CrossMark

Phase-space Spectral Line Deconfusion in Intensity Mapping

Yun-Ting Cheng[1] , Tzu-Ching Chang[1,2,3] , and James J. Bock[1,2]

[1] California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA; ycheng3@caltech.edu
[2] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[3] Institute of Astronomy and Astrophysics, Academia Sinica, 1 Roosevelt Road, Section 4, Taipei, 10617, Taiwan

Abstract

Line intensity mapping (LIM) is a promising tool to efficiently probe the three-dimensional large-scale structure by mapping the aggregate emission of a spectral line from all sources that trace the matter density field. Spectral lines from different redshifts can fall in the same observed frequency and be confused, however, which is a major challenge in LIM. In this work, we develop a line deconfusion technique in map space capable of reconstructing the three-dimensional spatial distribution of line-emitting sources. If multiple spectral lines of a source population are observable in multiple frequencies, using the sparse approximation, our technique iteratively extracts sources along a given line of sight by fitting the LIM data to a set of spectral templates. We demonstrate that the technique successfully extracts sources with emission lines present at a few σ above the noise level, taking into account uncertainties in the source modeling and presence of continuum foreground contamination and noise fluctuations. As an example, we consider a Tomographic Ionized-carbon Mapping Experiment/CarbON C II line in post-rEionisation and ReionisaTiOn epoch (TIME/CONCERTO)-like survey targeting [C II] at the epoch of reionization, and reliably reconstruct the 3D spatial distribution of the CO interlopers and their luminosity functions at $0.5 \lesssim z \lesssim 1.5$. We also demonstrate a successful deconfusion for the Spectro-Photometer for the History of the Universe, Epoch of Reionization, and Ices Explorer (SPHEREx) mission in the near-infrared wavelengths. We discuss a formalism in which the reconstructed maps can be further cross-correlated with a (galaxy) tracer population to estimate the total interloper power. This technique is a general framework to extract the phase-space distribution of low-redshift interlopers, without the need of external information, for any line deconfusion problem.

Unified Astronomy Thesaurus concepts: Observational cosmology (1146); Large-scale structure of the universe (902); Diffuse radiation (383)

1. Introduction

Line intensity mapping (LIM) has emerged as a promising tool to study the three-dimensional large-scale structures by mapping a particular spectral line emission, and infers the line-of-sight distance of the emission sources from the frequency-redshift relation. LIM measures the aggregate emission of all sources to constrain the bulk properties of the galaxies; whereas in traditional galaxy surveys, only the brighter sources can be individually detected. This relatively low spatial resolution and point-source sensitivity requirement in LIM enables the use of small apertures to efficiently scan a large survey volume out to high redshifts.

Several spectral lines have been proposed for LIM survey. The 21 cm hyperfine emission from neutral hydrogen (Scott & Rees 1990; Madau et al. 1997; Chang et al. 2008; Wyithe & Loeb 2008), the CO rotational lines (Righi et al. 2008; Visbal & Loeb 2010; Carilli 2011; Gong et al. 2011; Lidz et al. 2011; Pullen et al. 2013; Breysse et al. 2014, 2016; Keating et al. 2015, 2016; Mashian et al. 2015; Li et al. 2016; Breysse & Rahman 2017; Fonseca et al. 2017; Chung et al. 2019), the [C II] 157.7 μm fine-structure line (Gong et al. 2012; Uzgil et al. 2014; Silva et al. 2015; Yue et al. 2015; Fonseca et al. 2017), and the Lyα emission line (Silva et al. 2013; Gong et al. 2014; Pullen et al. 2014; Comaschi & Ferrara 2016; Croft et al. 2016, 2018; Fonseca et al. 2017) are among the most studied lines in the LIM regime.

One of the main challenges in LIM is the astrophysical foreground contaminations, including the continuum emission and line interlopers. Although the continuum foregrounds are usually a few orders of magnitude brighter than the lines (a situation more severe for 21 cm than for other lines), their smooth spectral feature

can be used to distinguish from the line signals. This has been extensively studied in the context of 21 cm LIM (e.g., Furlanetto et al. 2006; Morales et al. 2006; Bowman et al. 2009; Chapman et al. 2012; Liu & Tegmark 2012; Parsons et al. 2012; Switzer et al. 2015). The line interlopers, which originate from sources residing in different redshifts emitting spectral lines in the same observed frequency channel, are another pressing issue for LIM experiments. The two most studied line deconfusion techniques, source masking and cross-correlation, typically rely on external data sets that trace the same cosmic volume: the masking technique makes use of a galaxy survey catalog to identify and remove bright interloper sources (Breysse et al. 2015; Silva et al. 2015; Yue et al. 2015; Sun et al. 2018), whereas cross-correlation of an LIM survey with an external (or internal) data set can help extract signals of interest (Lidz et al. 2009; Chang et al. 2010, 2015; Visbal & Loeb 2010; Gong et al. 2012, 2014; Masui et al. 2013; Silva et al. 2015; Croft et al. 2016; Chung et al. 2019). In addition to these two methods, Lidz & Taylor (2016) and Cheng et al. (2016) use the anisotropy of the interloper power spectrum arising from projection to the target line redshift to separate the lines. Gong et al. (2020) distinguish the lines from the same projection effect but using the multipole power spectrum. de Putter et al. (2014) propose to use angular fluctuations of the light to reconstruct the 3D source luminosity density.

Most of the existing line deconfusion methods (e.g., cross-correlation and power spectrum anisotropy) only extract the two-point statistics (power spectrum or correlation function) but lose the phase information of individual line maps, which are valuable for cosmological parameter constraints and systematics control in

the data. With individual line maps, one can extract information beyond two-point statistics in the non-Gaussian intensity maps, especially ones from the epoch of reionization (EoR). For example, Breysse et al. (2017) and Ihle et al. (2019) show that the one-point statistics of the intensity field can help constrain the luminosity function model. In addition, individual line maps can be used directly as density tracers for various cross-correlation, multitracer analysis, delensing of the cosmic microwave background (CMB), and perform consistency tests on different spatial regions with different foreground properties.

In this work, we develop a technique to extract individual line intensity maps from an LIM data set with blended interlopers. Using the fact that when multiple spectral lines emitted by a source are observable in an LIM survey, the redshift of the source can be pinned down by fitting to a set of spectral templates that are unique at each redshift. Without any external tracers or spectroscopic follow-up observations, individual line maps can be directly derived. For demonstration, we apply our technique to simulated data of an LIM survey targeting the EoR [C II] line with multiple low-redshift CO interlopers. In this case, the intensity field of the low-z CO lines ($0.5 \lesssim z \lesssim 1.5$) can be reconstructed since they can be detected in multiple spectral channels.

Kogut et al. (2015) first explored the map-space line deconfusion using the multiline wavelength information in the context of a pencil-beam spectroscopic survey. In this work, we explore the technique in the LIM regime that has a much lower sensitivity and spectral resolution. In this regime, our template-fitting-based technique can obtain the signal-to-noise ratio (S/N) of the desired signals by using the data from multiple frequency channels.

The recent work by Moriwaki et al. (2020) demonstrated the feasibility of LIM phase-space deconfusion with deep learning. They show that, in the absence of noise or foreground components, their algorithm can reconstruct the individual line maps that are mixed in the LIM data set. Their training data generation relies on the assumption of signal clustering and the line luminosity model, whereas in this work, we develop the line deconfusion technique. It only makes use of the spectral feature of the lines, which is more robust against the model uncertainty and the noise.

This paper is organized as follows. First we introduce the model and the survey parameters we used to generate the mock LIM survey data in Section 2. Then, Section 3 describes our line deblending technique. Section 4 presents the results on the fiducial setup. In Section 5, we present the performance of the technique with more practical considerations, and discuss its applications and extensions. The conclusion are given in Section 6. Throughout this paper, we consider a flat ΛCDM cosmology with $n_s = 0.97$, $\sigma_8 = 0.82$, $\Omega_m = 0.26$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.69$, and $h = 0.68$, consistent with the measurement from Planck (Planck Collaboration et al. 2016).

2. Mock Light Cone Construction

For each spatial pixel in an LIM survey, there are a set of spectral channel measurements. Hereafter, the term "light cone" refers to the collection of the spectral measurements in a single pixel. The line deconfusion method introduced in this work is performed on a pixel-by-pixel basis, which only utilizes the spectral information in an individual line of sight (light cone), without taking into account the spatial clustering information, which we leave for future work.

We test our deconfusion technique on simulated light cones. Since the clustering information is not relevant to our technique,

we generate light cones that are based solely on the spectral line luminosity function models and not on the clustering properties. Thus, all light cones are independent from one another, and we also ignore the line-of-sight clustering in this work. This allows for both the light cone construction and deconfusion by parallelization to speed up without affecting the quantification of performance.

As a demonstration of the technique, we assume an LIM experiment targeting the redshifted [C II] fine-structure emission from the EoR: the LIM data set contains multiple low-z CO rotational transitions from low redshifts as interlopers. We note that the technique can be readily applied to any line-confusion problem at other wavelengths.

2.1. Line Signal Models

We model the line emissions of the redshifted [C II] emission and five low-redshift CO J-transitions: {CO(2–1), CO(3–2), CO(4–3), CO(5–4), and CO(6–5)}. In reality, in the submillimeter spectral range of interest (generally in the $\sim$200–300 GHz range), there are higher CO J lines (that are fainter), Galactic and extragalactic dust continuum emissions (the cosmic infrared background), CMB radiation, and atmospheric emissions that can all contribute to the measurements. Since none of them will produce strong spectral features that impact the performance of our technique, we will not include them in the light cones. Instead, in Section 5.3, we will demonstrate that the continuum foreground mitigation in the data analysis process has a negligible impact on our technique performance.

We use the [C II] and CO luminosity function models provided by Popping et al. (2016), in which a semi-analytic model including the effect of radiative transfer was used to estimate the CO and [C II] luminosity functions as constrained by current observations. Here we adopt their fitted Schechter luminosity functions to construct our light cones.

2.2. Survey Parameters

We consider a mock experiment that has similar survey parameters as the two ongoing EoR [C II] LIM experiments, TIME (Crites et al. 2014) and CONCERTO (Lagache et al. 2018).

The mock survey covers 200–305 GHz with 70 evenly spaced spectral channels ($\delta\nu = 1.5$ GHz), and the $\Omega_{\rm pix} = 0.43^2$ arcmin^2 pixel size. We assume the instrument noise is white and has a Gaussian distribution, with four different per-pixel noise levels of standard deviation $\sigma_n = \{10^3, 5 \times 10^3, 10^4, 5 \times 10^4\}$ Jy sr^{-1}. These values are comparable to the range of expected instrument noise in TIME and CONCERTO.
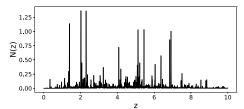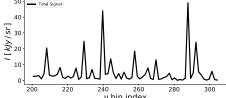
2.3. Light Cone Generation

Based on the assumed luminosity function models and survey parameters, we populate the light cones with sources drawn from random realizations of the Schechter function model. We first define a fiducial Schechter function[4] for all of the lines as a function of redshift,

$$\Phi(\ell) = \phi_*(\mathcal{L}/\mathcal{L}_*)^\alpha \, e^{-(\mathcal{L}/\mathcal{L}_*)}. \tag{1}$$

We discretize the luminosity and redshift into bins of $\Delta(\mathcal{L}/\mathcal{L}_*)$ in 100 luminosity bins in log-space from $\mathcal{L}/\mathcal{L}_* = 10^{-3}$ to 10,

---

[4] For the fiducial Schechter function, we choose the ϕ_* and α values of CO(1–0) for $z \leqslant 5$ and [C II] for $z > 5$ in Popping et al. (2016). We interpolate or extrapolate the Schechter function parameters to the desired redshift from the values of $z = [0, 1, 2, 3, 4, 6]$ given by Popping et al. (2016).
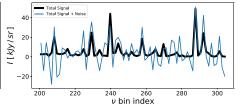
Figure 1. Steps of constructing a light cone. Left: $N(z)$, the effective number of ℓ_* sources in each redshift bin in this light cone. Middle: the signals from all six spectral lines and the sources in the 70 spectral channels. Right: the mock observed light cone (blue) consists of the signals (middle panel) and the $\sigma_n = 10$ kJy Gaussian noise. The black line is the signal component (same as the middle panel) for reference.

and 2000 redshift bins of $\Delta z = 5 \times 10^{-4}$ in linear space from $z = 0$ to $z = 10$. The expectation value of the source counts N within each $\Delta \mathcal{L}$ and Δz bin is given by

$$\langle N \rangle = \Phi(\ell)[\Delta(\mathcal{L}/\mathcal{L}_*)]\left[\Omega_{\mathrm{pix}}D_A^2(z)\frac{d\chi}{dz}\Delta z\right], \qquad (2)$$

where the last square bracket is the co-moving volume of the voxel defined by spatial pixel Ω_{pix} and redshift bin Δz, D_A is the co-moving angular diameter distance, and χ is the co-moving distance.

