



Universiteit  
Leiden  
The Netherlands

## Hidden star formation in the early Universe

Leeuwen, I. F. van

### Citation

Leeuwen, I. F. van. (2026, April 9). *Hidden star formation in the early Universe*. Retrieved from <https://hdl.handle.net/1887/4301034>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4301034>

**Note:** To cite this publication please use the final published version (if applicable).

# 2 | CHARACTERISING THE CONTRIBUTION OF DUST-OBSCURED STAR FORMATION AT $z \geq 5$ USING 18 SERENDIPITOUSLY IDENTIFIED [CII] EMITTERS



I. F. van Leeuwen, R. J. Bouwens, P. P. van der Werf, J. A. Hodge, S. Schouws, M. Stefanon, H. S. B. Algera, M. Aravena, L. A. Boogaard, R. A. A. Bowler, E. da Cunha, P. Dayal, R. Decarli, V. Gonzalez, H. Inami, I. de Looze, L. Sommovigo, B. P. Venemans, F. Walter, L. Barrufet, A. Ferrara, L. Graziani, A. P. S. Hygate, P. Oesch, M. Palla, L. Rowland and R. Schneider, *Monthly Notices of the Royal Astronomical Society*, 534, 2062, 2024

## Abstract

---

We present a new method to determine the star formation rate (SFR) density of the Universe at  $z \gtrsim 5$  that includes the contribution of dust-obscured star formation. For this purpose, we use a [CII] ( $158 \mu\text{m}$ ) selected sample of galaxies serendipitously identified in the fields of known  $z \gtrsim 4.5$  objects to characterise the fraction of obscured SFR. The advantage of a [CII] selection is that our sample is SFR-selected, in contrast to a UV-selection that would be biased towards unobscured star formation. We obtain a sample of 23 [CII] emitters near star-forming (SF) galaxies and QSOs – three of which we identify for the first time – using previous literature and archival ALMA data. 18 of these serendipitously identified galaxies have sufficiently deep rest-UV data and are used to characterise the obscured fraction of the star formation in galaxies with SFRs  $\gtrsim 30 M_{\odot} \text{ yr}^{-1}$ . We find that [CII] emitters identified around SF galaxies have  $\approx 63$  per cent of their SFR obscured, while [CII] emitters around QSOs have  $\approx 93$  per cent of their SFR obscured. By forward modeling existing wide-area UV luminosity function (LF) determinations, we derive the intrinsic UV LF using our characterization of the obscured SFR. Integrating the intrinsic LF to  $M_{\text{UV}} = -20$  we find that the obscured SFRD contributes to  $> 3\%$  and  $> 10\%$  of the total SFRD at  $z \sim 5$  and  $z \sim 6$  based on our sample of companions galaxies near SFGs and QSOs, respectively. Our results suggest that dust obscuration is not negli-

gible at  $z \gtrsim 5$ , further underlining the importance of far-IR observations of the  $z \gtrsim 5$  Universe.

## 2.1 Introduction

One of the most exciting frontiers in extragalactic astronomy is understanding the build-up of stars in galaxies during the epoch of reionization. A common way to characterise this early build-up is through the volume-averaged star formation rate density (SFRD) of the Universe. Thanks to a wealth of multi-wavelength observations from both space-based and ground-based observatories, the SFRD at  $z \leq 3$  has already been characterised in detail (e.g., Madau & Dickinson 2014). The general consensus is that the star formation rate increases with cosmic time until it reaches a peak at  $z \sim 2$ , the so-called ‘Cosmic Noon’, and then decreases again to the present epoch.

Reaching further back in time to  $z > 3$ , observations that are available to constrain the SFRD have been mostly limited to the rest-UV (e.g., Madau & Dickinson 2014). However, far-IR probes are necessary due to the impact of dust attenuation in substantially obscuring UV light from bright star-forming (SF) galaxies in the early Universe. Surveying the Universe in the far-IR has been challenging due to the limited spatial resolution of earlier wide-area probes resulting in substantial source confusion from lower-redshift sources (e.g. Nguyen et al. 2010; Everett et al. 2020). Moreover, higher spatial resolution probes such as are now available with the Atacama Large Millimeter/submillimeter Array (ALMA) are limited by their relative small field-of-view (FOV) (e.g., Hodge & da Cunha 2020).

While some of the debate in the literature regarding the evolution of the SFRD has focused on the evolution at  $z > 8$  (e.g. McLeod et al. 2016; Oesch et al. 2018), substantial uncertainties also exist at  $z \sim 5-8$  from the contribution of dust-obscured galaxies. For example, several recent results (e.g. Casey et al. 2018b; Gruppioni et al. 2020; Fudamoto et al. 2020b; Zavala et al. 2021; Khusanova et al. 2021; Loiacono et al. 2021; Talia et al. 2021) have suggested that the obscured contribution of the SFRD at high redshift could be as large as  $\sim 20-60\%$  of the total SFRD. Moreover, recent work by Algera et al. (2023) has demonstrated that the dust-obscured star formation at  $z \sim 7$  could contribute as much as  $\sim 30-50\%$  of the SFRD, significantly more than earlier estimated based on UV data alone. There are many examples of galaxies at  $z \sim 7$  that are  $\gtrsim 90\%$  obscured both in observations (e.g. Marrone et al. 2018; Hygate et al. 2023) and in theoretical models (e.g. Ferrara et al. 2022).

As previously discussed, progress in quantifying the fraction of obscured star formation in the early Universe has been challenging due to the limited spatial resolution of earlier probes of far-IR light and the limited areas surveyed by ALMA. Recently, there have been efforts to overcome ALMA’s

small field-of-view and survey much wider areas. Perhaps the widest area examples of such surveys are MORA (2 mm Mapping Obscuration to Reionization with ALMA: Zavala et al. 2021; Casey et al. 2021) and ex-MORA (Long et al., in prep.), covering  $184 \text{ arcmin}^2$  and  $0.2 \text{ deg}^2$ , respectively, as well as the substantial A<sup>3</sup>COSMOS effort (Liu et al. 2019) where a  $280 \text{ arcmin}^2$  survey area is constructed by combining the significant amount of ALMA observations over the COSMOS field.<sup>1</sup> Wide-area surveys such as these can provide us with a useful unbiased probe of the prevalence of dusty star-forming galaxies in the distant Universe. Nevertheless, an important disadvantage to this approach is also the large observational cost requiring not only wide area surveys in the far-IR, but also a pursuit of line scans to determine redshifts for the identified dusty sources.

Here we present a new method to characterise the obscured star formation in the  $z > 4$  Universe, using sensitive ALMA observations targeting bright ISM cooling lines to search for star-forming galaxies in the neighborhoods of bright QSOs and massive galaxies. Bright star-forming galaxies are readily identifiable via line emission in bright ISM cooling lines such as [CII] ( $157.74 \mu\text{m}$ ). Given the clustering of galaxies, line emission from dusty star-forming galaxies tends to be found close to the targeted sources not only in frequency, but also in angular position on the sky. We will therefore also refer to these serendipitously identified galaxies as ‘companion’ galaxies. Previous studies by e.g. Decarli et al. (2017) and Venemans et al. (2020) have shown that quasars are especially good places to look for bright star-forming companions, with  $\sim 50\%$  of the quasars they examined showing at least one companion galaxy (see also Meyer et al. 2022). Massive galaxies ( $M_* \gtrsim 10^{8.5} M_\odot$ ) at high redshifts are also excellent targets to use for efficient searches for companion galaxies (Loiacono et al. 2021; Fudamoto et al. 2021). To date about 40 companion galaxies have been discovered at  $z \gtrsim 4$  with ALMA observations (e.g., Decarli et al. 2017; Trakhtenbrot et al. 2017; Miller et al. 2020; Venemans et al. 2020; Loiacono et al. 2021; Fudamoto et al. 2021).

An important advantage to identifying star-forming galaxies using far-IR line emission is the insensitivity to dust obscuration or the brightness of sources in the rest-frame UV. [CII] is an especially valuable line to use in this regard due to it being known to trace the total star formation rate in ‘normal’ star-forming galaxies (e.g. De Looze et al. 2014) and also thanks to its particular brightness. As a result, [CII]-selected sources can thus be selected entirely on the basis of their SFRs, and sensitive rest-UV observations of such sources can be used to determine the extent to which star formation in the high-redshift Universe is obscured or not. [CII] line emission is known to have multiple origins but is believed to arise mostly from photodissociation regions (PDRs) in high redshift galaxies (e.g. Hollenbach

---

<sup>1</sup><https://sites.google.com/view/a3cosmos>



& Tielens 1999; Wolfire et al. 2022). As [CII] has a low ionization potential (11.3 eV) and is therefore easy to excite, it traces a large fraction of the gas reservoir of a galaxy (Wolfire et al. 2003).

In demonstrating the power of our new methodology, we make use of a significant sample of 18 [CII]-selected companion galaxies to quantify the fraction of star formation in  $z \sim 4-7$  galaxies that is obscured by dust compared to their total SFR. These [CII]-emitting companion galaxies are identified in the neighborhoods of 161 [CII]-emitting galaxies and QSOs at  $z > 4$  and have deep enough rest-UV observations to determine if star formation in a source is obscured or not. These 18 galaxies have been found as companions to galaxies drawn from 75 [CII]-emitting galaxies in ALPINE (Le Fèvre et al. 2020; Béthermin et al. 2020), 25 [CII]-emitting galaxies in REBELS (Bouwens et al. 2022b; see also Schouws et al. 2024, in prep.), 27 [CII]-emitting QSOs studied by Venemans et al. (2020), and 34 [CII]-emitting galaxies and QSOs from the ALMA archive. Based on the fraction of obscured star formation (from comparing rest-UV and far-IR observations) in our [CII]-selected samples, we forward model wide-area UV luminosity function (LF) results to recover the intrinsic luminosity function. The intrinsic UV luminosity function differs from the observed LF in that it accounts for the impact of dust obscuration on both the brightness of galaxies in the rest-UV and their selectability. Finally, we compute a dust-corrected SFRD at  $z \gtrsim 5$ .

The organization of the paper is as follows. In Section 2.2, we present both the archival ALMA data we use in this analysis as well as the process by which we both reduce and calibrate that data. Additionally, we present our procedure for finding [CII]-emitting companion galaxies and how we estimate the star formation rates in the rest-UV and IR from the observations. In Section 2.3, we present the results, including the intrinsic luminosity function and the implied SFRD. In Section 2.4, we discuss both our results and the uncertainties and finally conclude with a summary of our findings in Section 2.5. In this work all magnitudes are given in the AB system (Oke & Gunn 1983). Moreover, we assume a Chabrier (2003) initial mass function (IMF) (across a mass range of 0.1-100  $M_{\odot}$ ) and adopt a flat  $\Lambda$ CDM cosmology with  $h = 0.7$ ,  $\Omega_M = 0.3$ , and  $\Omega_{\Lambda} = 0.7$ . For reference, 1" corresponds to 5.7 kpc at  $z = 6$  using this set of cosmological parameters.

