Draft version May 13, 2020
Typeset using LaTeX **twocolumn** style in AASTeX63

arXiv:2005.05341v1 [astro-ph.CO] 11 May 2020

Phase-Space Spectral Line De-confusion 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*
[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 Rd, 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 de-confusion 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 sigma 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 TIME/CONCERTO-like survey targeting [C II] at the Epoch of Reionization, and reliably reconstruct the 3D spatial distribution of the CO interlopers at $0.5 \lesssim z \lesssim 1.5$. We also demonstrate a successful de-confusion for the SPHEREx mission in the near-infrared wavelengths. Potentially, the reconstructed maps can be further cross-correlated with a (galaxy) tracer population to estimate the total interloper power in the linear clustering regime. 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 de-confusion problem. [a)]

Keywords: cosmology: theory – observations – large-scale structure of universe – diffuse radiation

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; Wyithe & Loeb 2008; Chang et al. 2008), the CO rotational lines (Righi et al. 2008; Visbal & Loeb 2010; Carilli 2011; Lidz et al. 2011; Gong et al. 2011; Breysse et al. 2014; Pullen et al. 2013; Mashian et al. 2015; Keating et al. 2015; Breysse et al. 2016; Li et al. 2016; Keating et al. 2016; Fonseca et al. 2017; Breysse & Rahman 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 Lyman-α emission line (Silva et al. 2013; Gong et al. 2014; Pullen et al. 2014; Comaschi & Ferrara 2016; Croft et al. 2016; Fonseca et al. 2017; Croft et al. 2018) are amongst the most studied lines in the LIM regime.

One of the main challenges in LIM are 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 sever for 21cm than for other lines), their smooth spectral feature can

ycheng3@caltech.edu

a) © 2020. All rights reserved.

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; Liu & Tegmark 2012; Parsons et al. 2012; Chapman et al. 2012; Switzer et al. 2015). The line interlopers, originate from sources residing in different redshifts emitting spectral lines in the same observed frequency channel, is another pressing issue for LIM experiments. The two most studied line de-confusion 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; Yue et al. 2015; Silva 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; Visbal & Loeb 2010; Gong et al. 2012, 2014; Chang et al. 2015; Silva et al. 2015; Chang et al. 2010; Masui et al. 2013; Croft et al. 2016; Chung et al. 2019). In addition to these two methods, Lidz & Taylor (2016); 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 de-confusion 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 parameters 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 statistic 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, multi-tracer analysis, de-lensing 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 a LIM data set with blended interlopers. Using the fact that when multiple spectral lines emitted by a source are observable in a 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 a 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 de-confusion using the multi-line 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 the spectral resolution. In this regime, our template-fitting based technique can gain the signal-to-noise ratio (SNR) 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 de-confusion with deep learning. They show that, in the absence of noise or foreground components, their algorithm can reconstruct the individual line map that are mixed in the LIM data set. Their training data generation relies on the assumption of the signal clustering and the line luminosity model, whereas in this work, we develop the line de-confusion technique only makes use of the spectral feature of the lines, that 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 Sec. 2. Then, Sec. 3 describes our line de-blending technique. Sec. 4 presents the results on the fiducial setup. In Sec. 5, we present the performance of the technique with more practical considerations, and discuss its applications and extensions. The conclusion are given in Sec. 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 a 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 de-confusion method introduced in this work is performed on the pixel-by-pixel basis, which only utilizes the spectral information in individual line-of-sight (light cone), without taking into account the spatial clustering information which we leave for future work.

We test our de-confusion technique on simulated light cones. Since the clustering information is not relevant to our technique, we generate light cones that only based on the spectral line luminosity function models but not the clustering properties. Thus, each light cone is independent of each other, and we also ignore the line-of-sight clustering in this work. This allows speeding up both light cone construction and de-confusion by parallelization without affecting the quantification of performance.

As a demonstration of the technique, we assume a 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), CO(6-5)}. In reality, in the sub-mm spectral range of interest (generally in the $\sim$ 200-300 GHz range), there are higher CO J lines (that are fainter), the Galactic and extragalactic dust continuum emissions (the Cosmic Infrared Background), the Cosmic Microwave Background 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 Sec. 5.3, we will demonstrate that the continuum foreground mitigation in the data analysis process has negligible impact on our technique performance.

We use the [C II] and CO luminosity function models provided by Popping et al. (2016), where they use a semi-analytic model including the effect of radiative transfer 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 on-going EoR [C II] line intensity mapping experiments, TIME (Crites et al. 2014) and CONCERTO (Lagache et al. 2018).

The mock survey covers 200 - 305 GHz with 70 evenly-spacing spectral channel ($\delta\nu$ = 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 $\sigma_n = \{10^3,\ 5\times10^3,\ 10^4,\ 5\times10^4\}$ Jy/sr. These values are comparable to the range of expected instrument noise in TIME and CONCERTO.

2.3. *Light Cones 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 [1] for all 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, and 2000 redshift bins of $\Delta z = 5\times10^{-4}$ 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

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

where the last square bracket is the comoving volume of the voxel defined by spatial pixel $\Omega_{\rm pix}$ and redshift bin Δz, D_A is the comoving angular diameter distance, and χ is the comoving 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 $\mathcal{L}_{\rm tot}(z)$. We define $N(z) \equiv \mathcal{L}_{\rm tot}/\mathcal{L}_*$ as the "effective number counts" per redshift bin. Fig. 1 left panel shows the $N(z)$ of one example light cone.

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

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

This light cone construction procedure assumes all the spectral lines have the same Schechter function parameters (ϕ_* and α) thus the same luminosity function shape, and the CO spectral line energy distribution (SLED) is also fixed. That is, by construction, all the sources have

[1] For the fiducial Schechter function, we choose the ϕ_* and α values of CO(1-0) for $z \le 5$ 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).

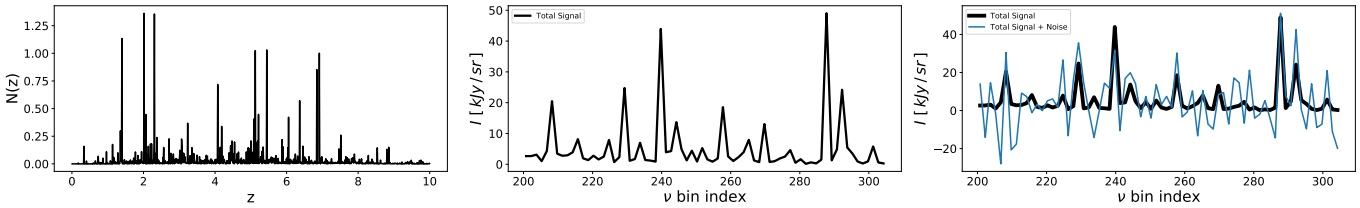

Figure 1. The steps of constructing a light cone. **Left:** $N(z)$, the effective number of ℓ_* sources in each redshift bins in this light cone. **Middle:** The signals from all six spectral lines and the sources in the 70 spectral channels. **Right:** The mock observed light cone (blue) consists of 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.

the same line luminosity ratio sets by the relative value of $\ell_*^{\rm line}$. In the main parts of this work, we use this fixed SLED model, and test the impact of adding the SLED variation on the performance of our technique in Sec. 5.1.

We further point out that, even 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 de-confusion technique, detailed in Sec. 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,

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

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

$\widetilde{A}$ is an $N_{\rm ch} \times N_{\rm z}$ matrix, where we have $N_{\rm ch} = 70$ spectral channels, and $N_{\rm z} = 2000$ redshift bins. 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,

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

where $\ell_*^{\rm line}$ is the line luminosity of ℓ_* source in the model, D_L is the luminosity distance, and $I_*^{\rm line}$ is defined as the observed line intensity of a $\ell_*^{\rm line}(z_j)$ source at z_j. Fig. 2 shows the $I_*^{\rm line}$ in our assumed model. Note that Eq. 3 assumes all 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 Sec. 5.1, we will show that our method also works in the realistic level of SLED variation.

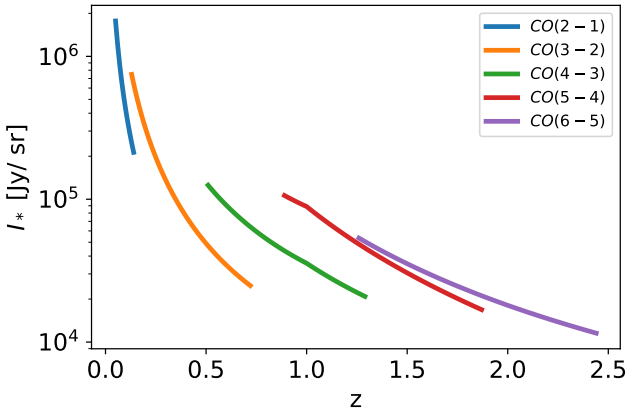

Figure 2. Intensity of the CO lines I_* from the sources of characteristic luminosity ℓ_*.

Eq. 3 can be written in the matrix form,

$$\mathbf{I} = \widetilde{\mathbf{A}}\mathbf{N} + \mathbf{n}, \tag{5}$$

where $\mathbf{I}$ and $\mathbf{n}$ are $N_{\rm ch}$-element column vectors, $\mathbf{N}$ is a N_z-element column vector, and $\widetilde{\mathbf{A}}$ is an $N_{\rm ch} \times N_{\rm z}$ matrix. Fig. 3 top panel shows the $\widetilde{\mathbf{A}}$ matrix in our model. The $\widetilde{\mathbf{A}}$ 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].

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

We further rewrite Eq. 3 by normalization and reducing the nuisance information. First, we normalize the columns in $\widetilde{\mathbf{A}}$ and move the normalization factor $\mathbf{I}_j^{\rm norm}$