For each $(\mathcal{L}, z)$ bin, we assign its source counts to a Poisson random number with expectation value $\langle N \rangle$, and then integrate along $\mathcal{L}$ to get the total luminosity in each redshift bin $\mathcal{L}_{\mathrm{tot}}(z)$. We define $N(z) \equiv \mathcal{L}_{\mathrm{tot}}/\mathcal{L}_*$ as the "effective number counts" per redshift bin. The left panel of Figure 1 shows the $N(z)$ of one example light cone.

Next, we assign the line luminosity signals in each redshift bin to $N(z)\ell_*^{\mathrm{line}}$, where ℓ_*^{line} is the ℓ_* value of the line in the Popping et al. (2016) model. We then project the line signals to their corresponding spectral channels to make the light cones. The middle panel of Figure 1 shows an example light cone spectrum from the $N(z)$ shown in the left panel.

Finally, we add a Gaussian random fluctuation with an rms σ_n value to each channel to account for the instrumental noise. The right panel of Figure 1 shows the same light cone with a $\sigma_n = 10$ kJy Gaussian noise.

This light cone construction procedure assumes that all of the spectral lines have the same Schechter function parameters (ϕ_* and α) and thus the same luminosity function shape; the CO spectral line energy distribution (SLED) is also fixed. That is, by construction, all of the sources have the same line luminosity ratio sets by the relative value of ℓ_*^{line}. In the main parts of this work, we use this fixed SLED model, and we test the impact of adding the SLED variation on the performance of our technique in Section 5.1.

We further point out that, even though the luminosity function has been sampled to the faint end ($10^{-3}\ell_*$), the light cone signals are still dominated by a few bright peaks. This indicates that the emission field can be well described by the few bright sources in the data. In other words, to extract the line emission field from a single line, one only needs to determine the redshift and luminosity of those bright sources. This is the main concept of our deconfusion technique, detailed in Section 3.

3. Methods

3.1. Formalism

The intensity of a light cone in frequency channel ν_i can be expressed as the linear combination of signals from all N_z

redshift bins and the noise n_i,

$$I(\nu_i) = \sum_{j=1}^{N_z} \widetilde{A}_{ij} N(z_j) + n_i, \qquad (3)$$

where $N(z_j)$ is the effective number of ℓ_* sources in redshift z_j (Section 2.3), and $\widetilde{A}_{ij}$ converts $N(z_j)$ to the observed intensity in channel ν_i.

$\widetilde{A}$ is an $N_{\mathrm{ch}} \times N_z$ matrix, where we have $N_{\mathrm{ch}} = 70$ spectral channels, and we use $N_z = 2000$ ($\Delta z = 5 \times 10^{-4}$) redshift bins. In principle, we can set Δz to infinitesimally small values; however, in practice the redshift resolution is limited by the instrument spectral resolution, as the SLEDs of nearby galaxies can be highly degenerate when their respective spectral lines fall in the same set of observed channels. Setting $N_z = 2000$ in fact gives a much finer redshift resolution compared to the spectral channel width. We will use this fine redshift resolution as a starting point, and discuss the strategy to reduce redshift bins and therefore remove redundant information at the end of Section 3.1.

Most of the elements in $\widetilde{A}_{ij}$ are zeros except for which the sources at z_j emit a spectral line at the observed frequency ν_i. In this case,

$$\widetilde{A}_{ij} \equiv I_*^{\mathrm{line}}(z_j) = \ell_*^{\mathrm{line}}(z_j)\frac{1}{4\pi D_L^2(z_j)\delta\nu_i\Omega_{\mathrm{pix}}}, \qquad (4)$$

where ℓ_*^{line} is the line luminosity of an ℓ_* source in the model, D_L is the luminosity distance, and I_*^{line} is defined as the observed line intensity of an $\ell_*^{\mathrm{line}}(z_j)$ source at z_j. Figure 2 shows the I_*^{line} in our assumed model.[5] Note that Equation (3) assumes that all of the sources at the same redshift have the same SLED, and we will first build our technique based on this assumption. In reality, the SLED varies across galaxy type. In Section 5.1, we will show that our method also works in the realistic level of SLED variation.

Equation (3) can be written in the matrix form,

$$\boldsymbol{I} = \widetilde{A}\boldsymbol{N} + \boldsymbol{n}, \qquad (5)$$

where $\boldsymbol{I}$ and $\boldsymbol{n}$ are N_{ch}-element column vectors, $\boldsymbol{N}$ is an N_z-element column vector, and $\widetilde{A}$ is an $N_{\mathrm{ch}} \times N_z$ matrix. The top panel of Figure 3 shows the $\widetilde{A}$ matrix in our model. The $\widetilde{A}$ matrix is mostly zeros, and the six curves from left to right are the six spectral lines, CO(2–1), CO(3–2), CO(4–3), CO(5–4), CO(6–5), and [C II].

---

[5] The glitch in CO(5–4) at $z \sim 1$ is due to a feature in the Schechter luminosity function model in Popping et al. (2016). In their CO(5–4) luminosity function fit at $z = 1$, a slightly higher L_* and and a lower α value are derived that cause the slight apparent discontinuity in redshift, even though the luminosity function does not show an abrupt change at $z = 1$.
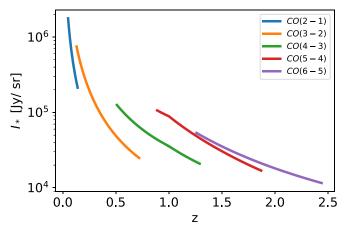
Figure 2. Intensity of the CO lines I_* from the sources of characteristic luminosity ℓ_*.

The goal of our line deconfusion technique is to solve for the source count vector N in a given observed light cone data I and a model $\widetilde{A}$ matrix. With the N solution, the intensity map of individual spectral lines I_{line} can be reconstructed by $I_{\text{line}} = \widetilde{A}_{\text{line}} N$, where $\widetilde{A}_{\text{line}}$ is $\widetilde{A}$ when only the target spectral line signals are turned on.

We further rewrite Equation (3) by normalization and reducing the nuisance information. First, we normalize the columns in $\widetilde{A}$ and move the normalization factor I_j^{norm} to their N element:

$$I = A\widetilde{N} + n, \qquad (6)$$

where

$$A_{ij} = \widetilde{A}_{ij} \big/ I_j^{\text{norm}},$$
$$\widetilde{N}_j = N_j \, I_j^{\text{norm}},$$
$$\sum_{i=1}^{N_{\text{ch}}} A_{ij}^2 = 1. \qquad (7)$$

Next, we will reduce the nuisance or redundant elements in Equation (6). The columns of A are the basis spanning the observed data space. In constructing $\widetilde{A}$, we simply design the columns to be equally spaced redshift bins as shown in the top panel of Figure 3. However, this natural basis is highly degenerate. To remove the nuisance information, we first discard the redshift bins that are zero vectors in $\widetilde{A}$. The sources in these redshifts do not emit lines in observable frequencies, and thus no information can be used to constrain their $N(z)$. Second, for the redshift bins containing only one CO line, their normalized columns in A are identical to other columns having a [C II] signal in the same frequency channel. In other words, given observed data I, we cannot distinguish the origin of the source with a single line emission. Therefore, we combine these identical columns into a single column. In conclusion, we keep the columns in A with redshift bins that can be observed with multiple spectral lines, plus an identity matrix for the redshift bins that only have a single detectable line.

For all of the redshifts that can be observed in multiple lines, we design the size of the redshift bins based on the following two competing considerations. On the one hand, the redshift bins have to be small enough to faithfully represent the emitting source distribution. On the other hand, finer redshift bins give larger $\widetilde{N}$ size, and therefore more unknown parameters to be solved. The information can be compressed by combining some neighboring redshift bins, which are highly degenerate, since they have signals in the same channels with similar amplitudes. Therefore, we design the redshift bins using the following procedures: we (1) generate A with fine redshift bins ($\Delta z = 5 \times 10^{-4}$ from $z = 0$ to 10),[6] (2) keep the columns with multiple lines, (3) then identify the group of neighboring columns that have signals in the same sets of channels, and (4) keep the medium bin and discard the others. With this process, we get 195 nondegenerate columns. Hereafter, "multiline redshift bins" refers to these 195 redshift bins that can be detected in multiple channels. Finally, we append the N_{ch}-sized identity matrix (that accounts for the single-line redshift bins) to these 195 columns to generate A. The bottom left panel in Figure 3 shows matrix A, which has size $N_{\text{ch}} \times N_z = 70 \times 265$ (195 multiline redshifts plus 70 columns of identity matrix). The bottom right panel shows the redshift of the 195 multiline bins, and their color labels are the pairs of detectable CO lines. We also define six broad bands from these pairs of lines by binning groups of channels. Table 1 lists the definition of the broad bands. Note that the CO(2–1) and CO(3–2) overlapping redshifts of $0.13 < z < 0.5$ can also be reconstructed, but they are only covered by four frequency channels, which makes it difficult to quantify the reconstruction performance with sufficient statistical power. Therefore, we ignore these redshifts in our analysis.

Our line deconfusion technique can only solve $\widetilde{N}$ in the 195 multiline redshift bins. The last 70 elements in $\widetilde{N}$ that correspond to the identity matrix in A are nuisance parameters, since they represent degenerate "single-line" signals from different redshifts—that is to say, we cannot reconstruct the signals in the single-line redshifts, which are the regions not covered by the lines in the bottom left panel of Figure 3 ($0.15 \leqslant z \leqslant 0.51$; $0.72 \leqslant z \leqslant 0.89$; $z \geqslant 1.87$). In the following analysis, we will only focus on the reconstruction of $\widetilde{N}$ in the 195 multiline redshift bins.

3.2. Sparse Approximation

The key step in our deconfusion technique is to solve for $\widetilde{N}$ in Equation (6), given the observed spectrum I and model A. This type of linear system has been extensively studied in the context of CMB map making, in which I, n, and $\widetilde{N}$ in Equation (6) can be analogized to the time-ordered data, time-stream noise, and the pointing matrix, respectively. However, contrary to the map-making problem, our system is an ill-posed problem as there are more unknown variables ($N_z = 265$) than the input data points ($N_{\text{ch}} = 70$). Thus the standard map-making algorithm (e.g., Stompor et al. 2002) cannot be applied.

In Equation (6), the columns of A form a basis for I, and the solution $\widetilde{N}$ is the linear combination coefficient. The columns of A form an over-complete basis, since the A matrix is only of rank N_{ch}, and thus the solution $\widetilde{N}$ is not unique. Indeed, for any given observed data I, there are infinite $\widetilde{N}$ that can perfectly fit the input.

Nevertheless, Equation (6) can be solved with the "sparse" condition, which means the preferred solution of $\widetilde{N}$ is the one with a small number of nonzero elements. With this constraint, we can solve Equation (6) with the following well-defined

---

[6] Note that with these fine bins, the same frequency bin can map to multiple redshift bins instead of a one-to-one mapping.
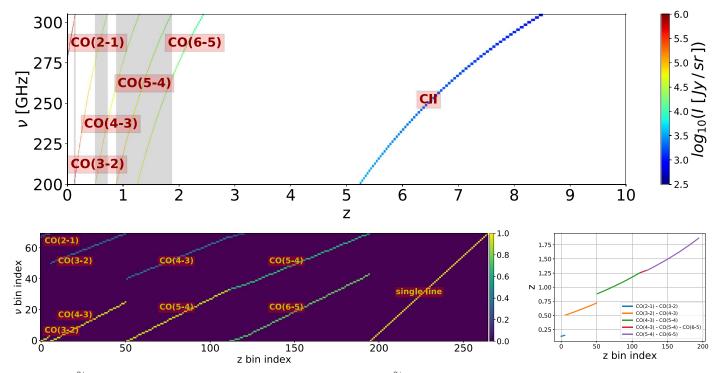
Figure 3. Top: $\widetilde{A}$ matrix with our model on 70 frequency channels and 2000 redshift bins. $\widetilde{A}$ are zeros (white) for the majority of the elements, and the six curves from left to right correspond to the six spectral lines, CO(2–1), CO(3–2), CO(4–3), CO(5–4), CO(6–5), and [C II]. The color scale indicates line intensities in the fiducial model. The gray shaded regions are the redshifts at which sources can be observed in multiple lines, and thus can be reconstructed in our technique. Bottom left: A matrix with $N_{ch} \times N_z = 70 \times 265$ size, which is the reduced and normalized $\widetilde{A}$. The color scale represents the intensities with the fiducial model SLED normalized within each column, i.e., for all column j, $\sum_{i=1}^{N_{ch}} A_{ij}^2 = 1$. Bottom right: the redshift of the 195 multiline redshift bins in A, which are the redshift bins that have multiple CO lines observable in our mock survey. The colors label the pairs of detectable CO lines. The redshifts not covered by multiple CO lines cannot be reconstructed with our technique ($0.15 \leqslant z \leqslant 0.51$; $0.72 \leqslant z \leqslant 0.89$; $z \geqslant 1.87$).