## 2.2 Methods

---

### 2.2.1 Observational data

In this study we pursue a systematic search for [CII]-emitting companion galaxies at similar redshifts as the bright [CII]-emitting targets. We make use of the [CII]-emitting companion galaxies identified around 75

[CII]-emitting galaxies from the ALPINE large program (Le Fèvre et al. 2020; Béthermin et al. 2020) at  $z = 4-6$ , 25 [CII]-emitting galaxies from the REBELS large program (Bouwens et al. 2022b) at  $z = 6.5-7.7$ , and 27 [CII]-emitting QSOs from Venemans et al. (2020) at  $z = 6.0-7.6$ .

Loiacono et al. (2021) report 12 [CII]-emitter candidates around UV-selected  $z = 4-6$  galaxies in the ALPINE survey, and Fudamoto et al. (2021) report 2 [CII]-emitting galaxies around UV-selected sources in REBELS. Around the 27 [CII]-emitting QSOs studied by Venemans et al. (2020), 27 line-emitting galaxies are identified, but only 20 of these appear to lie at frequencies  $< 2000$  km/s from the target [CII]-emitting source and thus are likely at  $z \sim 6-7$ . In addition to the dust-obscured companions of Fudamoto et al. (2021), another less obscured companion galaxy is found in the REBELS sample (Schouws et al. 2024, in prep.) that we present for the first time in this work and will also be included in our analysis.

In addition to making use of the [CII]-emitting companion galaxies already identified by searches over ALPINE, REBELS, and  $z > 6$  QSO samples, we have also executed a search for additional companion galaxies using archival ALMA data. For this search, we identify ALMA projects from the ALMA Science Archive<sup>2</sup> that have targeted the [CII] line in massive galaxies or quasars at  $z \geq 6$ , which lie in ALMA band 6 (211-275 GHz) from  $z \sim 8$  to  $z \sim 6$ . We focused on ALMA observations acquired at low or modest spatial resolution ( $\gtrsim 0.2$  arcsec) to avoid over resolving the [CII] emission, which can frequently be quite spatially extended (Fujimoto et al. 2019, 2020; Fudamoto et al. 2022). The examined observations have angular resolutions from  $\sim 0.2$  to  $\sim 1.2$  arcsec. A more extensive discussion on the resolutions of the observations can be found in Sec. 2.3.1. A complete list of examined projects and sources and their corresponding angular resolution can be found in Table 2.8 from Appendix 2.A.5. In total we examine the ALMA data cubes for 55 targets where [CII] is targeted. Of those 55, we find clear ( $\gtrsim 5\sigma$ ) [CII] detections in 34 targets. We search for companion galaxies in the archival observations of the 34 targets where [CII] is detected. In total 161 targets are examined for companion galaxies when we include the searches done in previous literature as described above.

We obtained the calibrated measurement sets produced by the ALMA observatory pipeline from the European ALMA Regional Centre. Before imaging the measurements sets, the data are time-averaged over a period of 30 seconds to reduce the total data volume (e.g. Schouws et al. 2024, in prep.). The flux loss after time-averaging the visibilities over 30 seconds is less than 1 percent.<sup>3</sup>

<sup>2</sup><https://almascience.nrao.edu/aq/>

<sup>3</sup>The average fraction of intensity that is maintained after time-averaging over a time interval  $\tau_a$  can be approximated by  $R_\tau = 1 - 1.08 \times 10^{-9} \text{ s}^{-2} (\frac{\theta}{\theta_{\text{HPBW}}})^2 \tau_a^2$ , with  $\frac{\theta}{\theta_{\text{HPBW}}}$  the distance to the phase center with respect to the half-power beamwidth. For details see Bridle & Schwab (1999).

To image the measurement sets we use the Common Astronomy Software Applications (CASA) software for ALMA data (CASA Team et al. 2022) (version 5.7.0). We use the TCLEAN task with natural weighting, to maximise our sensitivity to companion galaxies. We clean the data to a depth of  $2\sigma$  using ‘auto-multithresh’ (Kepley et al. 2020). The cell sizes of the images are set such that there are approximately 5 pixels per beam. Moreover, we use the *uvcontsub* task to subtract the continuum emission in the *uv*-space using the default zeroth order polynomial excluding the channels within  $3\times$  the FWHM of the [CII] line from the central source.

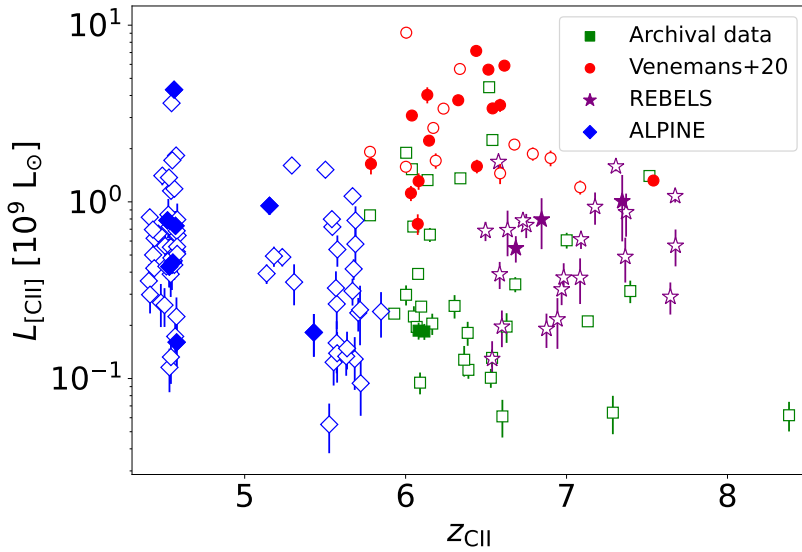
Figure 2.1 provides a convenient summary of the redshifts and [CII] luminosities of ALMA targets used for the construction of our [CII]-selected sample of companion galaxies. In Figure 2.2 the number of targets with companion galaxies as a function of their [CII] luminosity is shown. We see that fraction of galaxies with at least one companion galaxy increases with increasing luminosity.

### 2.2.2 Line-search algorithm

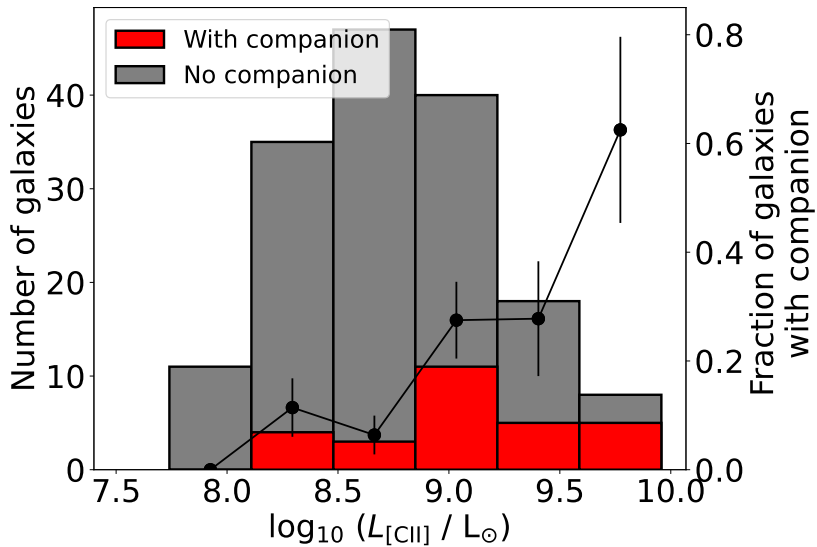
In the ALMA archival fields where we search for companion sources around [CII]-emitting galaxies and QSOs (green squares in Figure 2.1), we make use of the Matched Filtering in 3D (MF3D) code (Pavesi et al. 2018). Given that matched filtering works well for data cubes in which the emission is not very extended and thus well described by 2D Gaussians, MF3D is very effective in searching for line emission from high-redshift galaxies. To detect line emission MF3D uses 1D Gaussian frequency templates and 2D circular Gaussian spatial templates that are convolved with the ALMA data cubes.

The advantage of MF3D over normal matched filtering is that MF3D takes into account the spatial correlation present in noise in interferometric datacubes. The noise in the different channels of the data cubes are nevertheless uncorrelated and therefore the spectral width of the best fitting template matches approximately with the line width of the emission line. A useful reference for more details on the MF3D is Pavesi et al. (2018).

Detailed analyses have already been done on line identification with MF3D by Schouws et al. (in prep.) for the REBELS survey and it was found that lines with a signal-to-noise ratio (SNR)  $\geq 6.2$  showed a purity in excess of 95 percent. Purity (or fidelity) is defined as  $1 - \frac{N_{\text{neg}}}{N_{\text{pos}}}$ , so at SNR  $\geq 6.2$  for each positive peak with that SNR there are only 0.05 negative peaks. In order to decrease the possibility of selecting noise peaks, we set a constraint of SNR  $\geq 6.2$ . In addition to its identification of positive peaks, MF3D also identifies negative peaks arising from noise. To maximise the robustness of any candidate [CII] line we identify, we require the SNR of a potential galaxy candidate to be larger than the SNR of any negative peaks found in a data set. With SNR  $\geq 6.2$  we expect  $< 10\%$  of sources to be



**Figure 2.1:** The redshifts and [CII] luminosities of the 161 ALMA targets we use to search for a sample of [CII]-emitting companion galaxies. The [CII] luminosity of the archival sources is derived based on the peak flux and the redshift is from the peak frequency of [CII]. The red circles correspond to the 27 quasars with [CII] detections where Venemans et al. (2020) has already described a companion search, while the purple stars and blue diamonds correspond to the 25 [CII]-detected galaxies from REBELS (Bouwens et al. 2022b) and 75 [CII]-detected galaxies from ALPINE (Le Fèvre et al. 2020), where companion searches have been done (Fudamoto et al. 2021; Loiacono et al. 2021; see also Schouws et al. 2024, in prep.). Green squares indicate the 34 [CII]-detected ALMA targets where no searches for [CII]-emitting companion galaxies have thus far been reported in the literature, and we provide for the first time here. Filled markers indicate targeted galaxies with serendipitously detected companions, while open markers are targeted galaxies without detected companions.



**Figure 2.2:** The number of ALMA targets where a companion source is found (*red*) and not found (*grey*) vs. the [CII] luminosity of the primary ALMA target. The fraction of galaxies with at least one [CII]-emitting companion galaxy (*black line*) increases as a function of the [CII] luminosity of the target. The error on the fraction is given by the standard deviation of a binomial distribution.

affected by flux boosting in excess of  $3\sigma$  and expect a median flux boosting of  $< 21\%$  (see Decarli et al. 2020).

