–659. DOI: 10.1146/annurev-astro-032620-021859. arXiv: 2012.07960 [astro-ph.GA].

Ouchi, Masami et al. (June 2008). “The Subaru/XMM-Newton Deep Survey (SXDS). IV. Evolution of Ly$\alpha$ emitters from z = 3.1 to 5.7 in the 1 deg$^2$ Field: Luminosity functions and AGN.” In: *The Astrophysical Journal Supplement Series* 176.2, pp. 301–330. DOI: 10.1086/527673. arXiv: 0707.3161 [astro-ph].

Oxholm, Trevor et al. (June 2020). “The EXperiment for Large-Aperture Intensity Mapping (EXCLAIM).” In: *American Astronomical Society Meeting Abstracts #236*. Vol. 236. American Astronomical Society Meeting Abstracts, 244.02, p. 244.02.

Padmanabhan, Hamsa (Apr. 2018). “Constraining the CO intensity mapping power spectrum at intermediate redshifts.” In: *Monthly Notices of the Royal Astronomical Society* 475.2, pp. 1477–1484. DOI: 10.1093/mnras/stx3250. arXiv: 1706.01471 [astro-ph.GA].

Palanque-Delabrouille, Nathalie et al. (Mar. 2016). “The extended Baryon Oscillation Spectroscopic Survey: Variability selection and quasar luminosity function.” In: *Astronomy & Astrophysics* 587, A41, A41. DOI: 10.1051/0004-6361/201527392. arXiv: 1509.05607 [astro-ph.CO].

Park, Changbom et al. (Apr. 2007). "Environmental dependence of properties of galaxies in the Sloan Digital Sky Survey." In: *The Astrophysical Journal* 658.2, pp. 898–916. DOI: 10.1086/511059. arXiv: astro-ph/0611610 [astro-ph].

Paul, Sourabh et al. (Jan. 2023). "A first detection of neutral hydrogen intensity mapping on Mpc scales at $z \approx 0.32$ and $z \approx 0.44$." In: *arXiv e-prints*, arXiv:2301.11943, arXiv:2301.11943. DOI: 10.48550/arXiv.2301.11943. arXiv: 2301.11943 [astro-ph.CO].

Pavesi, Riccardo et al. (Sept. 2018). "The CO Luminosity Density at High-z (COLDz) Survey: A sensitive, large-area blind search for low-J CO emission from cold gas in the early universe with the Karl G. Jansky Very Large Array." In: *The Astrophysical Journal* 864.1, 49, p. 49. DOI: 10.3847/1538-4357/aacb79. arXiv: 1808.04372 [astro-ph.GA].

Peccei, Roberto D. and Helen R. Quinn (June 1977a). "CP conservation in the presence of pseudoparticles." In: *Phys. Rev. Lett.* 38 (25), pp. 1440–1443. DOI: 10.1103/PhysRevLett.38.1440. URL: https://link.aps.org/doi/10.1103/PhysRevLett.38.1440.

– (Sept. 1977b). "Constraints imposed by CP conservation in the presence of pseudoparticles." In: *Phys. Rev. D* 16 (6), pp. 1791–1797. DOI: 10.1103/PhysRevD.16.1791. URL: https://link.aps.org/doi/10.1103/PhysRevD.16.1791.

Pen, Ue-Li et al. (Mar. 2009). "First detection of cosmic structure in the 21-cm intensity field." In: *Monthly Notices of the Royal Astronomical Society: Letters* 394.1, pp. L6–L10. DOI: 10.1111/j.1745-3933.2008.00581.x. arXiv: 0802.3239 [astro-ph].

Peng, Ying-jie et al. (Sept. 2010). "Mass and environment as drivers of galaxy evolution in SDSS and zCOSMOS and the origin of the Schechter function." In: *The Astrophysical Journal* 721.1, pp. 193–221. DOI: 10.1088/0004-637X/721/1/193. arXiv: 1003.4747 [astro-ph.CO].

Pinetti, Elena (Mar. 2025). "First constraints on QCD axion dark matter using James Webb Space Telescope observations." In: arXiv: 2503.11753 [hep-ph].

Planck Collaboration et al. (Sept. 2020). "Planck 2018 results. VI. Cosmological parameters." In: *Astronomy & Astrophysics* 641, A6, A6. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209 [astro-ph.CO].