Table 1
Frequencies and Redshifts of the Six Defined Broad Bands

Name	Line	ν Bin Index	ν (GHz)	$\langle z \rangle$ ($z_{min} - z_{max}$)
J3 high	CO(3–2)	51–69 (19 bins)	200.75–227.75	0.61 (0.51–0.72)
J4 low	CO(4–3)	0–24 (25 bins)	268.25–304.25	0.61 (0.51–0.72)
J4 high	CO(4–3)	41–69 (29 bins)	200.75–242.75	1.09 (0.89–1.30)
J5 low	CO(5–4)	0–36 (37 bins)	250.25–304.25	1.09 (0.89–1.30)
J5 high	CO(5–4)	37–69 (33 bins)	200.75–248.75	1.59 (1.30–1.87)
J6 low	CO(6–5)	0–42 (43 bins)	241.25–304.25	1.56 (1.26–1.87)

Note. The CO(2–1) and CO(3–2) overlapping redshifts of $0.13 < z < 0.5$ can also be reconstructed, but they are only covered by four frequency channels, which makes it difficult to quantify the reconstruction performance with sufficient statistical power. Therefore, we ignore these redshifts in our analysis.

optimization problem:

$$\underset{\|\widetilde{N}\|_0}{\mathrm{argmin}} \frac{1}{N_{ch}} \|I - A\widetilde{N}\|_2^2 < \epsilon^2, \qquad (8)$$

where the ℓ_0-norm $\|\cdot\|_0$ is the number of nonzero elements, and ϵ sets the threshold of error tolerance of the fit.

This type of problem is known as "sparse approximation," which has been extensively studied in the context of signal processing and compressive sensing (Candes et al. 2006; Donoho 2006). The sparse approximation algorithms solve the sparse representation of the signal in a "dictionary" that is composed of a set of "atoms," and represent the signal in the

data in terms of the linear combination of a few atoms in the dictionary. In Equation (8), the dictionary is the matrix A, and the atoms are the column vectors of A.

The sparse approximation can only be applied if $k \equiv \|\widetilde{N}\|_0 \ll N_{ch}$. Note that the sparsity of the problem is quantify by k/N_{ch} but not k/N_z, since k/N_z can always be designed to be arbitrarily small by choosing a large basis (fine redshift bins). However, the degree of freedom in the solution is restricted by the input size N_{ch}, and thus the $k \ll N_{ch}$ condition prohibits the algorithm from using more parameters than the input degree of freedom to over-fit the data.

In general, LIM light cones are not sparse, since there is always a large number of faint sources in the typical luminosity function (e.g., Schechter function), so all of the elements in $\widetilde{N}$ are nonzero. However, as mentioned in Section 2.3, the light cone signals are dominated by only a few bright sources, and the intensity field can be well described by them. Consequently, the parameter k in our problem can be quantified by the "effective" number of these bright sources per voxel that contribute most of the emission. Following Cheng et al. (2019), we define the effective number N_{eff} as:

$$N_{eff}(z) \equiv \frac{\left(V_{vox} \int d\ell \, \Phi(\ell, z) \, \ell\right)^2}{V_{vox} \int d\ell \, \Phi(\ell, z) \, \ell^2}, \qquad (9)$$

where V_{vox} is the voxel size of the redshift bin. Note that Φ is the number of sources per luminosity per voxel, and therefore N_{eff} is dimensionless and is proportional to the voxel size. N_{eff} can be interpreted as the reciprocal of the effective shot noise in LIM, which is analogous to the $1/N$ shot noise in a galaxy
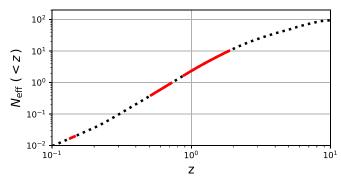
Figure 4. Redshift cumulated N_{eff} in our model (black dashed line). The red segments mark the multiline redshift ranges that have multiple observable CO lines. Thus, the deconfusion technique can be applied.

power spectrum. If the luminosity function Φ follows the Schechter function form, then N_{eff} is (approximately) the number of sources brighter than ℓ_*, which contributes the majority of the emission.

We can estimate k by the cumulative $N_{eff}(z)$ along the line of sight per light cone. Figure 4 shows the cumulative N_{eff} in our model. While $N_{eff} \sim 100$ from $z = 0$ to 10, the only relevant range is $z \lesssim 2.5$, where CO lines fall in the observed frequency range (see Figure 3, top panel). Above $z \sim 2.5$, only the high-z [C II] lines can be observed, but they are much fainter then the CO signals and the assumed noise level, so they can be treated as background fluctuations. Therefore, for $z \lesssim 2.5$, we find $N_{eff} \sim 10 \ll N_{ch} = 70$, so the sparse condition is qualified in our problem.

3.3. The Matching Pursuit Algorithm

We use the matching pursuit (MP) algorithm first introduced by Mallat & Zhang (1993) to solve for Equation (8). The MP algorithm iteratively selects an "atom" in the "dictionary" to project out part of the signals in the data, and keep track of the current solution of the signal and residual for the next step until the stopping criteria is met. In our case, the column vectors in A are the atoms that form the dictionary space. A detailed description of the MP algorithm is in Appendix A.

In each step of the MP algorithm, the selected atom is the one that has the maximum inner product with the residual. The S/N of the signals in each step is the ratio of that maximum inner product to the instrument noise level σ_n (see Appendix B for the proof). Therefore, if we set the stopping criteria to be the maximum inner product smaller than m times of σ_n, then this is an m-σ detection threshold on the signals (e.g., $m = 5$ for a 5σ detection). Note that the detection threshold here is based on the combined information in multiple spectral channels projected onto the dictionary space.

The choice of detection threshold "m" is a trade-off between the purity and completeness of the source extraction. Higher "m" values give a higher purity map, whereas lower "m" values pick out fainter sources at the cost of increased false detections from noise. The optimal value of "m" depends on the instrument sensitivity and the purpose of the reconstruction map. For example, to reconstruct the line luminosity function, one might use a higher threshold to reduce the false detections at the faint end; whereas to constrain the large-scale structure, a lower threshold is preferred to reduce the shot noise in the power spectrum. An analytical formalism to determine the optimal threshold is to make use of the Fisher information

framework (Cheng et al. 2019), where one calculates the expectation value of the desired observable (e.g., power spectrum) as a function of threshold for a given signal model and noise level, and estimates the threshold that optimizes the Fisher information. Alternatively, one can simply perform test simulations with different thresholds to determine the optimal value. Both approaches can provide guidance on choosing the optimal threshold for the problem at hand, and a detailed investigation is beyond the scope of this paper.

4. Results

We present line deconfusion results in the simple case where mock light cones and the template (A) are both generated from the same signal model (Section 2.3). We demonstrate that in this scenario, our technique is capable of extracting low-z CO signals in the presence of realistic instrumental noise. We discuss the robustness of the performance against uncertainties in the signal model and contamination from astrophysical foregrounds, and extend the application to spectral lines in different wavelengths in Section 5.

We quantify the reconstruction performance by computing two statistics on the true and reconstructed data: (1) the Pearson correlation coefficient (Section 4.2) and (2) the voxel intensity distribution (VID; Section 4.3). The former quantifies the phase-space information, whereas the latter captures the one-point statistics that describes the distribution of voxel intensities.

Finally, we present results with a variety of instrument noise levels σ_n and the reconstruction threshold m. For each test, we use 2500 mock light cones to calculate the correlation coefficient and VID, and estimate errors with 100 noise realizations.

4.1. Visualization of Example Results

Figure 5 visualizes the reconstruction results on one of the spectrum bins (274 GHz) that contains five lines (CO(3–2), CO(4–3), CO(5–4), CO(6–5), and [C II]) in the observable band. We place the 400 light cones into a 20×20 pixel map, and show the true and the reconstructed intensity fields. Note that the input signals do not exhibit spatial clustering because each light cone is generated independently. The top left panel shows the total signal intensity from all five lines, and the top right panel shows the observed intensity including the total line signals and a Gaussian instrument noise with $\sigma_n = 10^4$ Jy sr^{-1}. In this channel, three of the lines (CO(4–3), CO(5–4), and CO(6–5)) are in the multiline regime, so they can be reconstructed with our algorithm. The three bottom panels compare the true input to the reconstructed intensity maps for these lines with 10σ, 5σ, and 1σ reconstruction threshold as the MP algorithm stopping criteria.

The choice of threshold is a trade-off between the completeness and the purity in the reconstructed map. As shown in Figure 5, in the high threshold case (10σ), the MP algorithm only extracts a few bright sources that are above the reconstruction threshold. As the threshold decreases, more sources are reconstructed, at the cost of increased false positive detection from the noise fluctuations or the interlopers. This example provides a visual depiction of the reconstruction algorithm. To further quantify the reconstruction performance, we consider two summary statistics in the following sections.
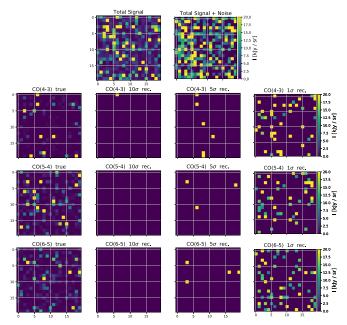
Figure 5. Visualization of the phase-space reconstruction with $\sigma_n = 10^4$ Jy sr^{-1}. We place the 400 toy model light cones into a 20×20 pixel map, and show the reconstruction results on one of the spectrum bins (274 GHz). The top left panel is the true signal intensity map from all of the spectral lines (CO (3–2), CO(4–3), CO(5–4), CO(6–5), and [C II]). The top right panel is the observed intensity map including line signals < (top left panel) and noise. The bottom three panels show the true input and the 10σ, 5σ, and 1σ reconstructed (right) emission field of the three spectral lines in the multiline regime where the signal can be reconstructed.

4.2. Pearson Correlation Coefficient

We quantify the reconstruction performance by the Pearson correlation coefficient between the true and the reconstructed maps in each channel. The Pearson correlation coefficient is defined by

$$r = \frac{\sum_{i=1}^{N_{lc}}(I_{true}^i - \langle I_{true}\rangle)(I_{rec}^i - \langle I_{rec}\rangle)}{\sqrt{\sum_{i=1}^{N_{lc}}(I_{true}^i - \langle I_{true}\rangle)^2}\sqrt{\sum_{i=1}^{N_{lc}}(I_{rec}^i - \langle I_{rec}\rangle)^2}}, \quad (10)$$

where $N_{lc} = 2500$ is the number of light cones, I_{true}^i and I_{rec}^i are the true and the reconstructed line intensity maps, respectively, at the ith light cone. Figure 6 shows the results of the correlation coefficient with a $\sigma_n = 10^4$ Jy sr^{-1} noise level and a 5σ reconstruction threshold. Our reconstructed map achieves $\sim80\%$ correlation with the true input map at $z \lesssim 1.5$.

Figure 7 shows the correlation coefficient r within a range of noise level σ_n and the reconstruction threshold m. For simplicity, we show the average r value of each broad band defined in Table 1. The key findings are summarized below.

1. At a fixed m value, r decreases as σ_n increases. This is because under the same purity (same m), the detection threshold $m\sigma_n$ is higher for higher σ_n, and thus fewer sources have been reconstructed.

2. The six defined bands correspond to three pairs of lines from different redshift bands (Table 1). The lines within each pair are strongly correlated because they are the signals from the same sources, and thus they are reconstructed in the same MP iteration. The pairs of lines in the same redshift are reconstructed in the same MP iteration, and as a result, they are highly correlated.
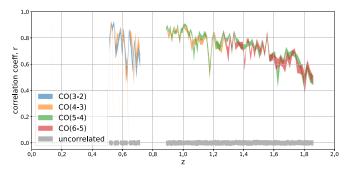

Figure 6. Pearson correlation coefficient r between the true and the reconstructed maps on 2500 light cones with $\sigma_n = 10^4$ Jy sr^{-1} 5σ reconstruction. The bands are the 1σ scatter of 100 noise realizations with the sample line signal. The gray bands are the r value with the white noise map for reference.

3. Because of the purity and completeness trade-off, the maximum correlation r happens at the intermediate threshold m (except for the lowest-noise $\sigma_n = 10^3$ Jy sr^{-1} case, discussed in the next enumerated point).

4. Correlation coefficient r has very low dependency on σ_{th} in the $\sigma_n = 10^3$ Jy sr^{-1} case. This can be understood by comparing the noise level σ_n to the quantity I_*, the intensity of the ℓ_* source in the Schechter function (Figure 2). At the redshift range in which we perform the reconstruction ($0.5 \lesssim z \lesssim 1.9$), $10^4 \lesssim I_* \lesssim 10^5$ Jy sr^{-1}, which indicates that $m\sigma_n < I_*$ for all of the m values considered in Figure 7 ($m = 1 \sim 10$). In the Schechter luminosity function, the sources $\gtrsim \ell_*$ contribute the majority of the information in the intensity field (Cheng et al. 2019), and thus the correlation coefficient r is not sensitive to the change in reconstruction threshold if $m\sigma_n \ll \ell_*$.