The width of the frequency templates used in our search ranges from  $100 \text{ km s}^{-1}$  to  $800 \text{ km s}^{-1}$ , and the width of the spatial templates ranges from 0 to 3 times the beam size. We define neighbours as serendipitously detected sources with similar redshifts as the targeted (quasar host) galaxy. We calculate the redshift of the candidates with the central frequency of the [CII] peak. We consider candidates that have a velocity difference with the target galaxy that is smaller than  $2000 \text{ km s}^{-1}$  (e.g. Venemans et al. 2020). For our final sample (Sec 2.3.2), we require sources to show  $\leq 500 \text{ km s}^{-1}$  velocity difference to minimise the impact of contamination from lower-redshift line-emitting sources on our scientific results. The velocity difference is defined by:

$$\Delta v = \frac{z_{\text{comp}} - z_{\text{target}}}{1 + z_{\text{target}}} c \quad (2.1)$$

where  $z_{\text{comp}}$  and  $z_{\text{target}}$  are the redshifts of the (potential) companion and target galaxy, respectively. At redshift  $z \sim 6$ , a velocity difference of  $2000 \text{ km s}^{-1}$  ( $500 \text{ km s}^{-1}$ ) corresponds to  $\Delta z \sim 0.05$  ( $\Delta z \sim 0.01$ ). We also apply this criteria to the companion galaxies originally identified by Loiacono et al. (2021), Venemans et al. (2020), and Fudamoto et al. (2021) to construct our sample of [CII]-emitting companion galaxies.

To further confirm that the source is at high redshift and the detected lines do not correspond to a CO transition at lower redshifts, we examine any available imaging data bluewards of the Lyman-alpha transition and look for possible detections. Due to hydrogen between the observer and the galaxy, photons with a rest-frame wavelength bluewards of  $1216 \text{ \AA}$  will be absorbed. The Lyman-alpha transition shifts to bluer wavelengths for sources at lower redshifts. Detections at bluer wavelengths than the Lyman-alpha transition at  $z > 5$  indicate that we are observing a low redshift source instead, meaning that the galaxy candidate is a foreground galaxy. We did not find detections in bluer bands for the candidate sources where available. It is worth remarking that this would have only small impact in removing dusty, low-redshift galaxies from our selection due to the faintness of these sources at bluer wavelengths, and our requirement that companion galaxies show a small velocity relative to the primary targets would be the main way of minimising contamination from this type of source (see Sec 2.3.2 and Appendix 2.A.1).

To summarise, a peak found by MF3D is considered to be a potential neighbour when:

- (i) The SNR is equal to or larger than 6.2.
- (ii) There are no negative peaks with an absolute SNR value greater than or equal to the SNR of the peak.

- (iii) The velocity difference between the central source and the peak is smaller than or equal to  $2000 \text{ km s}^{-1}$  ( $500 \text{ km s}^{-1}$  for the final sample).
- (iv) The galaxy is not detected in filters bluewards of the Lyman-alpha transition.

### 2.2.3 Moment-0 and continuum maps

For line candidates found by MF3D that meet the requirements of a neighbour galaxy the moment-0 and continuum maps are made in an iterative way with CASA. An initial moment-0 map is created with the task *immoments*, where we include the frequencies within  $2 \times$  the full width at half maximum (FWHM) of the line. As a starting point we use the FWHM and central frequency of the peak as found with by MF3D. For each source a signal-to-noise weighted spectrum is created only including the pixels that have values higher than  $3\sigma$  in the initial moment-0 map. A Gaussian (or double Gaussian if there is a double peak) is fitted to the spectrum and the FWHM and central frequency are extracted. A new moment-0 map is created with the updated values for the FWHM and the central frequency. This process is repeated and new spectra are made until the values for the FWHM and the central frequency converge. With these final values the definite moment-0 map is created. We also use these values to create the continuum maps. In this case we use TCLEAN on the measurements sets with *specmode='mfs'*. For the continuum map only the channels outside  $2 \times$  FWHM of the line are used. The [CII] and continuum fluxes are calculated from the moment-0 map and the continuum map, respectively. We fit the sources iteratively with the *imfit* task of CASA until the flux converges. Flux values quoted in this work are integrated fluxes. For companion sources that are not in the phase centre of the ALMA observation, and for which *uvcontsub* will not work properly, we make sure to subtract any significant ( $3\sigma$ ) continuum from the [CII] flux we obtain.

### 2.2.4 Rest-UV photometry

We use rest-UV observations from the Hubble Space Telescope (HST), the UltraVISTA survey (UVISTA), the Dark Energy Camera Legacy Survey (DECaLS), the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) and Hyper Suprime-Cam (HSC) on the Subaru Telescope. We focus on rest-UV observations at a rest-frame wavelength of  $1500 \text{ \AA}$  at  $z \sim 4-8$ . The approximate sensitivity of the relevant observations is summarised in Table 2.1.

A large fraction of the companion galaxies found around ALPINE galaxies are located over the COSMOS field, and therefore we can make use of the deep multi-wavelength observations available there. We use the photometry

code MOPHONGO (Labbé et al. 2006, 2010a,b, 2013, 2015) to perform this photometry. MOPHONGO uses high-resolution data to perform aperture photometry on low-resolution, deep optical data. Moreover, MOPHONGO models both flux from the sources themselves and neighboring objects in the same field such that it can subtract the light originating from nearby neighbors. Therefore, the aperture used with MOPHONGO can be larger and we adopt an aperture diameter of 2 arcsec. We have used the same observational data for the sources in the COSMOS field as used in Stefanon et al. (2019).

For the companion galaxies of REBELS we use the UV measurements in the J-filter of UltraVISTA in COSMOS provided by Fudamoto et al. (2021) and VISTA VIDEO in XMM-LSS (Stefanon et al, in prep.).

The companion sources we utilise from Venemans et al. (2020) are exclusively observed with HST, some in multiple IR (rest-UV) filters. For these sources the HST observations are stacked in order to increase the signal to noise ratio and to increase the number of detections. This is done by performing an inverse variance weighted stack of the images, where the variance is based on the uncorrelated RMS noise in the image (typically the noise starts behaving as uncorrelated when the image is rebinned by a factor of 10).

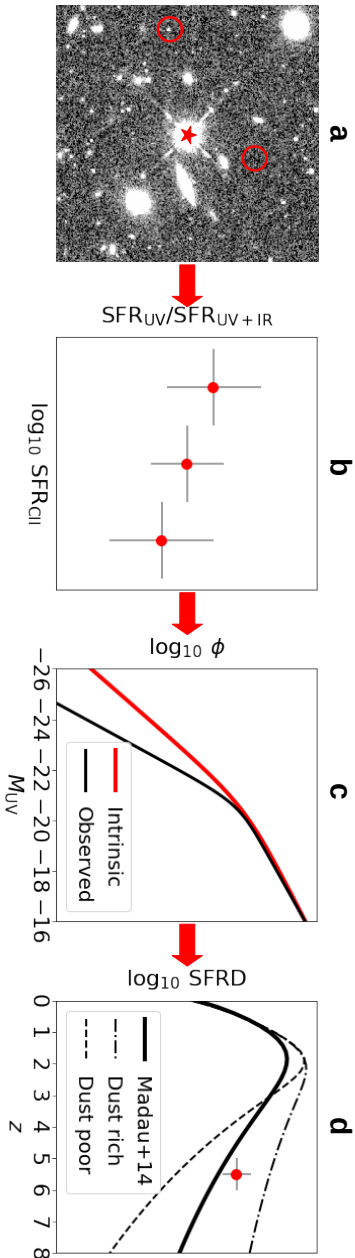
To extract the UV fluxes of the sources with exclusively HST data we cannot use MOPHONGO and use Source Extractor (SExtractor) (Bertin & Arnouts 1996) instead. SExtractor can detect and deblend objects in astronomical images and measure the corresponding fluxes. The algorithm starts by estimating the background signal. This allows the program to compute the relative flux and magnitude of objects. SExtractor has a double imaging mode where an image with the positions of the objects can be supplied. SExtractor then provides the relative magnitudes and fluxes of the source within an aperture diameter of 0.8 arcsec. We choose this aperture as a compromise between encapsulating all the flux of a source and excluding the flux of neighbouring sources. We make sure that we do the necessary aperture corrections for the HST IR filters as the encircled energy in an aperture of 0.8" diameter is  $\approx 0.84^4$ .

Moreover, for the galaxies found within this work that do not have HST imaging and are not in the COSMOS field we use the data from DECaLS and perform aperture photometry with SExtractor as explained for the Venemans et al. (2020) sources.

We opted to exclude companions (5 in total) where the radial distance to the QSO target is  $< 15$  kpc ( $\sim 2.6''$ ) as QSOs are particularly bright at rest-UV wavelengths. Our motivation for excluding these sources is due to potential blending between flux from the companion source and the QSO

---

<sup>4</sup><https://www.stsci.edu/hst/instrumentation/wfc3/data-analysis/photometric-calibration/ir-encircled-energy>



**Figure 2.3:** A schematic illustration of our procedure for estimating the contribution that obscured star formation makes to the SFR density at  $z \gtrsim 5$ . **a)** Use ALMA observations around bright galaxies and QSOs to select a sample of companion galaxies selected purely through a [CII] emission line (§2.3.1). **b)** Leverage observations of these [CII]-emitting companion galaxies in the rest-UV to quantify the unobscured fraction of the SFR (§2.3.2: red solid circles). **c)** Use a forward modeling procedure to infer the intrinsic LF (red line) that reproduces the observed UV LF (black line) when we apply the obscured fraction of SFR that we find (§2.3.4). **d)** Calculate the SFRD at  $z \gtrsim 5$  by integrating the intrinsic luminosity function (§2.3.5). We will be able to see if [CII]-selected galaxies imply dust-obscuration fractions consistent with a more dust-poor (dashed line) or dust-rich (dash-dotted line) Universe (Casey et al. 2018a). For context, the right-most panel includes the Madau & Dickinson (2014) SFR density as a thick solid line.



**Table 2.1:** The depths of the filters of rest-UV imaging data used in this work.

*Notes.* <sup>1</sup> Depths from Stefanon et al. (2019).

<sup>2</sup> For an emission line galaxy with half-light radius of 0.45 arcsec.

<sup>3</sup> For a point source at  $z = 6$  with 30 minutes exposure time. Calculated with the HST ETC.

Filter name	$5\sigma$ depth
<b>Ground-based</b>	
HSC $z$	25.9 <sup>1</sup>
CFHTLS $z$	25.2 <sup>1</sup>
SSC $z^+$	25.0 <sup>1</sup>
UVISTA $J$	25.4 <sup>1</sup>
DECaLS $z$	22.5 <sup>2</sup>
<b>HST</b>	
F125W	27.8 <sup>3</sup>
F140W	28.0 <sup>3</sup>
F160W	27.8 <sup>3</sup>

target. For one companion (P167-13C1), we use the measurements from Mazzucchelli et al. (2019) that do PSF modeling and therefore can reliably extract the flux from the companion that is close ( $<15$  kpc) to the quasar host galaxy.

