CONCERTO: Extracting the power spectrum of the [C II] emission line

M. van Cuyck, N. Ponthieu, G. Lagache, A. Beelen, M. Béthermin, A. Gkogkou, M. Aravena, A. Benoit, J. Bounmy, M. Calvo, et al.

A&A 676, A62 (2023)
https://doi.org/10.1051/0004-6361/202346270

Astronomy & Astrophysics

CONCERTO: Extracting the power spectrum of the [C$_{II}$] emission line

M. Van Cuyck[1], N. Ponthieu[2,1], G. Lagache[1], A. Beelen[1], M. Béthermin[1,3], A. Gkogkou[1], M. Aravena[4], A. Benoit[5], J. Bounmy[6], M. Calvo[5], A. Catalano[6], F. X. Désert[2], F.-X. Dupé[1], A. Fasano[1], A. Ferrara[7], J. Goupy[5], C. Hoarau[6], W. Hu[1], J.-C Lambert[1], J. F. Macías-Pérez[6], J. Marpaud[6], G. Mellema[8], A. Monfardini[5], and A. Pallottini[7]

[1] Aix-Marseille Univ., CNRS, CNES, LAM, Marseille, France
e-mail: mathilde.van-cuyck@lam.fr
[2] Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France
[3] Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France
[4] Núcleo de Astronomía, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Av. Ejército 441, Santiago, Chile
[5] Univ. Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France
[6] Univ. Grenoble Alpes, CNRS, LPSC/IN2P3, 38000 Grenoble, France
[7] Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
[8] The Oskar Klein Centre, Department of Astronomy, Stockholm University, AlbaNova 10691 Stockholm, Sweden

Received 28 February 2023 / Accepted 20 May 2023

ABSTRACT

Context. CONCERTO is the first experiment to perform a [C$_{II}$] line intensity mapping (LIM) survey on the COSMOS field to target $z > 5.2$. Measuring the [C$_{II}$] angular power spectrum allows us to study the role of dusty star-forming galaxies in the star formation history during the epochs of Reionization and post-Reionization. The main obstacle to this measurement is the contamination by bright foregrounds: the dust continuum emission and atomic and molecular lines from foreground galaxies at $z \lesssim 3$.

Aims. We evaluate our ability to retrieve the [C$_{II}$] signal in mock observations of the sky using the Simulated Infrared Dusty Extragalactic Sky (SIDES), which covers the mid-infrared to millimetre range. We also measure the impact of field-to-field variance on the residual foreground contamination.

Methods. We compared two methods for dealing with the dust continuum emission from galaxies (i.e. the cosmic infrared background fluctuations): the standard principal component analysis (PCA) and the asymmetric re-weighted penalized least-squares (arPLS) method. For line interlopers, the strategy relies on masking low-redshift galaxies using the instrumental beam profile and external catalogues. As we do not have observations of CO or deep-enough classical CO proxies (such as L_{IR}), we relied on the COSMOS stellar mass catalogue, which we demonstrate to be a reliable CO proxy for masking. To measure the angular power spectrum of masked data, we adapted the P of K EstimatoR (POKER) from cosmic infrared background studies and discuss its use on LIM data.

Results. The arPLS method achieves a reduction in the cosmic infrared background fluctuations to a sub-dominant level of the [C$_{II}$] power at $z \sim 7$, a factor of >70 below our fiducial [C$_{II}$] model. When using the standard PCA, this factor is only 0.7 at this redshift. The masking lowers the power amplitude of line contamination down to 2×10^{-2} Jy^2 sr^{-1}. This residual level is dominated by faint undetected sources that are not clustered around the detected (and masked) sources. For our [C$_{II}$] model, this results in a detection at $z = 5.2$ with a power ratio [C$_{II}$]/(residual interlopers) = 62 ± 32 for a 22% area survey loss. However, at $z = 7$, [C$_{II}$]/(residual interlopers) = 2.0 ± 1.4, due to the weak contrast between [C$_{II}$] and the residual line contamination. Thanks to the large area covered by SIDES-Uchuu, we show that the power amplitude of line residuals varies by 12–15% for $z = 5.2$–7, which is less than the field-to-field variance affecting [C$_{II}$] power spectra.

Conclusions. We present an end-to-end simulation of the extragalactic foreground removal that we ran to detect the [C$_{II}$] at high redshift via its angular power spectrum. We show that cosmic infrared background fluctuations are not a limiting foreground for [C$_{II}$] LIM. On the contrary, the CO and [C$_I$] line contamination severely limits our ability to accurately measure the [C$_{II}$] angular power spectrum at $z \gtrsim 7$.

Key words. galaxies: star formation – galaxies: high-redshift – galaxies: ISM – large-scale structure of Universe – cosmology: observations

1. Introduction

The Epoch of Reionization (EoR) started with the birth of the first stars and the formation of the first galaxies, whose ultraviolet light ionized the neutral hydrogen of the intergalactic medium. As time went by, the bubbles of ionized gas increased in number, expanding and merging, eventually reionizing the whole Universe. While we are beginning to understand the overall chronology of the EoR, the full story of reionization remains to be written. However, recent and planned observational advances promise progress in this respect over the coming decade.

The *Planck* satellite has provided an accurate measurement of the Thomson scattering optical depth, τ, using the cosmic microwave background (CMB) angular power spectrum (APS; Planck Collaboration Int. XLVII 2016). It shows that the

Universe was reionized at 10% at $z_{10\%} = 10.4 \pm 1.8$. On the other end of the reionization epoch, studies of the ionized intergalactic medium through quasar and gamma-ray burst absorption in the Lyman-α and Lyman-β hydrogen transitions set the late end at $z \approx 6$ (e.g., Chornock et al. 2013; Eilers et al. 2019; Wang et al. 2020; Bosman et al. 2022). As of today, the total ionizing photon budget required to reionize the Universe is still not probed by the deep surveys that have constrained the faint-end slope of the ultraviolet luminosity function in the early Universe (e.g., Ellis et al. 2013; Oesch et al. 2015; Carilli et al. 2016; Livermore et al. 2017; Drake et al. 2017). This may change in the near future with deep observations from the *James Webb* Space Telescope (JWST).

The CarbON CII line in post-rEionization and ReionizaTiOn (CONCERTO) project (CONCERTO Collaboration 2020) seeks to constrain the dust-obscured star formation history and its spatial distribution at high redshifts ($z > 5.2$). To that aim, CONCERTO is the first experiment to conduct a line intensity mapping (LIM) survey of the [CII] line at the end of the EoR and during the post-reionization era. A LIM survey consists of mapping the surface brightness in a given field of view as a function of position and observed frequency (see Kovetz et al. 2017 and Bernal & Kovetz 2022 for a review). Targeting an emission line then allows the redshift information to be recovered. Line intensity mapping is naturally sensitive to faint sources, missed by classical galaxy surveys, that potentially contribute the most to the ionizing photon budget (e.g., Finkelstein et al. 2012; Dayal et al. 2013; Atek et al. 2015). The CONCERTO project then studies surface brightness fluctuations in the Fourier space using the APS.

The [CII] 157.7 µm line is a line of choice for LIM at high redshifts. It is one of the brightest far-infrared lines and is unattenuated by the dust content of galaxies. Furthermore, it has been shown by both observations (e.g., Gullberg et al. 2015; Schaerer et al. 2020) and theoretical works (e.g., Pallottini et al. 2017; Lagache et al. 2018) that [CII] luminosity is a reliable tracer of star formation in the redshift range of interest. [CII] emission from high-redshift galaxies is dominated by the emission from photo-dissociation regions, which are regions of cool neutral gas penetrated by ionizing ultraviolet radiations from hot young stars (Pineda et al. 2014) that surround molecular clouds (Vallini et al. 2017; Pallottini et al. 2022). In addition, the transparent millimetre atmospheric window enables access from the ground to the [CII] emission originating at $z \gtrsim 4$. The main challenge for any kind of LIM experiment remains the foreground contamination, which can completely dominate the signal of interest. For the CONCERTO [CII] survey, the brightest foreground comes from fluctuations in the cosmic infrared background (CIB). The CIB is the line-of-sight-integrated emission of all dusty galaxies over the entire history of the Universe, from 8 to 1000 µm (Lagache et al. 2000). Further, there is a second foreground due to interloping emission lines, which are atomic and molecular emission lines redshifted in the same frequency band. For the [CII] survey in the CONCERTO frequency range, the main interlopers are the CO rotational ladder and the two [CI] fine structure lines from foreground galaxies at $z \lesssim 3$. In this paper we present in detail a deconfusion method designed for the [CII] survey conducted with CONCERTO.

The treatment of continuum foregrounds has mainly been studied in the context of 21 cm LIM. Most of the techniques rely on both its spectral smoothness and its strong fluctuations to reconstruct and subtract it (McQuinn et al. 2006; Furlanetto et al. 2006; Morales et al. 2006; Chang et al. 2010; Parsons et al. 2012; Liu & Tegmark 2012; Switzer et al. 2015; Yue et al. 2015) or used to avoid the contaminated angular scales in the (3D) Fourier space (Chapman et al. 2016). Regarding the interlopers, the cross correlation between the data and an alternative tracer probing the same cosmic volume can be used to disentangle the signal of interest from the interlopers. This option has been used for popular lines in intensity mapping (Lidz et al. 2009; Visbal & Loeb 2010; Masui et al. 2013; Croft et al. 2016; Chung et al. 2019) as well as for the [CII] line (Gong et al. 2012; Switzer et al. 2019; Padmanabhan et al. 2022; Chung 2023; Pullen et al. 2023) below $z < 3$. However, there are not enough ancillary probes tracing the EoR at $z > 6$ to cross correlate with the [CII] LIM data (but see Chang et al. 2015; Comaschi et al. 2016). A variety of deconfusion methods are thus under study. By exploiting the anisotropies in the APS of interlopers projected on the [CII] frame (Cheng et al. 2016; Lidz & Taylor 2016; Gong et al. 2020), it is possible to separate out the interloper contribution in the power spectrum, but a high sensitivity and a large-area survey are required to detect these anisotropy features. Cheng et al. (2020) propose a deconfusion method in the phase space for reconstructing the one-point statistics of CO-contaminated data, which can only be used if the CO spectral line energy distribution (SLED) in galaxies varies by less than 20% and with a reasonable noise level. A deep learning approach for de-blending H_α at $z = 1.3$ from [OIII] at $z = 2$ was also investigated in Breysse et al. (2015); it achieved a 91% precision but was not applied to the [CII] case.

Among the interloper separation methods, masking is conceptually the simplest one. The [CII] signal is not correlated with the signal from interlopers as they originate from lower redshifts. Hence, removing interloper-contaminated pixels should give an unbiased measurement of the [CII] power spectrum by lowering the interlopers' amplitude in the total APS. Masking efficiency has been studied through its effect on the clustering and shot noise power spectrum levels (Gong et al. 2014; Sun et al. 2018). However, due to the lack of realistic far-infrared to millimetre sky simulations to apply this method to, many side effects are poorly studied, such as the correlation between faint and high-mass galaxies, the correlation between the signal and the mask, and the mixing of k modes induced by the mask. Yue et al. (2015) applied the masking to mock maps, but they did not investigate the systematics as cosmic variance or residual contamination.

In this paper we study all the steps for an end-to-end component separation method applied to the [CII] LIM survey conducted with CONCERTO in order to extract the [CII] APS. We are interested in the detection significance and accuracy of the [CII] APS reconstruction that can be achieved in a typical CONCERTO field after foreground removal. We also study the effect of the field-to-field (FtF) variance on the masking results for CONCERTO-like field sizes. This variance affects both the interloper and the [CII] signals.

For this, we used Simulated Infrared Dusty Extragalactic Sky (SIDES), a newly developed realistic empirical simulation of the dusty infrared-millimetre sky up to $z = 7$, which is fully described in Béthermin et al. (2017, 2022) and Gkogkou et al. (2023), hereafter B17, B22, and G22, respectively. The component separation methods are applied directly to these mock observations, as they would be applied to the real CONCERTO observations. For the CIB contamination (continuum emission), our component separation is based on two different approaches. The first is principal component analysis (PCA) adapted to LIM data (e.g., Bigot-Sazy et al. 2015; Yohana et al. 2021), and the second is a baseline estimator called asymmetric reweighed penalized least squares (arPLS; Baek et al. 2015).

