

# The power spectrum of extended [C II] haloes around high redshift galaxies

Meng Zhang,<sup>1,2,3</sup> Andrea Ferrara<sup>3★</sup> and Bin Yue<sup>1,4</sup>

<sup>1</sup>National Astronomical Observatories, Chinese Academy of Sciences, 20A, Datun Road, Chaoyang District, Beijing 100101, China

<sup>2</sup>School of Astronomy and Space Science, University of Chinese Academy of Sciences, No.1 Yanqihu East Rd, Huairou District, Beijing 101408, China

<sup>3</sup>Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

<sup>4</sup>Key Laboratory of Radio Astronomy and Technology, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100101, China

Accepted 2023 September 25. Received 2023 September 12; in original form 2023 April 14

## ABSTRACT

Atacama Large Millimeter/sub-millimeter Array (ALMA) observations have detected extended ( $\simeq 10$  kpc) [C II] haloes around high-redshift ( $z \gtrsim 5$ ) star-forming galaxies. If such extended structures are common, they may have an impact on the line intensity mapping (LIM) signal. We compute the LIM power spectrum including both the central galaxy and the [C II] halo, and study the detectability of such signal in an ALMA LIM survey. We model the central galaxy and the [C II] halo brightness with a Sérsic + exponential profile. The model has two free parameters: the effective radius ratio  $f_{R_c}$ , and the central surface brightness ratio,  $f_{\Sigma}$ , between the two components. [C II] haloes can significantly boost the LIM power spectrum signal. For example, for relatively compact [C II] haloes ( $f_{\Sigma} = 0.4$ ,  $f_{R_c} = 2.0$ ), the signal is boosted by  $\simeq 20$  times; for more extended and diffuse haloes ( $f_{\Sigma} = 0.1$ ,  $f_{R_c} = 6.0$ ), the signal is boosted by  $\simeq 100$  times. For the ALMA ASPECS survey (resolution  $\theta_{\text{beam}} = 1.13$  arcsec, survey area  $\Omega_{\text{survey}} = 2.9$  arcmin<sup>2</sup>) the [C II] power spectrum is detectable only if the deL14d [C II]–SFR relation holds. However, with an optimized survey ( $\theta_{\text{beam}} = 0.232$  arcsec,  $\Omega_{\text{survey}} = 2.0$  deg<sup>2</sup>), the power spectrum is detectable for all the [C II]–SFR relations considered in this paper. Such a survey can constrain  $f_{\Sigma}$  ( $f_{R_c}$ ) with a relative uncertainty of  $\sim 15$  per cent ( $\sim 10$  per cent). A successful LIM experiment will provide unique constraints on the nature, origin, and frequency of extended [C II] haloes, and the [C II]–SFR relation at early times.

**Key words:** galaxies: formation – galaxies: high-redshift – dark ages, reionization, first stars – radio lines: galaxies.

## 1 INTRODUCTION

Cosmic reionization, as the last major phase transition of the Universe, is a direct consequence of the formation of the first luminous objects. Determining the properties of galaxies in the Epoch of Reionization (EoR) is crucial to understand structure formation and evolution (Dayal & Ferrara 2018). However, our progress is hampered by the limited number of EoR sources that can be directly accessed via targeted observations. Such limitation is particularly severe for faint galaxies, which are thought to be the primary sources of ionizing photons (Mitra, Choudhury & Ferrara 2015; Robertson et al. 2015; Castellano et al. 2016) but are far below the detection limits of the current surveys (Salvaterra et al. 2013).

Line intensity mapping (LIM) is emerging as an efficient tool to detect these faint galaxies. LIM measures the integrated emission of spectral lines from all galaxies and the intergalactic medium (see a review by Kovetz et al. 2017). By measuring fluctuations in the line emission from early galaxies, one could expect to obtain valuable physical insights about the properties of these sources and their role in reionization. Finally, LIM can also be used to put stringent constraints on cosmological models (Kovetz et al. 2019; Schaan & White 2021; Bernal & Kovetz 2022; Karkare et al. 2022). For example, Gong, Chen & Cooray (2020) proposed to use the multipole moments of

the redshift-space LIM power spectrum to constrain the cosmological and astrophysical parameters.

There are several spectral lines of interest for intensity mapping including HI 21 cm line (Chang et al. 2010; Salvaterra et al. 2013), CO rotational lines (Breyse, Kovetz & Kamionkowski 2014; Mashian, Sternberg & Loeb 2015; Li et al. 2016), bright optical emission lines such as Ly $\alpha$  and H $\alpha$  (Salvaterra et al. 2013; Pullen, Doré & Bock 2014; Comaschi & Ferrara 2016; Gong et al. 2017; Silva et al. 2018) and the far-infrared (FIR) fine-structure lines (Gong et al. 2012; Uzgil et al. 2014; Silva et al. 2015; Yue et al. 2015; Serra, Doré & Lagache 2016; Yue & Ferrara 2019).

Among these, the [C II] emission line with wavelength 157.7  $\mu\text{m}$ , corresponding to the  ${}^2P_{3/2} \rightarrow {}^2P_{1/2}$  forbidden transition of singly ionized carbon, is the brightest one in the FIR (Stacey et al. 1991), and a dominant coolant of the neutral interstellar medium (ISM; Stacey et al. 1991; Wolfire et al. 2003) and dense photodissociation regions (PDR; Hollenbach & Tielens 1999). Therefore, the [C II] line can be used to probe the properties of the ISM at high redshift (Capak et al. 2015; Knudsen et al. 2016; Pentericci et al. 2016; Carniani et al. 2017; Bakx et al. 2020; Matthee et al. 2020).

Notably, a tight relation between [C II] luminosity and star formation rate (SFR) is found from both observations (De Looze et al. 2014; Herrera-Camus et al. 2015; Schaerer et al. 2020) and simulations (Vallini et al. 2015; Olsen et al. 2017; Leung et al. 2020). Therefore, the [C II] emission lines can be used to trace the star formation across cosmic time, albeit at high redshift the scatter around the

★ E-mail: [andrea.ferrara@sns.it](mailto:andrea.ferrara@sns.it)

local relation increases (Carniani et al. 2018). Conveniently, at early epochs, the line is shifted into the (sub-)mm wavelength range, which is accessible to ground-based telescopes such as the Atacama Large Millimeter/sub-millimeter Array (ALMA).

With the advent of ALMA, a large number of galaxies with [C II] emission line at  $z > 4$  have been detected, boosting the studies of the obscured star formation at high redshift (Hodge & da Cunha 2020). Among these, a stacking analysis of ALMA observed galaxies (Fujimoto et al. 2019) discovered extended, 10 kpc scale [C II] haloes around high redshift galaxies, whose size is  $\simeq 5$  times larger than the UV size of the central galaxy. Similar results have been reported in the following studies both in subsequent stacking analysis (Ginolfi et al. 2020; Fudamoto et al. 2022) and individual galaxies (Fujimoto et al. 2020; Herrera-Camus et al. 2021; Akins et al. 2022; Lambert et al. 2022; Fudamoto et al. 2023).

The existence of extended [C II] haloes around normal star-forming galaxies opens new perspectives for early metal enrichment. However, the physical origin of [C II] haloes is still not clear. Possible scenarios include satellite galaxies, extended PDR or HII regions, cold streams, and outflows. These are discussed in Fujimoto et al. (2019). Indeed, the supernova-driven cooling outflow model explored by Pizzati et al. (2020, 2023) successfully produces the extended [C II] halo, their model also indicates that outflows are widespread phenomena in high- $z$  galaxies, but the extended [C II] haloes for low-mass galaxies are likely too faint to be detected with present levels of sensitivity. The outflow scenario is also supported by Ginolfi et al. (2020), who found outflow signatures in the stacking analysis of [C II] emission detected by ALMA in 50 main sequence star-forming galaxies at  $4 < z < 6$ . Fujimoto et al. (2020) also suggested that the star-formation-driven outflow is the most likely origin of the [C II] haloes. Finally, Herrera-Camus et al. (2021) found evidence of outflowing gas, which may be responsible for the production of extended [C II] haloes. On the other hand, Heintz et al. (2023) use [C II] emission as a proxy to infer the metal mass in the ISM of galaxies, and find that the majority of metals produced at  $z \gtrsim 5$  are confined to the ISM, disfavour efficient outflow processes at these redshifts, they claim extended [C II] haloes trace the extended neutral gas reservoirs of high- $z$  galaxies. In any case, if [C II] haloes are common in high redshift normal star-forming galaxies, they could leave a distinct imprint in LIM experiments.

In this paper, we aim to investigate the effects of extended [C II] haloes on the LIM signal and their detectability. We first construct the [C II] halo model and compute the intensity mapping power spectrum of high-redshift galaxy systems when both the central galaxy and extended [C II] halo are considered.<sup>1</sup> Then, we analyse the detectability of the signal in the ALMA intensity mapping survey and a proposed optimized survey. The paper is organized as follows: We outline our method in Section 2. Results, including the power spectrum signal and its signal-to-noise ratio (S/N) estimation, are presented in Section 3. We finally summarize the results and give a discussion in Section 4.

## 2 METHOD

In the following, we describe our model to derive the LIM signal when both central galaxies and [C II] haloes are considered. We start by constructing the extended [C II] halo model, and the relations

between central galaxy [C II] luminosity, extended [C II] halo luminosity, and dark matter halo mass. We then compute the intensity mapping power spectrum, including one-halo and two-halo terms, and the shot noise term. Finally, we introduce the method to estimate the S/N of the [C II] power spectrum, given the LIM survey parameters.

### 2.1 [C II] halo model

To compute the clustering term of the power spectrum, the first step is to model the [C II] radial surface brightness of the central galaxy and the extended halo. For this, we use a combined Sérsic + exponential model, where the central galaxy is described by the Sérsic model (Sérsic 1963; Sérsic 1968) while the extended [C II] halo is described by the exponential function (Fujimoto et al. 2019; Akins et al. 2022).

The [C II] radial surface brightness of the central galaxy writes

$$\Sigma_{\text{CII,g}}(R) = C_g \exp \left[ -b_n \left( \frac{R}{R_{e,g}} \right)^{1/n} \right], \quad (1)$$

where  $C_g$  is the central surface brightness,  $R$  is the projected distance to the source centre in the plane-of-sky,  $n$  is the Sérsic index, and  $R_{e,g}$  is the effective radius containing half of the integrated brightness. The term  $b_n$  is a function of  $n$ . It is obtained by solving the equation  $\Gamma(2n) = 2\gamma(2n, b_n)$ , where  $\Gamma(2n)$  is the Gamma function,  $\gamma(2n, b_n) = \int_0^{b_n} t^{2n-1} e^{-t} dt$  is the lower incomplete gamma function.

In the stacking analysis by Fujimoto et al. (2019), the central galaxies are well fitted by Sérsic model with  $n = 1.2$  and  $R_{e,g} = 1.1$  kpc. This is consistent with the rest-frame optical and UV sizes measured by Shibuya, Ouchi & Harikane (2015), who obtain a nearly constant value of  $R_{e,g}/R_{\text{vir}} = 1.0\text{--}3.5$  per cent for  $z = 0\text{--}8$ , where  $R_{\text{vir}}$  is the dark matter halo virial radius. Following these results, we adopt  $n = 1.2$  and fix  $R_{e,g} = 0.03R_{\text{vir}}$ .

For the extended [C II] halo, the surface brightness has the exponential form<sup>2</sup>

$$\Sigma_{\text{CII,h}}(R) = C_h \exp \left[ -b_1 \frac{R}{R_{e,h}} \right], \quad (2)$$

where  $b_1$  is the  $b_n$  for  $n = 1$ ,  $C_h$  is the surface brightness at the centre and the  $R_{e,h}$  is the effective radius of the [C II] halo. We further assume  $C_h = f_\Sigma C_g$  and  $R_{e,h} = f_{R_e} R_{e,g}$ , where the ratios  $f_\Sigma$  and  $f_{R_e}$  are two free parameters in our work.

With the above surface brightness profiles, the [C II] luminosity for the central galaxy

$$\begin{aligned} L_{\text{CII,g}} &= \int 2\pi R \Sigma_{\text{CII,g}}(R) dR \\ &= 2\pi C_g \frac{n\Gamma(2n)}{(b_n)^{2n}} R_{e,g}^2, \end{aligned} \quad (3)$$

and for the extended [C II] halo

$$\begin{aligned} L_{\text{CII,h}} &= \int 2\pi R \Sigma_{\text{CII,h}}(R) dR \\ &= 2\pi C_h \frac{\Gamma(2)}{(b_1)^2} R_{e,h}^2. \end{aligned} \quad (4)$$

Since restricted by the sensitivity of the telescope in previous observations, the observed [C II] luminosity in deriving the [C II]–SFR relation is likely dominated by central galaxies. Moreover, as noted by Fujimoto et al. (2019), to fully capture the extended [C II] haloes,

<sup>1</sup>Throughout the paper, we assume  $\Lambda$ CDM model with Planck Collaboration (2016) cosmological parameters:  $\Omega_m = 0.308$ ,  $\Omega_\Lambda = 1 - \Omega_m = 0.692$ ,  $\Omega_b = 0.048$ ,  $h = 0.678$ ,  $\sigma_8 = 0.815$ , and  $n_s = 0.968$ .