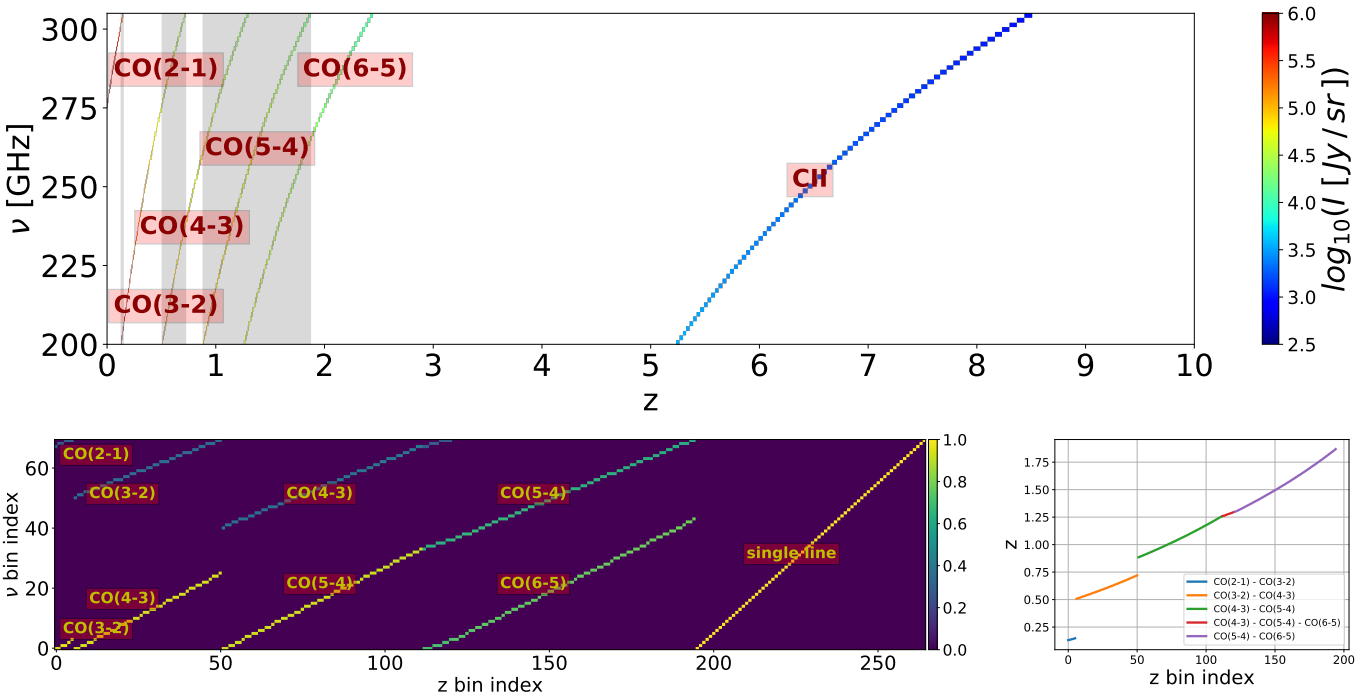

Figure 3. **Top:** $\widetilde{\mathbf{A}}$ matrix with our model on 70 frequency channels and 2000 redshift bins. $\widetilde{\mathbf{A}}$ are zeros (white) for the majority of the elements, and the six curves from left to right correspond to the six spectral lines, CO(2-1), CO(3-2), CO(4-3), CO(5-4), CO(6-5), and [C II].The color scale indicates line intensities in the fiducial model. The grey shaded regions are the redshifts that sources can be observed in multiple lines, and thus can be reconstructed in our technique. **Bottom Left: A** matrix with $N_{\rm ch} \times N_{\rm z} = 70 \times 265$ size, which is the reduced and normalized $\widetilde{\mathbf{A}}$. The color scale represents the intensities with the fiducial model SLED normalized within each column, i.e. for all column j, $\sum_{i=1}^{N_{\rm ch}} \mathbf{A}_{ij}^2 = 1.$, **Bottom Right:** The redshift of the 195 multi-line redshift bins in $\mathbf{A}$, 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 can not be reconstructed with our technique ($0.15 \leq z \leq 0.51$; $0.72 \leq z \leq 0.89$; $z \geq 1.87$).

to their $\mathbf{N}$ element:

$$\mathbf{I} = \mathbf{A}\widetilde{\mathbf{N}} + \mathbf{n}, \tag{6}$$

where

$$\begin{aligned} \mathbf{A}_{ij} &= \widetilde{\mathbf{A}}_{ij} \,/\, \mathbf{I}_j^{\rm norm}, \\ \widetilde{\mathbf{N}}_j &= \mathbf{N}_j \, \mathbf{I}_j^{\rm norm}, \\ \sum_{i=1}^{N_{\rm ch}} \mathbf{A}_{ij}^2 &= 1. \end{aligned} \tag{7}$$

Next, we will reduce the nuisance or redundant elements in Eq. 6. The columns of $\mathbf{A}$ are the basis spanning the observed data space. In constructing $\widetilde{\mathbf{A}}$, we simply design the columns to be equally-spacing redshift bins as shown in Fig. 3 top panel. However, this natural basis is highly degenerate. To remove the nuisance information, we first discard the redshift bins that are zero vectors in $\widetilde{\mathbf{A}}$. 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 $\mathbf{A}$ is identical to another column having [C II] signal in the same channe. In other words, given an observed data $\mathbf{I}$, we can not distinguish the origin of the source with a single line emission. Therefore, we combine these identical columns to a single one. In conclusion, we keep the columns in $\mathbf{A}$ with redshift bins that can be observed with multiple spectral lines, plus an identity matrix for the redshift bins that only have single detectable line.

For all 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 $\widetilde{\mathbf{N}}$ 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 in the following procedures: we first generate $\mathbf{A}$ with fine

Table 1. Frequencies and redshifts of the six defined broad bands

name	line	ν bin index	ν [GHz]	$\langle z \rangle (z_{\rm min} - z_{\rm max})$
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.

redshift bins ($\Delta z = 5\times10^{-4}$ from $z = 0$ to 10)[2], keep the columns with multiple lines, then identify the group of neighboring columns which have signals in the same sets of channels, and keep the medium bin and discard the others. With this process, we get 195 non-degenerate columns. Hereafter, "multi-line redshift bins" refers to these 195 redshift bins that can be detected in multiple channels. Finally, we append the $N_{\rm ch}$-sized identity matrix that account for the single-line redshift bins to these 195 columns generate $\mathbf{A}$. Fig. 3 bottom left panel shows $\mathbf{A}$ matrix which has size $N_{\rm ch} \times N_z = 70 \times 265$ (195 multi-line redshifts plus 70 columns of identity matrix). The bottom right panel are the redshift of the 195 multi-line bins, and their colors label are the pairs of detectable CO lines. We also define six broad bands from these pairs of lines by binning group of channels. Table 1 listed the definition of the broad bands. 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.

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

[2] Note that with this fine bins, the same frequency bin can map to multiple redshift bins instead of an one-on-one mapping.

3.2. *Sparse Approximation*

The key step in our de-confusion technique is to solve $\widetilde{\mathbf{N}}$ Eq. 6, given the observed spectrum $\mathbf{I}$ and model $\mathbf{A}$. This type of linear system has been extensively studied in the context of CMB map making, in which $\mathbf{I}$, $\mathbf{n}$, and $\widetilde{\mathbf{N}}$ in Eq. 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 ($N_{\rm ch} = 70$), and thus the standard map-making algorithm (e.g., Stompor et al. 2002) can not be applied.

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

Nevertheless, Eq. 6 can be solved with the "sparse" condition, which means the preferred solution of $\widetilde{\mathbf{N}}$ is the one with small number of non-zero elements. With this constraint, we can solve Eq. 6 with the following well-defined optimization problem,

$$\underset{\|\widetilde{\mathbf{N}}\|_0}{\rm argmin} \frac{1}{N_{\rm ch}} \left\| \mathbf{I} - \mathbf{A}\widetilde{\mathbf{N}} \right\|_2^2 < \epsilon^2, \tag{8}$$

where the ℓ_0-norm $\|\cdot\|_0$ is the number of non-zero 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" which is composed of a set of "atoms", and represent the signal in the data in terms of linear combination of a few atoms in the dictionary. In Eq. 8, the dictionary is the matrix $\mathbf{A}$, and the its atoms are the column vectors of $\mathbf{A}$.

The sparse approximation can only be applied if $k \equiv \left\|\widetilde{\mathbf{N}}\right\|_0 \ll N_{\rm ch}$. Note that the sparsity of the problem is quantify by $k/N_{\rm ch}$ but not k/N_z, since k/N_z can always designed to be arbitrarily small by choosing a large basis (fine redshift bins), but the degree of freedom in the solution is restricted by the input size $N_{\rm ch}$, and thus $k \ll N_{\rm ch}$ condition prohibits the algorithm to use 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 the elements in $\widetilde{\mathbf{N}}$ is non-zero. However, as mentioned in Sec. 2.3, the light cone signals are dominated by only a few bright sources, and the intensity field can be well-described with 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 $N_{\rm eff}$ as:

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

where $V_{\rm vox}$ is the voxel size of the redshift bin. Note that Φ is the number of sources per luminosity per voxel, and therefore $N_{\rm eff}$ is dimensionless and is proportional to the voxel size. $N_{\rm eff}$ can be interpreted as the reciprocal of the effective shot noise in LIM, which is an analogy to the $1/N$ shot noise in a galaxy power spectrum. If the luminosity function Φ follows the Schechter function form, then $N_{\rm eff}$ is (approximately) the number of sources brighter than ℓ_*, which contributes the majority of the emission.

We can estimate k by the cumulative $N_{\rm eff}(z)$ along the line of sight per light cone. Fig. 4 shows the cumulative $N_{\rm eff}$ in our model. While $N_{\rm eff} \sim 100$ from $z = 0$ to 10, the only relevant range is $z \lesssim 2.5$, where CO lines fall in the observed frequency range (see Fig. 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 $N_{\rm eff} \sim 10 \ll N_{\rm ch} = 70$, so the sparse condition is qualified in our problem.

3.3. *The Matching Pursuit Algorithm*

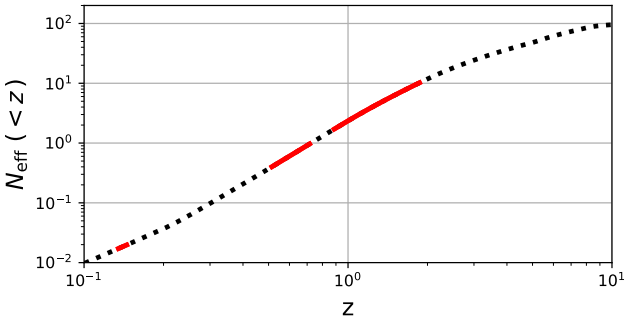

Figure 4. The redshift cumulated $N_{\rm eff}$ in our model (black dashed line). The red segments mark the multi-line redshift ranges that have multiple observable CO lines and thus the de-confusion technique can be applied.

We use the matching pursuit (MP) algorithm first introduced by Mallat & Zhang (1993) to solve for Eq. 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 $\mathbf{A}$ are the atoms that form the dictionary space. The 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 SNR 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 a m-σ detection threshold on the signals (e.g., $m = 5$ for a 5-σ 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 the higher purity map, whereas lower "m" values pick out fainter sources at the cost of increased false detection from noise. The optimal threshold can be determine with, for example, Fisher analysis (Cheng et al. 2019), but this is beyond the scope of this paper.

4. RESULTS

We present line de-confusion results in the simple case where mock light cones and the template ($\mathbf{A}$) are both generated from the same signal model (Sec. 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 robust-

ness 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 Sec. 5.

We quantify the reconstruction performance by computing two statistics on the true and reconstructed data: (1) the Pearson correlation coefficient (Sec. 4.2) and (2) voxel intensity distribution (VID) (Sec. 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 an Example Results*

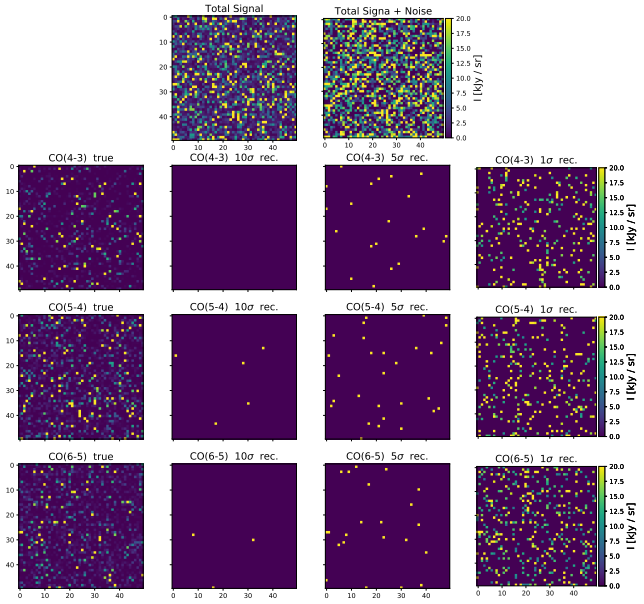

Figure 5. Visualization of the phase-space reconstruction with $\sigma_n = 10^4$ Jy/sr. We place the 2500 toy model light cones into a 50×50 pixel map, and show the reconstruction results on one of the spectrum bin (274 GHz). The top left panel is the true signal intensity map from all the spectral lines (CO(3-2), CO(4-3), CO(5-4), CO(6-5), [C II]). The top right panel is the observed intensity map including line signals (top left panel) and noise. The bottom three panels shows the true input and the 10-σ, 5-σ, and 1-σ reconstructed (right) emission field of the three spectral lines in the multi-line regime where the signal can be reconstructed with

Fig. 5 visualizes the reconstruction results on one of the spectrum bin (274 GHz) that contains five lines (CO(3-2), CO(4-3), CO(5-4), CO(6-5), [C II]) in the observable band. We place the 2500 light cones into a 50×50 pixel map, and show the true and the reconstructed intensity fields. Note that the input signals do not have the 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 is the observed intensity including the total line signals and a Gaussian instrument noise with $\sigma_n = 10^4$ Jy/sr. In this channel, three of the lines (CO(4-3), CO(5-4), CO(6-5)) are in the multi-line 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 10-σ, 5-σ, and 1-σ 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 Fig. 5, in the high threshold case (10-σ), the MP algorithm only extracts a few bright sources that are above the reconstruction threshold. As the threshold turning down, more sources are being reconstructed, at the cost of increased false positive detection from the noise fluctuations or the interlopers.