For the interloper emission lines, as CONCERTO carries out the [CII] survey on the widely studied Cosmic Evolution Survey (COSMOS; Scoville et al. 2007) field, we rely on a custom-made masking strategy. In this observational context, our main focus is on the 2D APS, which is the initial quantity derived from the data. The APS of each species can be directly compared, unlike the spherically averaged power spectra, which require projection onto the [CII] frame. Additionally, as detailed in B22, the lowest radial modes will not be explored, due to CONCERTO's low spectral resolution, and the majority of the captured signal will be predominantly present in the well-sampled transverse modes.

The paper is organized as follows. We first present the CONCERTO mock observation cubes from SIDES (Sect. 2). In Sect. 3 we address the removal of CIB anisotropy contamination. In Sect. 4 we present the masking strategy based on the use of the stellar mass, M_*, as a CO proxy and a beam width criterion. In Sect. 5 we present the masking results on the interloper maps alone and the interloper-contaminated [CII] maps. Finally, we summarize our results and conclude in Sect. 6. The validation of our APS estimator is detailed in Appendix A. Throughout the paper we assume a Planck Collaboration XIII (2016) cosmology and a Chabrier (2003) initial mass function, as in SIDES (B17).

2. SIDES simulation

The SIDES simulation, which mimics the CONCERTO extragalactic sky, is a central element of our study. SIDES accurately replicates a series of statistical properties, spanning from mid-infrared to millimetre wavelengths, including the redshift distributions and number counts of galaxies, line luminosity functions, and CIB fluctuations (see B17, B22, and G22 for a complete description of the simulation and for comparisons with observational constraints). It produces realistic LIM mock data constituting our input sky used to evaluate the component separation method efficiency.

In Sect. 2.1, we briefly summarize the important SIDES characteristics for our use. We further review the map and cube-making process in Sect. 2.2.

2.1. The SIDES catalogue

SIDES produces a catalogue of clustered galaxies with their sky coordinates and redshift (up to $z = 7$), together with a large set of physical parameters, as for example their star formation rate (SFR), stellar mass (M_*), or flux in several instrument bandpasses. All relations used to get these physical parameters are derived empirically from observational trends and include an appropriate scatter. First, from a dark matter (DM) simulation, SIDES produces a galaxy light cone by linking the stellar mass M_* to the DM halos' mass, using abundance matching. Two versions exist: one using the $1.4 \times 1.4\,\mathrm{deg}^2$ DM light cone from the Bolshoi-*Planck* simulation (B17, B22) and one using the $117\,\mathrm{deg}^2$ DM light cone from the Uchuu simulation (Ishiyama et al. 2021, G22). Next, galaxies are randomly assigned to star-forming or passive galaxies, according to the observed evolution of the star forming galaxy fraction by Davidzon et al. (2017). Passive-galaxy contributions are then neglected because they are reported to be very faint in the far-infrared/sub-millimetre (e.g., Whitaker et al. 2021). Star-forming galaxies are classified as main sequence (MS) or star-bursting (SB) galaxies according to a redshift-evolving fraction described in Béthermin et al. (2012). The SFR of each galaxy is derived using its redshift, M_*, and type (MS or SB).

Then, a SED is attributed to each galaxy based on its type, total infrared luminosity L_{IR} and mean radiation field intensity $\langle U \rangle$. L_{IR} is derived from the direct use of the Kennicutt (1998) conversion factor and SFR. As for the mean radiation field $\langle U \rangle$, it follows a redshift evolution described in Béthermin et al. (2017) and reflect the evolution of the dust temperature with redshift.

For CO, the luminosity of the $J = 1$ transition is first computed using the $L_{\mathrm{IR}} - L_{\mathrm{CO}(1-0)}$ correlation from Sargent et al. (2014). For SB galaxies, an offset of -0.46 dex is added to L_{IR}. Then, for the rest of the transitions, SLED templates from Bournaud et al. (2015) for MS and from Birkin et al. (2021) for SB are attributed. For MS, contributions to the final SLED of clumpy and diffuse SLEDs are defined by the relation between $\langle \mathrm{U} \rangle$ and CO(5-4)/CO(2-1) flux ratio from Daddi et al. (2015).

The two [CI] fine-structure lines are derived from empirical relations between line luminosity ratios. First the relation between [CI](1-0)/L_{IR} and CO(4-3)/L_{IR} is calibrated to derive $L_{[\mathrm{CI}](1-0)}$. Then, the relation between [CI](2-1)/[CI](1-0) and CO(7-6)/CO(3-4) is calibrated to derive $L_{[\mathrm{CI}](2-1)}$. The calibration of these relations is discussed in B22.

Finally, two models are implemented to compute the [CII] luminosity, following two $L_{[\mathrm{CII}]}$-SFR relations. The first one is the empirical relation calibrated in the local universe from De Looze et al. (2014), hereafter DL14. The second one is from Lagache et al. (2018), hereafter L18, obtained via a semi-analytical model, and containing a weak redshift dependence. Although the exact $L_{[\mathrm{CII}]}$-SFR relation remains debated at high redshifts, the recent observational results of the ALPINE survey (Schaerer et al. 2020) shows that high-redshift galaxies tend to follow the local $L_{[\mathrm{CII}]}$-SFR relation. For this reason, we take [CII] following the DL14 relation as our fiducial model.

2.2. Cube making

From the SIDES catalogue, we create spatial-spectral data cubes for each component: continuum (CIB), CO, [CI] and [CII]. The adopted pixel size is 5 arcseconds and the frequency axis ranges from 125 to 305 GHz with an absolute spectral resolution of 1 GHz. Voxels (the spatial-spectral elements of a cube) are populated with the corresponding intensities. Lastly, the beam smearing is applied to each frequency channel by convolution in Fourier space with a Gaussian beam kernel and periodic boundary conditions. The Gaussian beam profile in one frequency channel is

$$G_{\mathrm{beam}} = \frac{1}{2\pi\sigma^2} \mathrm{e}^{\frac{-r^2}{2\sigma^2}}, \tag{1}$$

where σ is the standard deviation of the Gaussian beam profile linked to the frequency ν of the channel by

$$\sigma = \frac{FWHM}{2\sqrt{2\ln(2)}} = \frac{1}{2\sqrt{2\ln(2)}} \frac{1.22c}{\nu D}, \tag{2}$$

with FWHM is the full with at half maximum of the instrument beam, c is the celerity of light and $D = 12$ m is the diameter of the APEX telescope where CONCERTO is installed. All the components, CIB, CO, [CI], and [CII],, are summed together to create a foreground-contaminated cube. This cube represents the mock observation of the sky as observed by CONCERTO. The aim of the deconfusion is to retrieve the [CII] signal embedded in the foreground-contaminated data. We deal with the two types of foreground one by one, starting with the continuum foreground.

3. Continuum removal

The CIB consists of a smooth, frequency-coherent baseline in the mock electromagnetic spectra, resulting from the sum of dust

continuum emission from galaxies along the line of sight (*los*) plus a part of continuum emission from the neighbouring *los* due to the beam smearing. The CIB has strong angular fluctuations that are the dominant contribution to the measured power spectrum (in particular, see Fig. 15 of B22). In this section we compare the efficiency of two different methods for continuum removal: the PCA and the asymmetric re-weighted penalized least squares (arPLS). The analysis was conducted on frequencies ranging from 125 to 305 GHz. For the [C_{II}] signal, as the SIDES simulation only covers the redshift range $z \leq 7$, the analysis was done only for $\nu > 237$ GHz. We first assessed the continuum removal on SIDES. SIDES contains both the CIB (mean continuum level) and its spectral fluctuations, which are about 10–20% of the CIB. But for CONCERTO, even if in theory it would be sensitive to the CIB, the data processing removes the mean levels. We thus also performed the continuum removal after subtracting the CIB at each frequency of the SIDES simulation (i.e. removing the mean level).

3.1. Principal component analysis

Principal component analysis is a non-parametric method of dimensionality reduction, based on constructing linear combinations of variables in the dataset. Geometrically, these new variables, the so-called principal components, represent the directions of maximum variance in the dataset. The principal components do not have a direct physical interpretation. They are only linked to the variance, not to an underlying physical origin. Most of the variance in the dataset is described by the first principal components, ranked by the amount of variance they carry. Dimensionality reduction corresponds to discarding principal components of lower significance, when projecting the dataset onto this new basis. As mentioned above, the CIB produces the strongest fluctuations at all frequencies in CONCERTO. It is therefore likely that its contamination is condensed in the first few principal components and can be accurately removed. We use for this work the PCA implementation from the package scikit-learn (Pedregosa et al. 2011). We create a matrix, X, that contains the collection of spectra from the foreground-contaminated mock cube of Sect. 2.2. The dimensions of X are $N_\nu \times N_\mathrm{p}$, where N_ν is the number of frequency channels and N_p is the number of pixels. For clarity, we use Greek symbols for the frequency indices and a Latin symbol for the spatial indices (p for pixel). The covariance matrix, C, of the spectral data reads

$$C_{\nu\nu'} = \tilde{X}_{\nu\mathrm{p}}\tilde{X}_{\mathrm{p}\nu'}, \tag{3}$$

where $\tilde{X}$ is defined as

$$\tilde{X}_{\nu\mathrm{p}} = X_{\nu\mathrm{p}} - \mu_\mathrm{p}, \tag{4}$$

and μ_p is the average in the pixel p over all frequencies. The covariance matrix, C, is square, symmetric, and has dimension N_ν. Next, we computed the eigenvector matrix, E, and the block matrix, λ, that contains the eigenvalues from the diagonalization of C, such that

$$C_{\nu\nu'} = E^{-1}_{\nu\alpha}\lambda_{\alpha\beta}E_{\beta\nu'}. \tag{5}$$

Eigenvectors and eigenvalues are sorted according to the power enclosed in each principal component. We select the k ($k < N_\nu$) first principal components and build the block matrix λ' with their corresponding eigenvalues. This allows the derivation of what is usually called the projection templates:

$$T_{\mathrm{p}\nu} = \tilde{X}^T_{\nu'\mathrm{p}}E^{-1}_{\nu'\alpha}\frac{1}{\sqrt{\lambda'_{\alpha\nu}}}, \tag{6}$$

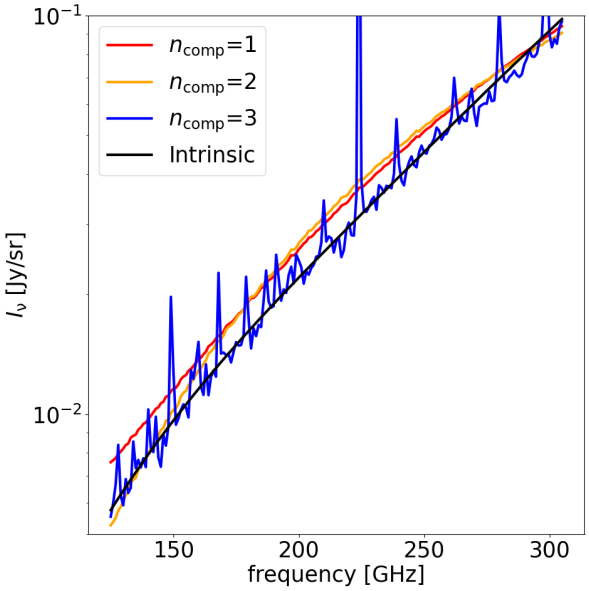

Fig. 1. Spectrum of CIB (continuum) emission in one pixel as a function of frequency (black line). The reconstructed spectra obtained with PCA using the first (red), the first two (orange), and the first three (blue) principal components are over-plotted.

and the continuum is therefore estimated as

$$\tilde{X}^{\mathrm{continuum}}_{\nu\mathrm{p}} = E^T_{\alpha\nu}\sqrt{\lambda}_{\alpha\beta}\,T^T_{\mathrm{p}\beta} + \mu_\mathrm{p}. \tag{7}$$

Eventually, the line spectrum is given by

$$X^{\mathrm{lines}}_{\nu\mathrm{p}} = X_{\nu\mathrm{p}} - \tilde{X}^{\mathrm{continuum}}_{\nu\mathrm{p}}. \tag{8}$$

Figure 1 shows a representative example of reconstructed continuum spectra with PCA. When using the first two components, the continuum is not well reconstructed, meaning that the continuum features are not totally contained in the first two principal components. When adding the third principal component, the reconstructed spectrum does get closer to the intrinsic continuum spectrum but line emission starts to show up. Hence, it is not possible to reconstruct more accurately the continuum by adding more principal components.