<sup>2</sup>Note that the Sérsic model equation (1) is reduced to exponential function when  $n = 1$ .

additional mechanisms not involved in previous simulations are required. Therefore, we consider the previously predicted [C II]–SFR relations (both by observations and by simulations) do not account for (or at least significantly underestimate) the extended [C II] halo contribution. For clarification, we shall use  $L_{\text{CII,SFR}}$ –SFR to denote the relations in the following of this work, where  $L_{\text{CII,SFR}}$  is the [C II] luminosity derived from the previously [C II]–SFR relations, see next subsection. We find the normalization by assigning  $L_{\text{CII,g}} \approx L_{\text{CII,SFR}}$ . Then

$$C_g = \frac{(b_n)^{2n}}{2\pi n \Gamma(2n) R_{e,g}^2} L_{\text{CII,SFR}}. \quad (5)$$

Then the [C II] halo luminosity is also derived from the  $L_{\text{CII,SFR}}$  via

$$L_{\text{CII,h}} = f_\Sigma f_{R_e}^2 \frac{(b_n)^{2n}}{(b_1)^2} \frac{\Gamma(2)}{n \Gamma(2n)} L_{\text{CII,SFR}}, \quad (6)$$

and when  $n = 1.2$  it is reduced to

$$L_{\text{CII,h}} \simeq 2.24 f_\Sigma f_{R_e}^2 L_{\text{CII,SFR}}. \quad (7)$$

We further assume that the profile is spherically symmetric, and perform a deprojection of the above surface density to obtain the 3D brightness profiles for the two components,  $\rho_{\text{CII,g}}(r)$  and  $\rho_{\text{CII,h}}(r)$ . This is given in Appendix A. The total 3D brightness profile is the sum of the two components

$$\rho_{\text{CII}}(r) = \rho_{\text{CII,g}}(r) + \rho_{\text{CII,h}}(r), \quad (8)$$

and we truncate it at  $R_{\text{vir}}$ . Then, the normalized Fourier transform of the profile in a dark matter halo of mass  $M_{\text{vir}}$ ,

$$u_{\text{CII}}(k|M_{\text{vir}}) = \frac{\int_0^{R_{\text{vir}}} dr 4\pi r^2 \frac{\sin kr}{kr} \rho_{\text{CII}}(r)}{\int_0^{R_{\text{vir}}} dr 4\pi r^2 \rho_{\text{CII}}(r)}, \quad (9)$$

which is used to compute the one-halo and two-halo terms of the power spectrum. The luminosity density profiles of the central galaxy and [C II] halo are shown in the left panel of Fig. 1. We take  $M_{\text{vir}} = 10^{12} M_\odot$  at  $z = 6$ ,  $f_\Sigma = 0.4$  and  $f_{R_e} = 2.0, 6.0, 10.0$ . In the right panel of Fig. 1 we show the Fourier transform of the profiles for different  $f_{R_e}$ .

## 2.2 The [C II]–dark matter halo mass relation

To derive the  $L_{\text{CII,SFR}}$ – $M_{\text{vir}}$  relation, we first use the abundance matching method to get SFR –  $M_{\text{vir}}$  relation. Following Yue et al. (2015), we start from the observed (dust-attenuated) galaxy UV luminosity function (LF), which can be described by a Schechter function (Schechter 1976)

$$\frac{dn}{dM_{\text{UV}}^{\text{obs}}} = 0.4 \ln(10) \Phi^* 10^{-0.4(M_{\text{UV}}^{\text{obs}} - M_{\text{UV}}^*) (\alpha + 1)} \times \exp[-10^{-0.4(M_{\text{UV}}^{\text{obs}} - M_{\text{UV}}^*)}], \quad (10)$$

where the  $M_{\text{UV}}^{\text{obs}}$  is the dust-attenuated absolute UV magnitude, and the redshift-dependent parameters are taken from Bouwens et al. (2015),

$$\begin{aligned} M_{\text{UV}}^* &= -20.95 + 0.01(z - 6), \\ \Phi^* &= 0.47 \times 10^{-3 - 0.27(z - 6)}, \\ \alpha &= -1.87 - 0.10(z - 6). \end{aligned} \quad (11)$$

Considering dust attenuation, the intrinsic absolute UV magnitude becomes

$$M_{\text{UV}} = M_{\text{UV}}^{\text{obs}} - A_{\text{UV}}, \quad (12)$$

where the UV dust attenuation is parameterized as

$$A_{\text{UV}} = C_1 + C_0 \beta \quad (A_{\text{UV}} \geq 0); \quad (13)$$

we set the coefficients  $C_0 = 2.10$  and  $C_1 = 4.85$ , following Koprowski et al. (2018). Finally, the UV spectral slope  $\beta$  is fitted by Bouwens et al. (2015)

$$\beta = \beta_0 + \frac{d\beta}{dM_0} (M_{\text{UV}}^{\text{obs}} - M_0), \quad (14)$$

with a redshift dependence given by

$$\begin{aligned} \beta_0 &= -1.97 - 0.06(z - 6), \\ \frac{d\beta}{dM_{\text{UV}}} &= -0.18 - 0.03(z - 6). \end{aligned} \quad (15)$$

The intrinsic UV LF is related to the observed UV LF by

$$\frac{dn}{dM_{\text{UV}}} = \frac{dn}{dM_{\text{UV}}^{\text{obs}}} \frac{dM_{\text{UV}}^{\text{obs}}}{dM_{\text{UV}}}. \quad (16)$$

By assuming that the most massive galaxies occupy the most massive dark matter haloes, we use the abundance matching method

$$\int_{M_{\text{UV}}} \frac{dn}{dM_{\text{UV}}} dM_{\text{UV}} = \int_{M_{\text{vir}}} \frac{dn}{dM_{\text{vir}}} dM_{\text{vir}}, \quad (17)$$

to derive the intrinsic absolute UV magnitude and dark matter halo mass relation. To get the SFR– $M_{\text{vir}}$  relation, we assume that the intrinsic UV luminosity scales with the SFR (Kennicutt 1998)

$$\text{SFR} = K_{\text{UV}} L_{\text{UV}}, \quad (18)$$

where  $K_{\text{UV}} = 0.7 \times 10^{-28} M_\odot \text{ yr}^{-1} / (\text{erg s}^{-1} \text{ Hz}^{-1})$  (Bruzual & Charlot 2003) for a Chabrier (2003) stellar initial mass function, metallicity in the range  $0.005$ – $0.4 Z_\odot$ , and stellar age  $> 100$  Myr; we ignore the scatter on  $K_{\text{UV}}$  as discussed by Yue & Ferrara (2019).

The final step to get the  $L_{\text{CII,SFR}}$ – $M_{\text{vir}}$  relation involves the  $L_{\text{CII,SFR}}$  – SFR relation. This can be parameterized as

$$\begin{aligned} \log L_{\text{CII,SFR}} &= \log \bar{L}_{\text{CII,SFR}} \pm \sigma_L \\ &= \log A + \gamma \log \text{SFR} \pm \sigma_L, \end{aligned} \quad (19)$$

where SFR is in units of  $M_\odot \text{ yr}^{-1}$ , and  $L_{\text{CII,SFR}}$  in units of  $L_\odot$ . We adopt six different  $L_{\text{CII,SFR}}$ –SFR relations to compute the power spectrum in our model. These relations are described in the following.

By combining a semi-analytical model of galaxy formation with the photoionization code CLOUDY (Ferland et al. 2013, 2017) to compute the [C II] luminosity for a large number of galaxies at  $z \geq 4$ , Lagache, Cousin & Chatzikos (2018) reproduced the  $L_{\text{CII,SFR}}$  – SFR relation observed in high- $z$  star-forming galaxies. They found a mild evolution of the  $L_{\text{CII,SFR}}$  – SFR relation (hereafter L18) with redshift from  $z = 4$  to  $z = 7.6$ ,

$$\log L_{\text{CII,SFR}} = (7.1 - 0.07z) + (1.4 - 0.07z) \log(\text{SFR}), \quad (20)$$

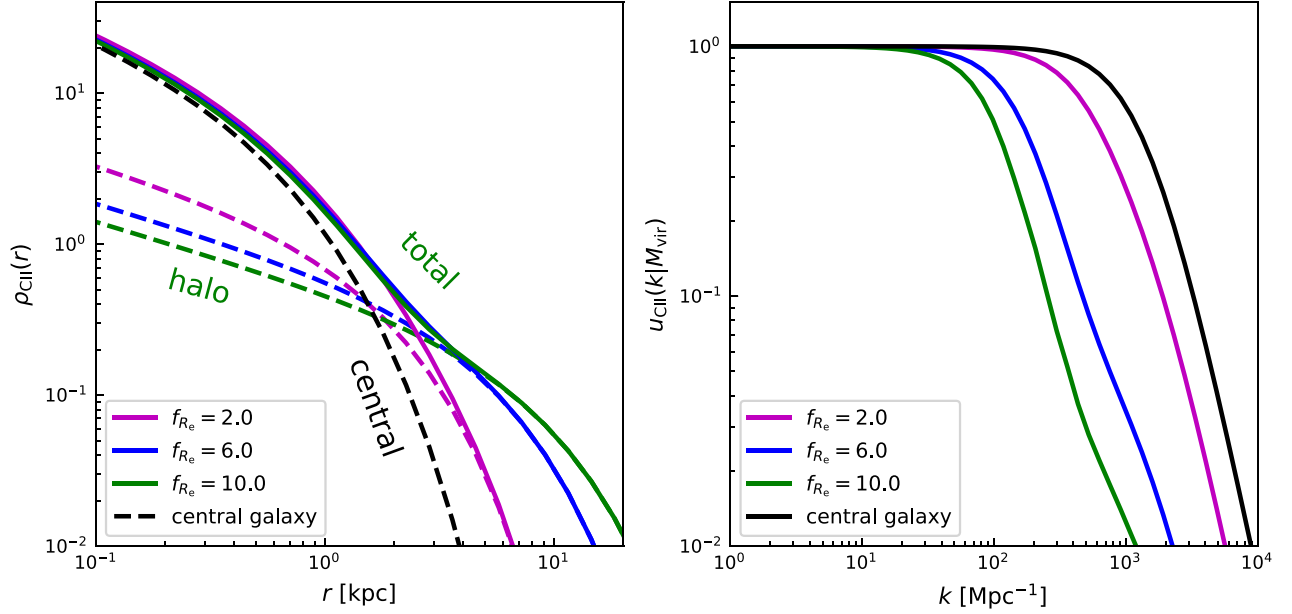
with a scatter of about 0.5 dex.

De Looze et al. (2014) analysed the  $L_{\text{CII,SFR}}$  – SFR relation for low-metallicity dwarf galaxies based on Herschel observations (Madden et al. 2013) and found a local relation with  $\log A = 7.16$ ,  $\gamma = 1.25$ , and 0.5 dex scatter (hereafter deL14d), i.e.

$$\log L_{\text{CII,SFR}} = 7.16 + 1.25 \log \text{SFR}, \quad (21)$$

De Looze et al. (2014) also obtained  $\log A = 7.22$ ,  $\gamma = 0.85$ , and 0.3 dex scatter (hereafter deL14z) for high- $z$  samples with  $z = 0.59$ – $6.60$ .

$$\log L_{\text{CII,SFR}} = 7.22 + 0.85 \log \text{SFR} \quad (22)$$



**Figure 1.** *Left:* The radial profile of a [C II] halo in dark matter halo with mass  $M_{\text{vir}} = 10^{12} M_{\odot}$ . We fix [C II] halo parameter  $f_{\Sigma} = 0.4$  and vary the effective radius ratios  $f_{R_e}$ . The black dashed line is the profile of the central galaxy. The green, blue, and magenta dashed lines are the profiles of the [C II] halo part with  $f_{R_e} = [2.0, 6.0, 10.0]$ , respectively. The solid lines represent the sum of the central galaxy and [C II] halo components. *Right:* The corresponding normalized Fourier transform of the [C II] halo profile with different effective radius ratios  $f_{R_e}$ .

Olsen et al. (2017) obtained  $\log A = 6.69$  and  $\gamma = 0.58$ , with a scatter of  $\sim 0.15$  dex (here after O17) by combining cosmological zoom simulations of galaxies with SÍGAME (Olsen et al. 2015) to model [C II] emissions from 30 main-sequence galaxies at  $z \sim 6$ ,

$$\log L_{\text{CII,SFR}} = (6.69 \pm 0.10) + (0.58 \pm 0.11) \log \text{SFR}. \quad (23)$$

This is consistent with Leung et al. (2020) who predict  $\log A = 6.82$  and  $\gamma = 0.66$ , based on cosmological hydrodynamics simulations using the SIMBA suite plus radiative transfer calculations via an updated version of SÍGAME.

Vallini et al. (2015) combined high-resolution, radiative transfer cosmological simulations with a subgrid multiphase model of the ISM to model the [C II] emissions from diffuse neutral gas and PDRs. By considering a physically motivated metallicity, they found  $L_{\text{CII,SFR}} - \text{SFR}$  relation both depends on the SFR and metallicity and can be fitted by

$$\log L_{\text{CII,SFR}} = 7.0 + 1.2 \log(\text{SFR}) + 0.021 \log(Z) + 0.012 \log(\text{SFR}) \log(Z) - 0.74 \log^2(Z). \quad (24)$$

For  $Z = 0.2 Z_{\odot}$ , we obtain  $\log A = 6.62$  and  $\gamma = 1.19$  (here after V15).