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

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

where $N_{\rm lc} = 2500$ is the number of light cones, $I^i_{\rm true}$ and $I^i_{\rm rec}$ are the true and the reconstructed line intensity map at the i-th light cone. Fig. 6 shows the results of correlation coefficient with a $\sigma_n = 10^4$ Jy/sr noise level and a 5-σ reconstruction threshold. Our reconstructed map achieves $\sim 80\%$ correlation with the true input map at $z \lesssim 1.5$.

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

- At the 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.

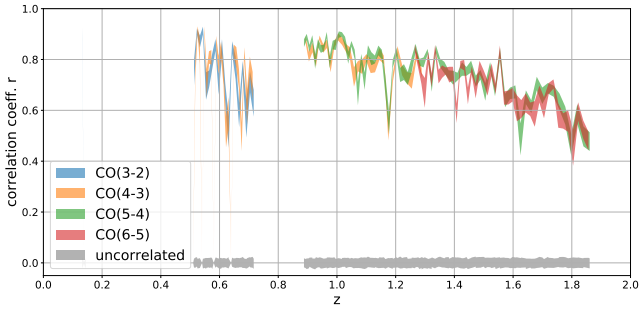

Figure 6. Pearson correlation coefficient r between the true and the reconstructed maps on 2500 light cones with $\sigma_n = 10^4$ Jy/sr 5-σ reconstruction. The bands are the 1-σ scatter of 100 noise realizations with the sample line signal. The grey bands are the r value with the white noise map for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

- The six defined bands correspond to three pair 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 being reconstructed in the same MP iteration. The pair of lines in the same redshift are reconstructed in the same MP iteration, and as a result, they are highly correlated.
- 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 case, discussed in the next bullet).
- Correlation coefficient r has very low dependency on $\sigma_{\rm th}$ in $\sigma_n = 10^3$ Jy/sr case. This can be understood by comparing the noise level σ_n to the quantity I_*, the intensity of the ℓ_* source in Schechter function (Fig. 2). At the redshift range that we perform the reconstruction ($0.5 \lesssim z \lesssim 1.9$), $10^4 \lesssim I_* \lesssim 10^5$ Jy/sr, which indicates that $m\sigma_n < I_*$ for all the m values considered in Fig. 7 ($m = 1 \sim 10$). In Schechter luminosity function, the sources $\gtrsim \ell_*$ contributes majority of the information in the intensity field (Cheng et al. 2019), and thus the correlation r is not sensitive to the change in reconstruction threshold if $m\sigma_n \ll \ell_*$.

4.3. *Voxel Intensity Distribution (VID)*

The Pearson correlation coefficient traces the phase-space variations between the true and reconstructed maps, but 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, voxel intensity distribution (VID). Note we do not directly compare the mean intensity of each map since we only reconstruct the bright sources that are above the threshold and neglecting all the faint sources in the reconstructed map, so the mean intensity is not expected be faithfully recovered.

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

Fig. 8 compares the VID of the true and reconstructed maps in the six broad bands with noise level $\sigma_n = 10^4$ Jy / sr and reconstruction threshold $m = 4$. The grey lines are the VID of the total observed maps that includes signals from all 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 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 Schechter function.

5. DISCUSSION

5.1. *Model Uncertainty*

For results presented in Sec. 4, the light cone signals and the dictionary template $\mathbf{A}$ are both generated from the same assumed signal model (Sec.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 variation and bias levels (at 20%, 50%, 100%) to the mock data, and run the reconstruction with the same dictionary template $\mathbf{A}$. We detail the definition of variation and bias below. Daddi et al. (2015) measured multiple CO lines of 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 $L^{\rm line}$:

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

where $L^{\rm line}_{fid}$ is the fiducial luminosity from our model, and δ_L is a zero-mean Gaussian random variable with

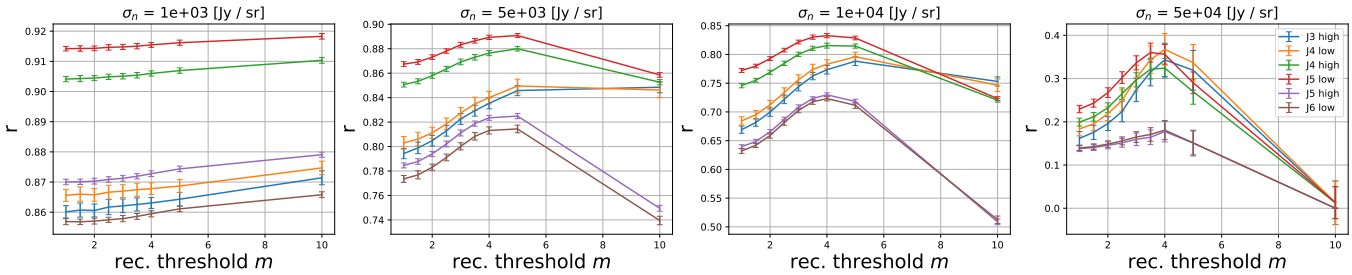

Figure 7. Average correlation coefficient r in each broadbands (Table 1) with a range of noise level σ_n and the reconstruction threshold m. The error bars are the r.m.s. of 100 noise realizations.

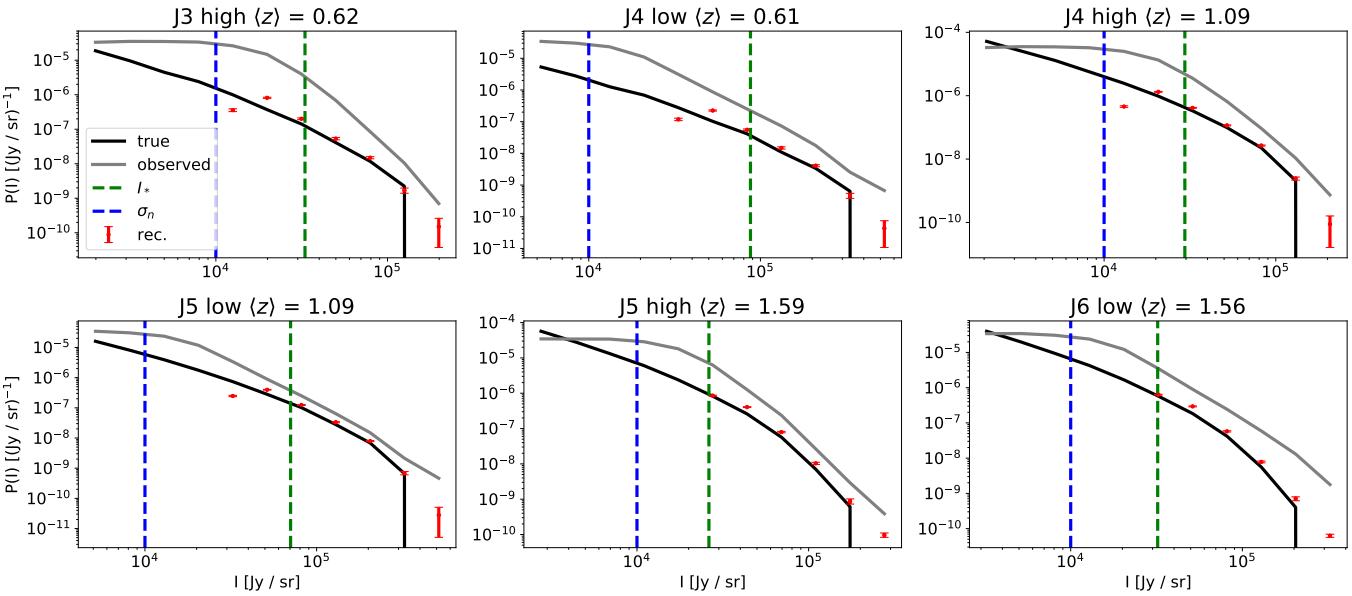

Figure 8. VID of the true input (black) and the reconstructed (red) maps in the six broad bands with noise level $\sigma_n = 10^4$ Jy /sr and reconstruction threshold $m = 4$. The error bars are the r.m.s. of 100 noise realizations. The grey curves are the VID of the total observed map that includes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

a standard deviation of 0.2, 0.5, 1.0 for the 20%, 50%, 100% variation cases, respectively.

Fig. 9 shows the reconstructed correlation coefficients r with a 4-σ reconstruction threshold ($m = 4$) with different SLED variations. We see that introducing 20% SLED variations in the model has a very mild impact on the reconstructed correlation; whereas with 50% (100%) variation, the correlation coefficient drops by about 10% ($\sim$ 50%). Fig. 10 compares the VID. We see that the SLED fluctuation has negligible impact on the VID reconstruction in the 20%, 50%, and 100% variation cases. Hence, we conclude that our technique is robust against realistic level of CO SLED fluctuations.

5.1.2. *Model offset*

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

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

where b_L is a constant offset that we assign to the model lines, and is applied to all 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), CO(6-5) to ensure the SLED ratio between neighboring lines are 20%, since our algorithm is only sensitive to the line ratio between two neighboring CO lines.

Fig. 11 and Fig. 10 shows the reconstructed correlation coefficients r and VID with a 4-σ with three SLED bias levels. Similar to Fig. 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.

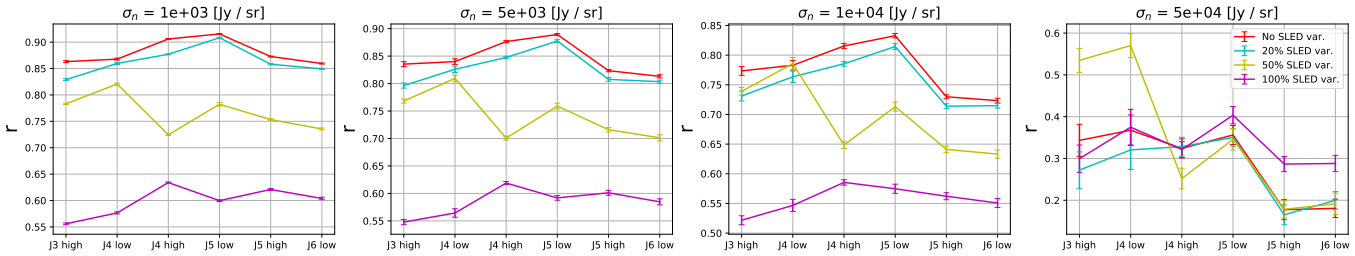

Figure 9. Comparing r of no SLED variation (red), 20% variation (cyan), 50% variation (yellow), and 100% variation (purple) with 4-σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the r.m.s. of 100 noise realizations of all the spectral bins in each band.