Figure 2 shows the effect of the reconstructed continuum subtraction on the continuum power spectrum amplitude, using the first two principal components. Around $z = 5.2$, the residual continuum is 29 times lower than our fiducial [C_{II}] model and 7 times lower than the L18 [C_{II}] model. But the higher the redshift, the lower the contrast. At $z = 7$, the residual continuum is at a similar level than the fiducial [C_{II}] model and 10 times higher than the L18 model.

When removing the CIB (i.e. the mean level of each frequency channel of the cube), the lines contaminate the continuum estimate starting from the third principal component, as previously. However, removing the CIB decreases the continuum variance, leading to a larger residual continuum power, which is a factor of 11 above the lines signal for $\nu > 237$ GHz. Hence, PCA does not have the cleaning efficiency required up to redshift 7, and we turned to another method often quoted in the literature that fits the continuum with slowly varying functions.

3.2. Advanced baseline estimator: arPLS

We test in this section the advanced baseline estimator named arPLS, introduced by Baek et al. (2015) and based a Whittaker-smoothing-based algorithm. Originally developed to serve in a

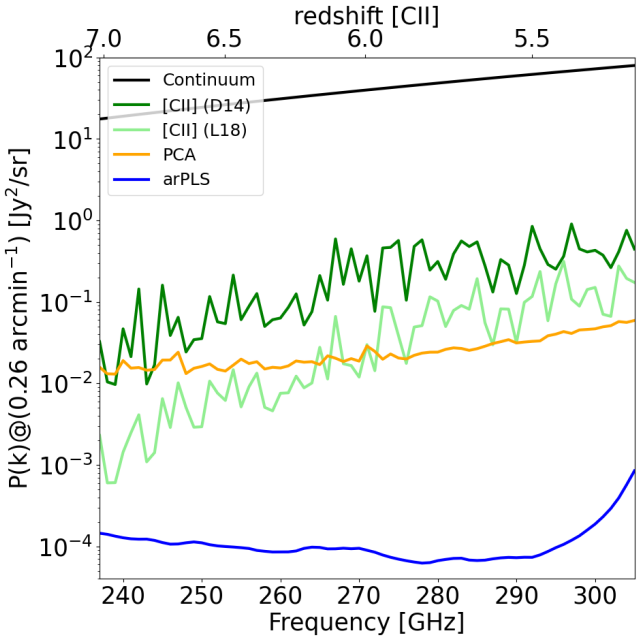

Fig. 2. APS amplitude as a function of frequency for the continuum (solid black line), the [C_{II}] generated with the DL14 or L18 SFR-$L_{[C_{II}]}$ relations (green and light green lines, respectively), the residual continuum obtained with PCA (Sect. 3.1) using the first two principal components (orange line), and the residual continuum obtained with arPLS (Sect. 3.2, blue line). The power spectrum amplitude is averaged at $k = 0.26\,\mathrm{arcmin}^{-1}$ in frequency channels of 1 GHz width.

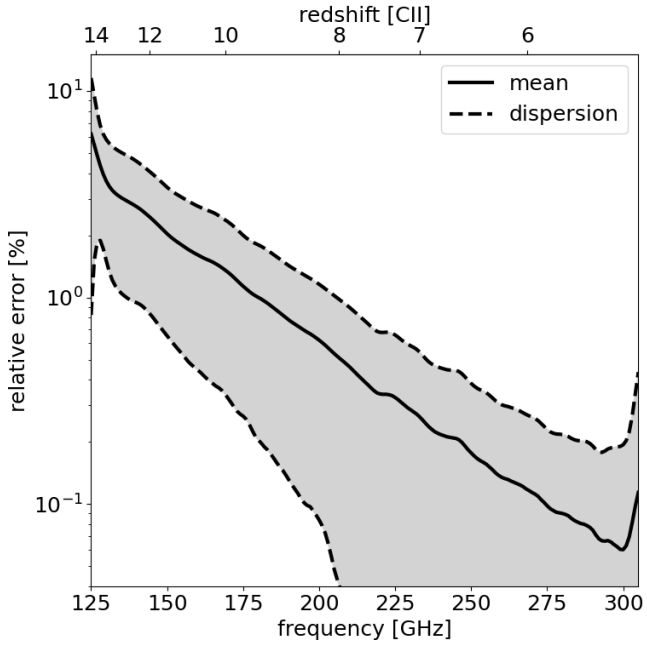

Fig. 3. Relative error of arPLS to the 'true' continuuum as a function of frequency. The mean is taken over the spectra of all the los of the mock cube.

context of Raman spectroscopy, this method was applied to time series in an astronomical context by Zeng et al. (2021). Here, we adapted the use of arPLS to spectral vectors. This estimator makes use of the frequency-coherent feature of CIB. It adjusts a slowly varying function to the smooth component of a (electromagnetic) spectrum under the assumption that it varies slowly with frequency. This smooth component (e.g., the black smooth baseline in the spectrum shown in Fig. 1) is the spectrum of CIB, the sum of all dust continuum emission. Moreover, as the frequency-coherent part of the spectrum is fitted, all other frequency-coherent contamination present as Galactic dust and the CMB are fitted simultaneously. Galaxies are different for each *los* of the mock cube and can drastically change the slope of the baseline so we need to run the baseline estimator on a pixel-by-pixel basis. The smoothed function is fitted to the data by iteratively minimizing a penalized least squares function:

$$S(b) = (y - b)^T W(y - b) + Rb^T D^T Db, \tag{9}$$

where y is the spectrum to be fitted, b the smooth baseline to be found, W the block weight matrix with the weights w for each data point, and D the difference matrix. It is a quasi blind method as the only parameter to be set is the regularization parameter R controlling the trade-off between the agreement of b to the data and its smoothness. In this work, we set it to 1. To find the line peak regions, weights are asymmetrically given to data points:

$$w_i = \begin{cases} f(y_i - b_i, m, \sigma),\ y_i \geq b_i \\ 1,\ y_i < b_i, \end{cases} \tag{10}$$

with

$$f(d^-, m, \sigma) = \frac{1}{1 + \exp(2(d - (-m + 2\sigma))/\sigma)}, \tag{11}$$

where $d^- = y^- - b$ represents the distance between the current estimate of b and the data points that are positioned exclusively below it. The m and σ are the mean and spread y^- of d.

This weighting scheme is designed to be robust to reasonable noise level blurring the baseline. Here, we first want to test the ability of this estimator to retrieve the baseline of the simulated LIM electromagnetic spectra along the CONCERTO frequency band, without any instrument noise.

Figure 3 shows the relative error between the continuum fitted with arPLS and the intrinsic continuum (from the CIB cube). While the baseline is slightly overestimated at all frequencies, the mean error stays below 0.3% for $\nu \geq 237$ GHz ($z_{[C_{II}]} \leq 7$).

Figure 2 shows that the power amplitude of continuum is brought below the fiducial [C_{II}] by a factor of >72 and below the L18 [C_{II}] by a factor >4, for $z = 5.-72$. This allows a clear detection of the [C_{II}] power even in the fainter [C_{II}] case of L18.

For $\nu \leq 237$ GHz, the mean relative error and its dispersion become larger as there are more and more CO lines. Around 130 GHz, at the end of the frequency band, the accuracy decreases even more since there are fewer neighbouring frequency channels to correlate with. Despite this, the fluctuations of the residual continuum are about 27 times lower than the CO power amplitude and permit a further investigation of the CO power spectrum at those frequencies (below 237 GHz).

Finally, we tested arPLS after CIB removal. In that case, the residual power continuum has the same amplitude as before (blue line in Fig. 2) but it is noisier, with a standard deviation of 10 times larger. In any case, the residual contamination is still lower than the L18 [C_{II}] power spectrum above 237 GHz.

We conclude that this method is appropriate for removing the continuum emission from galaxies for CONCERTO, incidentally better than a standard PCA, and is worth being investigated with noise realizations. In what follows, we detail the interloper component separation method.

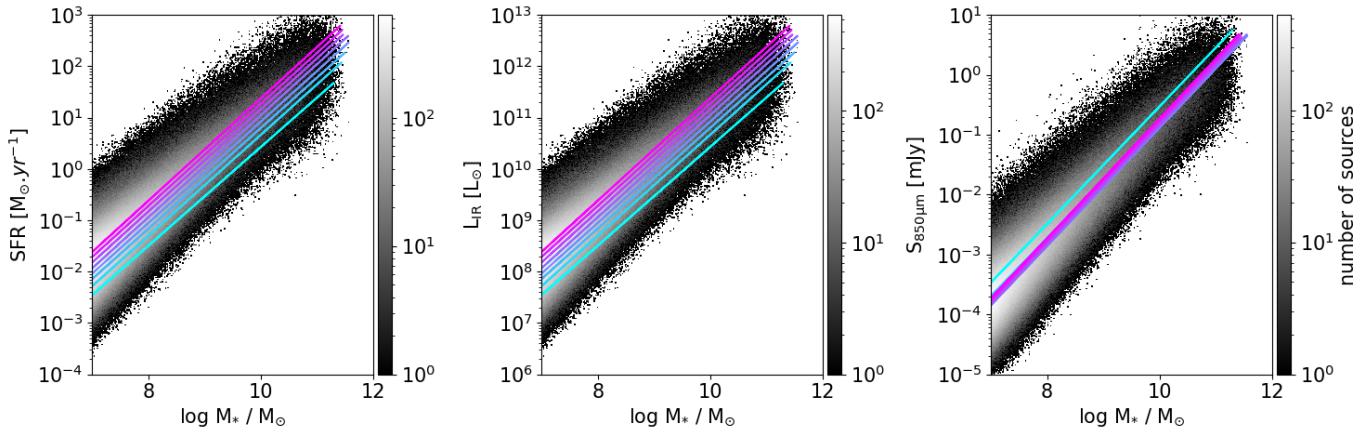

Fig. 4. Stellar mass distribution of sources as a function of different CO proxies available in SIDES. The stellar mass distribution of SIDES sources is shown in grey scale as a function of different proxies up to $z < 3.5$ (from left to right): $M_* - SFR$, $M_* - L_{\rm IR}$, and $M_* - S_{850}$. The best fits for each relation in each redshift bin (given in Table 1) are over plotted, from pink for the first bin ($0 < z < 0.35$) to cyan for the last bin ($2.75 < z < 3.5$).

4. Removing the contamination of line interlopers

Once the continuum power amplitude is successfully lowered, the main source of contamination becomes the line interlopers, that is to say, the emission lines redshifted in the same frequency band as [CII]. Dominant interlopers are the CO rotational ladder and two [CI] fine structure lines from foreground galaxies at $z < 3$. The [OI] 145 µm, [NII] 121.9 µm and NII 205.2 µm lines are also present in the CONCERTO frequency range but their intensities are negligible with respect to the [CII] intensity (Silva et al. 2015). Consequently, only CO and [CI] interlopers are considered.

The LIM survey exploits the confusion regime from the far-IR to (sub-)millimetre, where the beam smearing is significant. Because of the confusion regime and together with the expected instrument noise level for CONCERTO (CONCERTO Collaboration 2020) we do not expect to detect lines from individual foreground galaxies. It becomes thus impossible to blindly mask the brightest pixels using a fix intensity threshold. Therefore, to select interloper-dominated voxels, we need to rely on external catalogues of the COSMOS field (Ilbert et al. 2013; Laigle et al. 2016; Nayyeri et al. 2017; Weaver et al. 2022) providing accurate positions and redshifts of the foreground sources. In this section we present the masking strategy based on the ancillary catalogue of detected sources in the COSMOS field. We first justify our choice of using an external catalogue in stellar mass M_* and its use as a CO proxy to select foreground sources. We then detail the masks we build from this ancillary catalogue.

4.1. Stellar mass as a CO power proxy

For typical redshifts of foreground sources in COSMOS ($z \sim 1-2$), the selection wavelengths of the COSMOS catalogues trace the stellar emission, that traces the stellar mass, rather than the ultraviolet, that traces the unobscured SFR. Therefore, the stellar mass is usually more complete in those catalogues than the SFR, from which moreover the obscured contribution is often missing.