### 2.2.5 Star formation rates

The [CII] luminosity is calculated directly from the [CII] flux using:

$$L_{[\text{CII}]} / L_{\odot} = 1.04 \times 10^{-3} \nu_{[\text{CII}], \text{obs}} F_{[\text{CII}]} D_L^2 \quad (2.2)$$

where  $F_{[\text{CII}]}$  is the [CII] flux in  $\text{Jy km s}^{-1}$ ,  $\nu_{[\text{CII}], \text{obs}}$  the central frequency of the [CII] emission in GHz and  $D_L$  the luminosity distance in Mpc (Solomon et al. 1992). De Looze et al. (2014) derived a relation for the total SFR of  $z \sim 0$  galaxies based on their [CII] luminosity. The relation we use in this work is based on the starburst sample from De Looze et al. (2014) as this has been found to agree best with high redshift sources (e.g. Kohandel et al. 2019; Pallottini et al. 2022):

$$\log_{10} (\text{SFR} / M_{\odot} \text{ yr}^{-1}) = -7.06 + 1.00 \times \log_{10}(L_{[\text{CII}]}) \quad (2.3)$$

The  $L_{[\text{CII}]}$ –SFR relation does not appear to evolve significantly from  $z \sim 4$ –7 to  $z \sim 0$  (Schaerer et al. 2020; Carniani et al. 2018, 2020), and so we will use this relation to calculate the total SFR of sources in our analysis.

With the relative UV magnitudes and redshifts of the sources we calculate the absolute magnitudes and the UV luminosity:

$$L_{\text{UV}} / L_{\odot} = 10^{\frac{M_{\text{UV}} - M_{\text{sun}}(\lambda)}{-2.5}} \quad (2.4)$$

where  $M_{\text{sun}}(\lambda)$  is the absolute magnitude of the Sun as measured with a filter with wavelength  $\lambda \approx 1500 \text{ \AA}$  under the assumption that  $\nu L_\nu = \text{constant}$ . To convert the UV luminosities to star formation rates we use the following relation from Madau & Dickinson (2014):

$$\text{SFR}_{\text{UV}}/M_\odot \text{ yr}^{-1} = 0.63 \times K_{\text{UV}} \times L_{\text{UV}} \quad (2.5)$$

$K_{\text{UV}}$  ( $= 1.15 \times 10^{-28} M_\odot \text{ yr}^{-1} \text{ erg}^{-1} \text{ s Hz}$ ) is the conversion factor from UV luminosity to SFR, while the factor 0.63 is needed to convert from a Salpeter to a Chabrier IMF.

The continuum flux of the sources allows us to constrain the IR luminosity. To estimate the IR luminosity we assume a SED of the dust and integrate the emission between the rest-frame wavelengths of 8 and 1000  $\mu\text{m}$ . We assume that the dust is heated by star formation only. The dust SED can then be described by a modified blackbody with emissivity index  $\beta_{\text{em}} = 2.03$  and  $T_{\text{dust}} = 46$  Kelvin following the approach of Inami et al. (2022)<sup>5</sup>. As flux is always measured relative to the CMB we correct the IR luminosity for this. At higher redshifts, the CMB temperature is higher and contributes more significantly to the flux that is measured. The CMB correction factor is given by:

$$f_{\text{CMB}} = \frac{F_{\text{int}} + F_{\text{CMB}}}{F_{\text{int}}} = 1 - \frac{B_\nu(T_{\text{CMB}})}{B_\nu(T_{\text{dust}})} \quad (2.6)$$

where  $F_{\text{int}}$  corresponds to the intrinsic flux of the source,  $F_{\text{CMB}}$  the flux measured originating from the CMB,  $B_\nu$  the emission of a blackbody as a function of temperature and  $T_{\text{CMB}} = (2.73 \text{ K}) \times (1 + z)$  (da Cunha et al. 2013). The IR luminosity is then calculated by the following equation:

$$L_{\text{IR}}/L_\odot = 7.70 \times 10^3 \frac{S_{\text{cont},158\mu\text{m}}}{f_{\text{CMB}}} \frac{D_L^2}{1 + z_c} \quad (2.7)$$

with  $S_{\text{cont},158\mu\text{m}}$  is the continuum flux in mJy measured at a rest frame frequency of  $\sim 158 \mu\text{m}$ ,  $z_c$  the redshift corresponding to the central frequency of the [CII] emission line and  $D_L$  the luminosity distance (analogous to Venemans et al. (2020)). There is a systematic uncertainty of a factor of  $\approx 2$ -3 on this luminosity as the shape of the SED is unknown (Venemans et al. 2018).<sup>6</sup> The IR SFR is calculated similarly to the UV SFR using the

<sup>5</sup>Assuming Milky Way-like dust and using the median dust temperature for the REBELS galaxies that is found with the approach of Sommovigo et al. (2022b). Sommovigo et al. (2022b) furthermore find the median temperature of the ALPINE galaxies to be 48 K, and therefore our assumption of the dust temperature is also consistent with the ALPINE sample.

<sup>6</sup>We adopt a systematic uncertainty of a factor of 3, as we observe this systematic scatter in the IR luminosity when we vary the dust temperature by 10K and  $\beta_{\text{em}}$  between 1.5-2.5.

following equation:

$$\text{SFR}_{\text{IR}}/M_{\odot} \text{ yr}^{-1} = 0.94 \times K_{\text{IR}} \times L_{\text{IR}} \quad (2.8)$$

$K_{\text{IR}} = 3.88 \times 10^{-44} M_{\odot} \text{ yr}^{-1} \text{ erg}^{-1} \text{ s}$  and  $L_{\text{IR}}$  corresponds to the integrated luminosity from 8 to 1000  $\mu\text{m}$  (Murphy et al. 2011). The factor 0.94 is used to convert from a Kroupa to a Chabrier IMF.

## 2.3 Results

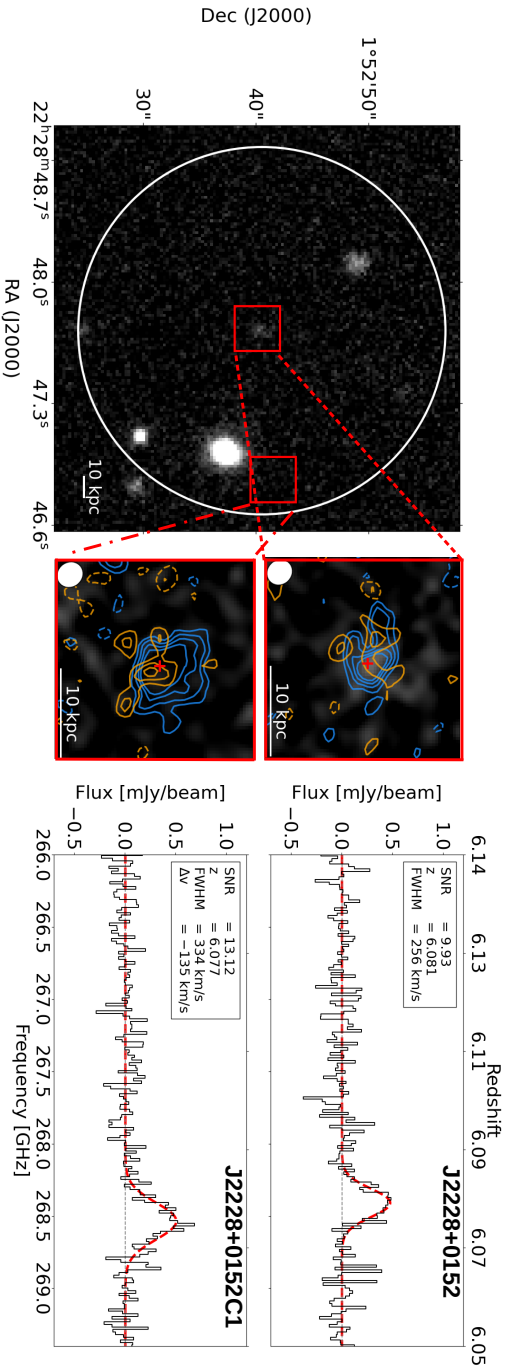
In this section, we make use of a substantial sample of 18 [CII]-selected companion galaxies (§2.3.1) with sensitive rest-UV imaging to characterise the fraction of SFR in  $z \sim 4$ -8 star-forming galaxies that is obscured (§2.3.2). We then make use of the obscuration fraction to derive the intrinsic LF at  $z \sim 5$ -6 that includes the obscured UV light (§2.3.4). Finally, we estimate the contribution dust-obscured galaxies make to the SFR density at  $z \gtrsim 5$  (§2.3.5). Figure 2.3 provides a schematic representation of the methodology we use to infer the dust-corrected SFR density at  $z \geq 4$  from the present analysis.

### 2.3.1 Newly identified candidate companion galaxies

The primary objective of this work is to better characterise the obscured fraction of star formation in the early Universe using a [CII]-selected sample of companion galaxies at  $z > 4$ . To do so, we not only make use of [CII]-selected companion sources already identified in various large data sets, e.g. ALPINE and REBELS, but we also conduct searches for [CII]-emitting companion galaxies in the archival ALMA data around  $z > 6$  targets showing robust [CII] lines. In this work, we analyse the ALMA cubes for 34 separate sources showing [CII] emission.

With the line-searching algorithm MF3D we found 2 candidate neighbour galaxies in the archival data and a third previously unreported candidate from REBELS. The first two sources were identified as part of ALMA project 2017.1.00541.S with UV imaging observations available from the DE-CaLS survey. We summarise the properties of the three neighbour galaxies in Table 2.2.

One of the companion galaxies was identified around quasar host galaxy J2228+0152, at  $z_{\text{Ly}\alpha} = 6.08$ . We find the quasar host galaxy to reside at  $z_{\text{[CII]}} = 6.081$  consistent with Matsuoka et al. (2018). For this quasar a neighbour candidate is found at  $z_{\text{[CII]}} = 6.077$ , which we will call J2228+0152C1. The left panel of Fig. 2.4 shows the position of the neighbour galaxy to the central quasar host galaxy and a zoomed-in image on the location of peaks found by MF3D (red crosses). On the right, a section of the corresponding SNR-weighted spectrum of the peaks is shown



**Figure 2.4:** An illustration of a newly discovered [CII]-emitting companion galaxy (lower panels) around ALMA QSO target J2228+0152 (upper panels). The leftmost panel shows an image in the rest-UV ( $z$ -filter from the DECaLS survey) of both the QSO and the [CII]-emitting companion galaxy, along with the FOV of ALMA (white circle). The upper and lower left panels present zoomed-in images of the QSO and companion galaxy. The solid blue and orange contours indicate those regions detected at  $2\sigma$ ,  $3\sigma$ ,  $4\sigma$ ,  $5\sigma$  in the moment-0 and continuum map, respectively (negative contours are shown with the dashed lines). The red crosses indicate the peak of the [CII] emission found with MF3D and the ellipses in the left lower corners indicate the ALMA beam size. The upper and lower rightmost panels show extracted spectra (binned at  $20 \text{ km s}^{-1}$ ) for the QSO and companion. The dashed lines show Gaussian fits to the detected [CII] lines. The SNR reported here is the SNR as found by MF3D.



with a Gaussian fit to the [CII] line. J2228+0152C1 has a more significant [CII] line than the central galaxy but is not detected in any of the filters of the DECaLS survey. The  $3\sigma$  detection this companion galaxy shows in the dust-continuum suggests this source is highly dust-obscured. In Table 2.6 the aperture fluxes are shown as measured with SExtractor along with their corresponding magnitudes.