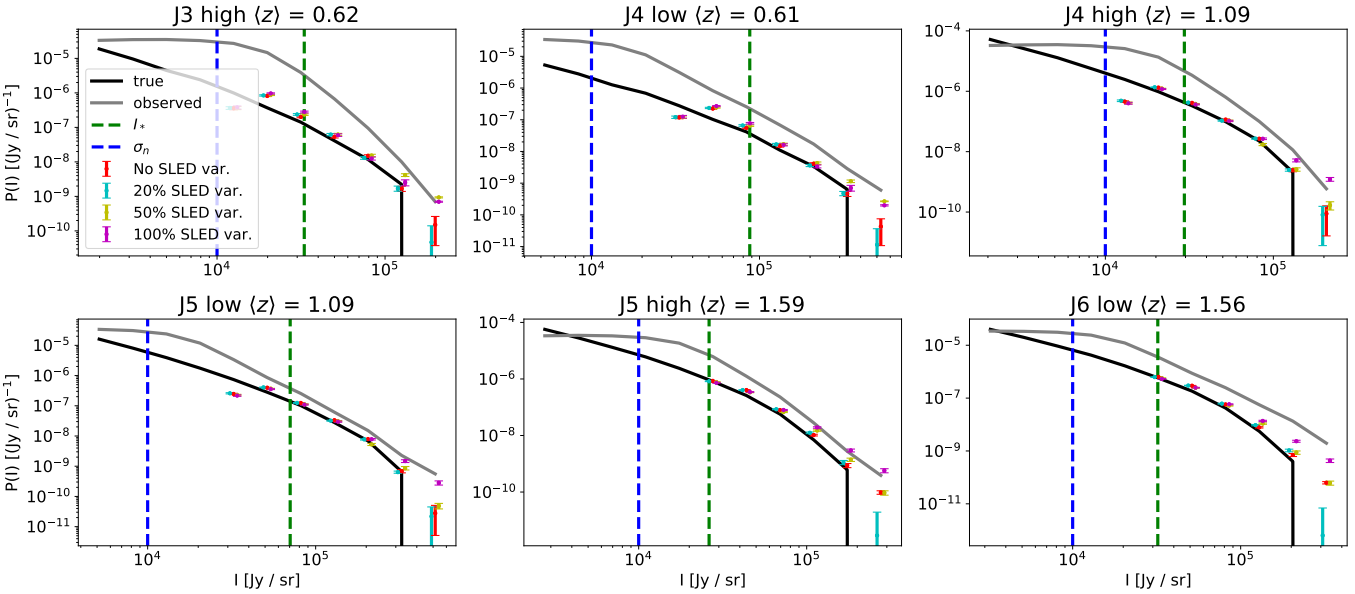

Figure 10. VID of no SLED variation (red), 20% variation (cyan), 50% variation (yellow), 100% variation (purple), and the true (black) maps in the six broad band with noise level $\sigma_n = 10^4$ Jy /sr and reconstruction threshold $m = 4$. The error bars are the r.m.s. of 100 noise realizations. The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

5.2. *Application to LIM in Other Wavelengths*

The technique developed in the work is not restricted to the [C II] and CO's line 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 SPHEREx-like survey.

SPHEREx is an on-going NASA MIDEX mission to conduct an all-sky near-infrared spectro-imaging survey (Doré et al. 2014, http://spherex.caltech.edu). SPHEREx will carry out the first all-sky spectral survey at wavelengths between 0.75 and 5 μm with 96 spectral channels and a 6.2″ 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 line intensity mapping regime.

The line signal model is described in Appendix C. We generate a near-infrared LIM mock data with a 6.2″ × 6.2″ pixel size and a 5σ point source sensitivity of $m_{AB} = 22$ (similar depth as the SPHEREx deep fields), and run our de-confusion algorithm on the mock light cones. Fig. 13 shows results of the correlation coefficients between the true and a 3σ-threshold reconstructed intensity maps. The reconstructed map achieves $\sim 80\%$ correlation with the true input map at $z \lesssim 3$, and decreases towards higher redshifts.

To account for uncertainties in modeling the SLED, we apply realistic level of line luminosity variations from Moustakas et al. (2006). We apply 10%, 50% and 100% SLED variation (Eq.11) to the line luminosity ratio of H_α/H_β, [O II]/H_α, and [O III]/[O II], respectively. Fig. 14 shows the correlation coefficients between the true and a 3σ threshold reconstructed intensity maps. Comparing with the fixed SLED case (Fig. 13), only the H_β line shows a significant decrease in the performance. The intensity map of the three brighter lines (H_α, [O

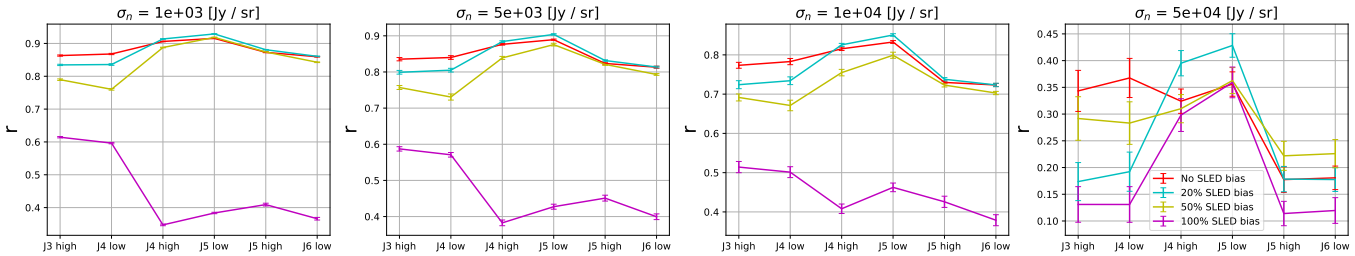

Figure 11. Comparing r of no SLED bias (red), 20% bias (cyan), 50% bias (yellow), and 100% bias (purple) with 4-σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the r.m.s. of 100 noise realizations of all the spectral bins in each band.

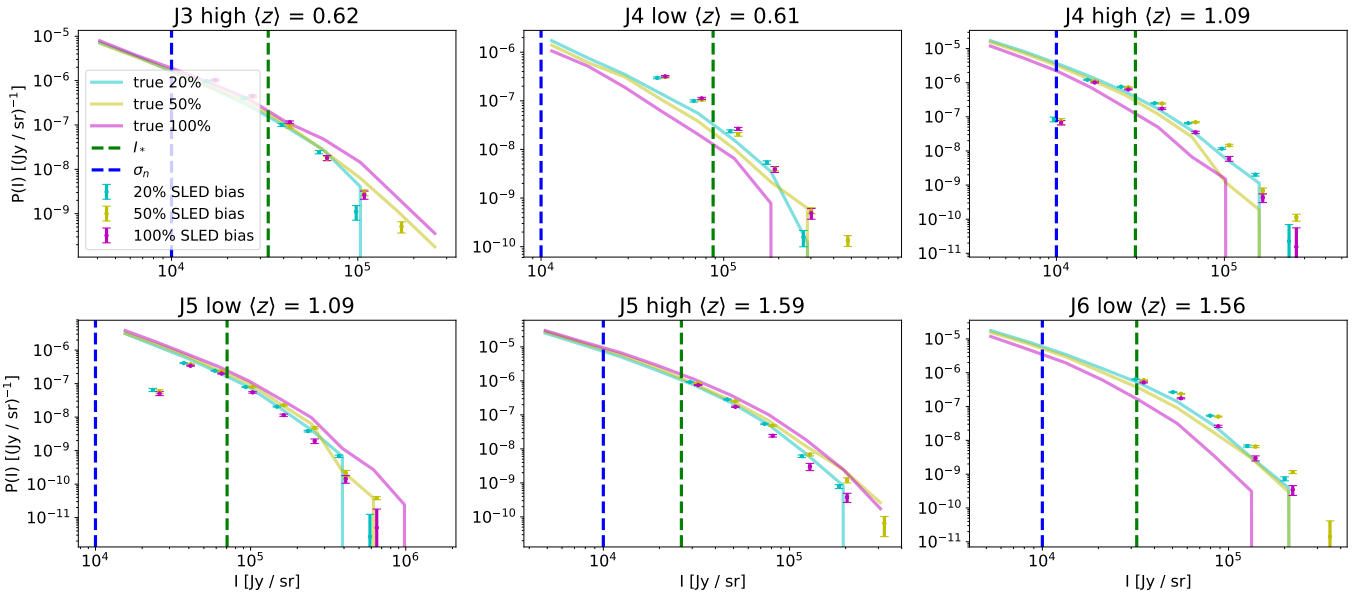

Figure 12. VID of no SLED bias (red), 20% bias (cyan), 50% bias (yellow), 100% bias (purple), and the true maps (solid lines) in the six broad band with noise level $\sigma_n = 10^4$ Jy /sr and reconstruction threshold $m = 4$. The error bars are the r.m.s. of 100 noise realizations. The grey curves are the VID of the total observed map that includes signals from all the lines and noise. For reference, the blue and green dashed lines mark the noise level σ_n and I_*.

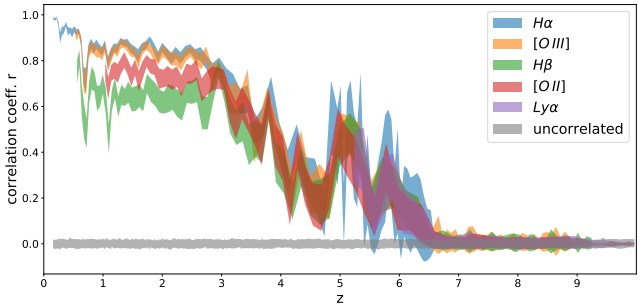

Figure 13. Pearson correlation coefficient r between the true and the reconstructed maps on 2500 light cones for a SPHEREx-like mock data. The bands are the r.m.s. of the value in 100 noise realizations with the sample line signal. The grey bands are the correlation coefficient with the uncorrelated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

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 (redshifed) 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 contin-

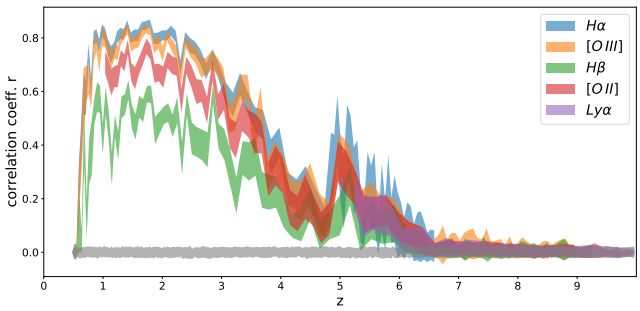

Figure 14. Pearson correlation coefficient r between the true and the reconstructed maps on 2500 light cones for a SPHEREx-like mock data with realistic level of SLED variation. The bands are the r.m.s. of the value in 100 noise realizations with the sample line signal. The grey bands are the correlation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

uum 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[3]. 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, 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 different in each light cone, since each light cone contain different galaxies with different continuum spectrum. We fit and subtract a first order polynomial function to the spectrum of each light cone before running the reconstruction.