Popping, Gergö et al. (Feb. 2019). "The art of modelling CO, [C I], and [C II] in cosmological galaxy formation models." In: *Monthly Notices of the Royal Astronomical Society* 482.4, pp. 4906–4932. DOI: 10.1093/mnras/sty2969. arXiv: 1805.11093 [astro-ph.GA].

Preskill, John, Mark B. Wise, and Frank Wilczek (Jan. 1983). "Cosmology of the invisible axion." In: *Physics Letters B* 120.1-3, pp. 127–132. DOI: 10.1016/0370-2693(83)90637-8.

Pullen, Anthony R. and Christopher M. Hirata (June 2013a). "Systematic effects in large-scale angular power spectra of photometric quasars and implications for constraining primordial non-Gaussianity." In: *Publications of the Astronomical Society of the Pacific* 125.928, p. 705. DOI: 10.1086/671189. arXiv: 1212.4500 [astro-ph.CO].

Pullen, Anthony R. et al. (May 2013b). "Cross-correlations as a cosmological carbon monoxide detector." In: *The Astrophysical Journal* 768.1, 15, p. 15. DOI: 10.1088/0004-637X/768/1/15. arXiv: 1211.1397 [astro-ph.CO].

Ramsey, Lawrence W. et al. (Aug. 1998). "Early performance and present status of the Hobby-Eberly Telescope." In: *Advanced Technology Optical/IR Telescopes VI*. Ed. by Larry M. Stepp. Vol. 3352. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pp. 34–42. DOI: 10.1117/12.319287.

Regis, Marco et al. (2021). "Searching for light in the darkness: Bounds on ALP dark matter with the optical MUSE-faint survey." In: *Phys. Lett. B* 814, p. 136075. DOI: 10.1016/j.physletb.2021.136075. arXiv: 2009.01310 [astro-ph.CO].

Rennie, Thomas J. et al. (July 2022). "COMAP Early Science. VI. A first look at the COMAP Galactic Plane Survey." In: *The Astrophysical Journal* 933.2, 187, p. 187. DOI: 10.3847/1538-4357/ac63c8. arXiv: 2111.05932 [astro-ph.GA].

Richardson, Jonathan et al. (Aug. 2012). "The halo occupation distribution of SDSS quasars." In: *The Astrophysical Journal* 755.1, 30, p. 30. DOI: 10.1088/0004-637X/755/1/30. arXiv: 1203.4570 [astro-ph.CO].

Riechers, Dominik A. et al. (Sept. 2007). "Detection of emission from the CN radical in the Cloverleaf quasar at z=2.56." In: *The Astrophysical Journal* 666.2, pp. 778–783. DOI: 10.1086/520335. arXiv: 0705.3338 [astro-ph].

Riechers, Dominik A. et al. (Sept. 2011). "CO($J = 1 \rightarrow 0$) in $z > 2$ quasar host galaxies: No evidence for extended molecular gas reservoirs." In: *The Astrophysical Journal Letters* 739.1, L32, p. L32. DOI: 10.1088/2041-8205/739/1/L32. arXiv: 1106.2553 [astro-ph.CO].

Riechers, Dominik A. et al. (Feb. 2019). "COLDz: Shape of the CO luminosity function at high redshift and the cold gas history of the Universe." In: *The Astrophysical Journal* 872.1, 7, p. 7. DOI: 10.3847/1538-4357/aafc27. arXiv: 1808.04371 [astro-ph.GA].

Riechers, Dominik A. et al. (June 2020). "VLA-ALMA Spectroscopic Survey in the Hubble Ultra Deep Field (VLASPECS): Total cold gas masses and CO line ratios for z = 2-3 main-sequence galaxies." In: *The Astrophysical Journal Letters* 896.2, L21, p. L21. DOI: 10.3847/2041-8213/ab9595. arXiv: 2005.09653 [astro-ph.GA].

Righi, Mattia, Carlos Hernández-Monteagudo, and Rashid A. Sunyaev (Oct. 2008). "Carbon monoxide line emission as a CMB foreground: Tomography of the star-forming universe with different spectral resolutions." In: *Astronomy & As-*

*trophysics* 489.2, pp. 489–504. DOI: 10.1051/0004-6361:200810199. arXiv: 0805.2174 [astro-ph].

Rivera-Thorsen, Thøger E. et al. (May 2015). "The Lyman Alpha Reference Sample. V. The impact of neutral ISM kinematics and geometry on Ly$\alpha$ escape." In: *The Astrophysical Journal* 805.1, 14, p. 14. DOI: 10.1088/0004-637X/805/1/14. arXiv: 1503.01157 [astro-ph.GA].