4.3. VID

The Pearson correlation coefficient traces the phase-space variations between the true and reconstructed maps, but it cannot distinguish a systematic constant offset, i.e., if the reconstructed line signals are systematically lower or higher than the true input. Therefore, we check the consistency of the reconstructed and true input maps using the one-point statistics, VID. Note that we do not directly compare the mean intensity of each map since we only reconstruct the bright sources that are above the threshold and neglect all faint sources in the reconstructed map, so the mean intensity is not expected be faithfully recovered.

The VID of an LIM map contains information beyond the power spectrum, and is valuable for LIM targeting a late-time universe where the large-scale structure is highly non-Gaussian and cannot be fully described by two-point statistics. For example, Breysse et al. (2017) showed that the VID can constrain the luminosity function model parameters, and Ihle et al. (2019) demonstrated that a joint analysis of the power spectrum and VID improved the constraining power on the source luminosity function.

Figure 8 compares the VID of the true and reconstructed maps in the six broad bands with noise level $\sigma_n = 10^4$ Jy sr^{-1} and reconstruction threshold $m = 4$. The gray lines compose the VID of the total observed map, which includes signals from all of the lines and noise. The black lines are the VID of the target line maps, and the red data points are the VID of the reconstructed target line map. While the VID of the total
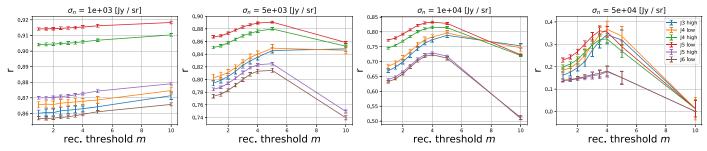
Figure 7. Average correlation coefficient r in each broad band (Table 1) within a range of noise level σ_n and the reconstruction threshold m. The error bars are the rms of 100 noise realizations.
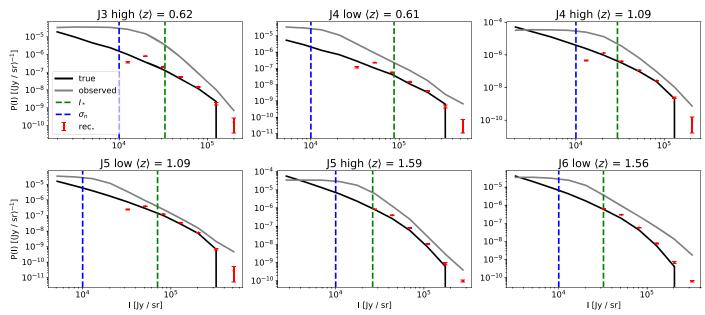

Figure 8. VID of the true input (black) and the reconstructed (red) maps in the six broad bands with noise level $\sigma_n = 10^4$ Jy sr^{-1} and reconstruction threshold $m = 4$. The error bars are the rms of 100 noise realizations. The gray curves are the VID of the total observed map, which includes signals from all of the lines and noise. For reference, the blue and green dashed lines mark the noise level σ_n and I_*, respectively.

observed map is one to two orders of magnitude above that of the target line signal, our reconstruction technique can faithfully recover the VID of the signal to slightly below the I_*-scale, the characteristic source luminosity in the Schechter function.

In this realistic survey setup with the assumed signal model, we show that our method can successfully reconstruct the VID of the CO signal down to $\sim \ell_*$ scales. This provides a strong constraint on the CO luminosity function, as both are 1D statistics of the intensity field and are closely related. The CO luminosity function at various redshifts provides valuable insight on the formation and evolution of galaxies across cosmic time. Specifically, CO is a tracer of H_2 gas in the interstellar medium, and can therefore be used to study the evolution of the molecular gas content and their distribution as a function of time (Walter et al. 2014; Decarli et al. 2019; Riechers et al. 2019). We note though that the expected S/N on the luminosity function depends on the assumed model.

5. Discussion

5.1. Model Uncertainty

For the results presented in Section 4, the light cone signals and the dictionary template A are both generated from the same assumed signal model (Section 2). However, in reality, the

variation in SLED across galaxies will affect the reconstruction performance. To test how the SLED uncertainties affect the reconstruction, we apply three different SLED model variations and bias levels (at 20%, 50%, and 100%) to the mock data, and run the reconstruction with the same dictionary template A. We detail the definition of variation and bias below. Daddi et al. (2015) measured multiple CO lines of ultraluminous infrared galaxies (ULIRGs) at $z \sim 1.5$, and estimated a $\sim$20% variation on the CO SLED ratio for their sample. Therefore, an assumed 50% or 100% variation can be more extreme than realistic variations.

5.1.1. SLED Model Variation

First, we test the case with SLED model variations. For each line of each source in the mock light cone, we assign a line luminosity L^{line}:

$$L^{\text{line}} = L^{\text{line}}_{\text{fid}}(1 + \delta_L), \qquad (11)$$

where $L^{\text{line}}_{\text{fid}}$ is the fiducial luminosity from our model, and δ_L is a zero-mean Gaussian random variable with standard deviations of 0.2, 0.5, and 1.0 for the 20%, 50%, and 100% variation cases, respectively.

Figure 9 shows the reconstructed correlation coefficients r with a 4σ reconstruction threshold ($m = 4$) with different SLED variations. We see that introducing 20% SLED variations in the
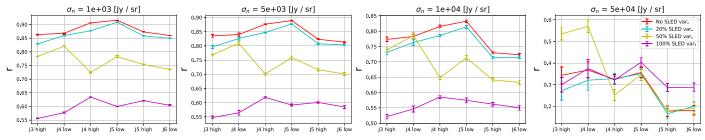
Figure 9. Comparing r of the no SLED variation (red), 20% variation (cyan), 50% variation (yellow), and 100% variation (purple) with the 4σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the rms of 100 noise realizations of all of the spectral bins in each band.
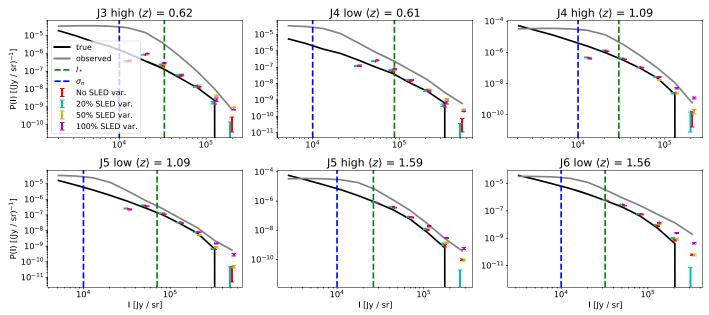

Figure 10. VID of the no SLED variation (red), 20% variation (cyan), 50% variation (yellow), and 100% variation (purple), and the true (black) maps in the six broad bands with noise level $\sigma_n = 10^4$ Jy sr^{-1} and reconstruction threshold $m = 4$. The error bars are the rms of 100 noise realizations. The gray curves are the VID of the total observed map, which includes signals from all of the lines and noise. For reference, the blue and green dashed lines mark the noise level σ_n and I_*.

model has a very mild impact on the reconstructed correlation; whereas with 50% (100%) variation, the correlation coefficient drops by about 10% ($\sim$50%). Figure 10 compares the VID. We see that the SLED fluctuation has a negligible impact on the VID reconstruction in the 20%, 50%, and 100% variation cases. Hence, we conclude that our technique is robust against a realistic level of CO SLED fluctuations.

5.1.2. Model Offset

In addition to SLED variation, we also test whether our model gives biased estimates of the average SLED by assigning line luminosity $L^{\rm line}$ as

$$L^{\rm line} = L^{\rm line}_{\rm fid}(1 + b_L), \qquad (12)$$

where b_L is a constant offset that we assign to the model lines, and it is applied to all of the sources in the mock light cones. For the 20% bias level, we apply $b_L = (+0.1, -0.1, +0.1, -0.1)$ for the four CO lines CO(3–2), CO(4–3), CO(5–4), and CO(6–5) to ensure the SLED ratio between neighboring lines is 20%, since our algorithm is only sensitive to the line ratio between two neighboring CO lines.

Figures 11 and 12 show the reconstructed correlation coefficients r and VID with a 4σ with three SLED bias levels.

Similar to Figure 9, the correlation only drops significantly when the bias is tuned to 100%. Therefore, our technique is also robust against a realistic level of potential CO SLED bias.

5.2. Application to LIM in Other Wavelengths

The technique developed in this work is not restricted to the [C II] and CO lines' blending problem. It can in principle be applied to a range of LIM experimental setups. As a demonstration, we apply our method to reconstruct near-infrared lines in a Spectro-Photometer for the History of the Universe, Epoch of Reionization, and Ices Explorer (SPHEREx)-like survey.

SPHEREx is an ongoing NASA MIDEX mission to conduct an all-sky near-infrared spectro-imaging survey (Doré et al. 2014).[7] SPHEREx will carry out the first all-sky spectral survey at wavelengths between 0.75 and 5 μm with 96 spectral channels and a $6\rlap{.}''2$ pixel size. Lyα (121.6 nm), Hα (656.3 nm), Hβ (486.1 nm), [O II] (372.7 nm), and [O III] (500.7 nm) are the five prominent lines detectable by SPHEREx across a range of redshifts especially in the LIM regime.

The line signal model is described in Appendix C. We generate near-infrared LIM mock data with a $6\rlap{.}''2 \times 6\rlap{.}''2$ pixel size and a 5σ point-source sensitivity of $m_{\rm AB} = 22$ (similar depth as the
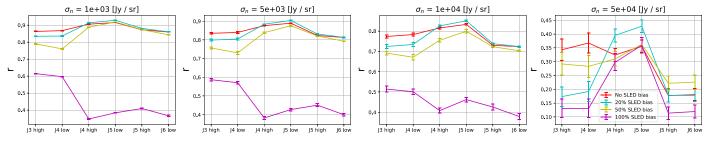
---

[7] http://spherex.caltech.edu

Figure 11. Comparison of r of the no SLED bias (red), 20% bias (cyan), 50% bias (yellow), and 100% bias (purple) with a 4σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the rms of 100 noise realizations of all of the spectral bins in each band.

SPHEREx deep fields), and run our deconfusion algorithm on the mock light cones. Figure 13 shows the results of the correlation coefficients between the true and 3σ-threshold reconstructed intensity maps. The reconstructed map achieves $\sim$80% correlation with the true input map at $z \lesssim 3$, and decreases toward higher redshifts, as the I_* of the lines approaches the noise level σ_n. At $z \sim 5$, the brightest line, Hα, is redshifted into SPHEREx bands with a high spectral resolution of $R \sim 130$ and suffers less signal dilation, resulting in an increase of S/N on a single source detection, and thus r slightly rebounds at this redshift.

To account for uncertainties in modeling the SLED, we apply a realistic level of line luminosity variations from Moustakas et al. (2006). We apply 10%, 50%, and 100% SLED variations (Equation (11)) to the line luminosity ratio of H$_\alpha$/Hβ, [O II]/Hα, and [O III]/[O II], respectively. Figure 14 shows the correlation coefficients between the true and 3σ-threshold reconstructed intensity maps. Comparing with the fixed SLED case (Figure 13), only the Hβ line shows a significant decrease in performance. The intensity map of the three brighter lines (Hα, [O II], and [O III]) can still be extracted with $\gtrsim$70% correlation compared to the true input.

We conclude that our algorithm can reasonably well reconstruct the phase-space LIM signal in a SPHEREx-like experiment, given the expected variation of (redshifted) optical line ratios. The technique can be generalized to different LIM experimental applications, and the reconstructions are fairly robust against uncertainties in the SLED modeling.

5.3. Foreground Subtraction

In addition to line interlopers, LIM data are subjected to strong continuum foreground from various sources. For the frequency range considered in this work ($\sim$200–300 GHz), the dominant foregrounds are the atmospheric emission, dust continuum, and the CMB; whereas for LIM in the near-infrared, e.g., SPHEREx (Doré et al. 2014), the zodiacal light and the galaxy stellar continuum are the dominant continuum foregrounds. Even though these foregrounds are brighter than the sought-after line signals, their spectral responses are expected to be smooth and are distinct from the spectral line features, so that the continuum foregrounds can be separated and mitigated, for example by a smooth function fit such as a low-order polynomial.[8] Here we test how the foreground mitigation process affects our line reconstruction results.

We consider two cases of foreground mitigation. First we emulate the foreground removal process in the presence of an approximately constant foreground in both the spatial and spectral dimensions, for example the zodiacal light. In this case,

---

[8] Some foreground components are also spatially smooth (e.g., zodiacal light), which can be filtered in the spatial domain as well.

before running the reconstruction, we subtract the mean value of the whole data cube, i.e., the mean intensity in $N_{\rm lc} \times N_\nu$ voxels. The second case is to emulate the continuum subtraction process of the galaxy stellar or dust continuum, which are expected to have smooth spectra but be different in each light cone, since each light cone contains different galaxies with different continuum spectra. We fit and subtract a first-order polynomial function to the spectrum of each light cone before running the reconstruction.