The results of a 4-σ-threshold reconstruction with different noise σ_n level are shown in Fig. 15 and Fig. 16. We can see compared to the no background subtraction case, the r value is even higher in these two tests. This is because our reconstruction only extract 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 have no impact on the reconstruction. In conclusion, the background subtraction in the LIM data reduction pipeline will not affect our line reconstruction technique.

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

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 de-confusion 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 $150\,L_\odot$ Mpc^{-2} (6×10^{-17} W m^{-2}) in each light cone. The source density in the catalog is ~ 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 a L_* galaxy at $z \sim 2$. According to Helgason et al. (2012), such a 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.

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 as usual until hitting the stopping criteria. The results of a 4-σ-threshold reconstruction with different noise levels are shown in Fig. 17. The reconstruction shows improved results for all noise levels. However, the huge improvement in the highest noise case cannot be interpreted as successful reconstruction of source intensities. At this high noise level, the signals are well 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.

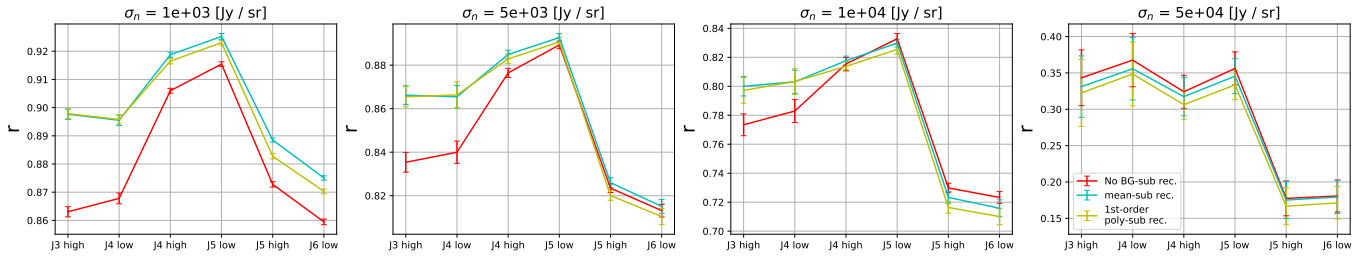

Figure 15. Comparing r of the reconstructed map with no background subtraction (red), subtracting the mean of the whole data cube (cyan), and subtracting a first order polynomial in spectral direction for each light cone (yellow), using 4-σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the r.m.s. of 100 noise realizations of all the spectral bins in each band.

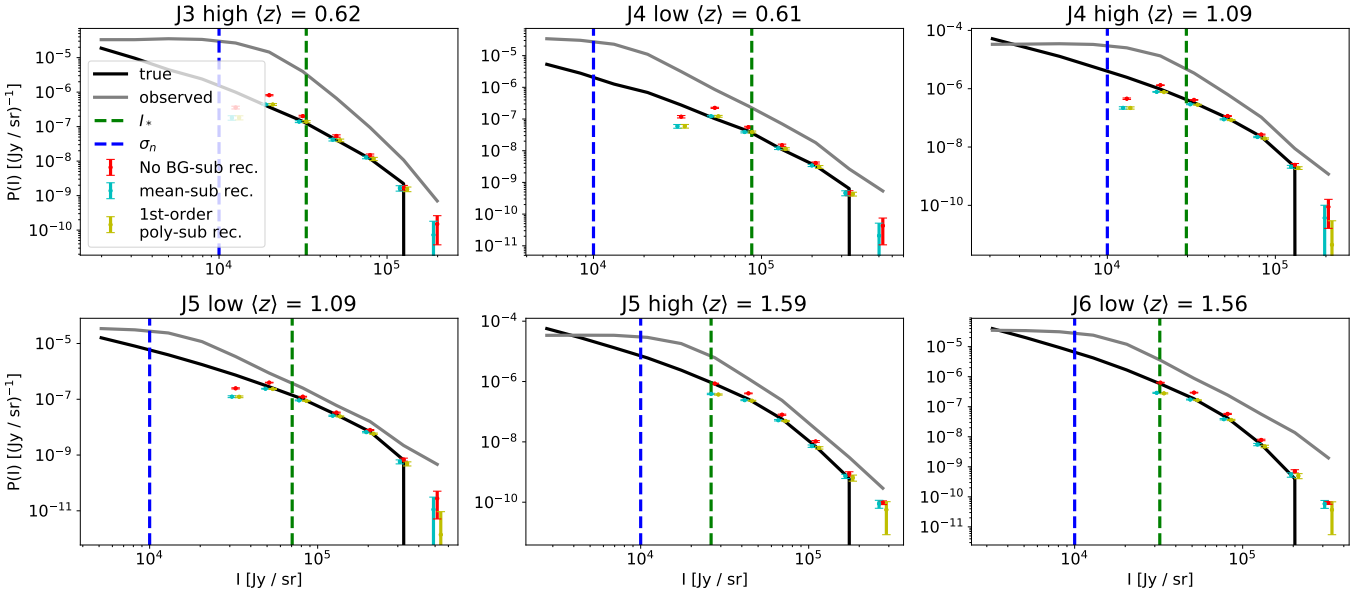

Figure 16. VID of no background subtraction (red),subtracting the mean of the whole data cube (cyan), subtracting a first order polynomial in spectral direction for each light cone (yellow), and the true (black) maps in the six broad band with noise level $\sigma_n = 10^4$ Jy /sr and reconstruction threshold $m = 4$. The error bars are the r.m.s. of 100 noise realizations. The grey curves are the VID of the total observed map that includes signals from all the lines and noise. For reference, the blue and green dashed lines mark the noise level σ_n and I_*.

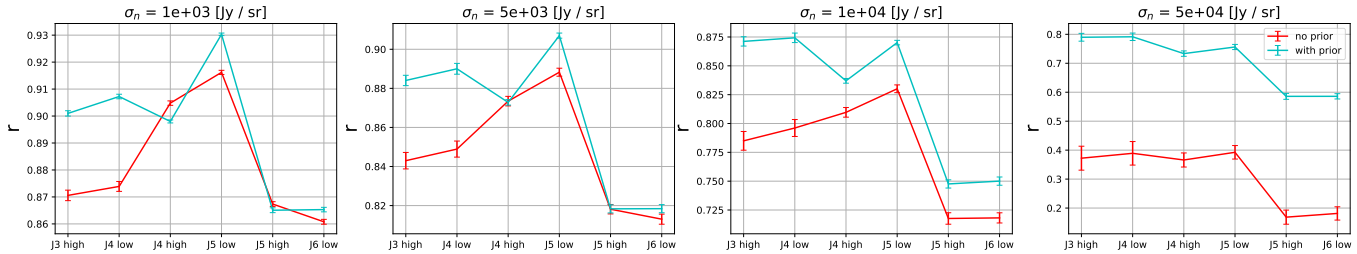

Figure 17. Comparing r of the reconstructed map with no external catalog prior (red), and utilizing the external catalog prior to fit the catalog sources before running the MP reconstruction (cyan), using 4-σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the r.m.s. of 100 noise realizations of all the spectral bins in each band.

Thus the improved correlation are 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-

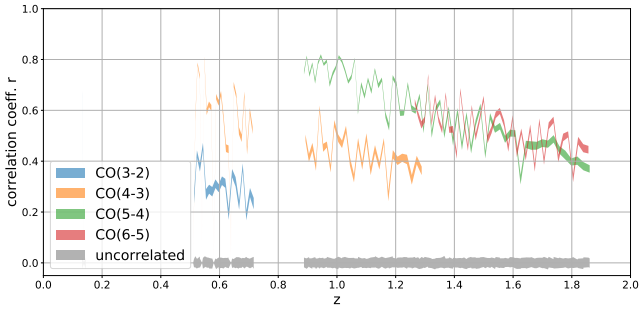

Figure 18. Pearson correlation coefficient r between the true line maps and the same maps adding the $\sigma_n = 10^{-4}$ Jy/sr noise on 2500 light cones. The bands are the r.m.s. of the value in 100 noise realizations with the sample line signal. The grey bands are the r value with the uncorrelated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

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.

Fig. 18 shows the results using $\sigma_n = 10^4$ Jy/sr. Comparing to Fig. 6, which has the same noise level, the case of single line plus noise have a lower correlation than the reconstruction in Fig. 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 have 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 better SNR on the fainter lines because of the better sensitivity on their brighter counterpart.

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 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

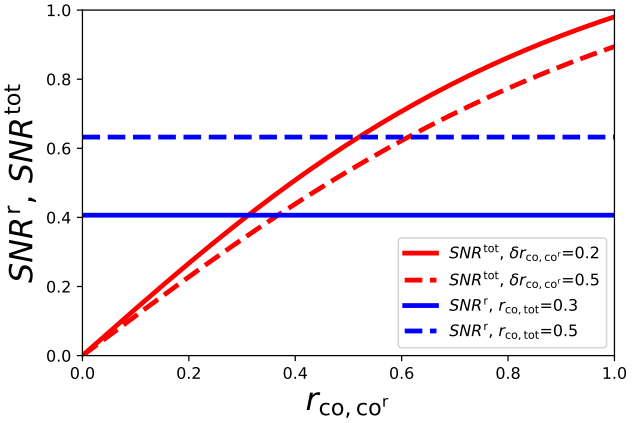

Figure 19. Comparing the SNR on cross power spectrum of one CO line and the external tracer, using the total observed map ($SNR^{\rm tot}$) and the reconstructed CO map ($SNR^{\rm r}$). We consider two different value of $\delta r_{\rm co,co^r}$ and $r_{\rm co,tot}$ that covers the range of realistic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

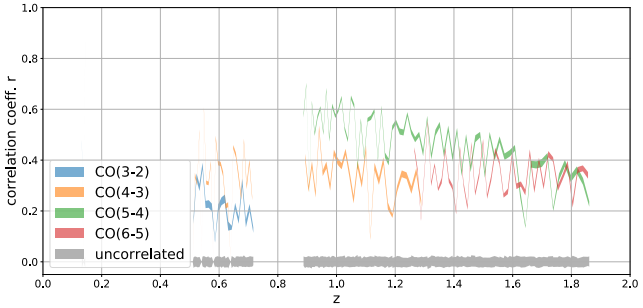

Figure 20. Pearson correlation coefficient r between the true input and the total observed maps on 2500 light cones with $\sigma_n = 10^4$ Jy/sr. The bands are the r.m.s. of the value in 100 noise realizations with the sample line signal. The grey bands are the r value with the uncorrelated white noise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

analysis, using 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 is given by

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

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. Eq. 13 can be

expressed in terms of the cross correlation coefficient: $r_{1,2} = P_{1,2}/\sqrt{P_1 P_2}$,

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

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 ($I_{\rm CO}$), other interloper lines ($I_{\rm interlopers}$), and the noise (I_n),

$$I_{\rm tot} = I_{\rm CO} + I_{\rm interlopers} + I_n. \tag{15}$$

The expectation value of the cross spectrum between the observed total map and galaxy is 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 $\langle P_{\rm g,tot}\rangle = \langle P_{\rm g,co}\rangle$.