Romano, Michael et al. (Apr. 2022). "The ALPINE-ALMA [CII] survey: The population of [CII]-undetected galaxies and their role in the $L_{[CII]}$-SFR relation." In: *Astronomy & Astrophysics* 660, A14, A14. DOI: 10.1051/0004-6361/202142265. arXiv: 2202.02334 [astro-ph.GA].

Ruz, Jaime et al. (Oct. 2025). "NuSTAR as an axion helioscope." In: *Physical Review Letters* 135.14, 141001, p. 141001. DOI: 10.1103/18sn-hxtb. arXiv: 2407.03828 [astro-ph.CO].

Saha, Akash Kumar et al. (Mar. 2025). "Shedding Infrared Light on QCD Axion and ALP Dark Matter with JWST." In: *arXiv e-prints*, arXiv:2503.14582, arXiv:2503.14582. DOI: 10.48550/arXiv.2503.14582. arXiv: 2503.14582 [hep-ph].

Sanders, David B. et al. (Feb. 1988a). "Ultraluminous infrared galaxies and the origin of quasars." In: *The Astrophysical Journal* 325, p. 74. DOI: 10.1086/165983.

Sanders, David B. et al. (May 1988b). "Warm ultraluminous galaxies in the IRAS survey: The transition from galaxy to quasar?" In: *The Astrophysical Journal Letters* 328, p. L35. DOI: 10.1086/185155.

Schechter, Paul (Jan. 1976). "An analytic expression for the luminosity function for galaxies." In: *The Astrophysical Journal* 203, pp. 297–306. DOI: 10.1086/154079.

Schöier, Fredrik et al. (Oct. 2010). *LAMDA: Leiden Atomic and Molecular Database*. Astrophysics Source Code Library, record ascl:1010.077. ascl: 1010.077.

Scoville, Nick Z. (2013). "Evolution of star formation and gas." In: *Secular Evolution of Galaxies*. Ed. by Jesús Falcón-Barroso and Johan H. Knapen, p. 491. DOI: 10.48550/arXiv.1210.6990.

Sevilla-Noarbe, Ignacio et al. (June 2021). "Dark Energy Survey Year 3 results: Photometric data set for cosmology." In: *The Astrophysical Journal Supplement Series* 254.2, 24, p. 24. DOI: 10.3847/1538-4365/abeb66. arXiv: 2011.03407 [astro-ph.CO].

Shapley, Alice E. (Sept. 2011). "Physical properties of galaxies from z = 2-4." In: *Annual Review of Astronomy & Astrophysics* 49.1, pp. 525–580. DOI: 10.1146/annurev-astro-081710-102542. arXiv: 1107.5060 [astro-ph.CO].

Shen, Xuejian et al. (Jan. 2020). "The bolometric quasar luminosity function at z = 0-7." In: *Monthly Notices of the Royal Astronomical Society* 495.3, pp. 3252–3275. DOI: 10.1093/mnras/staa1381. arXiv: 2001.02696 [astro-ph.GA].

Shibuya, Takatoshi et al. (June 2014). "What is the physical origin of strong Ly$\alpha$ emission? II. Gas kinematics and distribution of Ly$\alpha$ emitters." In: *The Astrophysical Journal* 788.1, 74, p. 74. DOI: 10.1088/0004-637X/788/1/74. arXiv: 1402.1168 [astro-ph.CO].

Shifman, Mikhail A., Arkady I. Vainshtein, and Vladimir I. Zakharov (Apr. 1980). "Can confinement ensure natural CP invariance of strong interactions?" In: *Nuclear Physics B* 166.3, pp. 493–506. DOI: 10.1016/0550-3213(80)90209-6.

Silva, Marta B. et al. (Nov. 2021). "Synergies between the COMAP CO Line Intensity Mapping mission and a Ly$\alpha$ galaxy survey: How to probe the early universe with voxel based analysis of observational data." In: *arXiv e-prints*, arXiv:2111.05354, arXiv:2111.05354. arXiv: 2111.05354 [astro-ph.GA].