The second companion galaxy identified is around the quasar host galaxy J1208-0200 found with the Subaru Telescope at a redshift  $z_{\text{Ly}\alpha} = 6.2$  (Matsuoka et al. 2018). Using the [CII]-line we found the quasar host galaxy to reside at  $z = 6.117$ , consistent with Izumi et al. (2019). The neighbour candidate, which we will call J1208-0200C1, is found at  $z_{\text{[CII]}} = 6.153$ . We show the quasar host galaxy and the neighbour with their corresponding spectrum in Fig. 2.12. While both galaxies are detected in [CII], only the quasar host galaxy is detected in the rest-UV. J1208-0200C1 is not visible in the rest-UV observations and therefore is likely highly dust-obscured.

In addition to the previously presented obscured companion galaxies by Fudamoto et al. (2021), another less obscured companion source is found in the REBELS sample (Schouws et al. 2024, in prep.). In contrast to the sources from Fudamoto et al. (2021) that are not detected in the rest-UV, the companion to REBELS-39 has a  $2\sigma$  detection in the J-band ( $0.23 \pm 0.11 \mu\text{Jy}$ ). In Fig. 2.13 we show this serendipitously identified source along with its target galaxy in a consistent manner as we have done for the sources from archival data.

### 2.3.2 [CII]-selected sample

With the present compilation of companion galaxies – including the newly identified candidates from §2.3.1 – our goal is to determine the fraction of SFR in galaxies that is obscured vs. the total SFR in galaxies. We do this as a function of the SFR given the well known dependence of dust obscuration on the stellar mass (e.g. Reddy et al. 2006; Pannella et al. 2009, 2015; Whitaker et al. 2018; Bouwens et al. 2020). This can be done by looking at the ratio between the unobscured (UV) star formation and the total star formation rate of the neighbouring galaxies found in [CII]. Our use of [CII]-selected galaxies ensures that we are not biased to unobscured galaxies.

Table 2.3 summarises the properties of the 23 neighbour candidates we examined. The full sample consists of the three new companion galaxy candidates at  $z > 6$  found in archival data and around REBELS-39 and the already known companion galaxies from Loiacono et al. (2021), Venemans et al. (2020), Fudamoto et al. (2021). We examined the companion galaxies that have rest-UV observations and have a velocity difference of  $\Delta v < 2000 \text{ km s}^{-1}$  with their central galaxies. In some cases galaxies are not detected in the rest-UV or IR continuum at  $3\sigma$ , as can be seen by the er-

**Table 2.2:** The properties of [CII]-emitting companion galaxies identified for the first time here.

Source	$z_{\text{[CII]}}$	RA (ICRS)	Dec (ICRS)	$\nu_c$ (GHz)	FWHM (km s <sup>-1</sup> )	$F_{\text{[CII]}}$ (Jy km s <sup>-1</sup> )	$S_{\text{cont}}$ (mJy)
J1208-0200C1	$6.1530 \pm 0.0003$	12:08:58.46	-02:00:32.85	$265.70 \pm 0.015$	$256 \pm 40$	$0.312 \pm 0.082$	$0.072 \pm 0.035$
J2228+0152C1	$6.0774 \pm 0.0002$	22:28:46.83	+01:52:41.55	$268.54 \pm 0.01$	$334 \pm 30$	$0.352 \pm 0.058$	$0.188 \pm 0.029$
REBELS-39C1	$6.8383 \pm 0.0004$	10:03:05.31	+02:18:44.62	$242.47 \pm 0.01$	$260 \pm 40$	$0.458 \pm 0.053$	$0.021 \pm 0.016$



**Table 2.3:** The properties of the sample of neighbouring galaxies included in our analysis.<sup>†</sup>

Source	S/N*	$z_{\text{[CII]}}$	$L_{\text{IR}}$ ( $10^{11} L_{\odot}$ )	SFR <sub>IR</sub> ( $M_{\odot} \text{ yr}^{-1}$ )	$L_{\text{UV}}$ ( $10^9 L_{\odot}$ )	SFR <sub>UV</sub> ( $M_{\odot} \text{ yr}^{-1}$ )	$L_{\text{[CII]}}$ ( $10^9 L_{\odot}$ )	SFR <sub>[CII]</sub> ( $M_{\odot} \text{ yr}^{-1}$ )	$\Delta v$ ( $\text{km s}^{-1}$ )	Used?
<b>This analysis</b>										
J1208-0200C1 <sup>a</sup>	6.5	6.15	$2.9^{+6.0}_{-2.3}$	$41^{+83}_{-33}$	61 ± 30	$7.3 \pm 3.4$	$0.305 \pm 0.080$	$24.9 \pm 13.2$	1546	
J2228+0152C1	13.1	6.08	$7.5^{+15.0}_{-2.2}$	$104^{+209}_{-51}$	$0.4 \pm 29.2$	$0.1 \pm 3.5$	$0.338 \pm 0.056$	$27.6 \pm 13.5$	-135	✓
REBELS-39C1	9.4	6.84	$1.0^{+2.2}_{-1.0}$	$14^{+14}_{-14}$	$410 \pm 196$	$61 \pm 29$	$0.521 \pm 0.060$	$42.7 \pm 20.3$	-252	✓
<b>Venemans et al. (2020)<sup>d</sup></b>										
J0100+2802C1	5.9	6.32	$1.1^{+3.3}_{-2.6}$	$15^{+47}_{-37}$	$69.6 \pm 73.2$	$13 \pm 13$	$0.813 \pm 0.132$	$66.6 \pm 32.5$	-112	✓
J0305-3150C1	6.0	6.61	$30^{+21.6}_{-21}$	$424^{+851}_{-293}$	$5.9 \pm 9.5$	$0.9 \pm 1.5$	$1.23 \pm 0.31$	$101 \pm 53$	-245	✓
J0842+1218C1	18.9	6.07	$6.2^{+12.6}_{-6.6}$	$86^{+176}_{-135}$	$5.3 \pm 7.0$	$1.0 \pm 1.2$	$1.73 \pm 0.12$	$142 \pm 66$	-431	✓
J0842+1218C2	7.5	6.06	$4.7^{+6.8}_{-3.8}$	$69^{+135}_{-53}$	$34.2 \pm 7.3$	$6.1 \pm 1.3$	$0.411 \pm 0.086$	$33.6 \pm 17.0$	-476	✓
J1319+0950C2	5.6	6.14	$7.5^{+15.2}_{-5.4}$	$105^{+212}_{-75}$	$9.6 \pm 11.9$	$1.6 \pm 2$	$0.341 \pm 0.078$	$27.9 \pm 14.4$	407	✓
P231-20C1	6.0	6.55	$1.4^{+1.8}_{-1.1}$	$19^{+25}_{-15}$	$-0.1 \pm 6.7$	$0.0 \pm 1.1$	$0.203 \pm 0.043$	$16.6 \pm 8.4$	-1615	✓
J2054-0005C1	5.8	6.04	$16^{+33}_{-12}$	$226^{+459}_{-170}$	$18 \pm 4.5$	$0.3 \pm 0.8$	$0.522 \pm 0.142$	$42.8 \pm 22.9$	-26	✓
J2100-1715C1	15.4	6.08	$145^{+291}_{-15.3}$	$2030^{+4060}_{-3170}$	$-6.7 \pm 12.3$	$-1.2 \pm 2.2$	$7.76 \pm 1.03$	$636 \pm 305$	-8	✓
P167-13C1 <sup>c</sup>	18.8	6.51	$7.6^{+5.3}_{-7.4}$	$106^{+74}_{-74}$	$62.4 \pm 25.0$	$11 \pm 4$	$1.34 \pm 0.09$	$109 \pm 51$	-118	✓
<b>Loiacono et al. (2021)</b>										
D_C_842313 <sup>a</sup>	35.5	4.54	$208^{+417}_{-139}$	$2910^{+5820}_{-1940}$	$94.2 \pm 15.7$	$14 \pm 2$	$4.56 \pm 0.14$	$373 \pm 172$	-829	
D_C_665626	12.5	4.58	$8.7^{+17.6}_{-6.1}$	$122^{+245}_{-85}$	$11.7 \pm 12.3$	$1.7 \pm 1.8$	$0.716 \pm 0.069$	$58.6 \pm 27.6$	-23	✓
v_c_5101209780	13.0	4.57	$1.6^{+3.9}_{-1.9}$	$22^{+26}_{-9}$	$301 \pm 16.3$	$45 \pm 2$	$1.56 \pm 0.18$	$128 \pm 61$	-56	✓
D_C_818760	10.3	4.57	$8.2^{+16.4}_{-5.6}$	$114^{+229}_{-78}$	$9.1 \pm 16.0$	$1.4 \pm 2.4$	$0.390 \pm 0.038$	$31.9 \pm 15.0$	241	✓
D_C_787780	15.5	4.51	$8.8^{+17.7}_{-3.4}$	$123^{+248}_{-87}$	$472 \pm 13.9$	$71 \pm 2$	$1.43 \pm 0.09$	$117 \pm 54$	-107	✓
D_C_873321	8.5	5.16	$-0.2^{+3.3}_{-3.3}$	$-3^{+47}_{-47}$	$72.2 \pm 16.8$	$9.7 \pm 2.3$	$1.86 \pm 0.23$	$152 \pm 73$	56	✓
D_C_378903	8.1	5.42	$6.3^{+12.9}_{-4.8}$	$89^{+180}_{-67}$	$49.3 \pm 14.2$	$6.4 \pm 1.8$	$0.637 \pm 0.049$	$52.2 \pm 24.4$	-294	✓
v_c_5100822662	6.6	4.52	$3.7^{+4.6}_{-3.0}$	$52^{+67}_{-37}$	$26.5 \pm 5.9$	$4 \pm 1$	$0.363 \pm 0.006$	$29.7 \pm 13.7$	-8	✓
D_C_859732	6.2	4.54	$-4.8^{+11.3}_{-6.9}$	$-67^{+158}_{-97}$	$-1.1 \pm 9.9$	$-0.2 \pm 1.5$	$0.59 \pm 0.13$	$48.3 \pm 24.7$	369	✓
<b>Fudamoto et al. (2021)</b>										
REBELS-12C1 <sup>b</sup>	8.2	7.35	$3.6^{+7.2}_{-2.6}$	$50^{+101}_{-36}$	$169 \pm 253$	$24 \pm 35$	$0.645 \pm 0.072$	$52.8 \pm 25.0$	118	
REBELS-29C1	10.3	6.68	$8.2^{+16.5}_{-5.6}$	$115^{+230}_{-77}$	$-50.7 \pm 67.7$	$-7.7 \pm 10.3$	$0.778 \pm 0.102$	$63.7 \pm 30.5$	-118	

*Notes.* <sup>†</sup> A galaxy is used to infer the obscured star formation rate only if the UV observations are deep enough ( $1\sigma$  depth would detect the UV if the galaxy is maximally 85 percent obscured) and it is likely to be at high redshift ( $\Delta v \leq 500 \text{ km s}^{-1}$ ). Only galaxies with a ✓ in the rightmost column are used to compute the unobscured fraction of the SFR as function of the total SFR.