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

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

$$\begin{aligned}\delta P_{\rm g,co}^{\rm tot} &= P_{\rm g,tot}\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{\rm g,tot}^2}\right)} \\ &= P_{\rm g,co}\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{\rm co,tot}^2}\right)}\end{aligned} \tag{16}$$

On the other hand, if we cross correlate the galaxies with the reconstructed CO map (CO$^{\rm r}$), the error is

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

However, $P_{\rm g,co^r}$ is a biased estimator of $P_{\rm g,co}$ because of the error in the reconstruction. If we assume the reconstructed map roughly preserves the same power as the true map, $P_{\rm co^r} \approx P_{\rm co}$, then we can write[4] $P_{\rm g,co^r} = r_{\rm g,co^r} P_{\rm g,co} = r_{\rm co,co^r} P_{\rm g,co}$. Therefore, we have to de-bias cross spectrum $P_{\rm g,co^r}$ by factor $1/r_{\rm co,co^r}$, $P_{\rm g,co} = P_{\rm g,co^r}/r_{\rm co,co^r}$. The value of $r_{\rm co,co^r}$ 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 $\delta r_{\rm co,co^r}$ due to the uncertainty in $r_{\rm co,co^r}$,

$$\begin{aligned}\delta P_{\rm g,co}^{\rm r} &= P_{\rm g,co}\sqrt{\left(\frac{\delta P_{\rm g,co^r}}{P_{\rm g,co^r}}\right)^2 + \left(\frac{\delta r_{\rm co,co^r}}{r_{\rm co,co^r}}\right)^2} \\ &= P_{\rm g,co}\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{\rm co,co^r}^2}\right) + \left(\frac{\delta r_{\rm co,co^r}}{r_{\rm co,co^r}}\right)^2}.\end{aligned} \tag{18}$$

In summary, the SNR on the galaxy-CO cross spectrum $P_{\rm g,co}$ using the total observed map and the reconstructed map are

$$\begin{aligned}SNR^{\rm tot} &= \frac{P_{\rm g,co}}{\delta P_{\rm g,co}^{\rm tot}} = \frac{1}{\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{\rm co,tot}^2}\right)}}, \\ SNR^{\rm r} &= \frac{P_{\rm g,co}}{\delta P_{\rm g,co}^{\rm r}} = \frac{1}{\sqrt{\frac{1}{2}\left(1 + \frac{1}{r_{\rm co,co^r}^2}\right) + \left(\frac{\delta r_{\rm co,co^r}}{r_{\rm co,co^r}}\right)^2}}.\end{aligned} \tag{19}$$

Note that the SNR in both cases converges to unity when r is unity and the de-bias error is zero, which is the limit of sample variance[5].

Fig. 19 shows the $SNR^{\rm tot}$ and $SNR^{\rm r}$ with two different value of $\delta r_{\rm co,co^r}$ and $r_{\rm co,tot}$ as a function of $r_{\rm co,co^r}$. According to the calculations in Sec. 5.1, the value of $r_{\rm co,co^r}$ ranges from ~ 0.7 to ~ 0.9 (except for the most noisy case ($\sigma_n = 5\times10^4$ Jy/sr). To estimate the realistic $r_{\rm co,tot}$ value, we calculate the correlation of the input CO line maps and the observed map with $\sigma_n = 10^4$ Jy/sr, the same noise level as in Fig. 6. The results are shown in Fig. 20. We get $r_{\rm co,tot}$ are around 0.2 to 0.5 on this case. With these range of parameters, Fig. 19 indicates that $SNR^{\rm r}$ is better than $SNR^{\rm tot}$, 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 a LIM dataset can be estimated

[4] $P_{\rm g,co^r} = r_{\rm g,co^r}\sqrt{P_g P_{\rm co^r}} \approx r_{\rm g,co^r}\sqrt{P_g P_{\rm co}} = r_{\rm g,co^r} P_{\rm g-co}$, where the last equality uses the assumption $r_{\rm g,co} = 1$.

[5] The power spectrum cross correlation coefficient can be derived from the Pearson correlation (Eq. 10) with a weighting on pixels. Since our pixels are generated and reconstructed independently with each others, the Pearson correlation coefficient here is an unbiased estimator of power spectrum correlation coefficient. Thus here we will use the value of Pearson correlation we derived for power spectrum correlation coefficient.

without bias. As an example, we write the observed intensity of a [CII] LIM dataset as

$$I^{\rm obs} = I_{\rm CO} + I_{\rm CII} + \delta_n^{\rm obs}, \tag{20}$$

where $I_{\rm CO}$ is the total CO interloper intensity, $I_{\rm CII}$ the [C II] signal, and $\delta_n^{\rm obs}$ 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:

$$I_{\rm CO}^{\rm rec} = \alpha I_{\rm CO_b} + \delta_n^{\rm rec}, \tag{21}$$

where $I_{\rm CO}^{\rm rec}$ is the reconstructed CO intensity and $I_{\rm CO_b}$ 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 $\delta_n^{\rm rec}$ describing random (or misidentified) fluctuations about the true CO intensity that is uncorrelated with the CO field.

The auto power spectrum of the reconstructed CO map is

$$\begin{aligned}\langle I_{\rm CO}^{\rm rec} I_{\rm CO}^{\rm rec}\rangle =& \bar{\alpha}^2 \langle I_{\rm CO_b} I_{\rm CO_b}\rangle + \langle \delta_n^{\rm rec}\delta_n^{\rm rec}\rangle \\ =& \bar{\alpha}^2 \langle I_{\rm CO_b}\rangle^2 b_{\rm CO_b}^2 P(k) + N^{\rm rec},\end{aligned} \tag{22}$$

where in the linear regime that we consider here $b_{\rm CO_b}$ is the cosmological clustering bias of the bright CO sources, $P(k)$ the matter density field, and $N^{\rm rec}$ the auto power spectrum of the $\delta_n^{\rm rec}$ 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:

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

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

where $\langle\alpha\rangle = \bar{\alpha}$, b_g is the bias of the galaxy tracer, $b_{\rm CO_b}$ 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.

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

$$\begin{aligned}\left\langle \widehat{I_{\rm CO} I_{\rm CO}} \right\rangle &= \langle I_{\rm CO}^{\rm rec} I_{\rm CO}^{\rm rec}\rangle \left(\frac{\langle g I_{\rm CO}^{\rm obs}\rangle}{\langle g I_{\rm CO}^{\rm rec}\rangle}\right)^2 \\ &= \left(\bar{\alpha}^2 \langle I_{\rm CO_b}\rangle^2 b_{\rm CO_b}^2 P(k) + N^{\rm rec}\right)\left(\frac{r\langle I_{\rm CO}\rangle b_{CO}}{\bar{\alpha} r_b \langle I_{\rm CO_b}\rangle b_{CO_b}}\right)^2 \\ &= \left(\langle I_{\rm CO_b}\rangle^2 b_{\rm CO_b}^2 P(k) + \frac{N^{\rm rec}}{\bar{\alpha}^2}\right)\left(\frac{r\langle I_{\rm CO}\rangle b_{CO}}{r_b \langle I_{\rm CO_b}\rangle b_{CO_b}}\right)^2\end{aligned} \tag{25}$$

On large scales, $r = r_b = 1$, as all tracers follow the dark matter distribution. If we carefully pick a reconstruction threshold such that the reconstruction of a subset of the CO population (bright CO sources) is nearly noiseless, $N^{\rm rec}/\bar{\alpha}^2 \to 0$, this estimator becomes an unbiased estimator of the full CO power spectrum on large scales:

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

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 estimate of the high-redshift [CII] power spectrum in the same LIM dataset, which is highly desirable.

We note that a high reconstruction threshold m, however, is required in order to achieve a (nearly) noiseless reconstruction. A high threshold tends to boost the shot noise in the reconstructed power spectrum, as well as uncertainties in the CO power spectrum estimator (Eq. 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: in the masking case the "cleaned map" is the observed map (including all the lines and instrument noise) masked using external CO catalogs; whereas the "cleaned map" derived from our algorithm is the observed map subtracting the CO reconstructed map.

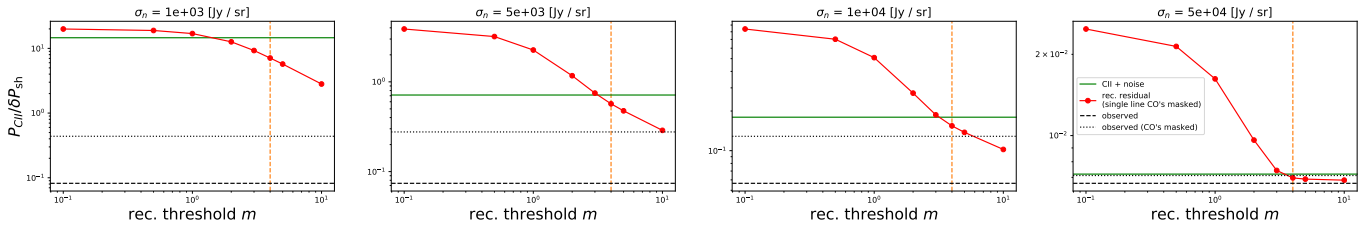

Figure 21. SNR on [C II] shot noise power spectrum after CO line foreground removal with masking (black dotted) and sparse reconstruction (red) with different instrument noise level σ_n and reconstruction threshold m. The orange dashed line marks the $m = 4$ threshold for reference. We also show the SNR before cleaning (black dashed), and the limiting case of the no CO signals (green) for comparison.

The external catalog considered here is the same as the one described in Sec. 5.4: a flux cut on CO(5-4) at the level of $150\,L_\odot$ Mpc^{-2} (6×10^{-17} W m^{-2}), which gives ~ 1.2 galaxies per light cone (0.43^2 arcmin^2 pixel solid angle) and corresponds to $m_{\rm AB} \sim 21.8$ 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 multi-line redshfit 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 the reconstructed single-line signals, we will over-subtract [C II] in the cleaned map. Therefore, for the single-line redshift bins ($0.15 \le z \le 0.51$; $0.72 \le z \le 0.89$; $z \ge 1.87$), we clean the data by masking voxels that contain external catalog sources.

As a figure of merit, we calculate the SNR 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

$$\delta P_{\rm sh} = \frac{P_{\rm tot}}{\sqrt{N_{\rm mode}}} = \frac{V_{\rm vox}\sigma^2_{\rm tot}}{\sqrt{N_{\rm mode}}}, \tag{27}$$

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, $\sigma^2_{\rm tot}$. $N_{\rm mode}$ 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

$$P_{\rm CII} = V_{\rm vox}\sigma^2_{\rm CII}, \tag{28}$$

where $\sigma^2_{\rm CII}$ is the voxel variance of [C II] signal map. The SNR of the [C II] shot noise power spectrum is then:

$$\frac{P_{\rm CII}}{\delta P_{\rm sh}} = \frac{\sigma^2_{\rm CII}}{\sigma^2_{\rm tot}}\sqrt{N_{\rm mode}}. \tag{29}$$

We test the foreground cleaning performance with the same set of 2500 mock light cones described in Sec. 4, and calculate the [C II] shot noise SNR using Eq. 29. The $\sigma^2_{\rm CII}$ is the variance of the input [C II] map, and $\sigma^2_{\rm tot}$ the variance of the cleaned map. For simplicity, $\sigma^2_{\rm CII}$ and $\sigma^2_{\rm tot}$ are the variance from all the frequency channels.