Simpson, James M. et al. (Nov. 2012). "The evolutionary connection between QSOs and SMGs: molecular gas in far-infrared luminous QSOs at z ~ 2.5." In: *Monthly Notices of the Royal Astronomical Society* 426.4, pp. 3201–3210. DOI: 10.1111/j.1365-2966.2012.21941.x. arXiv: 1206.4692 [astro-ph.CO].

Sinigaglia, Francesco et al. (Aug. 2022). "Optimizing spectral stacking for 21-cm observations of galaxies: accuracy assessment and symmetrized stacking." In: *Monthly Notices of the Royal Astronomical Society* 514.3, pp. 4205–4221. DOI: 10.1093/mnras/stac1584. arXiv: 2206.03300 [astro-ph.CO].

Smee, Stephen A. et al. (Aug. 2013). "The multi-object, fiber-fed spectrographs for the Sloan Digital Sky Survey and the Baryon Oscillation Spectroscopic Survey." In: *Astronomical Journal* 146.2, 32, p. 32. DOI: 10.1088/0004-6256/146/2/32. arXiv: 1208.2233 [astro-ph.IM].

Sobral, David et al. (June 2018). "The nature of luminous Ly $\alpha$ emitters at z ~ 2-3: maximal dust-poor starbursts and highly ionizing AGN." In: *Monthly Notices of the Royal Astronomical Society* 477.2, pp. 2817–2840. DOI: 10.1093/mnras/sty782. arXiv: 1802.10102 [astro-ph.GA].

Solomon, Philip M. et al. (Mar. 1997). "The molecular interstellar medium in ultraluminous infrared galaxies." In: *The Astrophysical Journal* 478.1, pp. 144–161. DOI: 10.1086/303765. URL: https://doi.org/10.1086%5C%2F303765.

Somerville, Rachel S., Gergö Popping, and Scott C. Trager (Nov. 2015). "Star formation in semi-analytic galaxy formation models with multiphase gas." In: *Monthly Notices of the Royal Astronomical Society* 453.4, pp. 4337–4367. DOI: 10.1093/mnras/stv1877. arXiv: 1503.00755 [astro-ph.GA].

Staniszewski, Zak et al. (Sept. 2014). "The Tomographic Ionized-Carbon Mapping Experiment (TIME) CII imaging spectrometer." In: *Journal of Low Temperature Physics* 176.5-6, pp. 767–772. DOI: 10.1007/s10909-014-1126-4.

Stanley, Flora et al. (Nov. 2019). "A spectral stacking analysis to search for faint outflow signatures in z ~ 6 quasars." In: *Astronomy & Astrophysics* 631, A78, A78. DOI: 10.1051/0004-6361/201834530. arXiv: 1908.11395 [astro-ph.GA].

Steidel, Charles C. et al. (Dec. 2018). "The Keck Lyman Continuum Spectroscopic Survey (KLCS): The emergent ionizing spectrum of galaxies at z ∼ 3." In: *The Astrophysical Journal* 869.2, 123, p. 123. DOI: 10.3847/1538-4357/aaed28. arXiv: 1805.06071 [astro-ph.GA].

Stein, George, Marcelo A. Alvarez, and J. Richard Bond (Feb. 2019). "The mass-Peak Patch algorithm for fast generation of deep all-sky dark matter halo catalogues and its N-body validation." In: *Monthly Notices of the Royal Astronomical Society* 483.2, pp. 2236–2250. DOI: 10.1093/mnras/sty3226. arXiv: 1810.07727 [astro-ph.CO].

Stone, Meredith A. et al. (Sept. 2024). "5-25 $\mu$m galaxy number counts from deep JWST data." In: *The Astrophysical Journal* 972.1, 62, p. 62. DOI: 10.3847/1538-4357/ad6308. arXiv: 2405.18470 [astro-ph.GA].

Strigari, Louis E. (Oct. 2013). "Galactic searches for dark matter." In: *Physics Reports* 531.1, pp. 1–88. DOI: 10.1016/j.physrep.2013.05.004. arXiv: 1211.7090 [astro-ph.CO].

Stutzer, Nils-Ole et al. (Nov. 2024). "COMAP Pathfinder – Season 2 results: II. Updated constraints on the CO(1–0) power spectrum." In: *Astronomy & Astrophysics* 691, A336, A336. DOI: 10.1051/0004-6361/202451123. arXiv: 2406.07511 [astro-ph.CO].