Figure 4 shows the distribution in M_* for galaxies $z \leq 3.5$ as a function of other quantities that can also serve as CO proxy available in SIDES: SFR, $L_{\rm IR}$ and flux at 850 µm S_{850}. From these distributions, we obtain the best fit relation for objects in each redshift bin of the completeness relation, taken from Laigle et al. (2016, hereafter L16) and reported in Table 1. The completeness relation of the COSMOS catalogues describes, as a function of redshift, the stellar mass down to which 90 % of sources are detected. Then from the best fits, we get equivalent depths to the stellar mass depths $M_*^{\rm L16}$. Results are listed in Table 1. One should reach these depths to get a catalogue as complete in these proxies as in stellar mass. In comparison, $L_{\rm IR}$ values inferred from the Multi-Band Imaging Photometer (MIPS)/*Spitzer* at 24 µm (Le Floc'h et al. 2005; Bavouzet et al. 2008) are $1 \times 10^{10}\,L_\odot$ at $z = 0.3$, $3 \times 10^{10}\,L_\odot$ at $z = 0.6$, and $1 \times 10^{11}\,L_\odot$ at $z = 1$. It is one order of magnitude above the $L_{\rm IR}$ depth needed. About SFR, no survey has such a completeness in both ultraviolet and IR to compute the total SFR. Geach et al. (2016) report the Submillimetre Common-User Bolometer Array 2 (SCUBA-2) 850 µm confusion limit to be 0.8 mJy beam^{-1} with a completeness below 10% at this flux. Completeness is close to 100% above 5 mJy beam^{-1}, which is largely above (between 1 and 3 orders of magnitude) the S_{850} requested depth.

To further motivate this masking strategy using SIDES, we look at the contribution of each stellar mass population to the CO power spectrum. Especially, we want to make sure that (1) most of the CO power is produced by sources above the L16 completeness relation, so that they can actually be detected and masked in real data, and (2) that CO power is produced by a reasonable amount of sources, so that masking them would not cost too much in survey area.

Figure 5 shows the cumulative contribution to the CO power amplitude per increasing logarithmic step of stellar mass at $z_{\rm [CII]} = 5.2$. By looking at the cumulative contribution rather than the non-cumulative contribution, we take into account for each stellar mass population (i.e. at each step) both its auto-correlation and its cross-correlation with the less massive populations. The cross-correlation contribution is actually important: the non-cumulative contributions (i.e. the auto-correlations alone) of all stellar mass populations represent 60% of the total CO power. The remaining 40% is due to cross-correlations between the different stellar masses. Figure 5 shows that the most massive galaxies, by their auto and cross correlations, contribute the most to the CO power. This is also the case at higher redshifts: at $z_{\rm [CII]} = 6.5$ and $z_{\rm [CII]} = 7$. The correspondence between [CII] redshifts and interloper redshifts is given in Table 2.

For the three redshifts, galaxies more massive than $\geq 10^{8.1}\,M_\odot$ represent only 22–25% of the total object number but produce most of the CO interloping power (99.7 %). On the other hand, 75–77% of objects less massive than $10^{8.1}\,M_\odot$ are responsible for

Table 1. Equivalent depth in SFR, $L_{\rm IR}$, and $S_{850\,\mu\rm m}$ to the $M_*^{\rm L16}$ depth of the COSMOS completeness relation, obtained with the best fits from Fig. 4.

Redshift	$\log M_*^{\rm L16}/M_\odot$	SFR $[M_\odot.{\rm yr}^{-1}]$	$L_{\rm IR}$ $[L_\odot]$	$S_{850\,\mu\rm m}$ [mJy]
0-0.35	8.1	$(4.0 \pm 1.3) \times 10^{-2}$	$(4.0 \pm 1.3) \times 10^{8}$	$(4.1 \pm 1.6) \times 10^{-3}$
0.35–0.65	8.7	$(2.5 \pm 0.8) \times 10^{-1}$	$(2.5 \pm 0.8) \times 10^{9}$	$(9.1 \pm 3.2) \times 10^{-3}$
0.65–0.95	9.1	$(8.3 \pm 2.7) \times 10^{-1}$	$(8.3 \pm 2.7) \times 10^{9}$	$(1.8 \pm 0.6) \times 10^{-2}$
0.95–1.3	9.3	1.7 ± 0.6	$(1.7 \pm 0.6) \times 10^{10}$	$(2.7 \pm 0.9) \times 10^{-2}$
1.3–1.75	9.7	5.7 ± 1.9	$(5.7 \pm 1.9) \times 10^{10}$	$(6.9 \pm 2.4) \times 10^{-2}$
1.75–2.25	9.9	$(1.2 \pm 0.4) \times 10^{1}$	$(1.2 \pm 0.4) \times 10^{11}$	$(1.2 \pm 0.4) \times 10^{-1}$
2.25–2.75	10.0	$(1.9 \pm 0.6) \times 10^{1}$	$(1.9 \pm 0.6) \times 10^{11}$	$(1.7 \pm 0.6) \times 10^{-1}$
2.75–3.5	10.1	$(3.0 \pm 1.0) \times 10^{1}$	$(3.0 \pm 1.0) \times 10^{11}$	$(2.3 \pm 0.8) \times 10^{-1}$

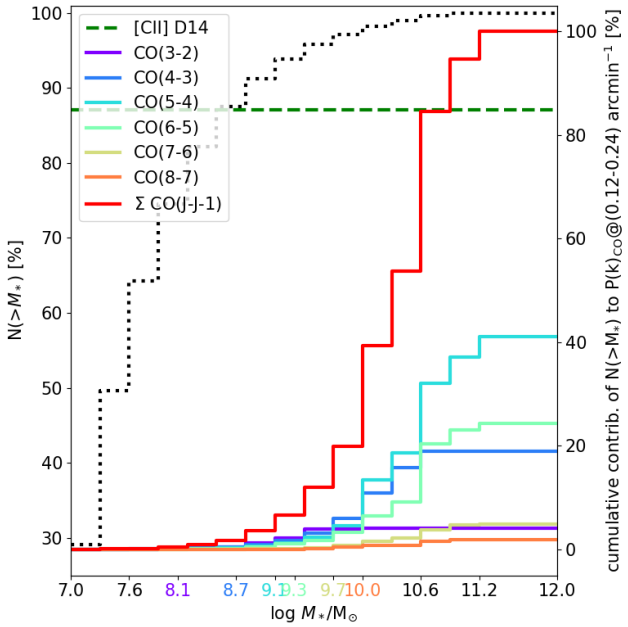

Fig. 5. Contribution to the total CO APS at $z = 5.2$ of each stellar mass bin and CO transition. The left axis and the dotted black curve show the cumulative number count of CO emitters per bin of stellar mass (in percent) redshifted in the frequency channel 305 ± 0.5 GHz ($z_{[\rm C_II]} = 5.2$). The right axis shows the cumulative contribution (in percent) to the total CO power amplitude. Coloured curves correspond to each CO transition, and the red curve corresponds to the sum of all CO transitions. The 90% stellar mass completeness in COSMOS2015 for each CO line is indicated with the same colour code on the x-axis. For comparison, the [$\rm C_{II}$] level (in percent) is represented with the dashed green line. Power amplitudes are averaged for $0.12 < k < 0.24\,{\rm arcmin}^{-1}$.

Table 2. Emission redshift for the lines at the three observed frequencies.

Line	$\nu_{\rm ref}$ (GHz)	z(305 GHz)	z(253 GHz)	z(237 GHz)
CO(3-2)	345.81	0.13	0.37	0.46
CO(4-3)	461.08	0.51	0.82	0.95
CO(5-4)	576.36	0.89	1.28	1.43
CO(6-5)	691.63	1.27	1.73	1.92
CO(7-6)	806.9	1.65	2.19	2.4
CO(8-7)	922.17	2.02	2.64	2.89
[$\rm C_I$] (1-0)	492.16	0.61	0.95	1.08
[$\rm C_I$] (2-1)	809.34	1.65	2.20	2.41
[$\rm C_{II}$]	1900.54	5.23	6.51	7.02

Table 3. Masked intensity per source and survey area loss at the three observed frequencies, corresponding to $z_{[\rm C_II]} = 5.2$, 6.5, and 7 for the masks built with different r_σ.

r_σ	Masked intensity per source	Masked fraction [%] at 305 GHz	Masked fraction [%] at 253 GHz	Masked fraction [%] at 237 GHz
1.5	95%	7	11	14
2.0	99%	11	15	22
2.5	>99%	16	22	29
3.0	>99%	22	29	38

less than 0.5% of the total CO power. For example, we can highlight that the galaxy population with $10^{10.6}\,M_\odot < M_* \leq 10^{10.9}\,M_\odot$ is responsible for 24–31% of the CO power while they only represent 0.5–0.8% of all the sources. Hence, the massive sources above $10^{8.1}\,M_\odot$, despite being the less numerous, strongly dominate the contribution to the CO power and belong to the massive class of known objects. This makes masking an efficient method for decreasing the CO contamination by removing a few sources without catastrophically masking the survey area. Hence, the stellar mass M_* is a reliable CO proxy for the masking purpose.

To make a realistic masking simulation, we masked only sources in the SIDES catalogue that are above the COSMOS completeness relation. We kept the completeness taken from L16, which is deeper than that from Ilbert et al. (2013) or Nayyeri et al. (2017). Although the completeness reported in Weaver et al. (2022) is deeper, it only corresponds to a 70% completeness fraction, so the L16 relation represents the best compromise between completeness fraction and depth.

4.2. Effective masking

External catalogues provide the position of the sources that we want to mask. We then needed to define the extent of the region to mask around these positions. We based our choice on a beam width criterion. Voxels were masked around the location of the selected sources (and thus around CO or [$\rm C_I$] expected emission) within a radius, r_σ, expressed in units of the Gaussian beam width parameter, σ (Eq. (2)). This beam varies with the observed frequency and so does the exact masked area. Table 3 lists the fraction of masked intensity per source and the masked fraction of the maps for masks built with different values of r_σ. Figure 6 displays examples of these masks.

Estimating the APS on masked data is not straightforward because the mask induces aliasing. Several estimators have been developed to overcome this difficulty in the context of the CMB

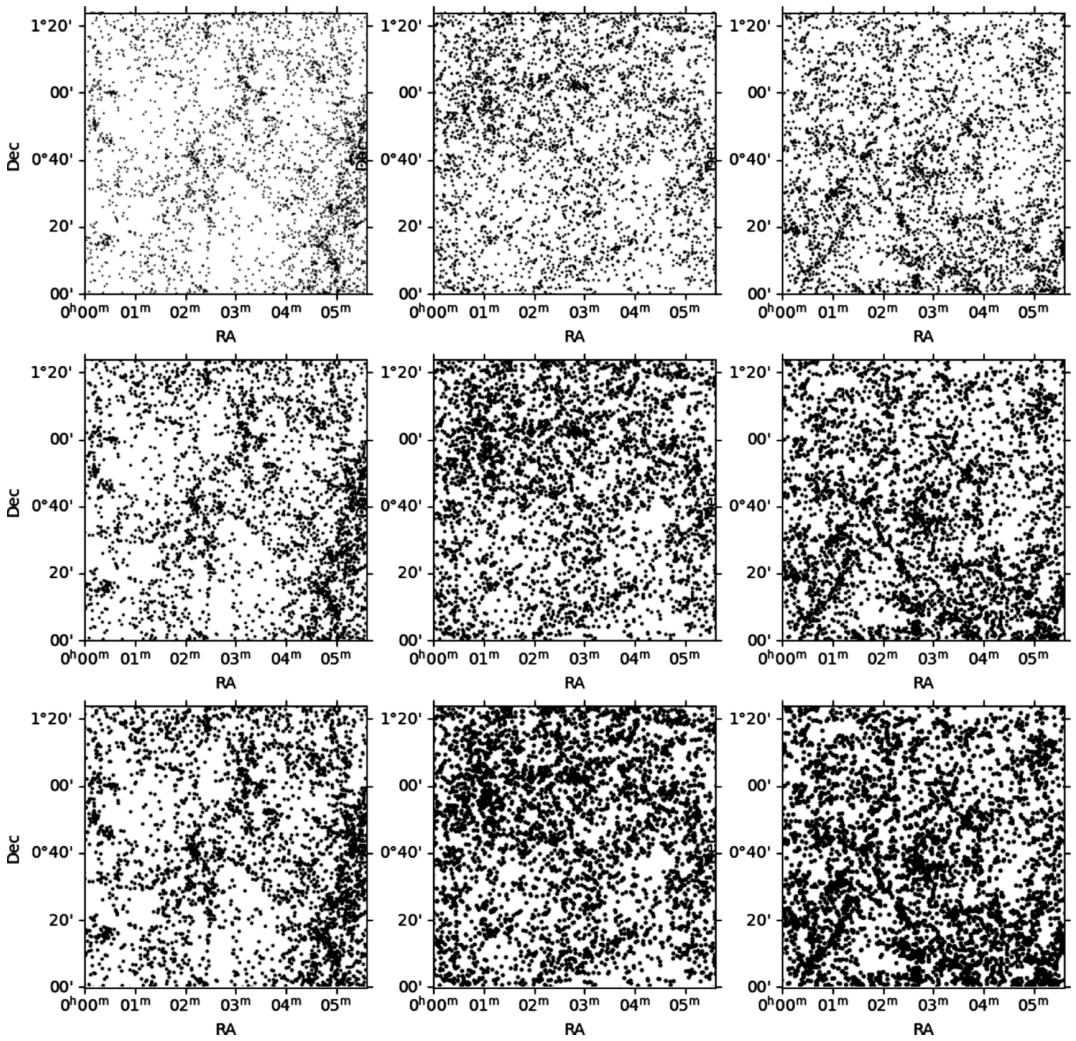

Fig. 6. Masks obtained following the strategy detailed in Sect. 4 using (from left to right) a masking radius, r_σ, of 1.5σ, 2.5σ, and 3σ, respectively. From the top to the bottom, the rows correspond to 305 GHz ($z_{[\mathrm{C_{II}}]} = 5.2$), 253 GHz ($z_{[\mathrm{C_{II}}]} = 6.5$), and 237 GHz ($z_{[\mathrm{C_{II}}]} = 7$), respectively. Masked data appear in black.

measurements. In this work, we used the P of K EstimatoR (POKER; Ponthieu et al. 2011), which was adapted to the case of sky patches of a few square degrees, with arbitrary high angular resolution and complex masks Planck Collaboration XVIII (2011), Planck Collaboration XXX (2014). A detailed discussion on the performance and validation of POKER in the context of this work and the derivation of error bars on APS are presented in Appendix A.