Schaerer et al. (2020), using 118 galaxies at  $4.4 < z < 5.9$  from the ALMA ALPINE Large Program (Le Fèvre et al. 2020), obtain

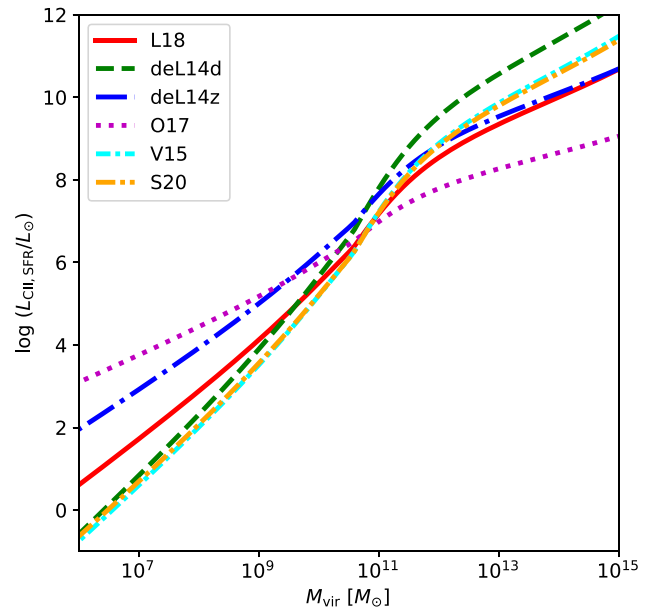
$$\log L_{\text{CII,SFR}} = (6.61 \pm 0.2) + (1.17 \pm 0.12) \log \text{SFR}, \quad (25)$$

with a scatter of  $\sim 0.28$  dex (here after S20).

The parameters of the six  $L_{\text{CII,SFR}} - \text{SFR}$  relations adopted are summarized in Table 1. Since the observed  $\sigma_L$  involves both the observational uncertainties and the intrinsic scatters, and the simulated  $\sigma_L$  might be limited by the method of modelling the input physical properties scattering, for simplicity in this work we ignore the  $\sigma_L$  and simply use  $\log L_{\text{CII,SFR}} = \log A + \gamma \log \text{SFR}$  in the following. We plot the corresponding  $L_{\text{CII,SFR}} - M_{\text{vir}}$  relations in Fig. 2.

**Table 1.** Summary of  $L_{\text{CII,SFR}} - \text{SFR}$  relation parameters adopted in our model.

Model	$\log A$	$\gamma$	$\sigma_L$	Redshift	References
L18	6.68	0.98	0.60	6	Lagache et al. (2018)
deL14d	7.16	1.25	0.5	Local	De Looze et al. (2014)
deL14z	7.22	0.85	0.3	6.6	De Looze et al. (2014)
O17	6.69	0.58	0.15	6	Olsen et al. (2017)
V15	6.62	1.19	0.4	6.6	Vallini et al. (2015)
S20	6.61	1.17	0.28	4.4–5.9	Schaerer et al. (2020)



**Figure 2.** The  $L_{\text{CII,SFR}} - M_{\text{vir}}$  relation obtained from Abundance Matching method for six  $L_{\text{CII,SFR}} - \text{SFR}$  relations.



### 2.3 Modelling the power spectrum

Since galaxies are biased and discrete tracers of the dark matter density fluctuations, basically the LIM power spectrum comprises clustering and shot noise components (Kovetz et al. 2017). The clustering term describes the large-scale clustering nature of galaxies and the shot noise term originates from the Poisson fluctuations of galaxy numbers. In the halo model framework (Cooray & Sheth 2002), the clustering term can be split into one-halo term and two-halo term, which arise from the correlation within haloes and between haloes, respectively (Moradinezhad Dizgah et al. 2022). The LIM power spectrum writes:

$$P_{\text{CII}}(k, z) = P_{\text{CII}}^{\text{CL}}(k, z) + P_{\text{CII}}^{\text{SN}}(z) \\ = P_{\text{CII}}^{\text{1h}}(k, z) + P_{\text{CII}}^{\text{2h}}(k, z) + P_{\text{CII}}^{\text{SN}}(z), \quad (26)$$

where the shot noise  $P_{\text{CII}}^{\text{SN}}(z)$  is generally independent of  $k$ . The one-halo term  $P_{\text{CII}}^{\text{1h}}(k, z)$  and the two-halo term  $P_{\text{CII}}^{\text{2h}}(k, z)$  depend on the large-scale structure, the normalized Fourier transform of the [C II] halo profile and the  $L_{\text{CII,SFR}}-M_{\text{vir}}$  relation.

Analogous to the halo model (Cooray & Sheth 2002) which assumes the correlation is tightly related to the (density or light) distribution profile within dark matter haloes, the one-halo term<sup>3</sup> (Moradinezhad Dizgah et al. 2022) is

$$P_{\text{CII}}^{\text{1h}}(k, z) = \int dM_{\text{vir}} \frac{dn}{dM_{\text{vir}}} u_{\text{CII}}(k|M_{\text{vir}})^2 \left[ \frac{L_{\text{CII}}(M_{\text{vir}}, z)}{4\pi D_L^2} y(z) D_c^2 \right]^2 \\ = \left( \frac{c}{4\pi \nu_{\text{CII}} H(z)} \right)^2 \int dM_{\text{vir}} \frac{dn}{dM_{\text{vir}}} \\ \times [L_{\text{CII}}(M_{\text{vir}}, z) u_{\text{CII}}(k|M_{\text{vir}})]^2, \quad (27)$$

and two-halo term (Moradinezhad Dizgah et al. 2022) is

$$P_{\text{CII}}^{\text{2h}}(k, z) \\ = \left( \int dM_{\text{vir}} \frac{dn}{dM_{\text{vir}}} u_{\text{CII}}(k|M_{\text{vir}}) \right)^2 \left[ \frac{L_{\text{CII}}(M_{\text{vir}}, z)}{4\pi D_L^2} y(z) D_c^2 \right]^2 \\ = \left( \int dM_{\text{vir}} \frac{dn}{dM_{\text{vir}}} L_{\text{CII}}(M_{\text{vir}}, z) u_{\text{CII}}(k|M_{\text{vir}}) b_{\text{SMT}}(M_{\text{vir}}, z) \right)^2 \\ \times \left( \frac{c}{4\pi \nu_{\text{CII}} H(z)} \right)^2 P(k, z). \quad (28)$$

$D_L$  is the luminosity distance and  $D_c$  is the comoving distance to redshift  $z$ .  $y(z) = c(1+z)/[v_0 H(z)]$  is the derivative of the comoving radial distance with respect to the observed frequency  $\nu_0$ , which is related to the rest-frame frequency of [C II] line as  $\nu_0 = \nu_{\text{CII}}/(1+z)$ ;  $c$  is the light speed and  $H(z)$  is the Hubble parameter. We adopt the Sheth-Tormen form (Sheth & Tormen 1999) of the halo mass function  $dn/dM_{\text{vir}}$ , the Eisenstein & Hu (1999) form of the linear matter power spectrum  $P(k, z)$ , and the Sheth, Mo & Tormen (2001) form of dark matter halo bias  $b_{\text{SMT}}(M_{\text{vir}}, z)$ .  $L_{\text{CII}}(M_{\text{vir}}, z) = L_{\text{CII,g}} + L_{\text{CII,h}}$  is the [C II] luminosity of the dark matter halo with mass  $M_{\text{vir}}$ , including the central galaxy and the extended [C II] halo.

Finally, the Poissonian shot noise, arising from the discrete nature of dark matter haloes, can be computed as (Uzgil et al. 2014)

$$P_{\text{CII}}^{\text{SN}}(z) = \left( \frac{c}{4\pi \nu_{\text{CII}} H(z)} \right)^2 \int dM_{\text{vir}} \frac{dn}{dM_{\text{vir}}} L_{\text{CII}}^2(M_{\text{vir}}, z) \quad (29)$$

<sup>3</sup>The one halo term formula of equation (27) tends to be a constant at large scale (low- $k$ ), this is an unphysical behaviour (Schaan & White 2021). Some attempts have been proposed to solve the problem (Cooray & Sheth 2002; Baldauf et al. 2013), but it is still an open issue.

The integration of equations (27, 28, and 29) is performed between  $M_{\text{min}}$  and  $M_{\text{max}}$ . We set  $M_{\text{min}} = 10^8 M_{\odot}$ , which is roughly the atomic cooling threshold (the virial mass corresponding to virial temperature  $10^4$  K) at redshift 6. This is the minimum mass that can sustain persistent star formation activity. Generally, the number density of dark matter haloes with mass  $M_{\text{vir}}$  can be approximated as  $\bar{n} \sim M_{\text{vir}} \frac{dn}{dM_{\text{vir}}} = \frac{dn}{d \ln M_{\text{vir}}}$ . Suppose the survey volume is  $V_{\text{survey}}$ , then in such volume the mean number of dark matter haloes with mass  $M_{\text{vir}}$  is just

$$\bar{n} V_{\text{survey}} \sim \frac{dn}{d \ln M_{\text{vir}}} V_{\text{survey}}. \quad (30)$$

In  $V_{\text{survey}}$ , the probability to find at least one dark matter halo with mass  $\sim M_{\text{vir}}$  is  $G(\geq 1|\bar{n} V_{\text{survey}})$ , where  $G$  is the cumulative Poisson probability with mean value  $\bar{n} V_{\text{survey}}$ . We obtain  $M_{\text{max}}$  by solving the equation  $G(\geq 1|\bar{n} V_{\text{survey}}) = 0.5$ , which means in the volume  $V_{\text{survey}}$  the contribution from dark matter haloes whose existence probability smaller than 50 per cent is excluded. This is to avoid such bright and rare objects biasing our results. When  $V_{\text{survey}} = 10^6 \text{ Mpc}^3$ ,  $M_{\text{max}} \sim 2.2 \times 10^{12} M_{\odot}$  at redshift 6.

Let's interpret the one-halo term, two-halo term, and shot noise of the [C II] power spectrum more physically. Since the mean [C II] specific intensity is (Gong et al. 2011)

$$\bar{I}_{\text{CII}}(z) = \int dM_{\text{vir}} \frac{dn}{dM_{\text{vir}}} \left[ \frac{L_{\text{CII}}(M_{\text{vir}}, z)}{4\pi D_L^2} y(z) D_c^2 \right], \quad (31)$$

so the one-halo term can be interpreted as the auto-correlations of the [C II] light distribution within dark matter haloes, weighed by the square of their contribution to the mean [C II] specific intensity. On the other hand, at scales much larger than the dark matter halo size, the two-halo term has the approximation form

$$P_{\text{CII}}^{\text{2h}}(k, z) \approx \bar{I}_{\text{CII}}^2(z) \bar{b}_{\text{SMT}}^2 P(k, z), \quad (32)$$

where  $\bar{b}_{\text{SMT}}$  is the mean halo bias. So the two-halo term actually describes the large-scale power spectrum of the dark matter haloes weighted by the square of mean [C II] specific intensity. Regarding the shot noise, it has the approximation

$$P_{\text{CII}}^{\text{SN}}(z) \approx \bar{I}_{\text{CII}}^2(z) \frac{1}{\bar{n}}, \quad (33)$$

where  $\bar{n}$  is the number density of dark matter haloes that have a dominant contribution in the shot noise. Clearly, this approximation shows that shot noise is actually the Poisson fluctuations of the dark matter haloes, weighted by the square of the mean [C II] specific intensity.

### 2.4 The signal-to-noise ratio estimation

For a LIM survey, the smallest and largest accessible  $k$  modes for probing the fluctuations are determined by the survey volume and the resolution (angular resolution  $\delta\theta$  and frequency resolution  $\delta\nu_0$ ). The survey volume is determined by the redshift of the target, the survey area  $\Omega_{\text{survey}}$ , and the bandwidth coverage  $\Delta\nu_0$ .

Along the line of sight, the spatial resolution corresponds to the frequency resolution (Bull et al. 2015),

$$\Delta r_{\parallel, \text{min}} \approx \frac{c(1+z)}{H(z)} \frac{\delta\nu_0}{\nu_0}, \quad (34)$$

while the largest spatial across corresponds to the bandwidth,

$$\Delta r_{\parallel, \text{max}} \approx \frac{c(1+z)}{H(z)} \frac{\Delta\nu_0}{\nu_0}. \quad (35)$$

The tangential spatial resolution corresponds to the synthesized beam size  $\Omega_{\text{beam}}$

$$\Delta r_{\perp, \min} \approx D_c(z) \sqrt{\Omega_{\text{beam}}}, \quad (36)$$

and the largest spatial across corresponds to the survey area  $\Omega_{\text{survey}}$

$$\Delta r_{\perp, \max} \approx D_c(z) \sqrt{\Omega_{\text{survey}}}. \quad (37)$$

Then, along and perpendicular to the line of sight, the smallest and largest  $k$  modes that can be detected by the survey are (Uzgil et al. 2019),

$$k_{\parallel, \min} = \frac{2\pi}{\Delta r_{\parallel, \max}} \text{ and } k_{\parallel, \max} = \frac{1}{2} \frac{2\pi}{\Delta r_{\parallel, \min}}, \quad (38)$$

$$k_{\perp, \min} = \frac{2\pi}{\Delta r_{\perp, \max}} \text{ and } k_{\perp, \max} = \frac{1}{2} \frac{2\pi}{\Delta r_{\perp, \min}}. \quad (39)$$

Finally, the survey detects the isotropic power spectrum in the range  $k_{\min} < k < k_{\max}$ , for which  $k_{\min} \approx \sqrt{k_{\parallel, \min}^2 + k_{\perp, \min}^2}$  and  $k_{\max} \approx \sqrt{k_{\parallel, \max}^2 + k_{\perp, \max}^2}$ .