Sun, Guochao et al. (Apr. 2018). "A foreground masking strategy for [C II] intensity mapping experiments using galaxies selected by stellar mass and redshift." In: *The Astrophysical Journal* 856.2, 107, p. 107. DOI: 10.3847/1538-4357/aab3e3. arXiv: 1610.10095 [astro-ph.GA].

Sun, Guochao et al. (July 2021). "Probing cosmic reionization and molecular gas growth with TIME." In: *The Astrophysical Journal* 915.1, 33, p. 33. DOI: 10.3847/1538-4357/abfe62. arXiv: 2012.09160 [astro-ph.GA].

Tacconi, Linda J. et al. (May 2013). "Phibss: Molecular gas content and scaling relations in z ~1-3 massive, main-sequence star-forming galaxies." In: *The Astrophysical Journal* 768.1, 74, p. 74. DOI: 10.1088/0004-637X/768/1/74. arXiv: 1211.5743 [astro-ph.CO].

Thélie, Emilie et al. (Feb. 2026). "Reionization bubbles from real-space cross correlations of line intensity maps." In: *arXiv e-prints*, arXiv:2602.12277, arXiv:2602.12277. DOI: 10.48550/arXiv.2602.12277. arXiv: 2602.12277 [astro-ph.CO].

Timlin, John D. et al. (May 2018). "The clustering of high-redshift (2.9 ≤ z ≤ 5.1) quasars in SDSS Stripe 82." In: *The Astrophysical Journal* 859.1, 20, p. 20. DOI: 10.3847/1538-4357/aab9ac. arXiv: 1712.03128 [astro-ph.GA].

Todarello, Elisa and Marco Regis (May 2025). "Bounds on axions-like particles shining in the ultra-violet." In: *Journal of Cosmology and Astroparticle Physics* 2025.5, 070, p. 070. DOI: 10.1088/1475-7516/2025/05/070. arXiv: 2412.02543 [hep-ph].

Todarello, Elisa et al. (May 2024). "Robust bounds on ALP dark matter from dwarf spheroidal galaxies in the optical MUSE-Faint survey." In: *Journal of Cosmology and Astroparticle Physics* 2024.5, 043, p. 043. doi: 10.1088/1475-7516/2024/05/043. arXiv: 2307.07403 [astro-ph.CO].

Torralba-Torregrosa, Alberto et al. (Oct. 2024). "The PAU Survey: The quasar Ly$\alpha$ and UV luminosity functions at 2.7 < z < 5.3." In: *Astronomy & Astrophysics* 690, A388, A388. doi: 10.1051/0004-6361/202451675. arXiv: 2407.19020 [astro-ph.GA].

Tozzi, Paolo et al. (Jan. 2000). "Radio signatures of H I at high redshift: mapping the end of the "dark ages"." In: *The Astrophysical Journal* 528.2, pp. 597–606. doi: 10.1086/308196. arXiv: astro-ph/9903139 [astro-ph].

Tytler, David and Xiao-Ming Fan (Mar. 1992). "Systematic QSO emission-line velocity shifts and new unbiased redshifts." In: *The Astrophysical Journal Supplement Series* 79, p. 1. doi: 10.1086/191642.

Umeda, Hiroya et al. (Apr. 2025). "SILVERRUSH. XIV. Ly$\alpha$ luminosity functions and angular correlation functions from 20,000 Ly$\alpha$ emitters at z ~ 2.2–7.3 from up to 24 deg$^2$ HSC-SSP and CHORUS surveys: Linking the postreionization epoch to the heart of reionization." In: *The Astrophysical Journal Supplement Series* 277.2, 37, p. 37. doi: 10.3847/1538-4365/adb1c0. arXiv: 2411.15495 [astro-ph.GA].

Uzgil, Bade D. et al. (Dec. 2019). "The ALMA Spectroscopic Survey in the HUDF: Constraining cumulative CO emission at $1 \lesssim z \lesssim 4$ with power spectrum analysis of ASPECS LP data from 84 to 115 GHz." In: *The Astrophysical Journal* 887.1, 37, p. 37. doi: 10.3847/1538-4357/ab517f. arXiv: 1911.00028 [astro-ph.GA].

Van Cuyck, Mathilde et al. (Aug. 2023). "CONCERTO: Extracting the power spectrum of the [$C_{II}$] emission line." In: *Astronomy & Astrophysics* 676, A62, A62. doi: 10.1051/0004-6361/202346270. arXiv: 2306.01568 [astro-ph.CO].

van Haarlem, M. P. et al. (Aug. 2013). "LOFAR: The LOw-Frequency ARray." In: *Astronomy & Astrophysics* 556, A2, A2. doi: 10.1051/0004-6361/201220873. arXiv: 1305.3550 [astro-ph.IM].