The results of a 4σ-threshold reconstruction with different noise σ_n level are shown in Figures 15 and 16. We can see that, compared to the no background subtraction case, the r value is even higher in these two tests. This is because our reconstruction only extracts the bright lines, and the fainter lines act as a background for the MP algorithm. The signals from the fainter lines introduce not only fluctuations but also a bias in the data, since the line signals are always positive unlike the zero-mean noise. The reconstruction performance is improved after background subtraction because this bias level is also removed during this process. For the VID results, we see that there is no significant difference compared to the no background subtraction case. This is again due to the fact that the background level is much fainter than the brightness of the sources being extracted with our algorithm, so the background subtraction has no impact on the reconstruction. In conclusion, the background subtraction in the LIM data reduction pipeline will not affect our line reconstruction technique.

5.4. Prior with External Catalogs

Our analysis uses the LIM data itself without invoking any external information. In practice, initial LIM survey fields are designed to in part overlap with existing photometric or spectroscopic galaxy surveys, and thus there will be information provided by external galaxy catalogs to aid the line deconfusion problem. Ignoring redshift uncertainties of the external catalogs, one simple approach to incorporate the external information is to force the MP algorithm to first select the redshift bins that contain galaxies from the catalog. After iterating through the catalog sources, we then continue the normal MP procedure until hitting the stopping criteria.

To test the effect of including prior knowledge from external catalogs, we generate a mock catalog by selecting sources with CO(5–4) flux greater than $150\,L_\odot\,{\rm Mpc}^{-2}$ ($6 \times 10^{-17}\,{\rm W\,m}^{-2}$) in each light cone. The source density in the catalog is $\sim$1.2 per light cone (integrated along the line of sight) for a 0.43^2 arcmin^2 pixel solid angle. The flux cut corresponds to an L_* galaxy at $z \sim 2$. According to Helgason et al. (2012), such an L_* galaxy has an absolute magnitude $M_{\rm AB} \sim -23$ in the optical, which gives an apparent magnitude of $m_{\rm AB} \sim 21.8$, approximately the depth of the assumed optical catalog.
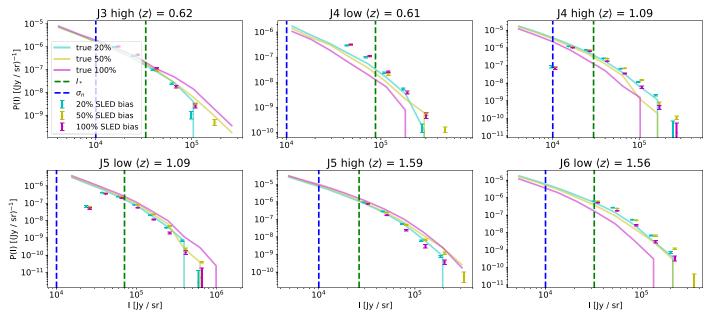
Figure 12. VID of the no SLED bias (red), 20% bias (cyan), 50% bias (yellow), and 100% bias (purple), and the true maps (solid lines) in the six broad bands with noise level $\sigma_n = 10^4$ Jy sr^{-1} and reconstruction threshold $m = 4$. The error bars are the rms of 100 noise realizations. The gray curves are the VID of the total observed map, which includes signals from all of the lines and noise. For reference, the blue and green dashed lines mark the noise level σ_n and I_*.
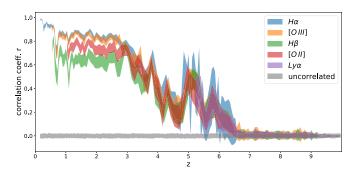

Figure 13. Pearson correlation coefficient r between the true and the reconstructed maps on 2500 light cones for SPHEREx-like mock data. The bands are the rms of the value in 100 noise realizations with the sample line signal. The gray bands are the correlation coefficient with the uncorrelated white noise map for reference.
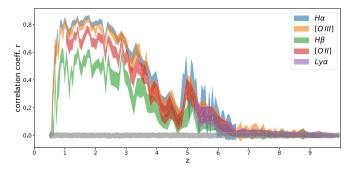

Figure 14. Pearson correlation coefficient r between the true and the reconstructed maps on 2500 light cones for SPHEREx-like mock data with a realistic level of SLED variation. The bands are the rms of the value in 100 noise realizations with the sample line signal. The gray bands are the correlation coefficient with the uncorrelated white noise map for reference.

With this mock external catalog, we identify the redshift bins containing the catalog sources, regardless of the noise level and threshold value. After projecting out these components, we run the MP algorithm on the residual data per usual until we hit the

stopping criteria. The results of a 4σ-threshold reconstruction with different noise levels are shown in Figure 17. The reconstruction shows improved results for all noise levels. However, the huge improvement in the highest-noise case cannot be interpreted as a successful reconstruction of source intensities. At this high noise level, the signals are buried under the noise, and when we fit the data with the catalog source redshift templates, the extracted components are dominated by noise rather than signal amplitude. Thus the improved correlation is merely due to the position information imposed by the external catalogs.

5.5. Comparing with the Limit of No Interlopers

In the case of no interloper lines (i.e., only single line emission in the data), the best estimator of the single-line intensity map is the observed map (regardless of foregrounds). To compare our reconstruction performance with this limiting case, we calculate the correlation coefficient between the input single-line map with the same single-line map plus the instrument noise.

Figure 18 shows the results using $\sigma_n = 10^4$ Jy sr^{-1}. Comparing to Figure 6, which has the same noise level, the single-line-plus-noise case has a lower correlation than the reconstruction in Figure 6. Especially for the fainter lines (e.g., CO(3–2) at $z \sim 0.6$ and CO(4–3) at $z \sim 1.1$), our reconstruction map has a much better correlation. This can be explained by the fact that in our algorithm, the sources are detected in the template space rather than in a single voxel. That is, if a source can be observed in two frequency channels, we extract the source by projecting the signals in these two channels to the template space, which is effectively combining the information from both channels. Consequently, we are able to achieve a better S/N on the fainter lines because of the greater sensitivity of their brighter counterpart.
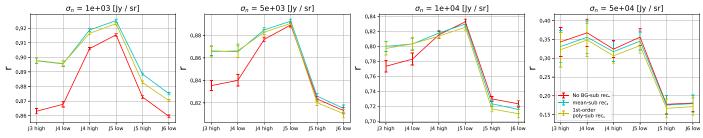
Figure 15. Comparing r of the reconstructed map with the no background subtraction (red), subtracting the mean of the whole data cube (cyan), and subtracting a first-order polynomial in the spectral direction for each light cone (yellow), using a 4σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the rms of 100 noise realizations of all of the spectral bins in each band.

5.6. Improving Cross-correlation Uncertainty

Another useful application of the reconstruction technique is for a more precise measurement of cross-correlation between LIM and other tracers. Cross-correlation analysis not only serves as a validation of a cosmological signal in LIM, since the cross-correlation is less susceptible to foreground contamination and other systematic effects, but also provides valuable astrophysical and cosmological information.

While cross-correlating the total observed LIM map with an external tracer gives an unbiased estimator of cross spectrum between the target line and the external tracer, the presence of continuum foregrounds and interlopers increases the error of this measurement. Here we present a simple argument that for cross-correlation analysis, using a reconstructed map instead of the total observed map can effectively reduce the error bars on the cross-power spectrum.

The cross-power spectrum errors between two fields δP_{12} in a single mode are given by

$$\delta P_{1,2}^2 = \frac{1}{2}(P_{1,2}^2 + \delta P_1 \delta P_2),\qquad(13)$$

where $P_{1,2}$ is the cross-power spectrum, and δP_1 and δP_2 are the errors on the auto power spectrum in the two fields. For a single k mode, $\delta P_1 = P_1$, $\delta P_2 = P_2$, where P_1 and P_2 are the total power spectra (including signals and noise) measured in two fields. Equation (13) can be expressed in terms of the cross-correlation coefficient: $r_{1,2} = P_{1,2}/\sqrt{P_1 P_2}$,

$$\delta P_{1,2}^2 = \frac{1}{2}P_{1,2}^2\left(1 + \frac{1}{r_{1,2}^2}\right).\qquad(14)$$

Say we have an external galaxy sample that traces one of the target CO lines, then we can write the total observed LIM data as the combination of target CO line (I_{CO}), other interloper lines ($I_{interlopers}$), and the noise (I_n),

$$I_{tot} = I_{CO} + I_{interlopers} + I_n.\qquad(15)$$

The expectation value of the cross spectrum between the observed total map and galaxy is the same as the cross spectrum with only the target CO, since the other components are not correlated with the large-scale structure at the same redshift, so $\langle P_{g,tot}\rangle = \langle P_{g,co}\rangle$.

For simplicity, we assume the galaxies are perfectly correlated with the target CO line field on the scale of interest, so $r_{g,co} = 1$. This implies the galaxy field and CO field always have the same correlation r with any given field x, $r_{g,x} = r_{co,x}$.

From Equation (14), the error on the galaxy-CO cross spectrum measured by cross-correlating the galaxy field with the total observed LIM data is

$$\delta P_{g,co}^{tot} = P_{g,tot}\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{g,tot}^2}\right)}$$

$$= P_{g,co}\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{co,tot}^2}\right)}.\qquad(16)$$

On the other hand, if we cross-correlate the galaxies with the reconstructed CO map (CO^r), the error is

$$\delta P_{g,co^r} = P_{g,co^r}\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{g,co^r}^2}\right)}.\qquad(17)$$

However, P_{g,co^r} is a biased estimator of $P_{g,co}$ because of the error in the reconstruction. If we assume that the reconstructed map roughly preserves the same power as the true map, $P_{co^r} \approx P_{co}$, then we can write[9] $P_{g,co^r} = r_{g,co^r}P_{g,co} = r_{co,co^r}P_{g,co}$. Therefore, we have to de-bias cross spectrum P_{g,co^r} by factor $1/r_{co,co^r}$, $P_{g,co} = P_{g,co^r}/r_{co,co^r}$. The value of r_{co,co^r} cannot be directly inferred from the data, so we have to estimate it by simulating the possible range of signals; this introduces an extra error term to the $\delta r_{co,co^r}$ due to the uncertainty in r_{co,co^r},

$$\delta P_{g,co}^r = P_{g,co}\sqrt{\left(\frac{\delta P_{g,co^r}}{P_{g,co^r}}\right)^2 + \left(\frac{\delta r_{co,co^r}}{r_{co,co^r}}\right)^2}$$

$$= P_{g,co}\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{co,co^r}^2}\right) + \left(\frac{\delta r_{co,co^r}}{r_{co,co^r}}\right)^2}.\qquad(18)$$

In summary, the S/N on the galaxy-CO cross spectrum $P_{g,co}$ using the total observed map and the reconstructed map is

$$S/N^{tot} = \frac{P_{g,co}}{\delta P_{g,co}^{tot}} = \frac{1}{\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{co,tot}^2}\right)}},$$

$$S/N^r = \frac{P_{g,co}}{\delta P_{g,co}^r} = \frac{1}{\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{co,co^r}^2}\right) + \left(\frac{\delta r_{co,co^r}}{r_{co,co^r}}\right)^2}}.\qquad(19)$$

---

[9] $P_{g,co^r} = r_{g,co^r}\sqrt{P_g P_{co^r}} \approx r_{g,co^r}\sqrt{P_g P_{co}} = r_{g,co^r}P_{g-co}$, where the last equality uses the assumption $r_{g,co} = 1$.
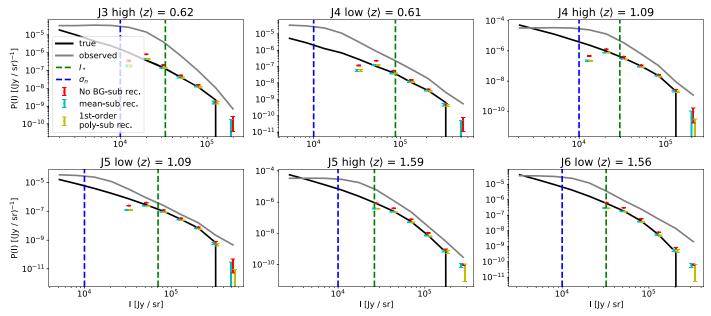
Figure 16. VID of no background subtraction (red), subtracting the mean of the whole data cube (cyan), subtracting a first-order polynomial in the spectral direction for each light cone (yellow), and the true (black) maps in the six broad bands with noise level $\sigma_n = 10^4$ Jy sr^{-1} and reconstruction threshold $m = 4$. The error bars are the rms of 100 noise realizations. The gray curves are the VID of the total observed map, which includes signals from all of the lines and noise. For reference, the blue and green dashed lines mark the noise level σ_n and I_*, respectively.

Note that the S/N in both cases converges to unity when r is unity and the de-bias error is zero, which is the limit of sample variance.[10]

Figure 19 shows the S/N^{tot} and S/N^r with two different values of $\delta r_{\text{co,co}^r}$ and $r_{\text{co,tot}}$ as a function of $r_{\text{co,co}^r}$. According to the calculations in Section 5.1, the value of $r_{\text{co,co}^r}$ ranges from ~0.7 to ~0.9 (except for the most noisy case, $\sigma_n = 5 \times 10^4$ Jy sr^{-1}). To estimate the realistic $r_{\text{co,tot}}$ value, we calculate the correlation of the input CO line maps and the observed map with $\sigma_n = 10^4$ Jy sr^{-1}, the same noise level as in Figure 6. The results are shown in Figure 20. We find that $r_{\text{co,tot}}$ are around 0.2–0.5 in this case. With this range of parameters, Figure 19 indicates that S/N^r is better than S/N$^{\text{tot}}$, which means that in our model, using the reconstructed map instead of the observed map in cross-correlation can reduce the uncertainty.