The variance of the power spectrum measured in the above LIM survey is

$$\sigma_p^2(k) = \frac{1}{N_m(k)} [P_{\text{CII}}(k, z) + P_N W_{\text{res}}^{-2}(k)]^2, \quad (40)$$

where  $W_{\text{res}}(k)$  is the window function that denotes the rapid decline of the measured power spectrum below the resolution (Seo et al. 2010; Lidz et al. 2011; Battye et al. 2013; Lidz & Taylor 2016; Bernal et al. 2019). In principle  $\Delta r_{\parallel, \min}$  can be very different from  $\Delta r_{\perp, \min}$ , so the window function for  $k_{\parallel}$  and  $k_{\perp}$  should be different (Lidz et al. 2011; Lidz & Taylor 2016; Bernal et al. 2019; Bernal & Kovetz 2022). In that case, at the small scale, the power spectrum is no longer isotropic and the cylinder power spectrum should be used. However, in this paper, we only investigate the isotropic power spectrum, we implicitly assume that the  $\Delta r_{\parallel, \min}$  are comparable to the  $\Delta r_{\perp, \min}$ . For this reason, we adopt the approximation (Battye et al. 2013)

$$W_{\text{res}}(k) \approx \exp \left[ -\frac{1}{2} k^2 D_c^2(z) \frac{\Omega_{\text{beam}}}{4} \right]. \quad (41)$$

In principle, there is another window function that describes the power spectrum decline beyond the finite survey volume (Bernal et al. 2019). In our calculations, we do not consider such a window function; instead, we consider the power spectrum at  $k > k_{\min}$  is fully measured and at  $k < k_{\min}$  is fully missed, for which  $k_{\min}$  is set by the survey volume. This is equivalent to adopting a sharp  $k$  cutoff step function as the window function. Although this may not be the real case, it will just have quite limited effects on the results, because the contributions to S/N and parameter constraints from  $k$  modes close to  $k_{\min}$  are small, see the results in Section 3.2.

In equation (40),  $P_N$  is the instrumental noise power spectrum.  $N_m(k)$  is the number of  $k$  modes sampled by a survey in the  $k$ -bin, which is estimated (Chung et al. 2020) by<sup>4</sup>

$$N_m(k) \approx 2\pi k^3 d \ln k \frac{V_{\text{survey}}}{(2\pi)^3}, \quad (42)$$

<sup>4</sup>Note that equation (42) may overestimate the number of  $k$  modes, particularly at small scales (large  $k$ ; Gong et al. 2017, 2020). Because it assumes that all  $k$  modes in the survey volume are independent. However, at small scales,  $k$  modes are crowded and there must be some degeneracies between them.

where  $d \ln k$  is the relative width of the selected  $k$ -bin, and  $V_{\text{survey}}$  is the survey volume,

$$V_{\text{survey}} \approx D_c^2(z) \Omega_{\text{survey}} \left[ \frac{c(1+z) \Delta v_0}{H(z) v_0} \right]. \quad (43)$$

Throughout this paper, we adopt bin width  $d \ln k = 0.2$ .

Then the S/N of each  $k$ -bin

$$\begin{aligned} \text{S/N} &= \frac{P_{\text{CII}}(k, z)}{\sigma_p(k)} \\ &= \frac{P_{\text{CII}}(k, z)}{\sqrt{N_m(k)} \sqrt{P_{\text{CII}}(k, z) + P_N W_{\text{res}}^{-2}(k)}}, \end{aligned} \quad (44)$$

and the total S/N is

$$(S/N)_{\text{tot}} = \left( \sum_i \frac{P_{\text{CII}}^2(k_i)}{\sigma_p^2(k_i)} \right)^{1/2}, \quad (45)$$

where the sum is performed for all  $k$ -bins.

If the instrumental noise flux is  $\sigma_N$ , the instrumental noise power spectrum is

$$P_N = \left( \frac{\sigma_N}{\Omega_{\text{beam}}} \right)^2 V_{\text{vox}}, \quad (46)$$

where

$$V_{\text{vox}} \approx D_c^2(z) \Omega_{\text{beam}} \left[ \frac{c(1+z) \delta v_0}{H(z) v_0} \right]. \quad (47)$$

is the comoving volume of the real space voxel.

Suppose the survey is carried out by an interferometer array with  $N_{\text{an}}$  antennas, each antenna has diameter  $d$  and system temperature  $T_{\text{sys}}$ , then the instrumental noise flux

$$\sigma_N = \frac{2k_B T_{\text{sys}}}{d^2 \sqrt{N_{\text{an}}(N_{\text{an}} - 1)/2\delta v_0 t_{\text{int}}}}, \quad (48)$$

where  $t_{\text{int}}$  is the integration time on source. Suppose each antenna has a primary beam

$$\Omega_{\text{primary}} \sim \left( \frac{\lambda_{\text{obs}}}{d} \right)^2, \quad (49)$$

and  $N_{\text{chn}}$  frequency channels. Then for each instantaneous pointing the array can survey a fraction  $f_{\text{survey}} = \frac{\Omega_{\text{primary}} N_{\text{chn}} \delta v_0}{\Omega_{\text{survey}} \Delta v_0}$  ( $N_{\text{chn}} \delta v_0 \leq \Delta v_0$ ) of the target volume. Therefore,

$$\begin{aligned} t_{\text{int}} &= t_{\text{obs}} f_{\text{survey}} \\ &= t_{\text{obs}} \frac{\Omega_{\text{primary}} N_{\text{chn}} \delta v_0}{\Omega_{\text{survey}} \Delta v_0} \quad (N_{\text{chn}} \delta v_0 \leq \Delta v_0), \end{aligned} \quad (50)$$

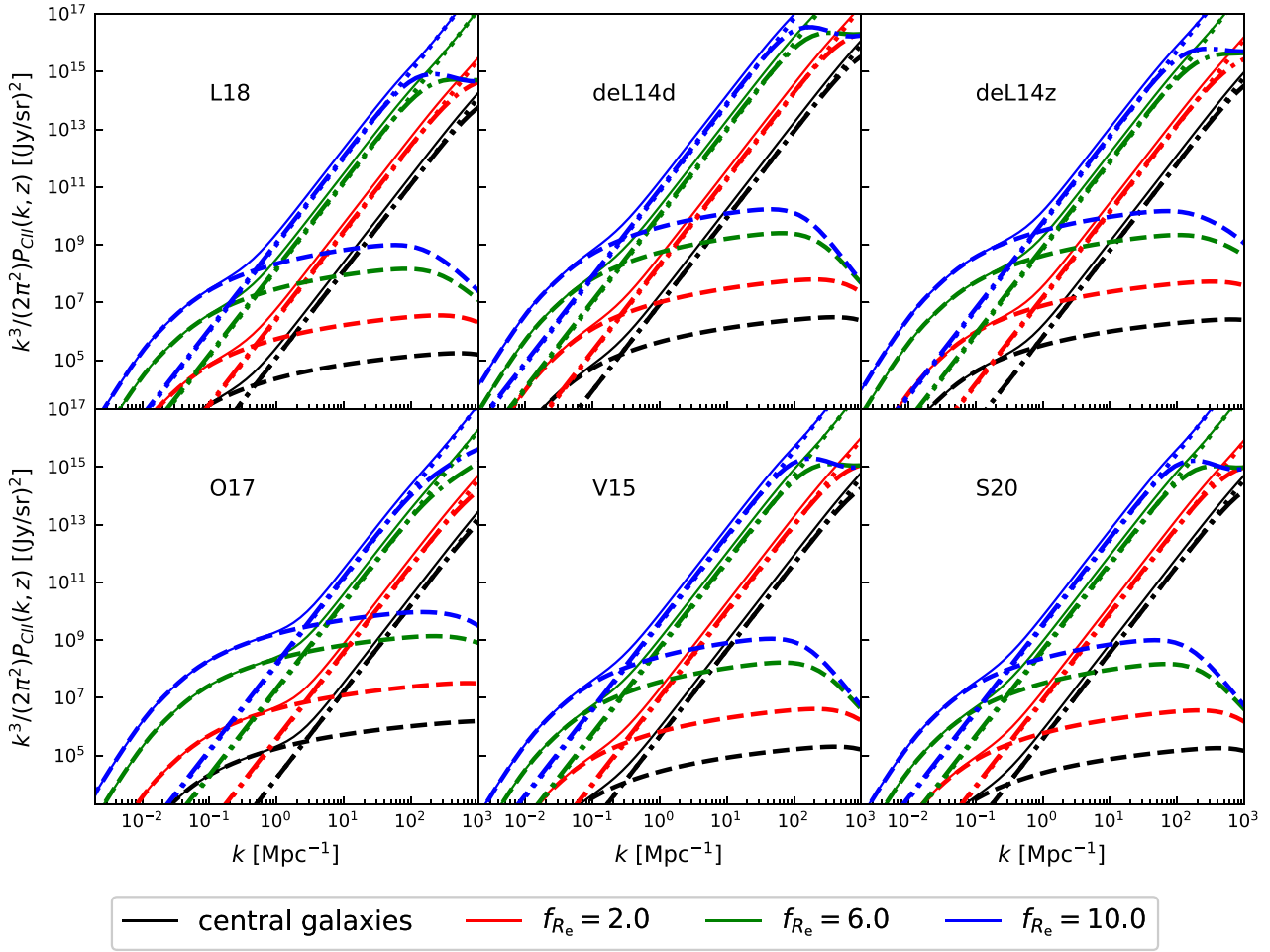
where  $t_{\text{obs}}$  is the total observation time.

If the interferometer has maximum baseline length  $b_{\text{max}}$ , then the Full Width at Half Maximum (FWHM) of the beam is  $\theta_{\text{FWHM}} = 1.2\lambda_{\text{obs}}/b_{\text{max}}$ , for Gaussian profile  $\theta_{\text{beam}} = \theta_{\text{FWHM}}/\sqrt{2 \ln 2}$ , hence  $\Omega_{\text{beam}} = (\theta_{\text{beam}})^2 = (\theta_{\text{FWHM}}/\sqrt{2 \ln 2})^2$ .

### 3 RESULTS

#### 3.1 [C II] power spectrum

In Fig. 3, we plot the intensity mapping power spectrum at  $z \sim 6$  for the six  $L_{\text{SFR, CII}} - \text{SFR}$  relations with effective radius ratio values  $f_{R_e} = [2.0, 6.0, 10.0]$ , representing conservative, moderate, and extreme cases, respectively. We fix  $f_{\Sigma} = 0.4$ , i.e. the central surface brightness of the halo represents  $\sim 30$  per cent of the total, which is a reasonable guess from the fitting results of the observational data (Akins et al. 2022). It clearly shows that [C II] haloes



**Figure 3.** Predicted [C II] power spectrum for six [C II]–SFR relations. The dash–dotted line, dashed line, dotted line, and thin solid line denote the one-halo, two-halo, shot noise terms, and the sum of them, respectively, in both panels. The dotted line (shot noise) overlaps with the dash–dotted line (one-halo term) at a large scale. The black line represents the [C II] power spectrum for the central galaxy, i.e. without [C II] halo considered. Red, green, and blue lines corresponding to  $f_{R_e} = [2.0, 6.0, 10.0]$ , with fixed  $f_{\Sigma} \sim 0.4$ .

largely boost the intensity of the power spectrum. Compared with the power spectrum of central galaxies, the signal is boosted by  $\sim 20$ ,  $10^3$ , and  $10^4$  times when  $f_{R_e} = 2.0, 6.0, 10.0$  independently on the specific form of  $L_{\text{CII,SFR}} - \text{SFR}$  relations. Moreover, [C II] haloes imprint a specific signature in the one-halo term at small scales. Since the central galaxy is more compact, the typical turnover scale of the one-halo term is  $\sim 10^3 \text{ Mpc}^{-1}$ . However, the [C II] halo is more extended and the typical turnover scale is shifted to  $\sim 100 \text{ Mpc}^{-1}$ . However, this feature is generally buried in the shot noise, which is the generally dominant component at the smallest scales. Fig. 4 illustrates instead the power spectrum dependence on  $f_{\Sigma}$  at a fixed value of  $f_{R_e} = 6.0$ . We show the results for  $f_{\Sigma} = 0.1$  and  $0.4$ , where the signal is strengthened by  $\sim 10^2$  and  $10^3$  times, respectively.