<sup>a</sup> S/N of the [CII] line as reported by MF3D. For the Venemans et al. (2020) sources this is the S/N found by their independent line-search method.

<sup>b</sup> Excluded from the analysis because the source lacks sufficiently deep imaging data to probe rest-UV SFRs.

<sup>c</sup> We have used the rest-UV fluxes from Mazzucchelli et al. (2019) as explained in 2.2.4.

<sup>d</sup> The only sources included in this list are those with sensitive rest-UV observations and which lie at least 15 kpc away from the central QSO target (§2.2.4). We have adopted the IR and [CII] fluxes as found in Venemans et al. (2020) to compute the luminosities and SFRs.



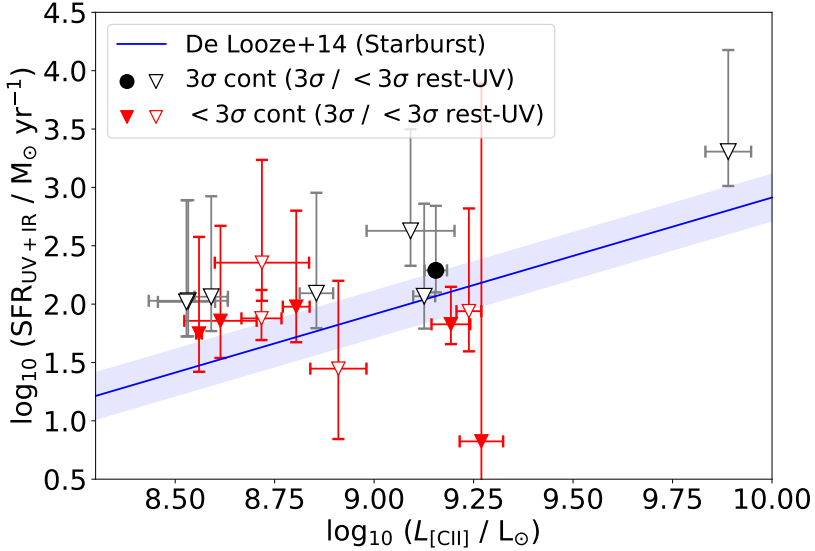
rors on their properties. Out of the 23 neighbour candidates, 11 (17) sources have  $3\sigma$  ( $2\sigma$ ) continuum detections. Moreover, out of the 23 sources, 7 (10) are detected with  $3\sigma$  ( $2\sigma$ ) in the rest-UV.

To constrain the obscured fraction, we only consider sources which feature UV observations sensitive enough to detect sources at  $1\sigma$  even if 85 percent of the UV light is obscured. Furthermore, we decide to restrict our analysis to those companion sources that have  $\Delta v \leq 500 \text{ km s}^{-1}$ . Using the full companion sample from Venemans et al. (2020) we quantify the probability of a companion galaxy being a lower redshift interloper as function of  $\Delta v$  in Appendix 2.A.1. We assume that sources with  $\Delta v > 2000 \text{ km s}^{-1}$  are contaminants from lower redshift (e.g. CO line emitters). We quantify that for a velocity difference of  $\leq 500 \text{ km s}^{-1}$  the probability of the neighbour candidate being at low redshift is low ( $\sim 13\%$ ). Applying all of the above criteria, we arrive at a sample of 18 galaxies.

We did not include the known systems of QSOs host galaxies and massive galaxies with companions such as PJ308-21 (Decarli et al. 2017, 2019) and BRI 1202-0725 (Carilli et al. 2013). They have been excluded from this study as our original focus was on companions at  $z > 6$ , using the Venemans et al. (2020) sample. Since we focus on companion galaxies selected by their [CII] emission, newly detected companions using JWST (e.g. Kashino et al. 2023; Wang et al. 2023) are likewise not included.

As a check on our analysis (also given the possibility of a small fraction of foreground interlopers which could enter our selection: Appendix 2.A.1), we compare the SFRs we compute based on flux measurements in the rest-UV and IR continuum and the [CII] luminosity based on the detected [CII] lines. We present the results of this comparison in Figure 2.5. In blue the starburst relation of De Looze et al. (2014) and the  $1\sigma$  dispersion is shown. The black circles indicate sources in our sample for which the far-IR continuum is detected ( $3\sigma$ ), the red squares correspond to sources with no such continuum detections. Finally, the filled data points correspond to galaxies detected in the rest-UV ( $3\sigma$ ), while the empty symbols indicate galaxies lacking a rest-UV detection. For most sources the error on the total SFR is dominated by the uncertainty introduced in the IR luminosity by the uncertain shape of the dust SED. We see from this figure that our sample agrees with the relation from De Looze et al. (2014), although for some data points we unfortunately have large uncertainties.

Since our sample of [CII]-emitting galaxies are selected from two different environments, around QSOs and SF galaxies, we examine whether the environment appears to have an impact on the obscured fraction of SFR. In Fig. 2.6 the cumulative distribution function is shown for the fraction of unobscured SFR. Fractions of unobscured SFR that have negative values are treated as having a value of 0 while fractions that are larger than 1 are treated as having a value of 1. We observe that the companion galaxies around QSOs in our sample appear to be significantly more obscured.



**Figure 2.5:** The total SFR ( $\text{SFR}_{\text{UV+IR}}$ ) as function of the [CII] luminosity of the 18 neighbour galaxies used in this work. In blue the starburst galaxy relation derived by De Looze et al. (2014) and the  $1\sigma$  dispersion is shown. The black filled and non-filled symbols indicate sources for which the dust continuum is detected ( $3\sigma$ ), the red symbols show the continuum non-detections. The filled data points are the galaxies that have a  $3\sigma$  rest-UV detection, empty symbols indicate galaxies undetected in the rest-UV. We include the uncertainty (factor of 3) on the IR luminosity due to the uncertain shape of the dust SED.

We therefore decide to create sub samples of the companion galaxies, separated by their environment. The SF galaxy companion sample consists of companions found around ALPINE and REBELS galaxies and has a mean redshift of  $z_{\text{SF}} = 5.14$ . The sample of companions around QSOs consists of the two companions shown in §2.3.1 and also companions from Venemans et al. (2020) and has a mean redshift of  $z_{\text{QSO}} = 6.21$ .

### 2.3.3 Obscured star formation

We have chosen to characterise the fraction of obscured star formation in galaxies as a function of their total SFRs as measured by [CII] – given our ability to readily quantify this SFR and not e.g. the stellar mass. In Figure 2.7, we show individual constraints on the obscured fraction as a function of the [CII]-derived SFRs for both the companions near QSO and SF galaxies. If a galaxy is not detected (not a  $3\sigma$  detection) in the continuum

and/or rest-UV we use the flux as measured (no upper limits)<sup>7</sup>.

Previous work by e.g. Bouwens et al. (2020); Fudamoto et al. (2020a); Bowler et al. (2024) on UV-selected galaxies show that the infrared excess (IRX) increases for larger stellar masses ( $M_*$ ). As the relation between the fraction of obscured or unobscured star formation and total SFR is unknown, we use the IRX- $M_*$  relation as a starting point. We can convert IRX to the fraction of unobscured star formation rate using:

$$\begin{aligned} \frac{\text{SFR}_{\text{UV}}}{\text{SFR}_{\text{UV+IR}}} &= \left(1 + \frac{\text{SFR}_{\text{IR}}}{\text{SFR}_{\text{UV}}}\right)^{-1} = \left(1 + C_1 \times \frac{L_{\text{IR}}}{L_{\text{UV}}}\right)^{-1} \\ &= (1 + C_1 \times \text{IRX})^{-1} \end{aligned} \quad (2.9)$$

with  $C_1 = K_{\text{IR}}/K_{\text{UV}}$  and  $\text{IRX} = L_{\text{IR}}/L_{\text{UV}}$ .

We then use the dependence of IRX on  $M_*$  ( $\text{IRX} \propto M_*^B$ ) and assume the best-fit main sequence relation ( $M_* \propto \text{SFR}_{\text{tot}}$ ) of Speagle et al. (2014) (eq. 29) to have an unobscured star formation rate fraction as a function of total SFR:

$$\frac{\text{SFR}_{\text{UV}}}{\text{SFR}_{\text{UV+IR}}} = (1 + A \times \text{SFR}_{[\text{CII}]}^B)^{-1} \quad (2.10)$$

where  $A$  and  $B$  are free parameters.

This relation is fitted to both the SF galaxy and QSO companion [CII]-selected samples. We find  $B$  to be consistent with zero as there is no clear dependence of the obscuration fraction on the total SFR for the sources in our sample. Therefore for simplicity we fix  $B$  to zero. Our fit can then be simplified by taking the mean of the (un)obscured fraction of the samples. The median (and 16th/84th percentiles) we find for the mean unobscured fraction of the SF galaxy and QSO sample when we generate  $10^6$  bootstrap samples are  $f_{\text{unobscured,SFG}} = 0.37_{-0.15}^{+0.17}$  and  $f_{\text{unobscured,QSO}} = 0.07_{-0.05}^{+0.05}$ . Adopting the functional form given in Eq. 2.10, when fixing  $B = 0$ , we find  $A_{\text{SFG}} = 1.6_{-0.8}^{+1.7}$  and  $A_{\text{QSO}} = 11.6_{-5.4}^{+15.6}$  (Table 2.4). The unobscured fractions for the individual galaxies together with the mean unobscured fractions are shown in Fig 2.7. Errors shown on the individual data points result from the formal errors on the star formation rates, the scatter on the IR luminosity introduced by the uncertain dust temperature and dust emissivity and the scatter on the [CII] SFR. The top axis shows stellar masses assuming the main sequence relation of Speagle et al. (2014) at  $z = 5$  and  $z = 6$ . As the galaxies in our sample do not all have a large amount of multi-wavelength observations, we could not securely determine their stellar masses.

For comparison, we also show the UV-selected galaxies of ALPINE with blue diamonds and the UV-selected galaxies of REBELS with purple crosses.

<sup>7</sup>We note that the median SNR of the dust continuum observations for our sample is  $\sim 3$ .