Vanderlinde, Keith et al. (Oct. 2019). "The Canadian Hydrogen Observatory and Radio-transient Detector (CHORD)." In: *Canadian Long Range Plan for Astronomy and Astrophysics White Papers*. Vol. 2020, 28, p. 28. doi: 10.5281/zenodo.3765414. arXiv: 1911.01777 [astro-ph.IM].

Verhamme, A. et al. (July 2018). "Recovering the systemic redshift of galaxies from their Lyman alpha line profile." In: *Monthly Notices of the Royal Astronomical Society* 478.1, pp. L60–L65. doi: 10.1093/mnrasl/sly058. arXiv: 1804.01883 [astro-ph.GA].

Vieira, Joaquin et al. (Sept. 2020). “The Terahertz Intensity Mapper (TIM): A next-generation experiment for galaxy evolution studies.” In: *arXiv e-prints*, arXiv:2009.14340, arXiv:2009.14340. DOI: 10.48550/arXiv.2009.14340. arXiv: 2009.14340 [astro-ph.IM].

Viero, Marco P. et al. (Dec. 2013). “HerMES: The contribution to the cosmic infrared background from galaxies selected by mass and redshift.” In: *The Astrophysical Journal* 779.1, 32, p. 32. DOI: 10.1088/0004-637X/779/1/32. arXiv: 1304.0446 [astro-ph.CO].

Viero, Marco P. et al. (Aug. 2015). “HerMES: Current cosmic infrared background estimates can be explained by known galaxies and their faint companions at z < 4.” In: *The Astrophysical Journal Letters* 809.2, L22, p. L22. DOI: 10.1088/2041-8205/809/2/L22. arXiv: 1505.06242 [astro-ph.GA].

Viero, Marco P. et al. (Oct. 2022). “The early Universe was dust-rich and extremely hot.” In: *Monthly Notices of the Royal Astronomical Society* 516.1, pp. L30–L34. DOI: 10.1093/mnrasl/slac075. arXiv: 2203.14312 [astro-ph.GA].

Wang, Hanyue et al. (2024). “Spectroscopic search for optical emission lines from dark matter decay.” In: *Physical Review D* 110.10, p. 103007. DOI: 10.1103/PhysRevD.110.103007. arXiv: 2311.05476 [astro-ph.CO].

Wang, Huiyuan et al. (Jan. 2018). “ELUCID. IV. Galaxy quenching and its relation to halo mass, environment, and assembly bias.” In: *The Astrophysical Journal* 852.1, 31, p. 31. DOI: 10.3847/1538-4357/aa9e01. arXiv: 1707.09002 [astro-ph.GA].

Wang, Junzhi, Zhiyu Zhang, and Yong Shi (Sept. 2011). “CS (5-4) survey towards nearby infrared bright galaxies.” In: *Monthly Notices of the Royal Astronomical Society* 416.1, pp. L21–L25. DOI: 10.1111/j.1745-3933.2011.01090.x. arXiv: 1106.2873 [astro-ph.CO].

Wang, Yun et al. (Mar. 2022). “The High Latitude Spectroscopic Survey on the Nancy Grace Roman Space Telescope.” In: *The Astrophysical Journal* 928.1, 1, p. 1. DOI: 10.3847/1538-4357/ac4973. arXiv: 2110.01829 [astro-ph.CO].

Wechsler, Risa H. and Jeremy L. Tinker (Sept. 2018). “The Connection Between Galaxies and Their Dark Matter Halos.” In: *Annual Review of Astronomy & Astrophysics* 56, pp. 435–487. DOI: 10.1146/annurev-astro-081817-051756. arXiv: 1804.03097 [astro-ph.GA].

Weinberg, Steven (Jan. 1978). “A new light boson?” In: *Physical Review Letters* 40 (4), pp. 223–226. DOI: 10.1103/PhysRevLett.40.223. URL: https://link.aps.org/doi/10.1103/PhysRevLett.40.223.

Weiss, Laurel H. et al. (May 2021). “The HETDEX survey: The Ly$\alpha$ escape fraction from 3D-HST emission-line galaxies at z $\sim$ 2.” In: *The Astrophysical Journal* 912.2, 100, p. 100. DOI: 10.3847/1538-4357/abedb9. arXiv: 2103.12126 [astro-ph.GA].