5.7. Estimating the Complete Interloper Population

Here we discuss a potential extension of the technique, which is capable of estimating the *total* power of an interloper population. The method uses an incompletely reconstructed interloper sample and an incomplete, external tracer of the interloper density field. Once the interloper contribution is fully quantified, the high-redshift signal of interest in an LIM data set can be estimated without bias. As an example, we write the observed intensity of a [C II] LIM data set as

$$I^{\text{obs}} = I_{\text{CO}} + I_{\text{C II}} + \delta_n^{\text{obs}}, \tag{20}$$

where I_{CO} is the total CO interloper intensity, $I_{\text{C II}}$ is the [C II] signal, and δ_n^{obs} is the instrumental noise. For simplicity, here

we only consider the contribution of one CO rotational line as the foreground.

Given a reconstruction threshold, we reconstruct the bright CO emissions using our technique as:

$$I_{\text{CO}}^{\text{rec}} = \alpha I_{\text{CO}_b} + \delta_n^{\text{rec}}, \tag{21}$$

where $I_{\text{CO}}^{\text{rec}}$ is the reconstructed CO intensity and I_{CO_b} is the intensity of bright CO sources in the reconstruction, which is a subset of the total CO population. There are two sources of error in the reconstruction: a multiplicative term α proportional to the bright CO intensity, where α can be greater or smaller than unity, and an additive term δ_n^{rec} describing random (or misidentified) fluctuations about the true CO intensity, which is uncorrelated with the CO field.

The auto power spectrum of the reconstructed CO map is

$$\langle I_{\text{CO}}^{\text{rec}} I_{\text{CO}}^{\text{rec}} \rangle = \bar{\alpha}^2 \langle I_{\text{CO}_b} I_{\text{CO}_b} \rangle + \langle \delta_n^{\text{rec}} \delta_n^{\text{rec}} \rangle$$
$$= \bar{\alpha}^2 \langle I_{\text{CO}_b} \rangle^2 b_{\text{CO}_b}^2 P(k) + N^{\text{rec}}, \tag{22}$$

where, in the linear regime that we consider here, b_{CO_b} is the cosmological clustering bias of the bright CO sources, $P(k)$ is the matter density field, and N^{rec} is the auto power spectrum of the δ_n^{rec} term as a noise bias.

Using an external galaxy sample g in the same redshift range as the CO interlopers, we cross-correlate g with the observed and reconstructed maps, respectively, and consider only linear clustering scales:

$$\langle g I_{\text{obs}} \rangle = r \langle I_{\text{CO}} \rangle b_{\text{CO}} b_g P(k), \tag{23}$$

$$\langle g I_{\text{CO}}^{\text{rec}} \rangle = \langle g \alpha I_{\text{CO}_b} \rangle = \bar{\alpha} r_b \langle I_{\text{CO}_b} \rangle b_{\text{CO}_b} b_g P(k), \tag{24}$$

where $\langle \alpha \rangle = \bar{\alpha}$, b_g is the bias of the galaxy tracer, b_{CO_b} is the bias of the bright CO population, and r (r_b) is the astrophysical stochastic cross-correlation parameter between the galaxy and CO (CO$_b$) populations.

---

[10] The power spectrum cross-correlation coefficient can be derived from the Pearson correlation (Equation (10)) with a weighting on pixels. Since our pixels are generated and reconstructed independently of each other, the Pearson correlation coefficient here is an unbiased estimator of the power spectrum correlation coefficient. Thus here we will use the value of the Pearson correlation we derived for the power spectrum correlation coefficient.
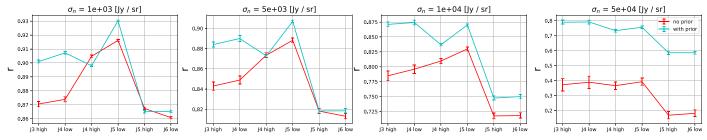
Figure 17. Comparing r of the reconstructed map with no external catalog prior (red), and utilizing the external catalog prior to fit the catalog sources before running the MP reconstruction (cyan), using a 4σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the rms of 100 noise realizations of all of the spectral bins in each band.
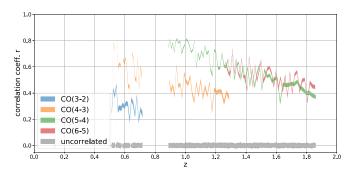

Figure 18. Pearson correlation coefficient r between the true line maps and the same maps adding the $\sigma_n = 10^{-4}$ Jy sr^{-1} noise on 2500 light cones. The bands are the rms of the value in 100 noise realizations with the sample line signal. The gray bands are the r value shown with the uncorrelated white noise map for reference.

Finally, we construct an estimator of the full CO power spectrum as

$$\widehat{\langle I_{CO} I_{CO} \rangle} = \langle I_{CO}^{rec} I_{CO}^{rec} \rangle \left(\frac{\langle gI_{CO}^{obs} \rangle}{\langle gI_{CO}^{rec} \rangle} \right)^2$$

$$= (\bar{\alpha}^2 \langle I_{CO_b} \rangle^2 b_{CO_b}^2 P(k) + N^{rec}) \left(\frac{r \langle I_{CO} \rangle b_{CO}}{\bar{\alpha} r_b \langle I_{CO_b} \rangle b_{CO_b}} \right)^2$$

$$= \left(\langle I_{CO_b} \rangle^2 b_{CO_b}^2 P(k) + \frac{N^{rec}}{\bar{\alpha}^2} \right) \left(\frac{r \langle I_{CO} \rangle b_{CO}}{r_b \langle I_{CO_b} \rangle b_{CO_b}} \right)^2. \quad (25)$$

If we can pick a reconstruction threshold such that the reconstruction of a subset of the CO population (bright CO sources) includes the majority of CO emitters, then $r \approx r_b$. Furthermore, if the reconstruction noise is negligible, $N^{rec}/\bar{\alpha}^2 \to 0$, then this estimator becomes an unbiased estimator of the full CO power spectrum on large scales:

$$\widehat{\langle I_{CO} I_{CO} \rangle} = \langle I_{CO}^{rec} I_{CO}^{rec} \rangle \left(\frac{\langle gI_{CO}^{obs} \rangle}{\langle gI_{CO}^{rec} \rangle} \right)^2 = \langle I_{CO} \rangle^2 b_{CO}^2 P(k). \quad (26)$$

In principle, this argument holds regardless of the luminosity limit of the CO population used for the reconstruction *and* regardless of the magnitude limit of the galaxy sample used for the cross-correlation estimate, as long as the galaxy and CO samples overlap spatially. This is potentially a powerful approach to access the entire interloper CO population without the need to identify the faint, undetected source contributions. Subtracting this CO estimate then provides an unbiased
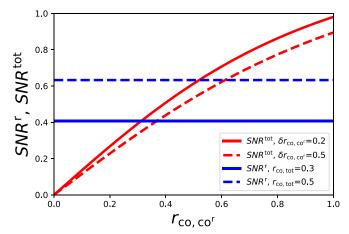

Figure 19. Comparing the S/N on the cross-power spectrum of one CO line and the external tracer, using the total observed map (S/N^{tot}) and the reconstructed CO map (S/N^r). We consider two different values of $\delta r_{co,co^r}$ and $r_{co,tot}$ that cover the range of realistic parameter values in our model. The value of $r_{co,co^r} \sim 0.7 - 0.9$ according to the results in Section 4, and therefore using the reconstructed map instead of the total observed map in cross-correlation can reduce the uncertainty (i.e., S/N^r > S/N^{tot}).
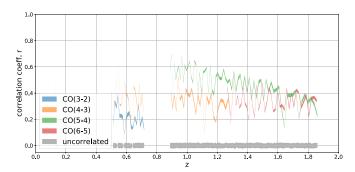

Figure 20. Pearson correlation coefficient r between the true input and the total observed maps on 2500 light cones with $\sigma_n = 10^4$ Jy sr^{-1}. The bands are the rms of the value in 100 noise realizations with the sample line signal. The gray bands are the r value shown with the uncorrelated white noise map for reference.

estimate of the high-redshift [C II] power spectrum in the same LIM data set, which is highly desirable.

With an external galaxy catalog along, one can also estimate the total CO power on large scales where $r \approx 1$, with the estimator: $\langle gI_{CO}^{obs} \rangle^2 / \langle gg \rangle$. However, the advantage of the Equation (25) estimator is that it only requires $r \approx r_b$, which can be valid on smaller scales where $r < 1$, as long as the reconstruction has sufficient quality in terms of the noise level purity and completeness. Therefore, with the reconstructed

map, one can potentially extract the total CO power to smaller scales.

We note that a high reconstruction threshold m is required in order to achieve a (nearly) noiseless reconstruction ($N^{\rm rec}/\bar{\alpha}^2 \to 0$). On the other hand, $r \approx r_b$ becomes invalid if the threshold is too high such that majority of the fainter sources is missed in the reconstruction. In addition, a high threshold tends to boost the shot noise in the reconstructed power spectrum, as well as uncertainties in the CO power spectrum estimator (Equation (25)). There is clearly a trade-off between the fidelity of the reconstructed signals and the uncertainty to estimate the desired signals. A detailed simulation is necessary to determine the optimal threshold to minimize the effect of bias from $N_{\rm rec}$ and the variance from the reconstruction shot noise, and evaluate the performance of the CO power spectrum estimator. We leave this investigation to future work.

5.8. Comparing Foreground Cleaning Capability with Masking

Our algorithm can serve as a foreground mitigation method for [C II] LIM measurement by identifying the bright CO foreground signals. In this section, we quantify this "foreground cleaning" performance using our algorithm and compare it with the masking method, where a "cleaned map" is obtained by masking out voxels that contain bright CO sources identified with an external source catalog.

In the following, we compare two cases: (1) in the masking case, the "cleaned map" is the observed map (including all of the lines and instrument noise) masked using external CO catalogs, and (2) the "cleaned map" derived from our algorithm is the observed map subtracting the CO reconstructed map.

The external catalog considered here is the same as the one described in Section 5.4: a flux cut on CO(5–4) at the level of $150\,L_\odot\,{\rm Mpc}^{-2}$ ($6 \times 10^{-17}\,{\rm W\,m}^{-2}$), which gives ∼1.2 galaxies per light cone (0.43^2 arcmin^2 pixel solid angle) and corresponds to a $m_{\rm AB} \sim 21.8$ threshold in the optical band. Note that this masking threshold is comparable to the "case A" masking in Sun et al. (2018), although here we consider a simpler model that ignores the scatter in the line luminosity model. For this masking scenario, the "cleaned map" is the observed map minus any voxels that contain the sources in the external catalog.

In our algorithm, the CO sources and their spectra are reconstructed and removed from the data iteratively, and the residual can be regarded as a "cleaned map" that is free of bright CO sources. However, we are only capable of cleaning the multiline redshift bins in our algorithm. The signal identified in the single-line redshift bins in our algorithm is the combination of the remaining CO and [C II] signals. If we remove all of the reconstructed single-line signals, we will over-subtract [C II] in the cleaned map. Therefore, for the single-line redshift bins ($0.15 \leqslant z \leqslant 0.51$; $0.72 \leqslant z \leqslant 0.89$; $z \geqslant 1.87$), we clean the data by masking voxels that contain external catalog sources.

As a figure of merit, we calculate the S/N of the respective [C II] shot-noise power spectra (since there is no clustering signal in our mock light cones) while including the residual CO as part of the noise contribution.

The error on the shot-noise power spectrum $\delta P_{\rm sh}$ in a map is

$$\delta P_{\rm sh} = \frac{P_{\rm tot}}{\sqrt{N_{\rm mode}}} = \frac{V_{\rm vox}\sigma_{\rm tot}^2}{\sqrt{N_{\rm mode}}}, \quad (27)$$

where $P_{\rm tot}$ is the total power spectrum of the cleaned map on the shot-noise scales, which is proportional to the total voxel variance in the map, $\sigma_{\rm tot}^2$. $N_{\rm mode}$ is the number of k-space modes used to measure the shot noise. $N_{\rm mode}$ is usually of the order of the total number of voxels, so we choose $N_{\rm mode} = 6000$, similar to the number of voxels in TIME. The shot-noise power of the [C II] signal is given by

$$P_{\rm C\,II} = V_{\rm vox}\sigma_{\rm C\,II}^2, \quad (28)$$

where $\sigma_{\rm C\,II}^2$ is the voxel variance of the [C II] signal map. The S/N of the [C II] shot-noise power spectrum is then:

$$\frac{P_{\rm C\,II}}{\delta P_{\rm sh}} = \frac{\sigma_{\rm C\,II}^2}{\sigma_{\rm tot}^2}\sqrt{N_{\rm mode}}. \quad (29)$$

We test the foreground cleaning performance with the same set of 2500 mock light cones described in Section 4, and calculate the [C II] shot-noise S/N using Equation (29). $\sigma_{\rm C\,II}^2$ is the variance of the input [C II] map, and $\sigma_{\rm tot}^2$ is the variance of the cleaned map. For simplicity, $\sigma_{\rm C\,II}^2$ and $\sigma_{\rm tot}^2$ are the variance from all of the frequency channels.