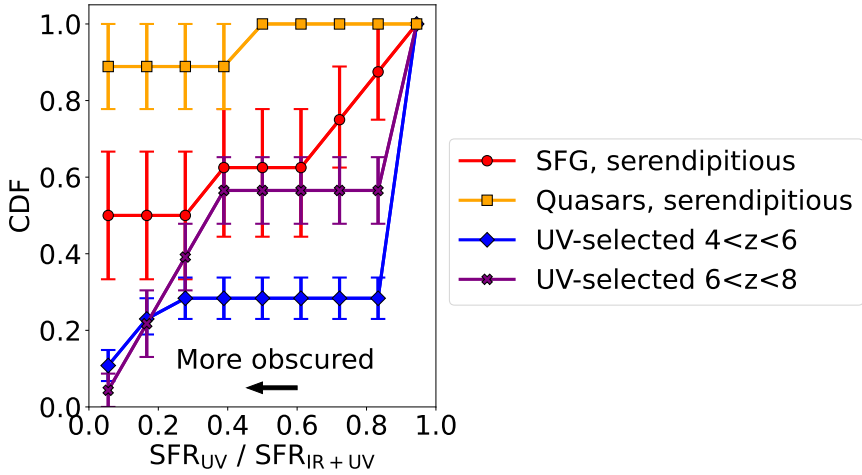
Also shown is the relation found by Mauerhofer & Dayal (2023) at  $z \sim 5 - 6$  with the DELPHI semi-analytic model baselined against ALPINE and REBELS observations. Moreover we plot the IRX- $M_*$  relations from Bouwens et al. (2020), Fudamoto et al. (2020a) and Bowler et al. (2024) using Eq. 2.9. The SF galaxy companion sample suggests an obscuration fraction ( $f_{\text{obs,median}} = 63\%$ ) similar to those expected from UV-selected samples. The [CII]-selected galaxies suggest a relation for galaxies around QSOs that implies that more SFR is obscured ( $f_{\text{obs,median}} = 93\%$ ) than what is inferred from UV-selected galaxies (as indicated by the relations from ASPECS, A3COSMOS, ALPINE and REBELS). The high obscuration fraction of the companion galaxies near QSOs is remarkable. Regardless of the origin of this increased dust attenuation, we clearly observe a difference between QSO and SF galaxy environments and hence we treat the samples separately in the rest of this paper.

In Appendix 2.A.1 we have used the Venemans et al. (2020) sample to estimate the fraction of low redshift contaminants we would expect in our sample. According to these calculations 13% of our sample of companion galaxies could correspond to lower redshift sources which happened to show CO line emission at approximately the same frequency as [CII] for the main target. We thus look at the effect this would have on the derived mean obscured fraction as shown in Fig. 2.7. Using the ASPECS sample at  $z \sim 1.3-2.5$  from Bouwens et al. (2020) as a reference low redshift sample, we would expect most low redshift emitters to have an IRX  $> 4$ , which according to Eq. 2.9 would correspond to  $\gtrsim 80\%$  of the star formation to be obscured. We have repeated our fitting procedure while we randomly remove one source from the SF and QSO companion sample that is  $> 80\%$  obscured to simulate the effect of excluding a potential contaminant from our sample. We find that the newly calculated unobscured fractions are within 20% of the unobscured fraction described above for the full sample and therefore within the errors on the mean unobscured fraction.

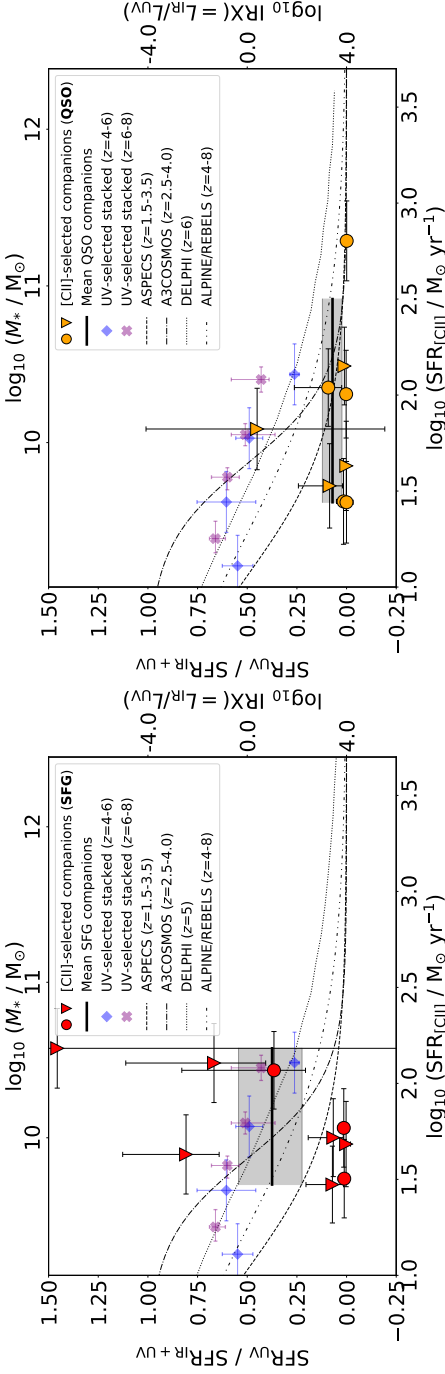
### 2.3.4 The Intrinsic UV Luminosity Function

Having established the unobscured star formation rate from the [CII]-selected neighbours, we use this characterization to recover the intrinsic UV luminosity function from the observed UV LF. The observed LF differs from the intrinsic LF in that both the luminosity of sources and their inclusion in the LF can be impacted by dust obscuration. For our reference UV LF from the observations, we have selected to use the double power law galaxy UV LF from Harikane et al. (2022) that utilises the very wide-area ( $\sim 300 \text{ deg}^2$ ) Hyper Suprime-Cam survey.

A double power law luminosity function is given by:

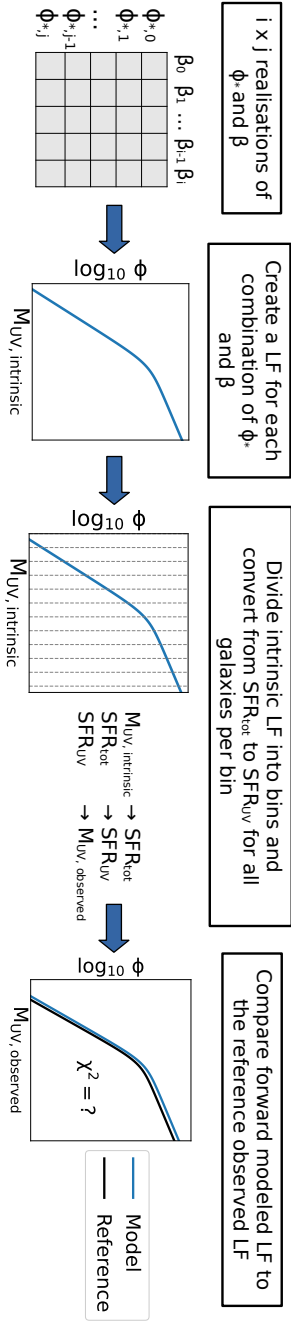


**Figure 2.6:** The cumulative distribution function (CDF) of the fraction of unobscured star formation rate for companion galaxies around SF galaxies (red) and QSOs (orange). The CDFs for the UV-selected samples of ALPINE (Le Fèvre et al. 2020) and REBELS (Bouwens et al. 2022b) are shown in blue and purple, respectively. The errors are obtained with bootstrapping. The companions around QSOs have lower SFRs in the rest-UV, suggesting that galaxies found in the neighborhoods around QSOs are dustier than around bright star-forming galaxies (but we stress that the current statistics are still small).



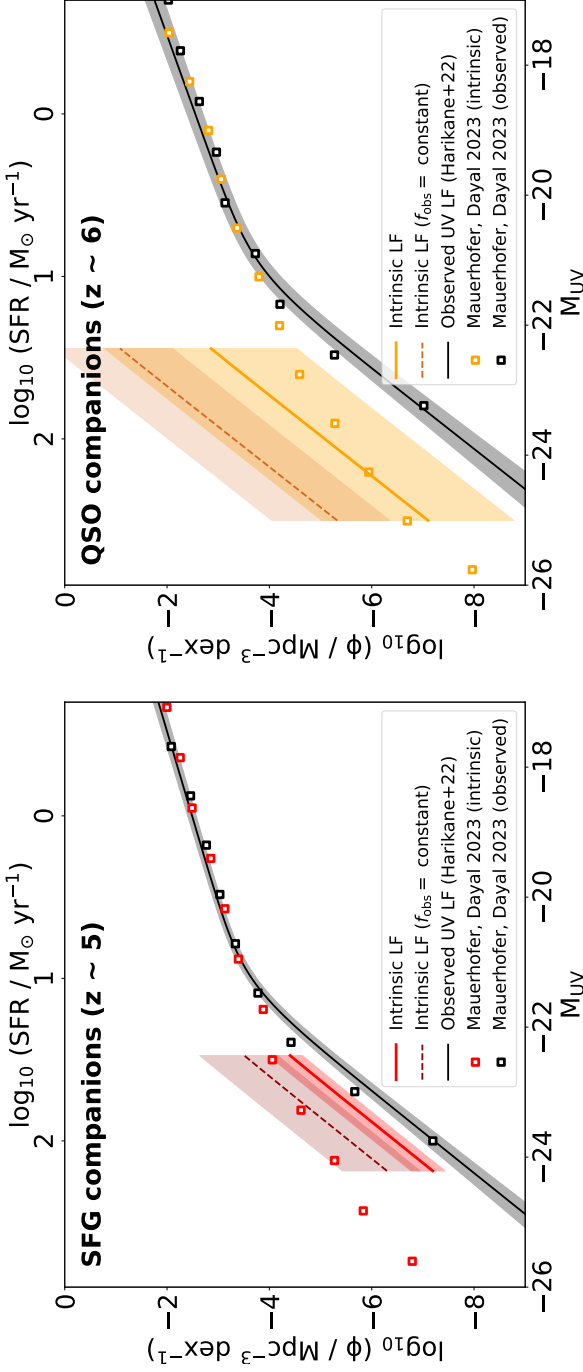
**Figure 2.7:** The unobscured fraction of the SFR as a function of the total SFR as measured from the [CII] emission from companions galaxies around SF galaxies (*left*) and companions galaxies around QSOs (*right*). The [CII]-selected galaxies are shown in red and orange for these samples, and the mean unobscured fraction by the black solid line. The width of the shaded grey region indicates the 16th and 84th percentiles of the mean fraction. Galaxies that are not detected in the far-IR are indicated by an upside down triangle. The blue crosses show the relation derived for the stacked UV-selected ALPINE galaxies (B  thermin et al. 2020), while the purple crosses show the relation based on the stacked REBELS sample (Algera et al. 2023). The dotted lines show the relation from Mauerhofer & Dayal (2023) from the DELPHI semi-analytic model calibrated with ALPINE and REBELS galaxies. Also shown are the IRX- $M_*$  relations from ASPECS (Bouwens et al. 2020), A<sup>3</sup>COSMOS (Fudamoto et al. 2020a) and ALPINE/REBELS (Bowler et al. 2024) assuming the main sequence relation from Speagle et al. (2014) for  $z = 5$  and  $z = 6$ . The top axis shows the stellar masses assuming the same relation from Speagle et al. (2014) at  $z = 5$  and  $z = 6$ .