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

According to Fig. 7, the optimal threshold that gives the maximum r value is $m \sim 4$. In Fig. 21, $m = 4$ is marked with 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), 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. $\sigma^2_{\rm tot} = \sigma^2_{\rm CII} + \sigma^2_n$. For a small reconstruction threshold m, the reconstructed SNR is better than this limit, which indicates that the reconstruction overfits and mis-identifies noise fluctuation as signal and remove them from the cleaned data. We see that for the two realistic noise levels at $m \sim 4$, the reconstructed SNR 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 SNR is close to the noise plus [C II] limit (green line), the reason is as follows: 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 SNR (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 SNR 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 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 of the line reconstruction for 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 $\mathbf{A}$ for every redshift bins 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 matching pursuit algorithm adopted in this work optimizes the ℓ_0 norm in Eq. 8, which is the direct sum of the non-zero elements in $\widetilde{\mathbf{N}}$. We can improve the algorithm by including prior information on the expected value of each element in $\widetilde{\mathbf{N}}$, which depends on the voxel size and source luminosity function. For instance, instead of using the matching pursuit algorithm, one can obtain the sparse solution by solving following the ℓ_1 norm regularization equation:

$$\underset{\widetilde{\mathbf{N}}}{\operatorname{argmin}} \left\| \mathbf{I} - \mathbf{A}\widetilde{\mathbf{N}} \right\|_2^2 + \lambda \left\| \mathbf{w} \cdot \widetilde{\mathbf{N}} \right\|_1 , \tag{30}$$

where the parameter λ determine the regularization strength for preventing overfitting, which have the similar effect as the stopping criteria in MP algorithm; the prior information on the number density of the sources in each redshift bins can be encoded in the weight vector $\mathbf{w}$ in this expression.

5.9.3. *Clustering Information*

In this work, we only perform the pixel-by-pixel line de-confusion 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.

6. CONCLUSION

We develop a spectral line de-confusion technique for LIM experiments, where different spectral lines emitted by sources at different redshifts can be observed in the same frequency channel and be confused. Unlike most of the previously proposed methods that decompose the line signals in the power spectrum space, we perform a phase-space de-confusion 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 a 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 into 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 matching pursuit algorithm.

As an example, we consider a LIM survey with similar survey parameters as the on-going 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, and the voxel intensity distribution of individual lines can also be correctly extracted to the ℓ_*-scale of the assumed Schechter luminosity function. The reconstruction performance is robust against 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 dataset 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 de-confusion problems in LIM experiments and enhance the science returns.

ACKNOWLEDGMENTS

We are grateful to helpful discussions with Chun-Lin Liu, Patrick Breysse, Matthieu Béthermin, Emmanuel Schaan, Matt Orr, the TIME collaboration, the Caltech ObsCos group, and the participants in the conference "Lines in the Large Scale Structure". Part of the research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

APPENDIX

A. THE MATCHING PURSUIT (MP) ALGORITHM

In this section, we describe the detailed steps in the 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 $\mathbf{f}$ and residual $\mathbf{R}$ for the next step until the solution meets the stopping criteria. In Eq. 8, for a given signal $\mathbf{I}$ and matrix $\mathbf{A}$, we define a set of vectors $\{\psi_i\}$ to be the column vectors in $\mathbf{A}$ (i.e. the atoms in dictionary). The MP algorithm works as follows:

1. Initialize at step $t = 0$: $\mathbf{f}_0 = \mathbf{0}$, $\mathbf{R}_0 = \mathbf{I}$, $\widetilde{\mathbf{N}}_0 = \mathbf{0}$.

2. Compute the inner product of $\mathbf{R}_0$ and ψ_i's:
$$\{u_{0i}\} = \{\langle \mathbf{R}_0, \psi_i \rangle\}\,. \tag{A1}$$

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 $\mathbf{N} = \mathbf{0}$. Else, proceed to step 5.

5. Update the current $\mathbf{f}_t$, $\mathbf{R}_t$, and record the amplitude of the new solution in γ-th element of vector $\widetilde{\mathbf{N}}_t$:
$$\mathbf{f}_{t+1} = \mathbf{f}_t + u_{t\gamma}\,\psi_\gamma, \tag{A3}$$
$$\mathbf{R}_{t+1} = \mathbf{R}_t - u_{t\gamma}\,\psi_\gamma, \tag{A4}$$
$$\widetilde{\mathbf{N}}_{t+1}(\gamma) = \widetilde{\mathbf{N}}_t(\gamma) + u_{t\gamma} \tag{A5}$$

6. Compute the inner product of $\mathbf{R}_{t+1}$ and ψ_i's:
$$\{u_{(t+1)i}\} = \{\langle \mathbf{R}_{t+1}, \psi_i \rangle\}\,. \tag{A6}$$

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 $\mathbf{N} = \widetilde{\mathbf{N}}_{t+1} \,/\, \mathbf{I}^{\mathrm{norm}}$.

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

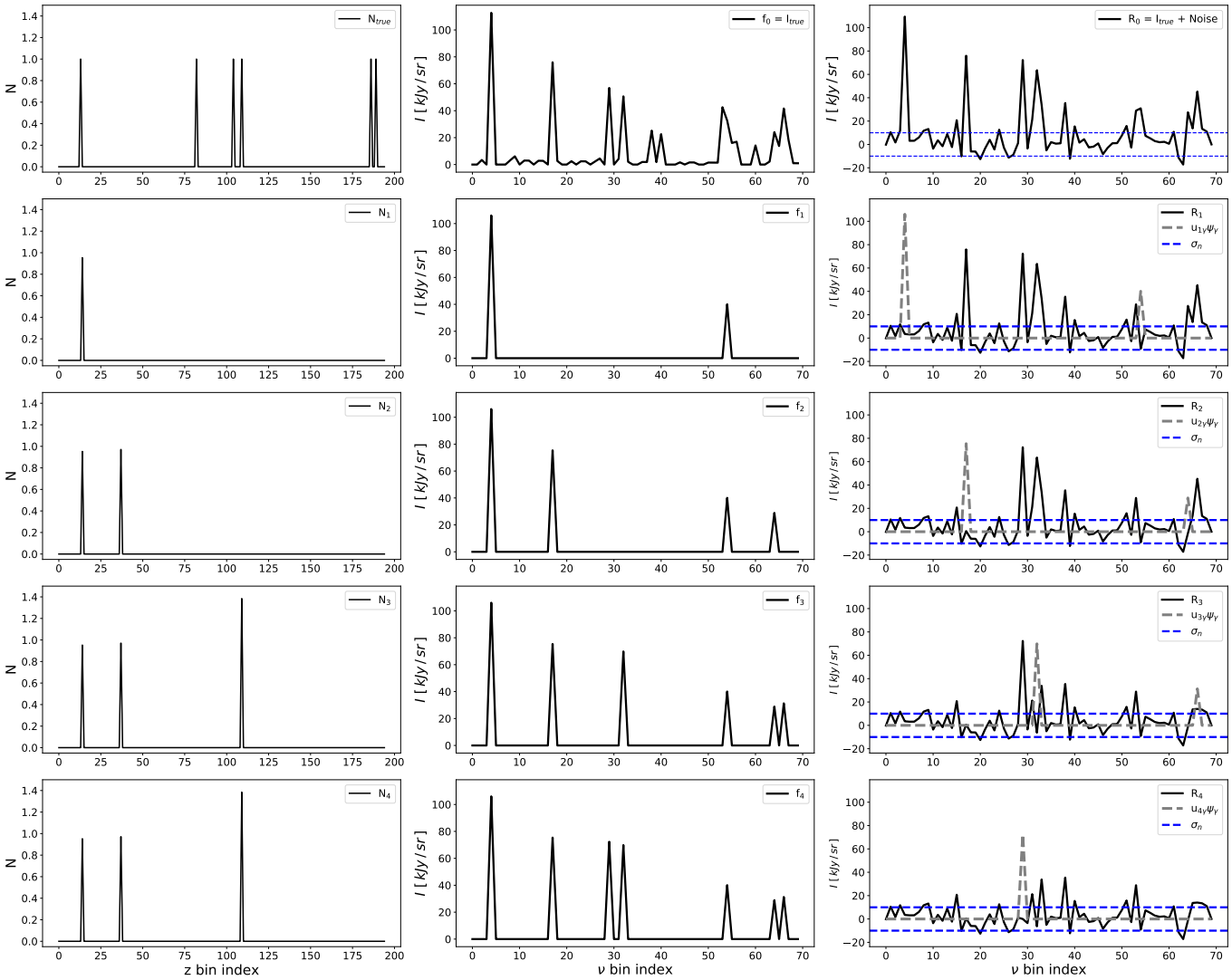

Figure 22. Illustration of how the MP algorithm solves for the source vector $\mathbf{N}$. See text for the detailed description.

Fig. 22 illustrates the steps of MP algorithm solving $\mathbf{N}$ 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 multi-line redshift bins at z = [0.54, 1.06, 1.20, 1.24, 1.79, 1.82], shown in the top left panel. Since $\mathbf{N}$ is the effective number of ℓ_* sources per redshift bin, the amplitudes in the six corresponding redshift bins is unity. The top middle panel is the total line signal in this light cone $\mathbf{I}_{\rm true}$, including the emission from those six sources as well as other sources in the single-line redshifts. The top right panel is the observed data after adding noise to the signal $\mathbf{I}_{\rm true}$, which is also the $\mathbf{R}_0$ vector in the first step of 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 grey dashed spectrum in the left panel of second row is the template signal extracted in this step ($u_{1\gamma}\psi_\gamma$). The updated the value of $\mathbf{N}_1$, $\mathbf{f}_1$, and $\mathbf{R}_1$ from step 5 is shown in the second row of Fig. 22.

Then we proceed to the second iteration. The MP algorithm selects the 37th column ($z = 0.65$). The grey dash 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 Fig. 22 shows the updated value of $\mathbf{N}_2$, $\mathbf{f}_2$, and $\mathbf{R}_2$ from the second iteration.

In the third iteration, the 109th ($z = 1.24$) column in the dictionary is selected. The fourth row of Fig. 22 shows the updated value of $\mathbf{N}_3$, $\mathbf{f}_3$, and $\mathbf{R}_3$ from this

iteration. Note that this time the MP algorithm picks up a correct redshift, while overestimates the amplitude by $\sim 40\%$.

In the fourth iteration, the algorithm selects the 224th columns from the dictionary. The 224th column is not in the multi-line redshift bins (first 195 columns), and thus the $\mathbf{N}_4$ in the bottom row of Fig. 22 remains unchanged, whereas $\mathbf{f}_4$ and $\mathbf{R}_4$ 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 Fig. 22 as the reconstruction results for this light cone.

In summary, in this example, two of the six multi-line redshift sources have been reconstructed, in addition to one mis-identified source. Comparing the final reconstructed light cone signal (bottom middle panel of Fig. 22) to the true input light cone (top middle panel of Fig. 22), we can see that the MP reconstruction captures the strong peaks in the data, and the remaining signals are close to the noise level.