We also explore the results in the  $f_{R_e} - f_{\Sigma}$  parameter space, as shown in Fig. 5. We plot the values of the total power spectrum,  $P_{\text{CII}}$ , at  $k = 0.3 \text{ Mpc}^{-1}$  by varying  $f_{\Sigma}$  and  $f_{R_e}$  simultaneously. To avoid repetition, we only show the result for the L18 relation. As we can see,  $P_{\text{CII}}$  is sensitive to both changes in  $f_{R_e}$  and in  $f_{\Sigma}$ .

### 3.2 Detectability of the [C II] halo signal

We now turn the discussion to the detectability of the predicted [C II] power spectrum signal in a LIM survey.

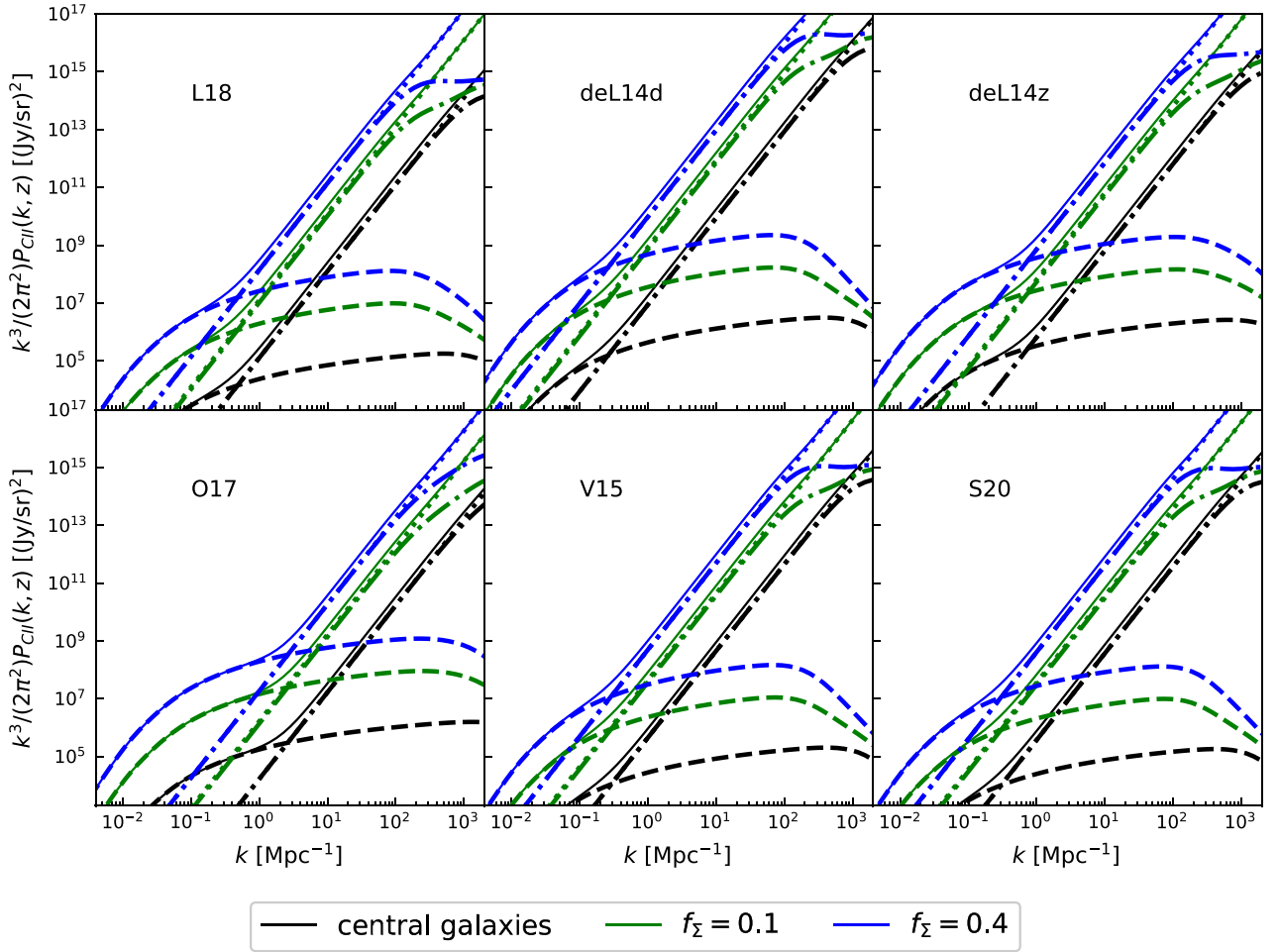
#### 3.2.1 Detectability in ALMA ASPECS Survey

The ALMA Large Program ASPECS (Aravena et al. 2020) surveying the Hubble Ultra Deep Field (HUDF) provides the first full frequency scan in Band 6, corresponding to the frequency window for [C II] emission from  $6 < z < 8$  galaxies, we first consider the detectability of [C II] power spectrum signal by such an experiment.

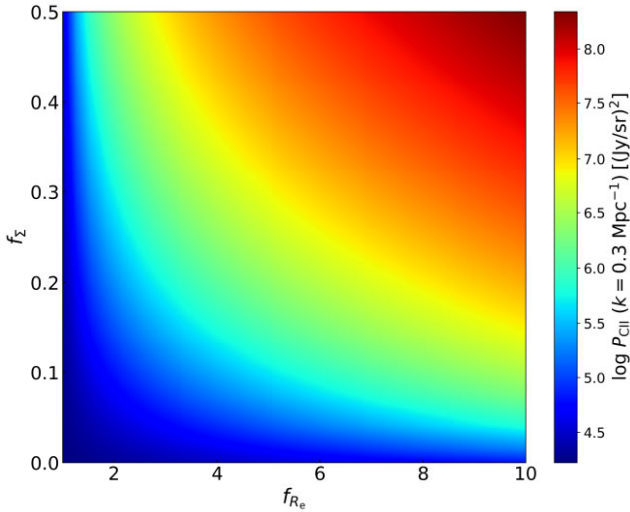
The ASPECS Band 6 data covers a total area of  $4.2 \text{ arcmin}^2$  in the HUDF, with  $2.9 \text{ arcmin}^2$  area within the 50 percent primary beam response (Decarli et al. 2020). The observed frequency range is 212–272 GHz, corresponding to the redshift range  $z = 5.99\text{--}7.97$  for the [C II] line. If we rebin the frequency channels by a factor of 8, as suggested by Uzgil et al. (2021), the spectral resolution is  $\delta\nu_0 = 62.5 \text{ MHz}$ . The synthesized beam size (the FWHMs of the pixel ellipse along its major and minor axes; see Uzgil et al. 2019) in the image cube is  $\Delta\theta_{\text{b,maj}} \times \Delta\theta_{\text{b,min}} = 1.6 \text{ arcsec} \times 1.1 \text{ arcsec}$  (Uzgil et al. 2021). Then

$$\Omega_{\text{beam}} = \frac{\theta_{\text{b,maj}} \times \theta_{\text{b,min}}}{2 \ln 2}. \quad (51)$$

From these survey parameters, we derive  $k_{\parallel,\text{min}} = 0.009 \text{ Mpc}^{-1}$ ,  $k_{\parallel,\text{max}} = 4.551 \text{ Mpc}^{-1}$ ,  $k_{\perp,\text{min}} = 1.503 \text{ Mpc}^{-1}$ , and  $k_{\perp,\text{max}} =$



**Figure 4.** As in Fig. 3 for different values of  $f_{\Sigma}$  and for the six different [C II]–SFR relations considered (see Table 1); we fix  $f_{R_e} = 6.0$  here.



**Figure 5.** The total [C II] halo power spectrum at  $k = 0.3 \text{ Mpc}^{-1}$  as a function of  $f_{\Sigma}$  and  $f_{R_e}$ , to avoid repetition we only show the result for L18 [C II]–SFR relation.

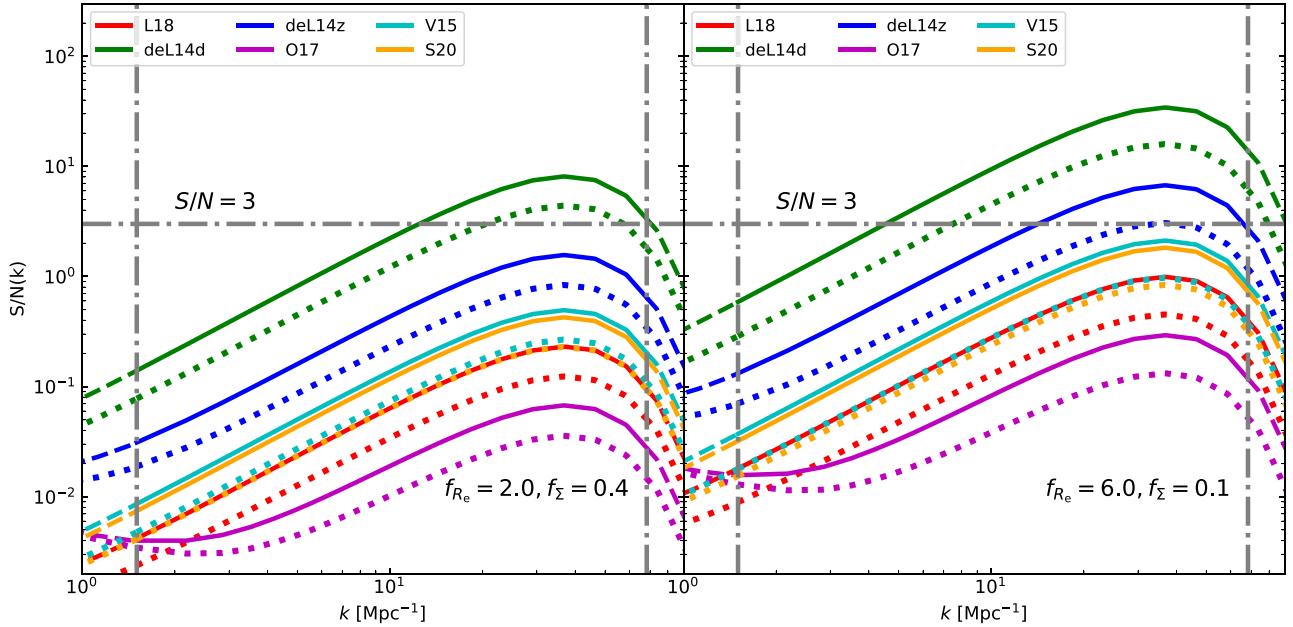
$68.147 \text{ Mpc}^{-1}$ , and we finally have  $k_{\min} = 1.503 \text{ Mpc}^{-1}$ ,  $k_{\max} = 68.299 \text{ Mpc}^{-1}$ . We summarize the ASPECS survey parameters derived for the [C II] power spectrum analysis in Table 2.

**Table 2.** The ASPECS survey parameters for S/N analysis.

$[\nu_0 - \Delta\nu_0, \nu_0]$	[212–272] GHz
$z$	[5.99, 7.97]
$\delta\nu_0$	62.5 MHz
$\Omega_{\text{survey}}$	2.9 arcmin <sup>2</sup>
$\Omega_{\text{beam}}$	1.6 arcsec $\times$ 1.1 arcsec
$\sigma_N$	0.30 mJy beam <sup>-1</sup>
$V_{\text{survey}}$	13 260 Mpc <sup>3</sup>
$\Delta r_{\parallel, \min}$	0.690 Mpc
$\Delta r_{\parallel, \max}$	662.701 Mpc
$\Delta r_{\perp, \min}$	0.046 Mpc
$\Delta r_{\perp, \max}$	4.180 Mpc
$k_{\parallel, \min}$	0.009 Mpc <sup>-1</sup>
$k_{\parallel, \max}$	4.551 Mpc <sup>-1</sup>
$k_{\perp, \min}$	1.503 Mpc <sup>-1</sup>
$k_{\perp, \max}$	68.299 Mpc <sup>-1</sup>
$k_{\min}$	1.503 Mpc <sup>-1</sup>
$k_{\max}$	68.299 Mpc <sup>-1</sup>
$P_N$	$1.48 \times 10^{11} \text{ (Jy/sr)}^2 \text{ Mpc}^3$

Next, we compute the S/N of the [C II] power spectrum for an ASPECS-like survey. The survey has noise flux  $\sigma_N = 0.30 \text{ mJy beam}^{-1}$  (Uzgil et al. 2021), which yields a surface brightness intensity sensitivity  $\sigma_N/\Omega_{\text{beam}} = 1.01 \times 10^7 \text{ Jy sr}^{-1}$ , and a noise





**Figure 6.**  $S/N$  for six different  $L_{\text{CII,SFR}}\text{--SFR}$  relations for the ALMA ASPECS Survey. The solid lines represent the  $S/N$  for the total power spectrum (including contributions from one-halo and two-halo terms as well as shot noise), while the dotted lines illustrate the  $S/N$  when the shot noise is excluded. The  $k_{\text{min}}$  and  $k_{\text{max}}$  are shown as grey dashed-dotted vertical lines. *Left panel:* case with  $f_{R_e} = 2.0$ ,  $f_{\Sigma} = 0.4$ . *Right panel:* same for  $f_{R_e} = 6.0$ ,  $f_{\Sigma} = 0.1$ .

power spectrum  $P_N = 1.48 \times 10^{11} \text{ (Jr/sr)}^2 \text{ Mpc}^3$  using equation (46).