Weiss, Laurel H. et al. (Feb. 2024). “Absorption troughs of Ly$\alpha$ emitters in HETDEX.” In: *The Astrophysical Journal* 962.2, 102, p. 102. DOI: 10.3847/1538-4357/ad1b51. arXiv: 2401.02490 [astro-ph.GA].

Weiss, Laurel H. et al. (Apr. 2025). “Using Ly$\alpha$ absorption to measure the intensity and variability of z $\sim$ 2.4 ultraviolet background light.” In: *The Astrophysical Journal* 983.1, 72, p. 72. DOI: 10.3847/1538-4357/adc0f9. arXiv: 2504.13253 [astro-ph.GA].

White, Martin et al. (Aug. 2012). “The clustering of intermediate-redshift quasars as measured by the Baryon Oscillation Spectroscopic Survey.” In: *Monthly Notices of the Royal Astronomical Society* 424.2, pp. 933–950. DOI: 10.1111/j.1365-2966.2012.21251.x. arXiv: 1203.5306 [astro-ph.CO].

Wilczek, Frank (Jan. 1978). “Problem of strong $P$ and $T$ invariance in the presence of instantons.” In: *Physical Review Letters* 40 (5), pp. 279–282. DOI: 10.1103/PhysRevLett.40.279. URL: https://link.aps.org/doi/10.1103/PhysRevLett.40.279.

Wolz, Laura, Chris Blake, and J. Stuart B. Wyithe (Sept. 2017). “Determining the H I content of galaxies via intensity mapping cross-correlations.” In: *Monthly Notices of the Royal Astronomical Society* 470.3, pp. 3220–3226. DOI: 10.1093/mnras/stx1388. arXiv: 1703.08268 [astro-ph.GA].

Wolz, Laura et al. (Mar. 2022). “H I constraints from the cross-correlation of eBOSS galaxies and Green Bank Telescope intensity maps.” In: *Monthly Notices of the Royal Astronomical Society* 510.3, pp. 3495–3511. DOI: 10.1093/mnras/stab3621. arXiv: 2102.04946 [astro-ph.CO].

Yang, Shengqi et al. (Apr. 2022). “An empirical representation of a physical model for the ISM [C II], CO, and [C I] emission at redshift $1 \leq z \leq 9$.” In: *The Astrophysical Journal* 929.2, 140, p. 140. DOI: 10.3847/1538-4357/ac5d57. arXiv: 2108.07716 [astro-ph.GA].

Yue, Bin and Andrea Ferrara (Dec. 2019). “Studying high-z galaxies with [C II] intensity mapping.” In: *Monthly Notices of the Royal Astronomical Society* 490.2, pp. 1928–1943. DOI: 10.1093/mnras/stz2728. arXiv: 1909.11656 [astro-ph.GA].

Zhang, Yechi et al. (Dec. 2021). “First HETDEX spectroscopic determinations of Ly$\alpha$ and UV luminosity functions at z = 2-3: Bridging a gap between faint AGNs and bright galaxies.” In: *The Astrophysical Journal* 922.2, 167, p. 167. DOI: 10.3847/1538-4357/ac1e97. arXiv: 2105.11497 [astro-ph.GA].

Zhang, Zhi-Yu et al. (Apr. 2014). “Dense gas tracers and star formation laws in active galaxies: APEX survey of HCN $J = 4 \rightarrow 3$, $HCO^+$ $J = 4 \rightarrow 3$, and CS $J = 7 \rightarrow 6$.” In: *The Astrophysical Journal Letters* 784.2, L31, p. L31. DOI: 10.1088/2041-8205/784/2/L31. arXiv: 1402.1039 [astro-ph.GA].

Zhitnitskii, Ariel P. (Feb. 1980). "Possible suppression of axion-hadron interactions." In: *Soviet Journal of Nuclear Physics (English Translation ); (United States)* 31:2. ISSN: ISSN SJNCA. URL: https://www.osti.gov/biblio/7063072.

Zubeldia, Íñigo et al. (Nov. 2021). "Understanding matched filters for precision cosmology." In: *Monthly Notices of the Royal Astronomical Society* 507.4, pp. 4852–4863. DOI: 10.1093/mnras/stab2461. arXiv: 2106.03718 [astro-ph.CO].