5. Interloper separation at $z_{[\mathrm{CII}]} = 5.2$, 6.5, and 7

In this section we assess the efficiency of masking at retrieving the underlying $[\mathrm{C_{II}}]_{(\mathrm{DL14})}$ APS in the interloper-contaminated mock LIM data (see Sect. 4.2). As mentioned above, the correction of the mask impact is taken care of by POKER, so we focus here on the residual contribution of CO and $[\mathrm{C_I}]$ left after masking on the COSMOS field. When using an APS measurement to derive cosmological parameters, one must also account for the FtF variance, namely the overall uncertainty on the relative contribution of CO and $[\mathrm{C_I}]$ to $[\mathrm{C_{II}}]$, in different parts of the sky. This particular point is addressed in Sect. 5.2, and we focus here on the measurement on COSMOS only.

In the following, power amplitudes and power ratios are given for the same k range (as in Table 4), where the clustering

Table 4. Power amplitude (in $\mathrm{Jy^2\,sr^{-1}}$) averaged between $k = 0.12$ and $0.24\,\mathrm{arcmin^{-1}}$ of the different components, before and after the corresponding component separation method is applied for foreground removal.

Frequency	305 GHz	253 GHz	237 GHz
	Intrinsic amplitudes		
$P(k)_{\mathrm{CIB}}$	1.1×10^{2}	3.6×10^{1}	2.4×10^{1}
$P(k)_{\mathrm{CO}}$	6.2×10^{-1}	4.1×10^{-1}	5.1×10^{-1}
$P(k)_{[\mathrm{CI}]}$	2.9×10^{-1}	1.1×10^{-1}	1.7×10^{-1}
$P(k)_{[\mathrm{CII}]}$	5.3×10^{-1}	6.0×10^{-2}	3.4×10^{-2}
	CIB residual amplitudes		
PCA	1.5×10^{-1}	2.8×10^{-2}	2.7×10^{-2}
arPLS	1.1×10^{-3}	1.5×10^{-4}	1.9×10^{-4}
	$r_\sigma = 3.0$ masked amplitudes		
$P(k)_{\mathrm{CO+[C_I]}}$	$(8.7 \pm 2.3) \times 10^{-3}$	$(1.1 \pm 0.4) \times 10^{-2}$	$(1.7 \pm 0.6) \times 10^{-2}$
$P(k)_{\mathrm{CO\ bright}}$	$(9.2 \pm 2.4) \times 10^{-5}$	$(2.0 \pm 0.7) \times 10^{-5}$	$(3.3 \pm 1.2) \times 10^{-5}$
$P(k)_{\mathrm{CO\ faint}}$	$(6.6 \pm 1.7) \times 10^{-3}$	$(7.7 \pm 2.5) \times 10^{-3}$	$(1.2 \pm 0.4) \times 10^{-2}$
$P(k)_{[\mathrm{CI}]\ \mathrm{bright}}$	$(5.3 \pm 1.4) \times 10^{-6}$	$(3.7 \pm 1.2) \times 10^{-6}$	$(3.1 \pm 1.1) \times 10^{-6}$
$P(k)_{[\mathrm{C_I}]\ \mathrm{faint}}$	$(1.9 \pm 0.5) \times 10^{-3}$	$(2.4 \pm 0.8) \times 10^{-3}$	$(3.6 \pm 1.3) \times 10^{-3}$
$P(k)_{[\mathrm{C_{II}}]}$	$(5.3 \pm 1.4) \times 10^{-1}$	$(7.1 \pm 2.3) \times 10^{-2}$	$(3.0 \pm 1.0) \times 10^{-2}$

Notes. Error bars on the CO, $[\mathrm{C_I}]$ and $[\mathrm{C_{II}}]$ amplitudes were computed using the dispersion measured on 54 independent fields (with $r_\sigma = 3.0$). Details are given in Appendix A.

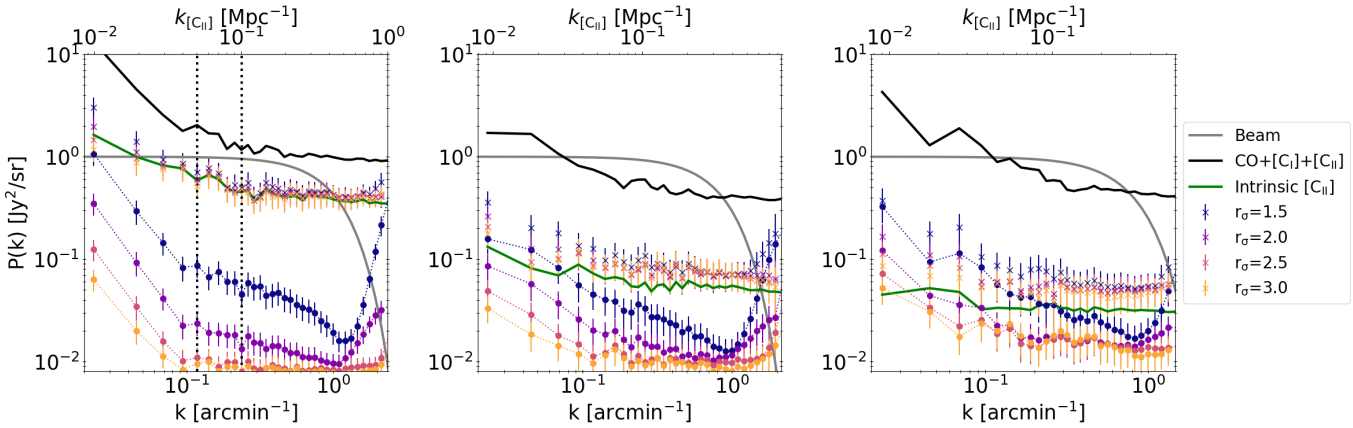

Fig. 7. APS as a function of spatial frequency, k. The solid black line shows the APS of the interloper-contaminated map (i.e. CO+[C$_I$] + [C$_{II}$]) with no mask. The solid green line shows the intrinsic [C$_{II}$] APS (from the unmasked contamination-free [C$_{II}$] map). Coloured crosses show the [C$_{II}$] APS estimates, obtained from masking the interloper-contaminated map with the different masks from $r_\sigma = 1.5$ to 3.0. Coloured points linked by dotted lines show the residual interlopers' APS measured on the masked interloper map (i.e. CO+[C$_I$] after masking) at (from left to right) $z = 5.2$, 6.5, and 7. The beam APS at the corresponding observed frequency is shown in grey. Vertical lines highlight the range $k = 0.12$–0.24 arcmin^{-1}.

dominates. This is because the clustering part of the APS should be easier to interpret than the shot noise, which is dominated by the few strongest sources just below the detection threshold (G22), thus lowering the statistical information.

Results are presented in Fig. 7 at $z_{[C_{II}]} = 5.2$, 6.5, and 7 for different r_σ. With $r_\sigma = 3$, the interloper power amplitude is smaller than 2×10^{-2} Jy^2 sr^{-1} for all three redshifts, that is, it is reduced by a factor of (100 ± 30), (50 ± 16), and (42 ± 15), for each redshift, respectively.

For [C$_{II}$], our fiducial model that follows the local DL14 L$_{[C_{II}]}$–SFR relation predicts a drop in power amplitude by a factor of 16 between $z = 5.2$ and 7. In this framework, masking with $r_\sigma = 2$ at $z_{[C_{II}]} = 5.2$ is enough to obtain a contrast between [C$_{II}$] and residual interlopers of [C$_{II}$]/(residual CO+[C$_I$]) = (30 ± 16) (i.e. 3% of residual contamination). The contrast goes up to (62 ± 32) with $r_\sigma = 3$. At $z_{[C_{II}]} = 6.5$, the [C$_{II}$] APS is overestimated. Even masking with $r_\sigma = 3$ does not achieved a better power ratio than [C$_{II}$]/(residual CO+[C$_I$]) = (5.5 ± 3.6) (i.e. 18% of residual contamination after masking). Finally, at $z_{[C_{II}]} = 7$, the contrast between [C$_{II}$] and residual interlopers is weak due to the low [C$_{II}$] APS, causing [C$_{II}$]/(residual CO+[C$_I$]) = 2.0 ± 1.4 (i.e. 50% of residual contamination after masking) with $r_\sigma = 3$.

These results raise several questions that we discuss in the next section regarding: how well the [C$_{II}$] APS is reconstructed, especially at redshift 5.2; the cause of the residual power spectra at $z_{[C_{II}]} = 6.5$ and $z_{[C_{II}]} = 7$; and whether increasing r_σ to >3 helps in reducing the residual contamination. To answer these questions, we study separately the masking of [C$_{II}$], free from contamination, the masking of the bright (i.e. above the mass threshold) sources and the masking of the faint (i.e. below the mass threshold) sources.

5.1. Interloper masking

The power spectra from the interloper-contaminated maps measured in the previous section are the sum of the power spectra of the three components ([C$_{II}$], the bright interlopers and the faint interlopers), all affected by the mask. In Fig. 8, we show what we obtain when using different masks on each component separately, for the three redshifts. Besides, Table 4 compiles the power amplitudes of each component.

We first focus on the APS behaviour when masking the faint and bright sources, as displayed in the second and third rows of Fig. 8. With $r_\sigma \leq 1.5$, the residual APS shows an overshoot at high k. This is due to the residuals on the edges of the bright sources ('wings') that are still dominating the contribution to the APS, when the mask is too narrow compared to the instrument beam. When masking at $r_\sigma \geq 2$, the residual contribution of bright sources drops below 10^{-3} Jy^2 sr^{-1} and becomes negligible compared to that of the faint sources that stays about 0.8×10^{-2} Jy^2 sr^{-1}. The final power ratio (with $r_\sigma = 3$) between faint and bright sources are (faint CO+[C$_I$])/(bright CO+[C$_I$]) = (91 ± 47), (470 ± 300), and (480 ± 330) at $z_{[C_{II}]} = 5.2$, 6.5, and 7, respectively. Interestingly, the third row shows how the APS of faint sources is also affected by the mask, contrary to what was expected in B22. Using SIDES, we indeed see that $r_\sigma = 3$ masking lowers the faint power amplitude by a factor of (1.9 ± 0.7)–(2.5 ± 0.6) at $z = 7$–5.2. An explanation for this could be that a fraction of faint sources is clustered around bright sources of the same redshift. Yet, if masking removes bright sources efficiently, it has only a moderate effect on the faint sources. Indeed, increasing r_σ from 1.5 to 3.0 increases the masked area by a factor of 3.1, 2.6, and 2.7 at $z = 5.2$, 6.5, and 7, but reduces the faint signal by only an extra 14%, 10%, and 1%, respectively, while in comparison the bright signal is reduced by over a thousand. As a result, enlarging the masked area around bright sources does not effectively remove more contamination from the faint sources and increases rapidly the loss of survey area.