We show the results of  $S/N$  for various  $L_{\text{CII,SFR}}\text{--SFR}$  relations with  $f_{\Sigma} = 0.4$ ,  $f_{R_e} = 2.0$  (left panel, represents a relatively compact [C II] halo) and  $f_{\Sigma} = 0.1$ ,  $f_{R_e} = 6.0$  (right panel, represents a relatively diffuse [C II] halo), respectively in Fig. 6. In the left panel, only the signal of the deL14d [C II]–SFR model is detectable (with  $S/N \gtrsim 3$ ) in the range  $10 \text{ Mpc}^{-1} \lesssim k \lesssim 70 \text{ Mpc}^{-1}$ . In the right panel, the model deL14d is detectable in  $4 \text{ Mpc}^{-1} \lesssim k \lesssim 70 \text{ Mpc}^{-1}$ , while the model deL14z is detectable in  $10 \text{ Mpc}^{-1} \lesssim k \lesssim 70 \text{ Mpc}^{-1}$ . This is because although the surface brightness of the [C II] halo is lower in the right panel, the total luminosity is larger as it scales with  $f_{R_e}^2$ . We additionally demonstrate the impact of the shot noise on  $S/N$  by dotted lines. When shot noise is excluded, the resulting  $S/N$  is reduced by a factor of  $\sim 2$ . This is because the ASPECS has a small area. The detected power spectrum is basically the one-halo clustering term plus the shot noise, the two-halo clustering term is almost negligible. While the shot noise is comparable to the one-halo clustering term, excluding it will reduce the shot noise by a factor  $\sim 2$ .

### 3.2.2 Optimal survey strategy

To optimally probe the extended [C II] halo signal, an ideal survey should be able to detect the signal with a resolution up to scale comparable to the extended [C II] halo size. On the other hand, to enhance the statistical significance of the signal, the survey should cover a sky area much larger than the ASPECS. Here we propose, by using ALMA 12-m antennas in an extended configuration with  $\sim 1000 \text{ m}$  baseline,  $N_{\text{an}} = 500$  and  $N_{\text{chn}} = 1000$ , to survey a total area of  $\Omega_{\text{survey}} \sim 2 \text{ deg}^2$  in frequency band [212–272] GHz with total observing time  $t_{\text{obs}} \sim 1000 \text{ hr}$ . The frequency resolution  $\delta\nu_0 = \Delta\nu_0/N_{\text{chn}} = 60 \text{ MHz}$ .

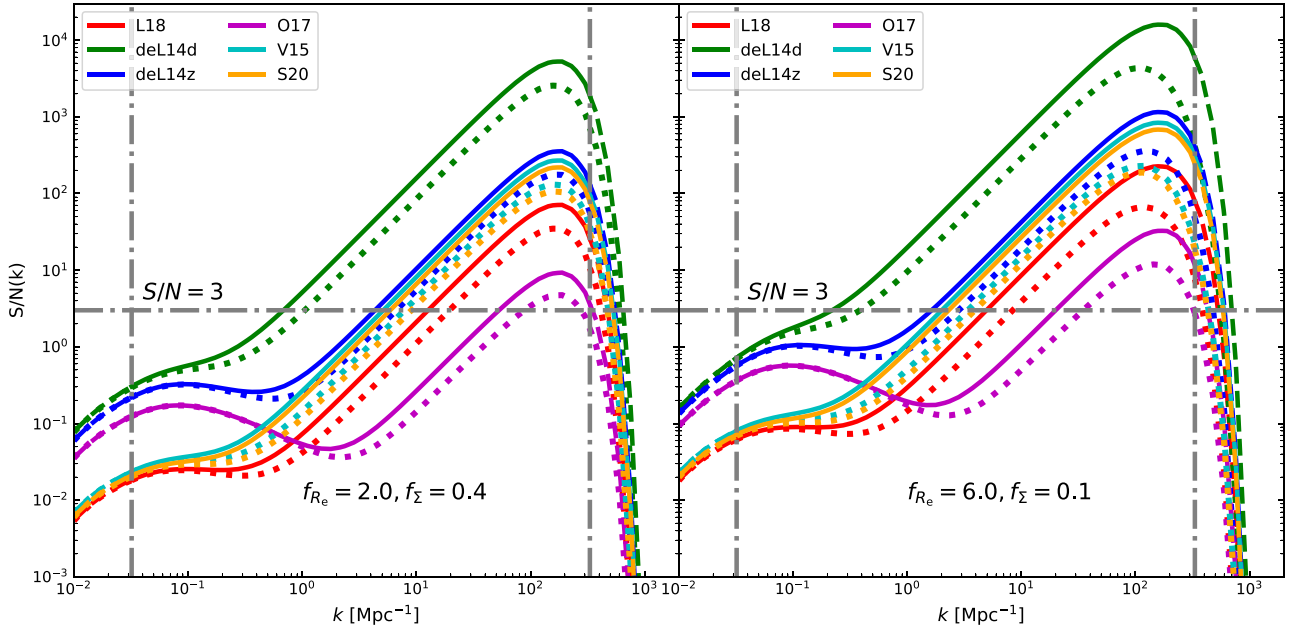
The system temperature<sup>5</sup> of ALMA at observing frequency  $\nu_0 = 272 \text{ GHz}$  is  $T_{\text{sys}} \sim 115 \text{ K}$ . Following the procedures in Section 2.4, we calculate the  $S/N$  and show them in Fig. 7. A summary of the parameters for this designed survey is given in Table 3.

For this optimized survey, when  $f_{\Sigma} = 0.4$  and  $f_{R_e} = 2.0$  (see the left panel of Fig. 7), the [C II] power spectra of the six  $L_{\text{CII,SFR}}\text{--SFR}$  relations all detectable with total  $S/N \gtrsim 10$ . For the more extended [C II] halo model with  $f_{\Sigma} = 0.1$  and  $f_{R_e} = 6.0$  (see the right panel of Fig. 7), the total  $S/N$ s of the power spectral are even higher. Similar to Fig. 6, we show the  $S/N$  if the shot noise is not involved by dotted lines. Different from the Fig. 6, when shot noise is excluded, the  $S/N$  does not change at  $k \lesssim 0.1 \text{ Mpc}^{-1}$ . However, at  $k \gtrsim 0.1 \text{ Mpc}^{-1}$  where the shot noise dominates over the clustering term, the  $S/N$  is reduced by a factor  $\sim 2$ , similar to Fig. 6. This is natural since at  $k \gtrsim 0.1 \text{ Mpc}^{-1}$ , the shot noise term is comparable to the one-halo clustering term.

### 3.3 Constraining the [C II] halo parameters

We have shown that when the contribution from extended [C II] haloes is considered, the [C II] power spectrum is boosted. However, this is mainly because the total [C II] luminosity of the system (central galaxy + [C II] halo) is larger. So there is a degeneracy between the [C II] halo contribution and the [C II]–SFR relation. [C II] haloes also change the shape of the [C II] power spectrum (one-halo term) at scales comparable to their size. However, the scales are so small, that generally, the shot noise is much larger than the one-halo term. Here, we investigate whether our designed optimal survey is able to put some conclusive constraints on the [C II] halo properties.

<sup>5</sup> $T_{\text{sys}}$  comes from the ALMA Sensitivity Calculator by setting the observing frequency  $\nu_0 = 272 \text{ GHz}$  and bandwidth  $\Delta\nu_0 = 60 \text{ GHz}$ .



**Figure 7.** As in Fig. 6 but for the optimal survey strategy by using ALMA. With 1000 m baseline, the largest  $k$  mode can be probed is  $k_{\max} = 331.43 \text{ Mpc}^{-1}$ . The signal for the six  $L_{\text{CII,SFR}}\text{--SFR}$  relations can be detected by this configuration in both two cases.

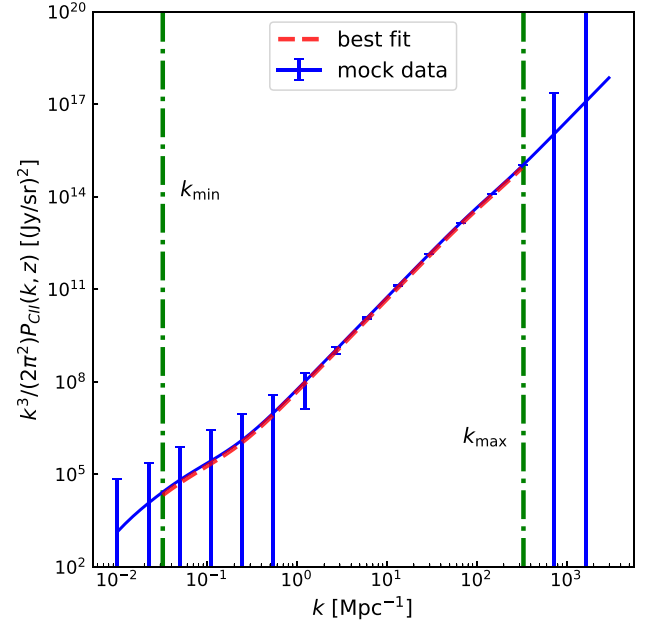
**Table 3.** Parameters of a designed optimal survey.

$d$	12 m
$T_{\text{sys}}$	115 K
$N_{\text{an}}$	500
$N_{\text{chn}}$	1000
$b_{\text{max}}$	1000 m
$\theta_{\text{beam}}$	0.232 arcsec
$[\nu_0 - \Delta\nu_0, \nu_0]$	[212,272] GHz
Redshift range	[6.0–8.0]
$\delta\nu_0$	60 MHz
$t_{\text{obs}}$	1000 h
$\Omega_{\text{survey}}$	2 deg <sup>2</sup>
$V_{\text{survey}}$	$2.88 \times 10^7 \text{ Mpc}^3$
$V_{\text{vox}}$	$5.96 \times 10^{-5} \text{ Mpc}^3$
$P_{\text{N}}$	$4.87 \times 10^{11} (\text{Jy sr}^{-1})^2 \text{ Mpc}^3$
$k_{\min}$	$0.032 \text{ Mpc}^{-1}$
$k_{\max}$	$331.43 \text{ Mpc}^{-1}$

We first generate the mock observed power spectrum  $P_{\text{CII}}^{\text{obs}}$  for our designed optimal survey from equation (26), by adopting  $\log A = 6.5$ ,  $\gamma = 1.2$ ,  $f_{\Sigma} = 0.1$ , and  $f_{R_e} = 6.0$  as the input [C II] halo parameters. Since in this paper we focus on constraining the [C II] halo parameters, we fix the cosmological parameters to reduce the freedoms. We adopt the cosmological parameters given by Planck Collaboration (2016). The mock uncertainties are from equation (40). We then obtain the forecast on constraining the [C II] halo parameters by minimizing

$$\chi^2(\Theta) = \sum_{i=1}^{N_{\text{bin}}} \frac{[P_{\text{CII}}(k_i, \Theta) - P_{\text{CII}}^{\text{obs}}(k_i)]^2}{\sigma_P^2(k_i)} \quad (52)$$