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

In Sec. A, the residual of step t $\mathbf{R_t}$ can be express in the linear combination of the dictionary atoms and noise:

$$\mathbf{R}_t = \sum_i c_i\, \psi_i + \mathbf{n}, \tag{B8}$$

where c_i's are the constant coefficient. Then we derive

$$\begin{aligned} u_{t\gamma} &\equiv \langle\, \mathbf{R}_t, \psi_\gamma \rangle \\ &= \sum_i c_i \langle\, \psi_i, \psi_\gamma \rangle + \langle\, \mathbf{n}, \psi_\gamma \rangle \\ &= \mathrm{const} + \sum_j n_j\, \psi_{\gamma j}. \end{aligned} \tag{B9}$$

The first term is not depend on the noise, so it's a constant term that does not attribute to the variance. Also note that $\langle\, \psi_i, \psi_\gamma \rangle \neq \delta_{i\gamma}$ since the dictionary $\{\psi_i\}$ are normalized but not orthogonal. With this expression, we can calculate the variance:

$$\begin{aligned} \langle u_{t\gamma} \rangle &= \mathrm{const} + \sum_j \langle n_j \rangle\, \psi_{\gamma j} = \mathrm{const} \\ \langle u_{t\gamma}^2 \rangle &= \mathrm{const}^2 + \sum_j \langle n_j^2 \rangle\, \psi_{\gamma j}^2 \\ &= \mathrm{const}^2 + \sigma_n^2 \sum_j \psi_{\gamma j}^2 \\ &= \mathrm{const}^2 + \sigma_n^2. \end{aligned} \tag{B10}$$

[6] 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 $z > 8$.

Therefore, we get

$$\mathrm{var}(u_{t\gamma}) = \langle u_{t\gamma}^2 \rangle - \langle u_{t\gamma} \rangle^2 = \sigma_n^2 \tag{B11}$$

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 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 the star formation activities, we model the signal in 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 halos.

We use the publicly available halo mass function calculator *HMFcalc* (Murray et al. 2013, http://hmf.icrar.org/) to obtain the halo mass function based on 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 [6]

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

$$\frac{SFR}{M_\odot/\mathrm{yr}} = 2.29 \times 10^{-41} \frac{L_{\mathrm{Ly}\alpha}}{\mathrm{erg/s}}. \tag{C12}$$

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

$$\frac{SFR}{M_\odot/\mathrm{yr}} = (7.9 \pm 2.4) \times 10^{-42} \frac{L_{H\alpha}}{\mathrm{erg/s}}, \tag{C13}$$

$$\frac{SFR}{M_\odot/\mathrm{yr}} = (1.4 \pm 0.4) \times 10^{-41} \frac{L_{[OII]}}{\mathrm{erg/s}}, \tag{C14}$$

$$\frac{SFR}{M_\odot/\mathrm{yr}} = (7.6 \pm 3.7) \times 10^{-42} \frac{L_{[OIII]}}{\mathrm{erg/s}}, \tag{C15}$$

and for Hβ line, we use the fixed line ratio $H\beta/H\alpha = 0.35$ (Osterbrock & Ferland 2006).

REFERENCES

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57, doi: 10.1088/0004-637X/770/1/57

Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2009, ApJ, 695, 183, doi: 10.1088/0004-637X/695/1/183

Breysse, P. C., Kovetz, E. D., Behroozi, P. S., Dai, L., & Kamionkowski, M. 2017, MNRAS, 467, 2996, doi: 10.1093/mnras/stx203

Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2014, MNRAS, 443, 3506, doi: 10.1093/mnras/stu1312

—. 2015, MNRAS, 452, 3408, doi: 10.1093/mnras/stv1476

—. 2016, MNRAS, 457, L127, doi: 10.1093/mnrasl/slw005

Breysse, P. C., & Rahman, M. 2017, MNRAS, 468, 741, doi: 10.1093/mnras/stx451

Candes, E. J., Romberg, J., & Tao, T. 2006, IEEE Trans. Inf. Theor., 52, 489, doi: 10.1109/TIT.2005.862083

Carilli, C. L. 2011, ApJL, 730, L30, doi: 10.1088/2041-8205/730/2/L30

Chang, T. C., Gong, Y., Santos, M., et al. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 4. https://arxiv.org/abs/1501.04654

Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Nature, 466, 463, doi: 10.1038/nature09187

Chang, T.-C., Pen, U.-L., Peterson, J. B., & McDonald, P. 2008, Physical Review Letters, 100, 091303, doi: 10.1103/PhysRevLett.100.091303

Chapman, E., Abdalla, F. B., Harker, G., et al. 2012, MNRAS, 423, 2518, doi: 10.1111/j.1365-2966.2012.21065.x

Cheng, Y.-T., Chang, T.-C., Bock, J., Bradford, C. M., & Cooray, A. 2016, ApJ, 832, 165, doi: 10.3847/0004-637X/832/2/165

Cheng, Y.-T., de Putter, R., Chang, T.-C., & Doré, O. 2019, ApJ, 877, 86, doi: 10.3847/1538-4357/ab1b2b

Chung, D. T., Viero, M. P., Church, S. E., et al. 2019, ApJ, 872, 186, doi: 10.3847/1538-4357/ab0027

Comaschi, P., & Ferrara, A. 2016, MNRAS, 455, 725, doi: 10.1093/mnras/stv2339

Crites, A. T., Bock, J. J., Bradford, C. M., et al. 2014, in Proc. SPIE, Vol. 9153, Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII, 91531W, doi: 10.1117/12.2057207

Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., Blomqvist, M., & Pieri, M. 2018, MNRAS, 481, 1320, doi: 10.1093/mnras/sty2302

Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., et al. 2016, MNRAS, 457, 3541, doi: 10.1093/mnras/stw204

Daddi, E., Dannerbauer, H., Liu, D., et al. 2015, A&A, 577, A46, doi: 10.1051/0004-6361/201425043

de Putter, R., Holder, G. P., Chang, T.-C., & Dore, O. 2014, ArXiv e-prints. https://arxiv.org/abs/1403.3727

Donoho, D. L. 2006, IEEE Trans. Inf. Theor., 52, 1289, doi: 10.1109/TIT.2006.871582

Doré, O., Bock, J., Ashby, M., et al. 2014, ArXiv e-prints. https://arxiv.org/abs/1412.4872

Fonseca, J., Silva, M. B., Santos, M. G., & Cooray, A. 2017, MNRAS, 464, 1948, doi: 10.1093/mnras/stw2470

Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhR, 433, 181, doi: 10.1016/j.physrep.2006.08.002

Gong, Y., Chen, X., & Cooray, A. 2020, arXiv e-prints, arXiv:2001.10792. https://arxiv.org/abs/2001.10792

Gong, Y., Cooray, A., Silva, M., et al. 2012, ApJ, 745, 49, doi: 10.1088/0004-637X/745/1/49

Gong, Y., Cooray, A., Silva, M. B., Santos, M. G., & Lubin, P. 2011, ApJL, 728, L46, doi: 10.1088/2041-8205/728/2/L46

Gong, Y., Silva, M., Cooray, A., & Santos, M. G. 2014, ApJ, 785, 72, doi: 10.1088/0004-637X/785/1/72

Helgason, K., Ricotti, M., & Kashlinsky, A. 2012, ApJ, 752, 113, doi: 10.1088/0004-637X/752/2/113

Ihle, H. T., Chung, D., Stein, G., et al. 2019, ApJ, 871, 75, doi: 10.3847/1538-4357/aaf4bc

Keating, G. K., Marrone, D. P., Bower, G. C., et al. 2016, ApJ, 830, 34, doi: 10.3847/0004-637X/830/1/34

Keating, G. K., Bower, G. C., Marrone, D. P., et al. 2015, ApJ, 814, 140, doi: 10.1088/0004-637X/814/2/140

Kennicutt, Robert C., J. 1998, Annual Review of Astronomy and Astrophysics, 36, 189, doi: 10.1146/annurev.astro.36.1.189

Kogut, A., Dwek, E., & Moseley, S. H. 2015, ApJ, 806, 234, doi: 10.1088/0004-637X/806/2/234

Lagache, G., Cousin, M., & Chatzikos, M. 2018, A&A, 609, A130, doi: 10.1051/0004-6361/201732019

Li, T. Y., Wechsler, R. H., Devaraj, K., & Church, S. E. 2016, ApJ, 817, 169, doi: 10.3847/0004-637X/817/2/169

Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, ApJ, 741, 70, doi: 10.1088/0004-637X/741/2/70

Lidz, A., & Taylor, J. 2016, ApJ, 825, 143, doi: 10.3847/0004-637X/825/2/143

Lidz, A., Zahn, O., Furlanetto, S. R., et al. 2009, ApJ, 690, 252, doi: 10.1088/0004-637X/690/1/252

Liu, A., & Tegmark, M. 2012, MNRAS, 419, 3491, doi: 10.1111/j.1365-2966.2011.19989.x

Ly, C., Malkan, M. A., Kashikawa, N., et al. 2007, ApJ, 657, 738, doi: 10.1086/510828

Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429

Mallat, S. G., & Zhang, Z. 1993, IEEE Transactions on Signal Processing, 41, 3397, doi: 10.1109/78.258082

Mashian, N., Sternberg, A., & Loeb, A. 2015, JCAP, 11, 028, doi: 10.1088/1475-7516/2015/11/028

Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJL, 763, L20, doi: 10.1088/2041-8205/763/1/L20

Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767, doi: 10.1086/506135

Moriwaki, K., Filippova, N., Shirasaki, M., & Yoshida, N. 2020, arXiv e-prints, arXiv:2002.07991. https://arxiv.org/abs/2002.07991

Moustakas, J., Kennicutt, Robert C., J., & Tremonti, C. A. 2006, ApJ, 642, 775, doi: 10.1086/500964

Murray, S. G., Power, C., & Robotham, A. S. G. 2013, Astronomy and Computing, 3, 23, doi: 10.1016/j.ascom.2013.11.001

Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of gaseous nebulae and active galactic nuclei

Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012, ApJ, 756, 165, doi: 10.1088/0004-637X/756/2/165

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13, doi: 10.1051/0004-6361/201525830

Popping, G., van Kampen, E., Decarli, R., et al. 2016, MNRAS, 461, 93, doi: 10.1093/mnras/stw1323

Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15, doi: 10.1088/0004-637X/768/1/15

Pullen, A. R., Doré, O., & Bock, J. 2014, ApJ, 786, 111, doi: 10.1088/0004-637X/786/2/111

Righi, M., Hernández-Monteagudo, C., & Sunyaev, R. A. 2008, A&A, 489, 489, doi: 10.1051/0004-6361:200810199

Scott, D., & Rees, M. J. 1990, MNRAS, 247, 510

Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1, doi: 10.1046/j.1365-8711.2001.04006.x

Silva, M., Santos, M. G., Cooray, A., & Gong, Y. 2015, ApJ, 806, 209, doi: 10.1088/0004-637X/806/2/209

Silva, M. B., Santos, M. G., Gong, Y., Cooray, A., & Bock, J. 2013, ApJ, 763, 132, doi: 10.1088/0004-637X/763/2/132

Stompor, R., Balbi, A., Borrill, J. D., et al. 2002, PhRvD, 65, 022003, doi: 10.1103/PhysRevD.65.022003

Sun, G., Moncelsi, L., Viero, M. P., et al. 2018, ApJ, 856, 107, doi: 10.3847/1538-4357/aab3e3

Switzer, E. R., Chang, T.-C., Masui, K. W., Pen, U.-L., & Voytek, T. C. 2015, ApJ, 815, 51, doi: 10.1088/0004-637X/815/1/51

Uzgil, B. D., Aguirre, J. E., Bradford, C. M., & Lidz, A. 2014, ApJ, 793, 116, doi: 10.1088/0004-637X/793/2/116

Visbal, E., & Loeb, A. 2010, JCAP, 11, 016, doi: 10.1088/1475-7516/2010/11/016

Wyithe, J. S. B., & Loeb, A. 2008, MNRAS, 383, 606, doi: 10.1111/j.1365-2966.2007.12568.x

Yue, B., Ferrara, A., Pallottini, A., Gallerani, S., & Vallini, L. 2015, MNRAS, 450, 3829, doi: 10.1093/mnras/stv933