We also applied the masking procedure to [C$_I$] and CO maps separately. Sun et al. (2018, hereafter S18) studied the masking of CO emitters having a magnitude $m_K^{AB} \lesssim 22$, which translates to a stellar mass cut evolving with redshift (see Fig. 7 in Sun et al. 2018). This lowers their CO power amplitude by a factor $\geq$100 at $z = 6.5$ and $k = (3$–$6) \times 10^{-2}$ h Mpc^{-1}, for a 8 % loss of survey area. With SIDES, at comparable redshift and spatial frequencies (corresponding to $k = 0.12$–0.24 arcmin^{-1} at $z = 6.5$), masking with this prescription lowers the CO power amplitude by only a factor of 20 for a 22% loss of survey area. This indicates that the difference in the CO power model between S18 and SIDES increases the power faint sources in our mock data. Faint sources left after masking contributes at least 5 times more in SIDES than in S18. Moreover, taking into account the beam smearing when masking increases significantly the masked fraction of the surveyed area. With our

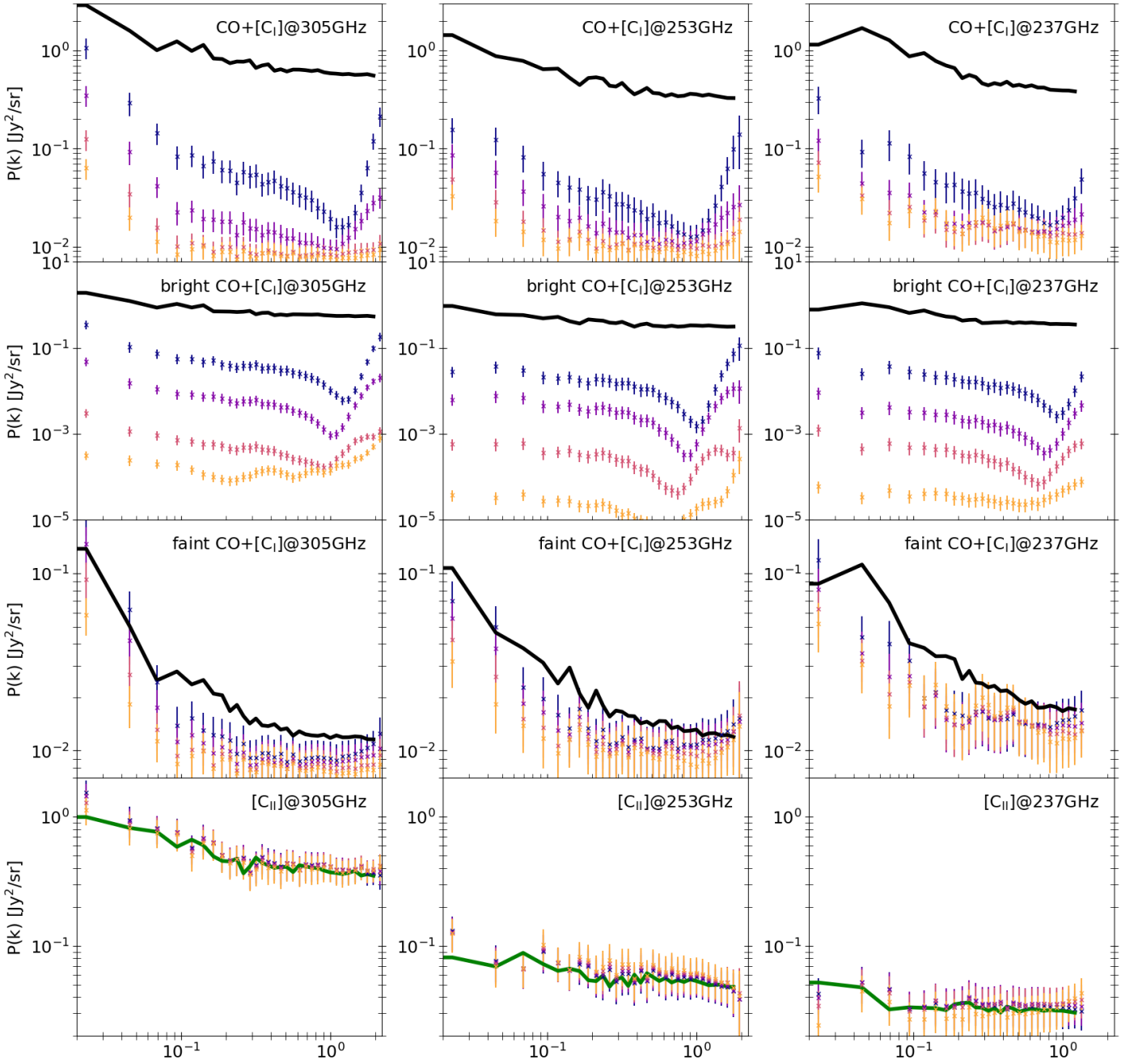

Fig. 8. APS (in $\mathrm{Jy^2\,sr^{-1}}$) as a function of spatial frequency k (in $\mathrm{arcmin^{-1}}$) for the three observed frequencies (from left to right): 305, 253, and 237 GHz. This is shown (from top to bottom) for each component separately: all the interlopers (CO+[C_I]), the bright interlopers (i.e. above the mass threshold), the faint interlopers, and [C_{II}]. The intrinsic power spectra of each (unmasked) component are shown with solid black lines, and with a solid green line for [C_{II}]. The power spectra obtained from the maps that are masked using r_σ = 1.5–3.0 are shown with coloured points (from blue to yellow, respectively).

strategy of masking all detected sources that have a magnitude $m_K^{\mathrm{AB}} < 24$ in Laigle et al. (2016), the CO power amplitude is lowered by a factor of 50, for a 29% loss of survey area (with $r_\sigma = 3$). Lastly, CO is still the dominant interloper after masking: the interlopers' power ratio is (residual CO/residual [C_I]) = (3.2± 2.1), on average for the three redshifts studied ($z_{[C_{II}]} = 5.2$, 6.5 and 7).

With the understanding of how masking affects the interlopers, we show in the last row of Fig. 8 the [C_{II}] APS reconstructed from the masked [C_{II}] maps of SIDES-Bolshoi. The reconstructed [C_{II}] power amplitude is consistent with the intrinsic one within 1σ. However, Appendix A details how the measured [C_{II}] APS (using the mask $r_\sigma = 3$) are biased low, on average by 1%, 7%, and 8% at $z_{[C_{II}]} = 5.2$, 6.5, and 7, respectively, for 54 Uchuu fields.

In conclusion, the masking strategy with $r_\sigma = 3$ is able to nicely reconstruct the [C_{II}] power spectrum from the SIDES-Bolshoi realization at $z_{[C_{II}]} = 5.2$, reaching a power ratio [C_{II}]/(residual CO) = (80 ± 42) and [C_{II}]/(residual [C_I]) = 280 ± 150, with a bias on the [C_{II}] APS of 1% due to the mask. At $z_{[C_{II}]} = 6.5$, the power ratios are [C_{II}]/(residual CO) = (7.8 ± 5.1) and [C_{II}]/(residual [C_I]) = (25 ± 16), with a 7 % bias on

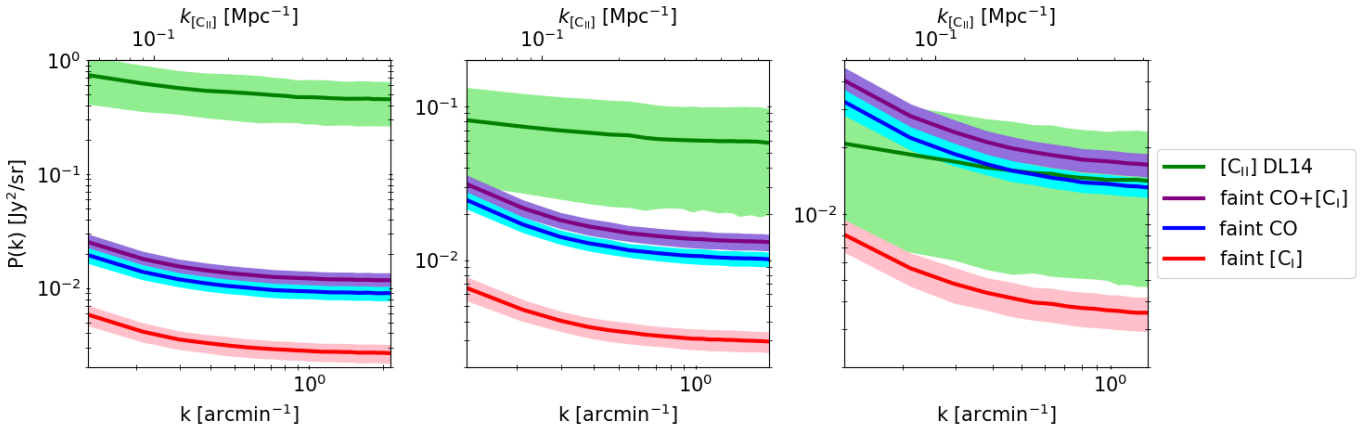

Fig. 9. Mean APS (solid lines) and the 1σ FtF variance interval (coloured area) for a 2 deg^2 field, for [C$_{II}$], faint CO, faint [C$_I$], and the sum of the last two (green, blue, red, and purple, respectively). This is shown, from left to right, for 300–305 GHz centred on $z_{[C_{II}]} = 5.2$, for 260–265 GHz centred on $z_{[C_{II}]} = 6.5$, and for 235–240 GHz centred on $z_{[C_{II}]} = 7$.

the [C$_{II}$] APS due to the mask. At $z_{[C_{II}]} = 7$, power ratios are [C$_{II}$]/(residual CO) = (2.7 ± 1.9) and [C$_{II}$]/(residual [C$_I$]) = (9.4 ± 6.6), with an 8 % bias on the [C$_{II}$] APS. For all redshifts, residual contamination becomes dominated by the faint sources for $r_\sigma \geq 2$ since such sources are moderately affected by the mask.

5.2. Field-to-field variance effects on residuals

The detection of the [C$_{II}$] power spectrum depends on its contrast compared to that of the (masked) faint sources, the latter dominating the residual interloper contamination as shown in the previous section. In addition, these measurements will also be affected by the FtF variance. Indeed, G22 shows that FtF variance cannot be neglected when making a forecast for a LIM experiment, especially for the [C$_{II}$] LIM survey conducted by CONCERTO on a field the size of COSMOS. We explore this here in more details. We use the Uchuu simulation and the variance model from G22 (their Eq. (8)) to determine the faint CO and faint [C$_I$] APS variance. This equation is used to compute a factor f to rescale the variance σ for a $\Omega = 1$ deg^2 field to the $\Omega = 2$ deg^2 field of the SIDES-Bolshoi simulation, assuming that the mean power spectrum μ does not change between the two field sizes:

$$f = \frac{\sigma(\Omega = 2\,\mathrm{deg}^2)}{\sigma(\Omega = 1\,\mathrm{deg}^2)}. \quad (12)$$

The values of f for the clustering term and the shot noise term are similar at the 1% level.

Figure 9 shows the mean APS and its associated 1σ FtF variance interval in a 2 deg^2 field, for [C$_{II}$] and the faint interlopers at $z_{[C_{II}]} = 5.2$, 6.5, and 7. The [C$_{II}$] power amplitude varies by 40% at $z_{[C_{II}]} = 5.2$ and by 60% at $z_{[C_{II}]} = 6.5$ and 7. The faint (CO + [C$_I$]) APS varies by only 12–15% from $z_{[C_{II}]} = 5.2$ to 7. Thus, the variance of faint sources does not have a strong impact on the contrast between residuals and [C$_{II}$] in comparison to the [C$_{II}$] variance. Table 5 lists the level of faint interloper contamination and its ratio to the 1σ value of the [C$_{II}$] FtF variance. As a result, residual contamination is smaller than 20% of the [C$_{II}$] FtF variance up to 287 GHz ($z_{[C_{II}]} = 5.62$) and smaller than 50% of the FtF variance up to 253 GHz ($z_{[C_{II}]} = 6.5$).

Table 5. Ratio (in percent) of faint interloper contamination (i.e. 'unmasked faint CO+[C$_I$]') to the [C$_{II}$] power amplitude when the latter is evaluated at the mean and at the edges of the [C$_{II}$] ± 1σ FtF variance interval.

Frequency band	At $+1\sigma$	At mean	At -1σ	Faint interlopers /1σ
305 GHz	2%	3%	4%	7%
287 GHz	6%	8%	12%	21%
253 GHz	17%	27%	66%	45%
237 GHz	87%	139%	348%	232%

Notes. The last column gives (also in percent) the ratio between the faint interloper contamination and the 1σ value of the [C$_{II}$] FtF variance.