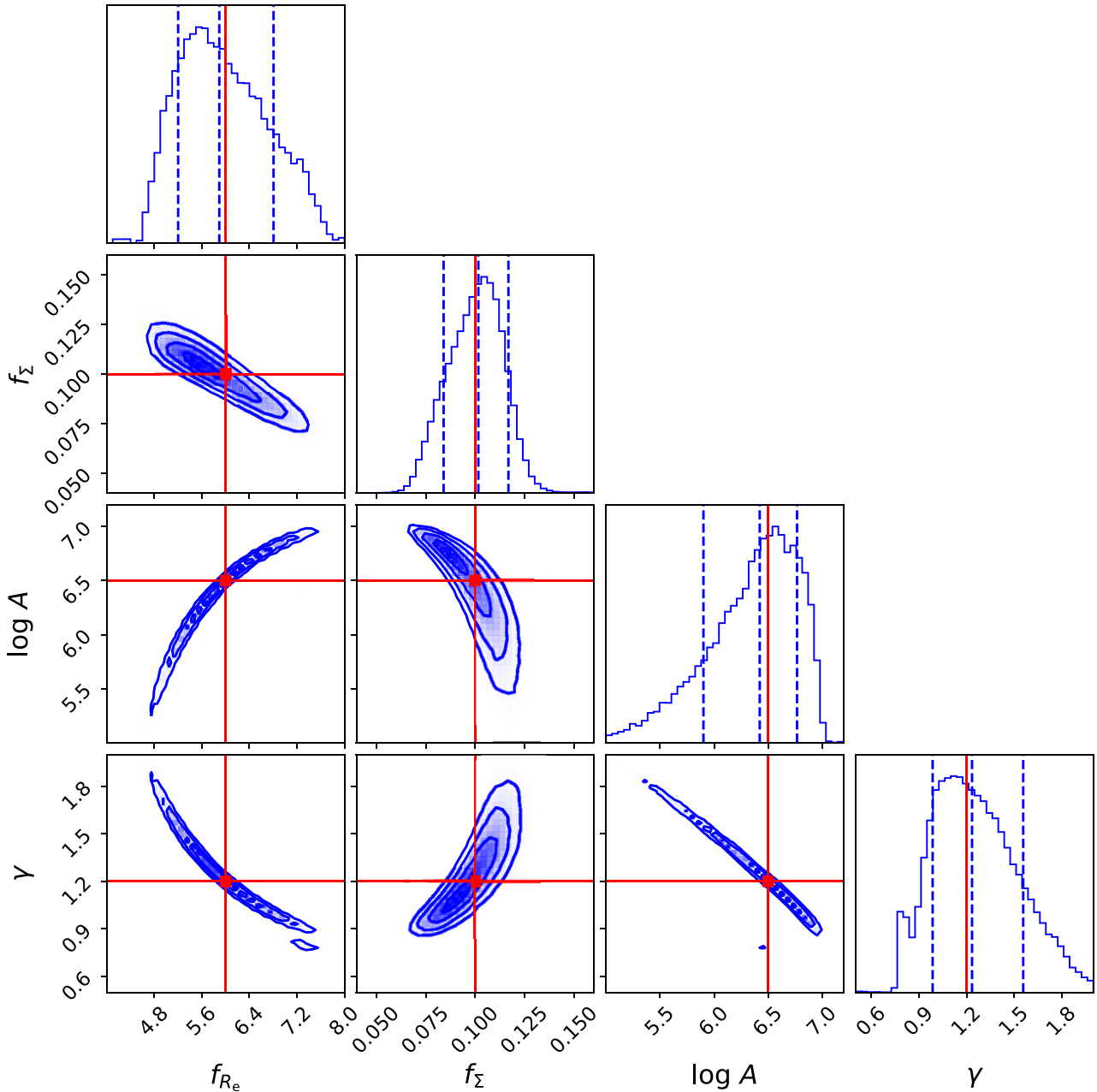
using the MCMC procedure (Goodman & Weare 2010), where  $\Theta = \{\log A, \gamma, f_{\Sigma}, f_{R_e}\}$  represents the  $L_{\text{CII,SFR}}\text{--SFR}$  relation and extended [C II] halo parameters;  $N_{\text{bin}}$  is the number of  $k$  bins between  $k_{\min}$



**Figure 8.** The mock observed power spectrum  $P_{\text{CII}}^{\text{obs}}(k)$  (solid line with error bars) and the best fitting power spectrum in  $k_{\min} < k < k_{\max}$  (dashed line). We mark  $k_{\min}$  and  $k_{\max}$  by two vertical dashed–dotted lines.

and  $k_{\max}$ , with width  $\ln k = 0.2$ . When performing the MCMC procedure, we set flat priors  $\log A \in [5.0, 9.0]$ ,  $\gamma \in [0.1, 2.0]$ ,  $f_{\Sigma} \in [0.0, 0.5]$ ,  $f_{R_e} \in [1.0, 10.0]$ .

In Fig. 8, we show the mock observed power spectrum with uncertainties and the best-fitting curve obtained from our MCMC fitting. The forecast on the constraints of [C II] halo parameters for our designed optimal survey is shown in Fig. 9. The marginalized parameters for [C II] halo are:  $f_{\Sigma} = 0.102^{+0.015}_{-0.018}$ ,  $f_{R_e} = 5.894^{+0.912}_{-0.693}$ . Indeed, it is feasible to distinguish the extended [C II] halo contribu-



**Figure 9.** The forecast of confidence levels of the  $L_{\text{CII,SFR}}$ –SFR relation and [C II] halo parameters obtained by the optimized survey. We mark the input parameters by the crosses of vertical and horizontal lines.

tion to the power spectrum in this designed optimal survey. However, we also check that the ASPECS survey is still hard to give conclusive constraints, because of the smaller S/N and lower resolution.

We also perform a MCMC fitting for which we set the shot noise as an independent free parameter, instead of a function of  $\Theta$ . We obtain  $f_{\Sigma} = 0.216^{+0.152}_{-0.109}$  and  $f_{R_e} = 5.917^{+0.683}_{-0.686}$ . The constraints become looser. This is not surprising, because treating the shot noise as an independent parameter increases the degree of freedom. Moreover, in our case, the one-halo term contributes most to the constraints on [C II] halo parameters because it contains the information of [C II] luminosity profile. However, the shot noise, if expressed as a function of  $\Theta$ , also provides some constraints on  $\Theta$ . Shot noise

is also a kind of useful signal, although using this term solely will suffer from heavy parameter degeneracy.

#### 4 CONCLUSIONS

In this paper, we predicted the foreground- and contamination-cleaned [C II] power spectrum signal at  $z \approx 6$ , when both central galaxies and extended [C II] haloes are considered, and investigate the detectability of such signal using ALMA ASPECS survey and a designed optimized survey. We modelled the [C II] luminosity profiles of the central galaxy and extended [C II] halo by a Sérsic + exponential profile, and derive the [C II]–dark matter halo mass relation by

matching the dark matter halo mass function with dust-corrected UV luminosity function of high-redshift galaxies. The main results are:

(i) The extended [C II] haloes around high redshift galaxies can significantly enhance the LIM signal compared with the signal produced by central galaxies alone, both in terms of the clustering signal and shot noise.

(ii) Our [C II] halo model has two free parameters: the effective [C II] halo/galaxy radius ratio,  $f_{R_e}$ , and the central surface brightness ratio,  $f_{\Sigma}$ . The luminosity of extended [C II] haloes is  $\sim 2.24 f_{\Sigma} f_{R_e}^2$  times the central galaxies. When the [C II] halo contribution is included, compared with the power spectrum from central galaxies only, the signal is boosted by a factor from  $\sim 20$  to  $\sim 10^3$  when  $f_{R_e}$  varies between 2.0 and 6.0, if  $f_{\Sigma} = 0.4$ . Given  $f_{R_e} = 6.0$ , the power spectrum is enhanced by  $\sim 100$  to  $\sim 10^3$  times when  $f_{\Sigma}$  changes from 0.1 to 0.4.

(iii) For a LIM experiment configured as the ALMA ASPECS Large Program (with resolution  $\theta_{\text{beam}} = 1.13$  arcsec and survey area  $\Omega_{\text{survey}} = 2.9$  arcmin<sup>2</sup>), the [C II] power spectrum signal is detectable ( $S/N \gtrsim 3$ ) for the deL14d  $L_{\text{CII,SFR}}\text{--SFR}$  relation with  $f_{\Sigma} = 0.4$  and  $f_{R_e} = 2.0$ , and for the deL14d, deL14z relations with  $f_{\Sigma} = 0.1$  and  $f_{R_e} = 6.0$ . So for a LIM experiment, the signal of more extended, low surface brightness [C II] haloes is more easily detected.

(iv) To optimally detect the signal, we proposed an optimized survey using more ALMA antennas and longer baselines. The survey has a higher resolution ( $\theta_{\text{beam}} \sim 0.232$  arcsec) and larger survey area ( $\Omega_{\text{survey}} \sim 2$  deg<sup>2</sup>). The resulting S/N is large enough so that the signal from six  $L_{\text{CII,SFR}}\text{--SFR}$  relations is detectable. We also predicted the confidence levels of the constraints on the  $L_{\text{CII,SFR}}\text{--SFR}$  relation and [C II] halo parameters by this optimized survey. These two halo parameters are degenerate, but we still expect to obtain meaningful constraints, with uncertainties of  $\sim 15$  per cent on  $f_{\Sigma}$  and  $\sim 10$  per cent on  $f_{R_e}$ .

In this work, we ignored the continuum foreground and the interloping lines, assuming they could be ideally removed and a clean [C II] signal plus pure instrumental white noise is obtained. In practice, however, this is quite a challenge. The IR continuum foreground from dust emission of extragalactic star-forming galaxies can be  $\gtrsim 10^3$  times higher than the [C II] LIM signal (e.g. Yue et al. 2015; Béthermin et al. 2022; Van Cuyck et al. 2023), while the interloping lines (for example the CO and [C I] lines from low redshift galaxies) can also be much higher than the [C II] line signal (e.g. Yue & Ferrara 2019; Béthermin et al. 2022; Van Cuyck et al. 2023). The Galactic dust emission and CMB also play a crucial role (Yue & Ferrara 2019). In our model, the [C II] LIM signal is enhanced because of the contribution from [C II] haloes. However, it is still far from enough to overcome the foreground/intervening lines. Many methods have been proposed to remove the foreground and interloping lines. We list some of them in Appendix B. We refer interested readers to the references therein.

To show the influence of foreground and interloping lines, one can make mock observations by simulations, and then apply the removal algorithms. However, it is beyond the scope of this work. Generally, if there are some residual foreground/intervening lines, particularly at a small scale, our constraints on the [C II] halo parameters will be biased.

Our predictions depend on the effective radius ratio, the central surface brightness ratio, and the choice of the  $L_{\text{CII,SFR}}\text{--SFR}$  relation. Therefore, detection of the [C II] halo signal in the LIM power spectrum can be used to verify whether [C II] haloes are ubiquitous in high-redshift galaxies, and provide information about [C II] halo size, and their possible relation with outflows carrying the emitting

material out of the main galaxy body. Finally, it can be potentially used to constrain the  $L_{\text{CII,SFR}}\text{--SFR}$  relation at high- $z$  as the amplitude of the power spectrum is sensitive to such quantity.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for the useful comments that helped us to improve this paper. This work is supported by the National SKA Program of China No. 2020SKA0110402. MZ acknowledges the financial support from the China Scholarship Council (CSC, No. 202104910322). AF acknowledges support from the ERC Advanced Grant INTERSTELLAR H2020/740120. Partial support from the Carl Friedrich von Siemens-Forschungspreis der Alexander von Humboldt-Stiftung Research Award is kindly acknowledged. We gratefully acknowledge the computational resources of the Center for High Performance Computing (CHPC) at SNS. We make use of SCIPY (Virtanen et al. 2020), NUMPY (Harris et al. 2020), MATPLOTLIB (Hunter 2007), EMCEE (Foreman-Mackey et al. 2013), and CORNER (Foreman-Mackey 2016) package for PYTHON to do the calculations and produce the plots.

## DATA AVAILABILITY

The data produced in this study are available from the corresponding author upon reasonable request.

## REFERENCES

- Akins H. B. et al., 2022, *ApJ*, 934, 64  
 Aravena M., Carilli C., Decarli R., Walter F., *ASPECS Collaboration.*, 2020, *The Messenger*, 179, 17  
 Baes M., Gentile G., 2011, *A&A*, 525, A136  
 Bakx T. J. L. C. et al., 2020, *MNRAS*, 493, 4294  
 Baldauf T., Seljak U., Smith R. E., Hamaus N., Desjacques V., 2013, *Phys. Rev. D*, 88, 083507  
 Battye R. A., Browne I. W. A., Dickinson C., Heron G., Maffei B., Pourtsidou A., 2013, *MNRAS*, 434, 1239  
 Bernal J. L., Kovetz E. D., 2022, *A&A Rev.*, 30, 5  
 Bernal J. L., Breyssse P. C., Gil-Marín H., Kovetz E. D., 2019, *Phys. Rev. D*, 100, 123522  
 Béthermin M. et al., 2022, *A&A*, 667, A156  
 Bigot-Sazy M. A. et al., 2015, *MNRAS*, 454, 3240  
 Binney J., Mamon G. A., 1982, *MNRAS*, 200, 361  
 Binney J., Tremaine S., 1987, *Galactic dyn.*. Princeton Univ. Press, Princeton  
 Bouwens R. J. et al., 2015, *ApJ*, 803, 34  
 Breyssse P. C., Kovetz E. D., Kamionkowski M., 2014, *MNRAS*, 443, 3506  
 Breyssse P. C., Kovetz E. D., Kamionkowski M., 2015, *MNRAS*, 452, 3408  
 Breyssse P. C., Kovetz E. D., Kamionkowski M., 2016, *MNRAS*, 457, L127  
 Breyssse P. C., Anderson C. J., Berger P., 2019, *Phys. Rev. Lett.*, 123, 231105  
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000  
 Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, *ApJ*, 803, 21  
 Capak P. L. et al., 2015, *Nature*, 522, 455  
 Carniani S. et al., 2017, *A&A*, 605, A42  
 Carniani S. et al., 2018, *MNRAS*, 478, 1170  
 Castellano M. et al., 2016, *ApJ*, 818, L3  
 Chabrier G., 2003, *PASP*, 115, 763  
 Chang T.-C., Pen U.-L., Bandura K., Peterson J. B., 2010, *Nature*, 466, 463  
 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  
 Chung D. T., Viero M. P., Church S. E., Wechsler R. H., 2020, *ApJ*, 892, 51  
 Comaschi P., Ferrara A., 2016, *MNRAS*, 455, 725  
 Cooray A., Sheth R., 2002, *Phys. Rep.*, 372, 1  
 Dayal P., Ferrara A., 2018, *Phys. Rep.*, 780–782, 1  
 De Looze I. et al., 2014, *A&A*, 568, A62