In Figure 21, the black dotted lines and the red lines show the [C II] shot-noise S/N with the cleaned map obtained from masking and from our algorithm, respectively. For reference, the black dashed lines are the [C II] shot-noise S/N of the observed map before cleaning, and thus the map includes the contribution from all of the lines and the instrument noise.

According to Figure 7, the optimal threshold that gives the maximum r value is $m \sim 4$. In Figure 21, $m = 4$ is marked with an orange dashed line, and we see that in the highest-noise case for this threshold, masking (black dotted line) using an external (deep) catalog performs slightly better than reconstruction (red line) because it is difficult for the MP algorithm to extract the signals from noisy data directly. For realistic noise levels (between $\sigma_n = 5 \times 10^3$ and 10^4 Jy sr^{-1}), the reconstruction outperforms masking.

We also compare this result with the limiting case where there are only [C II] and instrumental noise in the data (green line), i.e., $\sigma_{\rm tot}^2 = \sigma_{\rm C\,II}^2 + \sigma_n^2$. For a small reconstruction threshold m, the reconstructed S/N is better than this limit, which indicates that the reconstruction over-fits and misidentifies noise fluctuation as a signal and removes them from the cleaned data. We see that for the two realistic noise levels at $m \sim 4$, the reconstructed S/N is lower than this limit, and thus indicates that overfitting is not an issue at this threshold. Also note that in the highest-noise case, the masking S/N is close to the noise-plus-[C II] limit (green line), for the following reason: since the external catalog goes much deeper than the noise level, noise fluctuation dominates over the line signals after masking, and thus variance in the masked map is close to the noise variance.

Finally, we point out that the reconstructed [C II] shot-noise S/N (red line) converges to a constant instead of increasing with smaller threshold m values. This is because of the fact that in the low threshold limit, the reconstruction residual in the cleaned map is subdominant compared to the (masked) single-line redshift bin signals being added back to the residual, and thus the cleaned map S/N does not depend on the threshold value.

To sum up, for a realistic noise level, our reconstruction performs better than masking in terms of foreground cleaning capability, given our signal model and the external catalog considered in this work. We note that this conclusion depends
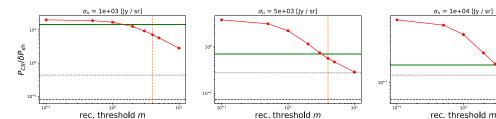
Figure 21. S/N on [C II] shot-noise power spectrum after CO line foreground removal with masking (black dotted) and sparse reconstruction (red) with different instrument noise level σ_n and reconstruction threshold m. The orange dashed line marks the $m = 4$ threshold for reference. We also show the S/N before cleaning (black dashed), and the limiting case of the no CO signals (green) for comparison.

on the line luminosity function model and the depth of the external catalog for masking, and we leave a more detailed analysis to future work.

5.9. Technique Extensions

In this section, we outline some directions for extending the current framework to further improve upon the line reconstruction in future work.

5.9.1. Template Generalization

Currently we use a single SLED template for all redshift bins in the reconstruction; this can be easily generalized to incorporate multiple spectral templates to account for the redshift evolution and SLED variation of the signals. In addition, the extension can help differentiate the emission from different types of galaxies that have different SLEDs. For instance, if we have two different SLED models for early- and late-type galaxies, respectively, we can incorporate them by having two columns in A for every redshift bin such that the reconstruction can infer not only the redshift and luminosity of the sources, but also their galaxy type from the SLED templates.

5.9.2. Alternative Sparse Approximation Algorithm

The MP algorithm adopted in this work optimizes the ℓ_0 norm in Equation (8), which is the direct sum of the nonzero elements in $\widetilde{N}$. We can improve the algorithm by including prior information on the expected value of each element in $\widetilde{N}$, which depends on the voxel size and source luminosity function. For instance, instead of using the MP algorithm, one can obtain the sparse solution by solving the following ℓ_1-norm regularization equation:

$$\underset{\widetilde{N}}{\mathrm{argmin}} \ \|I - A\widetilde{N}\|_2^2 + \lambda \|w \cdot \widetilde{N}\|_1, \qquad (30)$$

where the parameter λ determines the regularization strength for preventing overfitting, which has a similar effect as the stopping criteria in the MP algorithm; the prior information on the number density of the sources in each redshift bin can be encoded in the weight vector w in this expression.

5.9.3. Clustering Information

In this work, we only perform the pixel-by-pixel line deconfusion using the information in the spectral correlation due to the multiple lines emitted from the same source to reconstruct the signals. The clustering information of the galaxies, which is neglected in this work, could provide additional information on the emission field. We can generalize this

framework by incorporating the clustering information from the known galaxy two-point correlation, and perform the reconstruction on an ensemble of pixels to simultaneously fit for the spectral correlation and clustering. For example, if we have a theoretical model for the line-of-sight two-point correlation function of $N(z)$, $\xi_{\mathrm{th}}(N(z))$, we can add another ℓ_2-norm regularization term to Equation (30),

$$\lambda_{\mathrm{clus}} \|\xi(\widetilde{N}(z)) - \xi_{\mathrm{th}}(N(z))\|_2^2. \qquad (31)$$

This will enforce the algorithm to give higher priority to solutions close to the theoretical correlation function. Similarly, with an external catalog that traces the same large-scale structure, one can also constrain the algorithm with cross-correlation:

$$\lambda_{\mathrm{clus}}^x \|\xi^x(\widetilde{N}(z), \delta_{\mathrm{ext}}) - \xi_{\mathrm{th}}^x(N(z), \delta_{\mathrm{ext}})\|_2^2, \qquad (32)$$

where δ_{ext} is the density field of the external tracer, e.g., galaxy samples, and ξ^x is the cross-correlation between $N(z)$ and δ_{ext}. We leave further investigation that makes use of clustering information to future work.

6. Conclusion

We develop a spectral line deconfusion technique for LIM experiments, where different spectral lines emitted by sources at different redshifts can be observed in the same frequency channel and then confused. Unlike most of the previously proposed methods that decompose the line signals in the power spectrum space, we perform a phase-space deconfusion that reconstructs the individual line intensity maps, if multiple spectral lines of a redshifted source population are observable. The reconstructed line intensity maps are direct data products of an LIM experiment, and can be used to trace the underlying density field for various science applications.

Our method is based on the information that multiple spectral lines emitted by redshifted sources are mapped onto distinct observed frequencies, which give deterministic features in the observed spectrum that can be fitted by a template. With a set of spectral template models and assuming the sparse approximation, we fit the LIM data iteratively with the MP algorithm.

As an example, we consider an LIM survey with similar survey parameters as the ongoing EoR [C II] experiments, TIME, and CONCERTO. The intervening CO line intensity maps at $0.5 \lesssim z \lesssim 1.5$ can be extracted with our technique, since multiple CO rotational transitions are observable. We demonstrate that with the assumed signal model and realistic noise level, our reconstructed CO maps reach $\sim 80\%$ spatial correlation with the true maps. In addition, in our assumed signal model and realistic survey setup, the VID of individual

16

lines can be correctly extracted with a high S/N down to the ℓ_*-scale. The CO luminosity function derived from the VID measurement can provide information on galaxy formation and evolution as traced by the CO distribution across cosmic time. The reconstruction performance is robust against a realistic level of line ratio uncertainties and continuum foreground mitigation process.

In addition to probing the large-scale luminosity and density fields, the reconstructed line intensity maps can also be used for a variety of applications. As a demonstration, we show that using the reconstructed map instead of the original LIM data set can effectively reduce uncertainties in cross-correlation measurements, and improve the performance of interloper masking to reveal the high-redshift line emissions. Furthermore, given that the reconstructed intensity map, even if incomplete, traces the matter density on large scales, we construct an estimator capable of estimating the total interloper power. The estimator invokes the cross-correlation of the reconstructed map with an external density tracer such as galaxies in the linear clustering regime. This approach has the potential to fully specify the interloper and high-redshift source populations and warrants future investigation.

While we mainly discuss the application for an EoR [C II] LIM experiment in this paper, this technique is not restricted to this setup. We demonstrate that our technique can successfully extract redshifted optical line signals from a SPHEREx-like experiment in the near-infrared. The technique is a general framework that can be readily applied to mitigate line deconfusion problems in LIM experiments and enhance the science returns.

Appendix A
The Matching Pursuit Algorithm

In this section, we describe the detailed steps in the matching pursuit (MP) algorithm. The MP algorithm iteratively selects an atom in the dictionary to project out part of the signals in the data, and keep track of the current solution of the signal f and residual R for the next step until the solution meets the stopping criteria. In Equation (8), for a given signal I and matrix A, we define a set of vectors $\{\psi_i\}$ to be the column vectors in A (i.e., the atoms in the dictionary). The MP algorithm works as follows:

1. Initialize at step $t = 0$: $f_0 = 0$, $R_0 = I$, $\widetilde{N}_0 = 0$.
2. Compute the inner product of R_0 and ψ_i's:

$$\{u_{0i}\} = \{\langle R_0, \psi_i \rangle\}. \tag{A1}$$

3. Select the element γ to be updated by

$$\gamma = \underset{i}{\mathrm{argmax}}\ \{u_{0i}\}. \tag{A2}$$

4. If $u_{0\gamma}$ meets the stopping criteria, end the process and return $N = 0$. Else, proceed to step 5.
5. Update the current f_t, R_t, and record the amplitude of the new solution in γ-th element of vector $\widetilde{N}_t$:

$$f_{t+1} = f_t + u_{t\gamma}\ \psi_\gamma, \tag{A3}$$

$$R_{t+1} = R_t - u_{t\gamma}\ \psi_\gamma, \tag{A4}$$

$$\widetilde{N}_{t+1}(\gamma) = \widetilde{N}_t(\gamma) + u_{t\gamma}. \tag{A5}$$

6. Compute the inner product of R_{t+1} and ψ_i's:

$$\{u_{(t+1)i}\} = \{\langle R_{t+1}, \psi_i \rangle\}. \tag{A6}$$

7. Select the element γ to be updated by

$$\gamma = \underset{i}{\mathrm{argmax}}\ \{u_{(t+1)i}\}. \tag{A7}$$

8. If $u_{(t+1)\gamma}$ meets the stopping criteria, go back to step 5 for the next iteration. Else, proceed to step 9.
9. Return the final solution $N = \widetilde{N}_{t+1}\ /\ I^{\mathrm{norm}}$.

As described in Section 3.3, the stopping criteria is set by comparing $u_{t\gamma}$ with the noise σ_n. This follows the fact that var $(u_{t\gamma}) = \sigma_n^2$ (see Appendix B for the proof), so if we set $u_{(t+1)\gamma} < m\ \sigma_n$ in step 8, this is effectively setting an "m-σ" detection threshold (e.g., $m = 5$ for a 5σ detection).

Figure A1 illustrates the steps of the MP algorithm solving N of an example light cone. In this example, we set $\sigma_n = 10$ kJy, and the detection threshold $m = 5$. In this light cone, there are six ℓ_* sources in the multiline redshift bins at $z = [0.54, 1.06, 1.20, 1.24, 1.79, 1.82]$, shown in the top left panel. Since N is the effective number of ℓ_* sources per redshift bin, the amplitudes in the six corresponding redshift bins are equal to unity. The top middle panel shows the total line signal in this light cone I_{true}, including the emission from those six sources as well as that from other sources in the single-line redshifts. The top right panel shows the observed data after adding noise to the signal I_{true}, which is also the R_0 vector in the first step of the MP algorithm. The blue dashed lines mark the noise level $\pm\sigma_n = 10$ kJy for reference.

In the first iteration, the MP algorithm selects the 14th z-bin index ($z = 0.54$) with amplitude $u_{1\gamma} \sim 1$, so the $z = 0.54$ source is successfully extracted in this iteration. The gray dashed spectrum in the left panel of the second row is the template signal extracted in this step ($u_{1\gamma}\psi_\gamma$). The updated values of N_1, f_1, and R_1 from step 5 are shown in the second row of Figure A1.

Then we proceed to the second iteration. The MP algorithm selects the 37th column ($z = 0.65$). The gray dashed spectrum in the left panel of the third row is the template signal fitted in this step ($u_{2\gamma}\psi_\gamma$). However, there is no $z = 0.65$ source in the input, which means the MP algorithm misidentifies the emission from noise or other lines as the signal. The third row of Figure A1 shows the updated values of N_2, f_2, and R_2 from the second iteration.