6. Summary and conclusion

We tested a complete foreground deconfusion process for the [C$_{II}$] LIM survey conducted with CONCERTO using realistic mock data of the sub-millimetre sky provided by SIDES-Bolshoi (B17, B22). To build the SIDES mock data, a galaxy catalogue complete down to $10^7\,M_\odot$ and $z = 7$ was first obtained using abundance matching between stellar mass and DM halo mass in a simulated DM light cone. Galaxy properties were then derived from their type (MS or SB), redshift, and stellar mass using empirically calibrated relations with appropriate scatters. Finally, mock cubes for CIB, CO, [C$_I$], and [C$_{II}$] were produced. A foreground-contaminating mock cube was created by adding the separate foreground cubes to the [C$_{II}$] cube.

We first assessed the contamination by CIB (continuum) fluctuations. We tested the ability of PCA and arPLS to subtract the continuum emission for each *los*, by exploiting its (electromagnetic) frequency-coherent distinctive feature. Principal component analysis is not able to separate the continuum from molecular and atomic lines up to $z_{[C_{II}]} = 7$. The arPLS method fits all the frequency-coherent emissions with a precision of 0.3% in the mock spectra and sufficiently removes the continuum contamination for the two [C$_{II}$] models provided by SIDES up to $z = 7$. For $5.2 \leq z \leq 7$, the residual CIB power amplitudes are lower than that of our fiducial DL14 (L18) [C$_{II}$] model by a factor of 72 (4).

For the deconfusion with the line interlopers, we relied on a masking approach that uses an external COSMOS stellar mass

catalogue, which provides accurate sky positions and redshifts for foreground sources to mask interloper-contaminated voxels. We show that M_* is the deepest CO and [C I] proxy available in the COSMOS field. We also investigated its use as a CO proxy. The most contaminating sources ($M_* > 10^{8.1}\,M_\odot$) are the least numerous and are above the L16 completeness relation, allowing us to efficiently remove the interloper contamination with a limited loss of the surveyed area. The spatial extent of the masks is defined with a (frequency-evolving) beam width criterion. To test the masking on our mock data in a realistic way, we masked only sources above the COSMOS (deep) completeness relation described in L16.

For the SIDES-Bolshoi simulation, the power amplitude from bright CO and [C I] sources was lowered by four orders of magnitudes. Masking is also more efficient than expected in B22 as it also removes a fraction of the faint source signal at the same time, thanks to the clustering. The residual power amplitude left after masking is lower than 2×10^{-2} Jy^2 sr^{-1}, rather independent of frequency, and is dominated by the faint sources that are not masked (especially CO sources). The CO power amplitude is lowered by a factor of 50 at $z = 6.5$ with our masking, while it is lowered by a factor greater than 100 with a smaller masking depth in S18, indicating that the signal of sources below the mass threshold is stronger in SIDES than in S18.

In the framework of our model, which contains a strong decrease in the [C II] APS with redshift, we obtained the following results for the [C II] APS measurement.

At $z_{[\mathrm{C\,II}]} = 5.2$, the power ratio is [C II]/(residual CO+[C I]) = 62 ± 32 (2.5% residual contamination) for a 22 % surveyed area loss and 1% of bias due to the mask. At this given redshift, the [C II] power spectrum varies by about 40% due to the FtF variance. So, FtF variance is larger than any effect introduced by the masking or by residual contamination (which is ≤ 4% of FtF variance).

At $z_{[\mathrm{C\,II}]} = 6.5$, the residual contamination after masking is a factor of [C II]/(residual CO+[C I]) = 5.5 ± 3.6 below the [C II] power spectrum amplitude. This residual contamination and the reconstruction by POKER (see Appendix A) led to an overestimate of the [C II] power spectrum by a factor of 1.4 for a 29% loss of survey area. The variance estimate from SIDES-Uchuu gives a residual contamination by the faint sources that should not exceed half of the [C II] FtF variance at this redshift. At $z_{[\mathrm{C\,II}]} = 7$, SIDES predicts a drop in [C II] power spectrum amplitude by a factor of 16 with respect to $z_{[\mathrm{C\,II}]} = 5.2$ and an important FtF variance for a COSMOS-like field. Indeed, the measured residual contamination after masking is only a factor of 2.0 ± 1.4 below the [C II] power spectrum amplitude.

Hence, residual contamination is still important above $z_{[\mathrm{C\,II}]} > 6.5$, despite the fact that most (>98%) of the CO power has been masked. Masking even more sources to bring residues to a fainter level is tricky because the [C II] APS is already slightly affected by the mask (i.e. underestimated by 7–8% between $z = 6.5$ and 7; see Appendix A) and the cost per masked source in a surveyed area is lower frequencies (i.e. higher redshifts). To reach the [C II] signal at these redshifts, other or complementary methods for masking will be needed.

Our analysis makes use of the SIDES simulations (Béthermin et al. 2022), in which the [C II] power spectrum is predicted with a much lower amplitude than in earlier models (e.g., Serra et al. 2016). For the first generation of LIM experiments, such as CONCERTO, and given such a low-amplitude [C II] power spectrum, we expect the instrument noise to dominate. Van Cuyck et al. (in prep.) will introduce the noise generator of SIDES and investigate the effects of instrument white noise on different LIM observables, such as auto-power spectra, cross-power spectra between lines and galaxies, and cross-power spectra between different lines within the LIM dataset. We have shown that our method for removing the continuum emission does not create any systematic residuals, but this will have to be tested with instrument noise. On the other hand, we expect masking to be robust to noise contamination because the interlopers' signal is removed regardless of the noise level. The robustness of cross-correlation measurements between different lines within the LIM dataset in the presence of instrument noise is also worth investigating in the context of component separation. Such a cross correlation could allow for a direct estimate of interlopers' residual power amplitude after masking. The last effect that would need further checks for the masking technique is the inclusion of the catalogues' redshift uncertainties, reported to be $\delta z/(1+z) = 0.025$ below $z \leq 3$ for the COSMOS field (Weaver et al. 2022).

Finally, the CO-line emission, while being a strong foreground for the [C II] survey, is of great importance. The CO part of the CONCERTO survey can be used to probe the CO SLED and the gas content of galaxies at $z \leq 3$, including at cosmic noon ($2 \leq z \leq 3$), when star formation in galaxies statistically reaches its peak. The component separation methods and tools developed in this work for [C II] can be adapted and used to exploit CO surveys.

Acknowledgements. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 788212) and from the Excellence Initiative of Aix-Marseille University-A*Midex, a French "Investissements d'Avenir" programme. M.A. acknowledges support from FONDECYT grant 1211951, CONICYT + PCI + Max Planck Institute for Astronomy MPG 190030, CONICYT + PCI + REDES 190194 and ANID BASAL project FB210003.

References

Atek, H., Richard, J., Jauzac, M., et al. 2015, ApJ, 814, 69
Baek, S.-J., Park, A., Ahn, Y.-J., & Choo, J. 2015, Analyst, 140, 250
Bavouzet, N., Dole, H., Le Floc'h, E., et al. 2008, A&A, 479, 83
Bernal, J. L., & Kovetz, E. D. 2022, A&ARv, 30, 5
Béthermin, M., Daddi, E., Magdis, G., et al. 2012, ApJ, 757, L23
Béthermin, M., Wu, H.-Y., Lagache, G., et al. 2017, A&A, 607, A89
Béthermin, M., Gkogkou, A., Van Cuyck, M., et al. 2022, A&A, 667, A156
Bigot-Sazy, M. A., Dickinson, C., Battye, R. A., et al. 2015, MNRAS, 454, 3240
Birkin, J. E., Weiss, A., Wardlow, J. L., et al. 2021, MNRAS, 501, 3926
Bosman, S. E. I., Davies, F. B., Becker, G. D., et al. 2022, MNRAS, 514, 55
Bournaud, F., Daddi, E., Weiß, A., et al. 2015, A&A, 575, A56
Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2015, MNRAS, 452, 3408
Carilli, C. L., Chluba, J., Decarli, R., et al. 2016, ApJ, 833, 73
Chabrier, G. 2003, PASP, 115, 763
Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Nature, 466, 463
Chang, T. C., Gong, Y., Santos, M., et al. 2015, in Advancing Astrophysics with the Square Kilometre Array (AASKA14), 4
Chapman, E., Zaroubi, S., Abdalla, F. B., et al. 2016, MNRAS, 458, 2928
Cheng, Y.-T., Chang, T.-C., Bock, J., Bradford, C. M., & Cooray, A. 2016, ApJ, 832, 165
Cheng, Y.-T., Chang, T.-C., & Bock, J. J. 2020, ApJ, 901, 142
Chornock, R., Berger, E., Fox, D. B., et al. 2013, ApJ, 774, 26
Chung, D. T. 2023, Phys. Rev. D, 107, 023509
Chung, D. T., Viero, M. P., Church, S. E., et al. 2019, ApJ, 872, 186
Comaschi, P., Yue, B., & Ferrara, A. 2016, MNRAS, 463, 3193
CONCERTO Collaboration (Ade, P., et al.) 2020, A&A, 642, A60
Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., et al. 2016, MNRAS, 457, 3541
Daddi, E., Dannerbauer, H., Liu, D., et al. 2015, A&A, 577, A46
Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, A&A, 605, A70
Dayal, P., Dunlop, J. S., Maio, U., & Ciardi, B. 2013, MNRAS, 434, 1486
De Looze, I., Cormier, D., Lebouteiller, V., et al. 2014, A&A, 568, A62
Drake, A. B., Garel, T., Wisotzki, L., et al. 2017, A&A, 608, A6
Eilers, A.-C., Hennawi, J. F., Davies, F. B., & Oñorbe, J. 2019, ApJ, 881, 23
Ellis, R. S., McLure, R. J., Dunlop, J. S., et al. 2013, ApJ, 763, L7
Finkelstein, S. L., Papovich, C., Ryan, R. E., et al. 2012, ApJ, 758, 93

Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, Phys. Rep., 433, 181
Geach, J. E., Dunlop, J. S., Halpern, M., et al. 2016, MNRAS, 465, 1789
Gkogkou, A., Béthermin, M., Lagache, G., et al. 2023, A&A, 670, A16
Gong, Y., Cooray, A., Silva, M., et al. 2012, ApJ, 745, 49
Gong, Y., Silva, M., Cooray, A., & Santos, M. G. 2014, ApJ, 785, 72
Gong, Y., Chen, X., & Cooray, A. 2020, ApJ, 894, 152
Gullberg, B., De Breuck, C., Vieira, J. D., et al. 2015, MNRAS, 449, 2883
Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55
Ishiyama, T., Prada, F., Klypin, A. A., et al. 2021, MNRAS, 506, 4210
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kovetz, E. D., Viero, M. P., Lidz, A., et al. 2017, ArXiv e-prints [arXiv:1709.09066]
Lagache, G., Puget, J. L., Abergel, A., et al. 2000, in ISO Survey of a Dusty Universe, eds. D. Lemke, M. Stickel, & K. Wilke (Springer), 548, 81
Lagache, G., Cousin, M., & Chatzikos, M. 2018, A&A, 609, A130
Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24
Le Floc'h, E., Papovich, C., Dole, H., et al. 2005, ApJ, 632, 169
Lidz, A., & Taylor, J. 2016, ApJ, 825, 143
Lidz, A., Zahn, O., Furlanetto, S. R., et al. 2009, ApJ, 690, 252
Liu, A., & Tegmark, M. 2012, MNRAS, 419, 3491
Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, Am. Astron. Soc. Meeting Abst., 229, 141.09
Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJ, 763, L20
McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., & Furlanetto, S. R. 2006, ApJ, 653, 815
Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767
Nayyeri, H., Hemmati, S., Mobasher, B., et al. 2017, ApJS, 228, 7
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2015, ApJ, 808, 104
Padmanabhan, H., Breysse, P., Lidz, A., & Switzer, E. R. 2022, MNRAS, 515, 5813
Pallottini, A., Ferrara, A., Bovino, S., et al. 2017, MNRAS, 471, 4128
Pallottini, A., Ferrara, A., Gallerani, S., et al. 2022, MNRAS, 513, 5621
Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012, ApJ, 756, 165
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, J. Mach. Learn. Res., 12, 2825
Pineda, J. L., Langer, W. D., & Goldsmith, P. F. 2014, A&A, 570, A121
Planck Collaboration XVIII. 2011, A&A, 536, A18
Planck Collaboration XXX. 2014, A&A, 571, A30
Planck Collaboration XIII. 2016, A&A, 594, A13
Planck Collaboration Int. XLVII. 2016, A&A, 596, A108
Ponthieu, N., Grain, J., & Lagache, G. 2011, A&A, 535, A90
Pullen, A. R., Breysse, P. C., Oxholm, T., et al. 2023, MNRAS, 521, 6124
Sargent, M. T., Daddi, E., Béthermin, M., et al. 2014, ApJ, 793, 19
Schaerer, D., Ginolfi, M., Béthermin, M., et al. 2020, A&A, 643, A3
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Serra, P., Doré, O., & Lagache, G. 2016, ApJ, 833, 153
Silva, M., Santos, M. G., Cooray, A., & Gong, Y. 2015, ApJ, 806, 209
Sun, G., Moncelsi, L., Viero, M. P., et al. 2018, ApJ, 856, 107
Switzer, E. R., Chang, T. C., Masui, K. W., Pen, U. L., & Voytek, T. C. 2015, ApJ, 815, 51
Switzer, E. R., Anderson, C. J., Pullen, A. R., & Yang, S. 2019, ApJ, 872, 82
Vallini, L., Ferrara, A., Pallottini, A., & Gallerani, S. 2017, MNRAS, 467, 1300
Visbal, E., & Loeb, A. 2010, JCAP, 2010, 016
Wang, F., Davies, F. B., Yang, J., et al. 2020, ApJ, 896, 23
Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, ApJS, 258, 11
Whitaker, K. E., Williams, C. C., Mowla, L., et al. 2021, Nature, 597, 485
Yohana, E., Ma, Y.-Z., Li, D., Chen, X., & Dai, W.-M. 2021, MNRAS, 504, 5231
Yue, B., Ferrara, A., Pallottini, A., Gallerani, S., & Vallini, L. 2015, MNRAS, 450, 3829
Zeng, Q., Chen, X., Li, X., et al. 2021, MNRAS, 500, 2969