- Decarli R. et al., 2020, *ApJ*, 902, 110
- Eisenstein D. J., Hu W., 1999, *ApJ*, 511, 5
- Ferland G. J. et al., 2013, *RMxAA*, 49, 137
- Ferland G. J. et al., 2017, *RMxAA*, 53, 385
- Foreman-Mackey D., 2016, *J. Open Source Softw.*, 1, 24
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Fudamoto Y. et al., 2022, *ApJ*, 934, 144
- Fudamoto Y. et al., 2023, preprint (arXiv:2303.07513)
- Fujimoto S. et al., 2019, *ApJ*, 887, 107
- Fujimoto S. et al., 2020, *ApJ*, 900, 1
- Ginolfi M. et al., 2020, *A&A*, 633, A90
- Gong Y., Cooray A., Silva M. B., Santos M. G., Lubin P., 2011, *ApJ*, 728, L46
- Gong Y., Cooray A., Silva M., Santos M. G., Bock J., Bradford C. M., Zemcov M., 2012, *ApJ*, 745, 49
- Gong Y., Cooray A., Silva M. B., Zemcov M., Feng C., Santos M. G., Dore O., Chen X., 2017, *ApJ*, 835, 273
- Gong Y., Chen X., Cooray A., 2020, *ApJ*, 894, 152
- Goodman J., Weare J., 2010, *Commun. Appl. Math. Comput. Sci.*, 5, 65
- Harris C. R. et al., 2020, *Nature*, 585, 357
- Heintz K. E. et al., 2023, *A&A*, 678, A30
- Herrera-Camus R. et al., 2015, *ApJ*, 800, 1
- Herrera-Camus R. et al., 2021, *A&A*, 649, A31
- Hodge J. A., da Cunha E., 2020, *R. Soc. Open Sci.*, 7, 200556
- Hollenbach D. J., Tielens A. G. G. M., 1999, *Rev. Mod. Phys.*, 71, 173
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Karkare K. S., Moradinezhad Dizgah A., Keating G. K., Breyse P., Chung D. T., 2022, preprint (arXiv:2203.07258)
- Kennicutt Robert C. J., 1998, *ARA&A*, 36, 189
- Knudsen K. K., Richard J., Kneib J.-P., Jauzac M., Clément B., Drouart G., Egami E., Lindroos L., 2016, *MNRAS*, 462, L6
- Koprowski M. P. et al., 2018, *MNRAS*, 479, 4355
- Kovetz E. D. et al., 2017, preprint (arXiv:1709.09066)
- Kovetz E. et al., 2019, *BAAS*, 51, 101
- Lagache G., Cousin M., Chatzikos M., 2018, *A&A*, 609, A130
- Lambert T. S. et al., 2022, *MNRAS*, 518, 3183
- Le Fèvre O. et al., 2020, *A&A*, 643, A1
- Leung T. K. D., Olsen K. P., Somerville R. S., Davé R., Greve T. R., Hayward C. C., Narayanan D., Popping G., 2020, *ApJ*, 905, 102
- Li T. Y., Wechsler R. H., Devaraj K., Church S. E., 2016, *ApJ*, 817, 169
- Lidz A., Taylor J., 2016, *ApJ*, 825, 143
- Lidz A., Furlanetto S. R., Oh S. P., Aguirre J., Chang T.-C., Doré O., Pritchard J. R., 2011, *ApJ*, 741, 70
- Madden S. C. et al., 2013, *PASP*, 125, 600
- Makinen T. L., Lancaster L., Villaescusa-Navarro F., Melchior P., Ho S., Perreault-Levasseur L., Spergel D. N., 2021, *J. Cosmol. Astropart. Phys.*, 2021, 081
- Mashian N., Sternberg A., Loeb A., 2015, *J. Cosmol. Astropart. Phys.*, 2015, 028
- Matthee J., Sobral D., Gronke M., Pezzulli G., Cantalupo S., Röttgering H., Darvish B., Santos S., 2020, *MNRAS*, 492, 1778
- Mazure A., Capelato H. V., 2002, *A&A*, 383, 384
- Mitra S., Choudhury T. R., Ferrara A., 2015, *MNRAS*, 454, L76
- Moradinezhad Dizgah A., Nikakhtar F., Keating G. K., Castorina E., 2022, *J. Cosmol. Astropart. Phys.*, 2022, 026
- Moriwaki K., Yoshida N., 2021, *ApJ*, 923, L7
- Moriwaki K., Filippova N., Shirasaki M., Yoshida N., 2020, *MNRAS*, 496, L54
- Moriwaki K., Shirasaki M., Yoshida N., 2021, *ApJ*, 906, L1
- Olsen K. P., Greve T. R., Narayanan D., Thompson R., Toft S., Brinch C., 2015, *ApJ*, 814, 76
- Olsen K., Greve T. R., Narayanan D., Thompson R., Davé R., Niebla Rios L., Stawinski S., 2017, *ApJ*, 846, 105
- Pentericci L. et al., 2016, *ApJ*, 829, L11
- Pizzati E., Ferrara A., Pallottini A., Gallerani S., Vallini L., Decataldo D., Fujimoto S., 2020, *MNRAS*, 495, 160
- Pizzati E., Ferrara A., Pallottini A., Sommovigo L., Kohandel M., Carniani S., 2023, *MNRAS*, 519, 4608
- Planck Collaboration, 2016, *A&A*, 594, A13
- Prugniel P., Simien F., 1997, *A&A*, 321, 111
- Pullen A. R., Doré O., Bock J., 2014, *ApJ*, 786, 111
- Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, *ApJ*, 802, L19
- Salvaterra R., Maio U., Ciardi B., Campisi M. A., 2013, *MNRAS*, 429, 2718
- Schaan E., White M., 2021, *J. Cosmol. Astropart. Phys.*, 2021, 067
- Schaerer D. et al., 2020, *A&A*, 643, A3
- Schechter P., 1976, *ApJ*, 203, 297
- Seo H.-J., Dodelson S., Marriner J., Mcginnis D., Stebbins A., Stoughton C., Vallinotto A., 2010, *ApJ*, 721, 164
- Serra P., Doré O., Lagache G., 2016, *ApJ*, 833, 153
- Sérsic J. L., 1963, *Boletín de la Asociacion Argentina de Astronomia La Plata Argentina*, 6, 41
- Sérsic J. L., 1968, *Atlas de Galaxias Australes*
- Sheth R. K., Tormen G., 1999, *MNRAS*, 308, 119
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, 323, 1
- Shibuya T., Ouchi M., Harikane Y., 2015, *ApJS*, 219, 15
- Silva M., Santos M. G., Cooray A., Gong Y., 2015, *ApJ*, 806, 209
- Silva B. M., Zaroubi S., Kooistra R., Cooray A., 2018, *MNRAS*, 475, 1587
- Stacey G. J., Geis N., Genzel R., Lugten J. B., Poglitsch A., Sternberg A., Townes C. H., 1991, *ApJ*, 373, 423
- Sun G. et al., 2018, *ApJ*, 856, 107
- Uzgil B. D., Aguirre J. E., Bradford C. M., Lidz A., 2014, *ApJ*, 793, 116
- Uzgil B. D. et al., 2019, *ApJ*, 887, 37
- Uzgil B. D. et al., 2021, *ApJ*, 912, 67
- Vallini L., Gallerani S., Ferrara A., Pallottini A., Yue B., 2015, *ApJ*, 813, 36
- Van Cuyck M. et al., 2023, *A&A*, 676, A62
- Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
- Vitral E., Mamon G. A., 2020, *A&A*, 635, A20
- Wang X., Tegmark M., Santos M. G., Knox L., 2006, *ApJ*, 650, 529
- Wolfrum M. G., McKee C. F., Hollenbach D., Tielens A. G. G. M., 2003, *ApJ*, 587, 278
- Yue B., Ferrara A., 2019, *MNRAS*, 490, 1928
- Yue B., Ferrara A., Pallottini A., Gallerani S., Vallini L., 2015, *MNRAS*, 450, 3829
- Zhou X., Gong Y., Deng F., Zhang M., Yue B., Chen X., 2023, *MNRAS*, 521, 278

## APPENDIX A: PROFILE DEPROJECTION

The deprojection of the 2D surface density profile  $\Sigma(R)$  is the inversion of the Abel integral (Binney & Mamon 1982; Binney & Tremaine 1987)

$$\rho(r) = -\frac{1}{\pi} \int_r^{+\infty} \frac{d\Sigma}{dR} \frac{dR}{\sqrt{R^2 - r^2}}, \quad (\text{A1})$$

where  $r$  is the 3D radius,  $\rho(r)$  is the 3D density profile. For the Sérsic model of  $\Sigma(R)$ , the exact analytical expression of the above integral involves the Meijer G special function (Mazure & Capelato 2002; Baes & Gentile 2011) or Fox H function (Baes & Gentile 2011), both of them are complicated. Therefore, several analytical approximations are also proposed. They are described in detail in Vitral & Mamon (2020) and references therein. Here we choose to use the analytical approximation given by Prugniel & Simien (1997). In this model, one writes the dimensionless 3D density profile as

$$\tilde{\rho}(x) = \frac{b_n^{(3-p_n)n}}{n\Gamma[(3-p_n)n]} x^{-p_n} \exp[-b_n x^{1/n}] \quad (\text{A2})$$

where  $x = r/R_c$ ,  $p_n$  is a function depending on the index  $n$ ,

$$p_n = 1 - \frac{0.594}{n} + \frac{0.055}{n^2}. \quad (\text{A3})$$

Therefore, in our model, the 3D luminosity density profile for the central galaxies can be written as

$$\begin{aligned}\rho_{\text{CII,g}}(r) &= \frac{L_{\text{CII,g}}}{4\pi R_{\text{e,g}}^3} \tilde{\rho}(x) \\ &= \frac{L_{\text{CII,SFR}}}{4\pi R_{\text{e,g}}^3} \frac{(b_n)^{(3-p_n)n}}{n\Gamma[(3-p_n)n]} x^{-p_n} \exp[-b_n x^{1/n}].\end{aligned}\quad (\text{A4})$$

For the extended [CII] halo, since the exponential profile corresponds to the Sérsic model with  $n = 1$ , one can easily derive  $\rho_{\text{CII,h}}(r)$  by analogizing to equation (A4) and replacing  $L_{\text{CII,g}}$  with  $L_{\text{CII,h}}$ ,

$$\begin{aligned}\rho_{\text{CII,h}}(r) &= \frac{L_{\text{CII,h}}}{4\pi R_{\text{e,h}}^3} \frac{(b_1)^{(3-p_1)}}{\Gamma[(3-p_1)]} x^{-p_1} \exp[-b_1 x] \\ &= \frac{L_{\text{CII,SFR}}}{4\pi R_{\text{e,g}}^3} \frac{f_{\Sigma}}{f_{R_e}} \frac{1}{n} \frac{\Gamma(2)}{\Gamma(2n)} \frac{(b_n)^{2n} (b_1)^{(1-p_1)}}{\Gamma[(3-p_1)]} \\ &\quad \times x^{-p_1} \exp[-b_1 x].\end{aligned}\quad (\text{A5})$$

## APPENDIX B: CONTAMINATION REMOVAL

Here, we briefly introduce some foreground/interloping line removal methods for LIM in literature.

By taking the advantage that the foreground is generally smooth in frequency space (Wang et al. 2006), the most popular way is to fit the IR continuum of each line of sight by a polynomial, then subtract it from the line of sight spectrum to get the residual LIM signal (Yue et al. 2015).

Moreover, one can also separate the smooth continuum foreground and the fluctuating LIM signal through the standard principal component analysis (PCA; Bigot-Sazy et al. 2015; Yue et al. 2015; Van Cuyck et al. 2023) or the asymmetric re-weighted penalized least-squares method (Van Cuyck et al. 2023).

Recently, it has been demonstrated that deep learning has the potential to be a powerful tool for removing continuum foreground,

interloping lines, and noise, and finally reconstructing the 3D distribution of the target emission line (Moriwaki et al. 2020; Makinen et al. 2021; Moriwaki & Yoshida 2021; Moriwaki, Shirasaki & Yoshida 2021; Zhou et al. 2023).

If the LIM survey area overlaps with a galaxy survey, to migrate the foreground and interloping lines, one can also simply mask the voxels occupied by bright low redshift galaxies that are suspected contamination sources (Silva et al. 2015; Yue et al. 2015; Sun et al. 2018; Yue & Ferrara 2019; Béthermin et al. 2022; Van Cuyck et al. 2023), or blindly mask a small fraction of bright voxels (Breysse, Kovetz & Kamionkowski 2015).

For removing the interloping lines from low redshift galaxies, one can also take the advantage that their cylinder power spectrum is asymmetry if they are incorrectly identified as the target emission line (Cheng et al. 2016; Lidz & Taylor 2016; Yue & Ferrara 2019; Gong et al. 2020).

Breysse, Kovetz & Kamionkowski (2016) and Breysse, Anderson & Berger (2019) also proposed that the one-point statistics can migrate the continuum foreground and the interloping lines, and get the luminosity function of the target emission line.

Cheng, Chang & Bock (2020) proposed that, if multiple lines of sources are observed at multiple frequencies, then the target signal can be extracted by fitting the observations to a set of spectrum templates.

Finally, the cross-correlation between [CII] line and the 21 cm line can be used to distinguish the signal and contaminations (Gong et al. 2011; Silva et al. 2015), while the cross-correlation between foreground lines and galaxy surveys are also explored in Silva et al. (2015) in order to probe the intensity of the foregrounds.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.