In the third iteration, the 109th ($z = 1.24$) column in the dictionary is selected. The fourth row of Figure A1 shows the updated values of N_3, f_3, and R_3 from this iteration. Note that
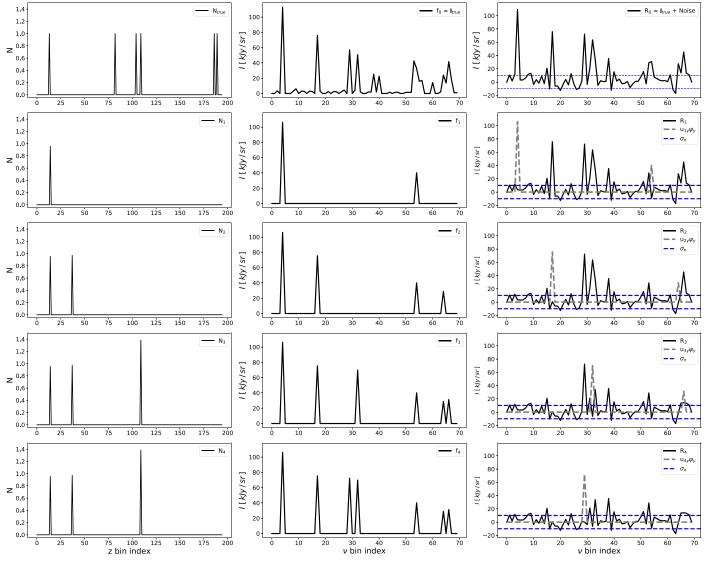
Figure A1. Illustration of how the MP algorithm solves for the source vector N. See the text for a detailed description.

this time the MP algorithm picks up a correct redshift, while it overestimates the amplitude by ~40 %.

In the fourth iteration, the algorithm selects the 224th column from the dictionary. The 224th column is not in the multiline redshift bins (first 195 columns), and thus the N_4 in the bottom row of Figure A1 remains unchanged, whereas f_4 and R_4 are updated with a single-peak signal.

In the fifth iteration, the stopping criteria in step 8 is met ($u_{5\gamma} < 5\sigma_n$), so the reconstruction terminates and returns the last row of Figure A1 as the reconstruction results for this light cone.

In summary, in this example, two of the six multiline redshift sources have been reconstructed, in addition to one misidentified source. Comparing the final reconstructed light cone signal (bottom middle panel of Figure A1) to the true input light cone (top middle panel of Figure A1), we can see that the MP reconstruction captures the strong peaks in the data, and the remaining signals are close to the noise level.

Appendix B
Proving $\mathrm{var}(u_{t\gamma}) = \sigma_n^2$

In Section A, the residual of step t R_t can be expressed in the linear combination of the dictionary atoms and noise:

$$R_t = \sum_i c_i \, \psi_i + n, \qquad (B1)$$

where c_i represents the constant coefficient. Then we derive

$$
\begin{aligned}
u_{t\gamma} &\equiv \langle R_t, \, \psi_\gamma \rangle \\
&= \sum_i c_i \langle \psi_i, \, \psi_\gamma \rangle + \langle n, \, \psi_\gamma \rangle \\
&= \mathrm{const} + \sum_j n_j \, \psi_{\gamma j}.
\end{aligned}
\qquad (B2)
$$

The first term is not depend on the noise, so it is a constant term that is not contributing to the variance. Also note that $\langle \psi_i, \, \psi_\gamma \rangle \neq \delta_{i\gamma}$ since the dictionary $\{\psi_i\}$ are normalized but

Cheng, Chang, & Bock

not orthogonal. With this expression, we can calculate the variance:

$$\langle u_{t\gamma} \rangle = \text{const} + \sum_j \langle n_j \rangle \, \psi_{\gamma j} = \text{const}$$

$$\langle u_{t\gamma}^2 \rangle = \text{const}^2 + \sum_j \langle n_j^2 \rangle \, \psi_{\gamma j}^2$$

$$= \text{const}^2 + \sigma_n^2 \sum_j \psi_{\gamma j}^2$$

$$= \text{const}^2 + \sigma_n^2. \tag{B3}$$

Therefore, we get

$$\text{var}(u_{t\gamma}) = \langle u_{t\gamma}^2 \rangle - \langle u_{t\gamma} \rangle^2 = \sigma_n^2. \tag{B4}$$

Appendix C
SPHEREx Line Signal Model

In this section, we describe the line signal model in SPHEREx wavelengths. We model five lines from $z = 0$ to 10 in the SPHEREx band: Lyα (121.6 nm), Hα (656.3 nm), Hβ (486.1 nm), [O II] (372.7 nm), and [O III] (500.7 nm).

Since the optical lines are associated with star formation activities, we model the signal with the following steps: we start with the halo mass function, and use the star formation rate (SFR)–halo mass (M) relation, and the SFR–line luminosity relation to paint the spectral line signals to each halo.

We use the publicly available halo mass function calculator *HMFcalc* (Murray et al. 2013)[11] to obtain the halo mass function based on the Sheth et al. (2001) model. For the SFR–M relation, we use the model from Behroozi et al. (2013), in which the SFR–M relation is derived based on several observational constraints.[12]

For the SFR–line luminosity relation, we assume a linear relation for all of the lines. For Lyα, we use the prescription provided by Fonseca et al. (2017) with their fiducial values: $\gamma_{\text{Ly}\alpha} = 1, f_{\text{esc}}^{\text{UV}} = 0.2, f_{\text{esc}}^{\text{Ly}\alpha} = 0.2, E_{\text{UV}} = 1.0$ in their Equations (8) and (15), and derive the conversion factor:

$$\frac{\text{SFR}}{M_\odot \, \text{yr}^{-1}} = 2.29 \times 10^{-41} \frac{L_{\text{Ly}\alpha}}{\text{erg s}^{-1}}. \tag{C1}$$

For other spectral lines, we adopt the relation from Kennicutt (1998) and Ly et al. (2007):

$$\frac{\text{SFR}}{M_\odot \, \text{yr}^{-1}} = (7.9 \pm 2.4) \times 10^{-42} \frac{L_{\text{H}\alpha}}{\text{erg s}^{-1}}, \tag{C2}$$

$$\frac{\text{SFR}}{M_\odot \, \text{yr}^{-1}} = (1.4 \pm 0.4) \times 10^{-41} \frac{L_{\text{[O II]}}}{\text{erg s}^{-1}}, \tag{C3}$$

$$\frac{SFR}{M_\odot \, \text{yr}^{-1}} = (7.6 \pm 3.7) \times 10^{-42} \frac{L_{\text{[O III]}}}{\text{erg s}^{-1}}, \tag{C4}$$

and for the Hβ line, we use the fixed line ratio Hβ/Hα = 0.35 (Osterbrock & Ferland 2006).

---

[11] http://hmf.icrar.org/
[12] The SFR(M, z) is downloaded from author's webpage (https://www.peterbehroozi.com/data.html). The model is only available at $0 < z < 8$, so we use $z = 8$ model for $z > 8$.

ORCID iDs

Yun-Ting Cheng ⓘ https://orcid.org/0000-0002-5437-0504
Tzu-Ching Chang ⓘ https://orcid.org/0000-0001-5929-4187

References

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57
Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2009, ApJ, 695, 183
Breysse, P. C., Kovetz, E. D., Behroozi, P. S., Dai, L., & Kamionkowski, M. 2017, MNRAS, 467, 2996
Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2014, MNRAS, 443, 3506
Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2015, MNRAS, 452, 3408
Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2016, MNRAS, 457, L127
Breysse, P. C., & Rahman, M. 2017, MNRAS, 468, 741
Candes, E. J., Romberg, J., & Tao, T. 2006, ITIT, 52, 489
Carilli, C. L. 2011, ApJL, 730, L30
Chang, T. C., Gong, Y., Santos, M., et al. 2015, in Proc. of Advancing Astrophysics with the Square Kilometre Array (AASKA14) (Giardini Naxos, Italy, 2014 June 8-13) 4, doi:10.22323/1.215.0004
Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Natur, 466, 463
Chang, T.-C., Pen, U.-L., Peterson, J. B., & McDonald, P. 2008, PhRvL, 100, 091303
Chapman, E., Abdalla, F. B., Harker, G., et al. 2012, MNRAS, 423, 2518
Cheng, Y.-T., Chang, T.-C., Bock, J., Bradford, C. M., & Cooray, A. 2016, ApJ, 832, 165
Cheng, Y.-T., de Putter, R., Chang, T.-C., & Doré, O. 2019, ApJ, 877, 86
Chung, D. T., Viero, M. P., Church, S. E., et al. 2019, ApJ, 872, 186
Comaschi, P., & Ferrara, A. 2016, MNRAS, 455, 725
Crites, A. T., Bock, J. J., Bradford, C. M., et al. 2014, Proc. SPIE, 9153, 91531W
Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., et al. 2016, MNRAS, 457, 3541
Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., Blomqvist, M., & Pieri, M. 2018, MNRAS, 481, 1320
Daddi, E., Dannerbauer, H., Liu, D., et al. 2015, A&A, 577, A46
de Putter, R., Holder, G. P., Chang, T.-C., & Dore, O. 2014, arXiv:1403.3727
Decarli, R., Walter, F., Gónzalez-López, J., et al. 2019, ApJ, 882, 138
Donoho, D. L. 2006, ITIT, 52, 1289
Doré, O., Bock, J., Ashby, M., et al. 2014, arXiv:1412.4872
Fonseca, J., Silva, M. B., Santos, M. G., & Cooray, A. 2017, MNRAS, 464, 1948
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhR, 433, 181
Gong, Y., Chen, X., & Cooray, A. 2020, ApJ, 894, 152
Gong, Y., Cooray, A., Silva, M., et al. 2012, ApJ, 745, 49
Gong, Y., Cooray, A., Silva, M. B., Santos, M. G., & Lubin, P. 2011, ApJL, 728, L46
Gong, Y., Silva, M., Cooray, A., & Santos, M. G. 2014, ApJ, 785, 72
Helgason, K., Ricotti, M., & Kashlinsky, A. 2012, ApJ, 752, 113
Ihle, H. T., Chung, D., Stein, G., et al. 2019, ApJ, 871, 75
Keating, G. K., Bower, G. C., Marrone, D. P., et al. 2015, ApJ, 814, 140
Keating, G. K., Marrone, D. P., Bower, G. C., et al. 2016, ApJ, 830, 34
Kennicutt, R. C. J. 1998, ARA&A, 36, 189
Kogut, A., Dwek, E., & Moseley, S. H. 2015, ApJ, 806, 234
Lagache, G., Cousin, M., & Chatzikos, M. 2018, A&A, 609, A130
Li, T. Y., Wechsler, R. H., Devaraj, K., & Church, S. E. 2016, ApJ, 817, 169
Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, ApJ, 741, 70
Lidz, A., & Taylor, J. 2016, ApJ, 825, 143
Lidz, A., Zahn, O., Furlanetto, S. R., et al. 2009, ApJ, 690, 252
Liu, A., & Tegmark, M. 2012, MNRAS, 419, 3491
Ly, C., Malkan, M. A., Kashikawa, N., et al. 2007, ApJ, 657, 738
Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429
Mallat, S. G., & Zhang, Z. 1993, ITSP, 41, 3397
Mashian, N., Sternberg, A., & Loeb, A. 2015, JCAP, 11, 028
Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJL, 763, L20
Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767
Moriwaki, K., Filippova, N., Shirasaki, M., & Yoshida, N. 2020, MNRAS Letters, 496, L54
Moustakas, J., Kennicutt, R. C. J., & Tremonti, C. A. 2006, ApJ, 642, 775
Murray, S. G., Power, C., & Robotham, A. S. G. 2013, A&C, 3, 23
Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei (Mill Valley, CA: University Science Books)
Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012, ApJ, 756, 165
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Popping, G., van Kampen, E., Decarli, R., et al. 2016, MNRAS, 461, 93
Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15
Pullen, A. R., Doré, O., & Bock, J. 2014, ApJ, 786, 111

19

Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, ApJ, 872, 7
Righi, M., Hernández-Monteagudo, C., & Sunyaev, R. A. 2008, A&A, 489, 489
Scott, D., & Rees, M. J. 1990, MNRAS, 247, 510
Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1
Silva, M., Santos, M. G., Cooray, A., & Gong, Y. 2015, ApJ, 806, 209
Silva, M. B., Santos, M. G., Gong, Y., Cooray, A., & Bock, J. 2013, ApJ, 763, 132
Stompor, R., Balbi, A., Borrill, J. D., et al. 2002, PhRvD, 65, 022003

Sun, G., Moncelsi, L., Viero, M. P., et al. 2018, ApJ, 856, 107
Switzer, E. R., Chang, T.-C., Masui, K. W., Pen, U.-L., & Voytek, T. C. 2015, ApJ, 815, 51
Uzgil, B. D., Aguirre, J. E., Bradford, C. M., & Lidz, A. 2014, ApJ, 793, 116
Visbal, E., & Loeb, A. 2010, JCAP, 11, 016
Walter, F., Decarli, R., Sargent, M., et al. 2014, ApJ, 782, 79
Wyithe, J. S. B., & Loeb, A. 2008, MNRAS, 383, 606
Yue, B., Ferrara, A., Pallottini, A., Gallerani, S., & Vallini, L. 2015, MNRAS, 450, 3829