Appendix A: Measuring the APS of a masked distribution of point sources with POKER

In this paper we make extensive use of POKER, an algorithm developed to compute the APS of a masked 2D signal on a few square degrees at arbitrary high angular resolution (Ponthieu et al. 2011). POKER was developed to measure the power spectrum of the CIB in the context of experiments such as the infrared astronomical satellite (IRAS) and *Planck* (Planck Collaboration XVIII 2011) for which the CIB could be seen as an ubiquitous diffuse signal and treated as a random Gaussian field. From the direct Fourier transform of the masked data (pseudo-power spectrum), POKER inverts the mask induced aliasing and provides an estimate of the underlying APS that is unbiased on average when the signal is independent of the mask. When the signal of interest is highly non Gaussian such as in this work, with point sources and voids in between, the outcome of the power spectrum derived on masked area must however be reconsidered, and this, actually regardless of the APS estimator.

For pedagogical purpose, we can start with the simplest example: a signal whose APS is constant for all angular modes. This APS can be that of a single point source or that of uniform white noise. In the case of uniform white noise, even a mask as complex as those shown in this work will leave some signal on the map from which the pseudo-power spectrum can be derived. In the case of a point source, if the source is masked, the pseudo-power spectrum is then zero and no measurement is possible. In this work, the signal can be approximated as resulting from a collection of N_s point sources. If the sources had all the same flux, the intrinsic power spectrum would simply be N_s times the uniform APS of single source of flux ϕ. The mask covers a fraction f_{mask} of the field but suppresses a fraction of sources $f_{sources}$ that is different because sources are not uniformly distributed, and loosely speaking, several sources can fall in the same hole. In this case, the pseudo-power spectrum is underestimated by $f_{sources}$ and corrected up by f_{mask} by POKER. Figure A.1 illustrates two different examples where the mask and the source locations are taken directly from a 114 arcmin^2 section of the SIDES-Bolschoi simulation and their flux is fixed to 1: depending on the effective (a priori unknown) clustering of the sources compared to the mask, the reconstructed APS is larger or smaller than the intrinsic [C$_{\rm II}$] APS. In this case, it is therefore impossible by construction to measure an unbiased APS in the presence of a mask, and the bias will depend on the clustering properties of the signal compared to the mask. We can extend this reasoning to the sum of several distributions of point sources, each distribution with a different flux. Each one of them will lead to an APS estimate biased by a different $f_{sources}/f_{mask}$ so that the final APS is some average of these ratios, weighted by the source flux. Last, the convolution by the point spread function adds another feature. Indeed, it gives a spatial extension to the initial point sources, so that their signal can be only partially masked and more degeneracies show up when computing the absolute level of the pseudo-power spectrum. It is therefore essential to test the APS derivations on realistic source and mask simulations to assess the magnitude of these potential biases.

Figure A.2 presents the results of this test, from which we built the error bars given in Sect.5.1. For each of the 54 fields of 2 deg^2 cut in the Uchuu simulation, we built a mask with $r_\sigma = 3$ following Sect. 4 and computed the ratio between the output (masked) and input [C$_{\rm II}$] APS. We then computed the average of this ratio over all the 54 fields and show that it is not biased low by more than 1%, 7%, and 8% on average for the three redshifts ($z_{[\rm CII]}$ = 5.2, 6.5, and 7), up to an angular scale of k_{lim} = 2.14, 1.94, 1.33 arcmin^{-1}, respectively. The bias increases with redshift because the masked area is larger, as it depends on the beam size. Error bars for an APS obtained using a mask at a given frequency are the dispersion of this ratio at this frequency. The obtained errors bars represent 25%, 33%, and 34% of the masked APS at $z_{[\rm CII]} = 5.2$, 6.5, and 7, respectively, and are also used for smaller r_σ masks. Error bars displayed in Fig. A.2 are normalized by the number of realizations. They are constant for all angular scales because they result from an effective overall scaling by the mask to flux ratio (see e.g., Fig. A.1) that varies from one field to another. For reference, the error bars that would come from sampling variance of a diffuse signal are displayed in blue and their amplitude decreases (as expected) with the increasing number of available angular modes at higher k.

At smaller angular scales, there is a systematic small-scale bias. In Fig. A.3 we check via simulations that this bias either damps or boosts the estimate depending on the relative size of the mask and the FWHM. Table A.1 gives the spatial frequency at which this bias between the measured and the expected APS becomes larger than 20% and 50%, in the interloper contaminated maps masked at $r_\sigma = 3.0$. These values are well above the beam cutoff, and for consistency, in all this work, we restrict to angular scales $k < k_{20\%}$.

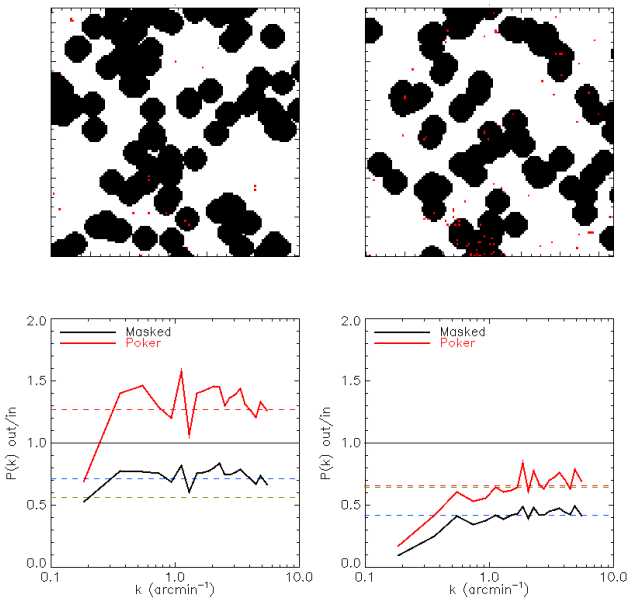

Fig. A.1. Illustration of the POKER correction for two different 114 arcmin^2 fields. *Top row*: CO+[C$_{\rm I}$] masks corresponding to the two fields, as well as the [C$_{\rm II}$] sources (red points). *Bottom row*: APS of the masked [C$_{\rm II}$] source maps, with or without correction of mask aliasing (labelled 'Poker' and 'Masked', respectively). On the left-hand side, the mask covers 44% of the sky patch (dashed blue line) but masks only 29% of the [C$_{\rm II}$] sources (dashed green line). POKER returns an output power spectrum that has the correct shape (flat, up to the sample variance) but that is overestimated by a factor of (1-0.29)/(1-0.44)=1.27. This is the dashed red line, and it matches the average level of the output APS. On the right-hand side, the mask covers only 36% of the sky patch but 58% of the sources. Hence, the output APS returned by POKER is underestimated by (1-0.58)/(1-0.36)=0.65.

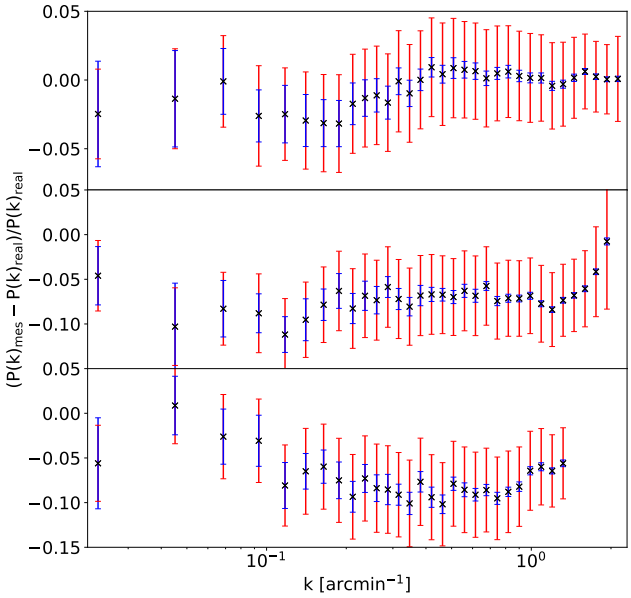

Fig. A.2. Average relative error (black crosses) between the [C$_{II}$] APS measured in the 54 Uchuu sub-fields masked with $r_\sigma = 3$ as a function of spatial frequency, k, at $z = 5.2$ (top), $z = 6.5$ (middle), and $z = 7$ (bottom). The mean errors over the k modes are <1%, 7%, and 8% for each redshift, respectively. The dispersion (red error bars) normalized by the square root of the number of realizations is also represented. The mean sizes of these normalized error bars over the k modes are 3%, 4%, and 5% for each redshift, respectively. For reference, the equivalent normalized error bars for a perfect Gaussian diffuse signal from the same number of Monte Carlo realizations are over-plotted in blue.

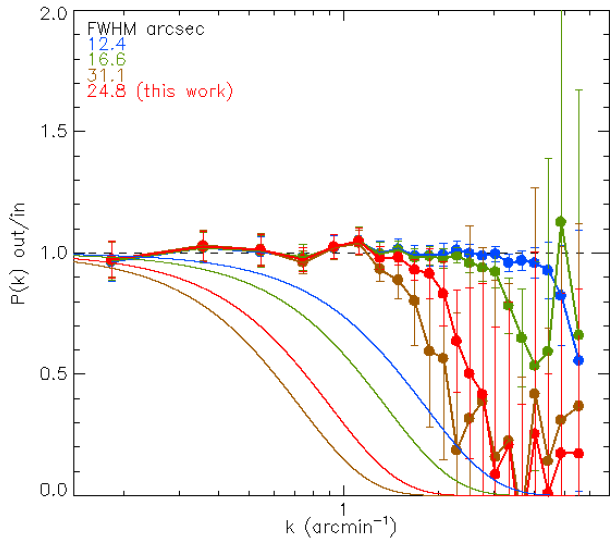

Fig. A.3. Output (masked) over input (intrinsic) [C$_{II}$] APS ratio for various beam sizes but with the same mask. The smaller the beam, the larger the k angular mode up to which the output APS equals the input one. The analytical beam is displayed in solid lines for reference and shows that the effect only appears at high angular scales, where the beam has mostly smeared out any signal.

Table A.1. Spatial frequency (in arcmin^{-1}) at which the pseudo power spectrum becomes biased by 20% and 50%, and the corresponding angular size, r, as a fraction of the FWHM.

frequency	$k_{20\%}$	$r_{20\%}$ [FWHM]	$k_{50\%}$	$r_{50\%}$ [FWHM]
305 GHz	2.14	1.36	2.35	1.24
253 GHz	1.94	1.24	2.14	1.36
237 GHz	1.33	1.71	1.46	1.55