**Figure 2.8:** A schematic overview of our forward modeling procedure that is used to recover the intrinsic UV LF. 1) We create  $i \times j$  realisations of  $\phi_*$  and  $\beta_{DP}$ , which serve as the input parameters of the intrinsic LF. 2) For every combination of  $\phi_*$  and  $\beta_{DP}$  we compute a model LF. 3) We segregate the model intrinsic LF across various bins in magnitudes and for all the galaxies in a bin we convert the total SFR to the UV SFR from the distribution we find in Fig. 2.7. 4) Finally, we compare the model observed LF to the reference observed LF from Harikane et al. (2022) and compute  $\chi^2$ .





**Figure 2.9:** The observed UV luminosity function at  $z \sim 5$  (left panel) and  $z \sim 6$  (right panel) from Harikane et al. (2022) (*black line*) and the intrinsic UV luminosity function we obtain with the dust obscuration correction procedure we describe in this paper based on the characteristics of companion galaxies around SF galaxies and QSOs (*red and orange line, respectively*). The width of the shaded regions (*shown in red and orange*) indicates the 16th and 84th percentiles of the observed and intrinsic LFs. Shown with dashed lines are the intrinsic LFs when instead of using the distribution of obscured fractions we adopt the mean obscuration instead (see Sec. 2.3.4). Clearly, the dust corrections we infer here substantially increase the volume density of intrinsically luminous sources at  $z \gtrsim 5$ . This is also shown by the predictions of the intrinsic and observed UV LF from Mauerhofer & Dayal (2023) that are calibrated with ALPINE and REBELS data at  $z \sim 5$  and  $z \sim 6$ . We only infer the intrinsic LF over the SFR range sampled by our companion galaxy sample.



$$\phi(M) = \frac{\phi_*}{10^{0.4(M-M_*)(\alpha_{\text{DP}}+1)} + 10^{0.4(M-M_*)(\beta_{\text{DP}}+1)}} \quad (2.11)$$

with  $\phi_*$  the normalization density,  $M_*$  the characteristic magnitude,  $\alpha_{\text{DP}}$  the faint-end slope and  $\beta_{\text{DP}}$  the bright-end slope. To recover the intrinsic UV LF, we use a forward modeling methodology. Fig. 2.8 shows a schematic overview of our procedure. The steps of our forward modeling procedure are the following:

- (i) We allow the parameters of the input intrinsic LF ( $\phi_*$  and  $\beta_{\text{DP}}$ ) to span a sufficient range of parameters such that we are able to reproduce the observed LFs with the distribution of obscured fractions of SFR.  $M_*$  and  $\alpha_{\text{DP}}$  are fixed to the determinations obtained by Harikane et al. (2022). We create 1000 x 100 realisations of  $\phi_*$  and  $\beta_{\text{DP}}$ .
- (ii) For each combination of  $\phi_*$  and  $\beta_{\text{DP}}$  we create the (intrinsic) LF. Then the intrinsic LF is divided up into various bins of magnitudes. For all galaxies within a bin we have the intrinsic magnitudes and therefore  $\text{SFR}_{\text{tot}}$ . To convert this to observed magnitudes and  $\text{SFR}_{\text{UV}}$  we apply the distribution of  $f_{\text{unobscured}}$  we find for our sample and use:  $\text{SFR}_{\text{UV}} = \frac{\text{SFR}_{\text{UV}}}{\text{SFR}_{\text{tot}}} \times \text{SFR}_{\text{tot}} = f_{\text{unobscured}} \times \text{SFR}_{\text{tot}}$ .
- (iii) We then sum the contribution from all dust-corrected galaxies in the previous step to arrive at a model UV LF and compare it to the observed UV LF of Harikane et al. (2022). We do this by obtaining the  $\chi^2$  value. To ensure that our derived LF results are fully guided by our observational constraints, we only infer the intrinsic LF for the SFR range that is covered by our sample of companion galaxies.

We repeat this procedure 1000 times while we randomly sample from the distribution of  $f_{\text{unobscured}}$  we find in Sec. 2.3.3 allowing  $f_{\text{unobscured}}$  to vary within its uncertainties. We then correct the SFR of each galaxy in the bin of the UV LF with each of the values in the distribution, while we make sure to conserve the number of sources in each bin according to the input intrinsic LF. When we sample from the distribution, we make sure to only correct the intrinsic UV LF with physical values of  $f_{\text{unobscured}}$ . This means that  $f_{\text{unobscured}} < 0$  is treated as 0 and  $f_{\text{unobscured}} > 1$  is treated as 1. Subsequently, this ensures that galaxies with  $f_{\text{unobscured}} = 0$  that would be non-detected in rest-UV observations, do not contribute to the UV LF. This method therefore takes into account that the intrinsic UV LF includes the correction on the luminosity of the sources as well as their inclusion in the UV LF according to their detectability. For each trial in these simulations, we save the input LF that has the smallest  $\chi^2$  value. Since we use the distribution of obscured fractions directly and therefore the obscuration does not depend on the SFR, we find that we can reproduce the observed

LF by only varying  $\phi_*$  and keeping  $\beta_{\text{DP}}$  constant to the value of Harikane et al. (2022). In Fig. 2.9 we show the intrinsic UV LF obtained by using the distribution of obscured fractions of SFR of the SFG and QSO companion samples with red and orange solid lines, respectively. The solid black lines show the observed UV LF from Harikane et al. (2022) at  $z \sim 5$  (left panel) and  $z \sim 6$  (right panel) and the shaded grey area shows the 16th and 84th percentiles of  $10^5$  realisations of the luminosity function as we vary the input parameters (e.g. faint-end slope) according to the uncertainties specified for the parameters. The intrinsic UV LF is a factor  $\sim 6$  and  $\sim 450$  larger than the observed UV LF based on the distribution of obscured fractions for SFG and QSO companions, respectively.

It is worthwhile comparing these results (utilizing the full distribution of obscured fractions) with what we would obtain if we assumed the same fixed obscured value for every source. We have performed this experiment with the obscured fraction of the SFG companion sample using both our forward-modeling code and analytically where

$$M_{\text{UV, intrinsic}} = M_{\text{UV, observed}} - 2.5 \times \log_{10}(f_{\text{unobs}}) \quad (2.12)$$

and find consistent results. We show the median intrinsic luminosity function in Fig. 2.9 with the dark red and dark orange dotted lines based on the mean dust obscuration from the SF galaxy and QSO companion sample, respectively. We see that for both samples this procedure results in significantly more galaxies at the bright end of the luminosity function than our primary procedure of sampling from the distribution of obscured fractions. The large differences are due to the fact that in our primary approach we account for the few high-SFR galaxies in our sample that are only subject to minimal dust obscuration rather supposing that all galaxies are subject to the same obscuration. This makes it possible to reproduce the observed LF with a much smaller overall correction and therefore the simplified approach presented in this paragraph could easily result in a SFRD that is substantially overestimated.

### 2.3.5 The SFRD at $z \gtrsim 5$

With the intrinsic luminosity function obtained from forward modeling we can calculate the SFRD at  $z \gtrsim 5$ . In this work our SF galaxy and QSO companion sample probe a limited range in SFRs,  $\approx 30 - 150$  and  $\approx 30 - 320 M_{\odot} \text{ yr}^{-1}$ , respectively. First we obtain the SFRD of these ranges by numerically integrating the LF of both samples that have been obtained using the distribution of  $f_{\text{unobscured}}$ . We obtain  $\log_{10}(\text{SFRD}/M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}) = -3.36^{+0.42}_{-0.24}$  for the SF galaxy and  $-1.85^{+2.08}_{-1.68}$  for the QSO companion sample. For comparison, if we would integrate the observed UV LF over the same range we obtain  $\log_{10}(\text{SFRD}/M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}) = -4.11^{+0.17}_{-0.18}$  and

$-4.51_{-0.33}^{+0.29}$  at  $z \sim 5$  and  $z \sim 6$ , respectively. If we compare the results from the observed and intrinsic LF solely integrated over the range of SFRs probed in this work, we find a median increase in density of  $\approx 6$  and  $\approx 500$ .

We also integrate to  $-17$  mag ( $0.20 M_{\odot} \text{ yr}^{-1}$ ), as is more commonly done and will make it easier to compare to literature results. When we integrate over magnitudes that are not probed by the SFRs in this work we assume the observed UV LF. Essentially we assume that at those magnitudes there is no dust obscuration and therefore this provides us with a lower limit on the SFRD. We obtain  $\log_{10}(\text{SFRD}/M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}) = -1.81_{-0.08}^{+0.09}$  for the SF galaxy and  $-1.55_{-0.40}^{+1.78}$  for the QSO companion sample. If we would integrate the observed UV LF to  $-17$  mag we obtain  $\log_{10}(\text{SFRD}/M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}) = -1.83_{-0.08}^{+0.08}$  and  $-1.94_{-0.16}^{+0.16}$  at  $z \sim 5$  and  $z \sim 6$ , respectively. The difference between the observed and intrinsic LF originates from the dust-obscured star formation, therefore the dust-obscured SFRD is  $\log_{10}(\text{SFRD}/M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}) = -3.44_{-0.33}^{+0.47}$  and  $-1.85_{-1.72}^{+2.08}$ , which is  $2_{-1}^{+4}\%$  and  $55_{-53}^{+44}\%$  of the total intrinsic SFRD according to the median values for the SF galaxy and QSO companion samples, respectively.

As the samples in this work probe the bright-end of the luminosity function, we also take a look at the fractional contribution that dust-obscured star formation provides to the SFRD integrated down to  $-20$  mag ( $\sim 3 M_{\odot} \text{ yr}^{-1}$ ). We obtain  $\log_{10}(\text{SFRD}/M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}) = -2.25_{-0.09}^{+0.10}$  for the SF galaxy and  $-1.77_{-0.74}^{+1.99}$  for the QSO companion sample. The contribution from dust-obscured star formation is  $7_{-4}^{+11}\%$  and  $84_{-74}^{+16}\%$  (or equivalently  $> 3\%$  and  $> 10\%$ ) of the total intrinsic SFRD integrated to  $-20$  mag leveraging results from the SF galaxy and QSO companion samples, respectively.

We note that the dust-obscured SFRDs are obtained from Monte Carlo sampling both the observed and intrinsic SFRD with their corresponding errors and subtracting the observed SFRD from the intrinsic SFRD to account for the asymmetric errors.

We present our results in Table 2.4 and Fig. 2.10 where they are also compared with the SFRD values from the literature. The solid black line indicates the SFRD evolution from Madau & Dickinson (2014) (MD14) and the dashed/dotted lines indicate the dust-rich and dust-poor models from Casey et al. (2018a). Both the intrinsic and observed SFRD inferred from the companions near SF galaxies agree with MD14. The companions near QSOs imply a dustier Universe with the intrinsic SFRD while the observed SFRD agrees with MD14.

We have shown in Sec. 2.3.4 that we obtain significantly larger intrinsic UV LF if we use the mean obscured fraction instead of the distribution. The corresponding obscured SFRD to these UV LFs are:  $\log_{10}(\text{SFRD}/M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}) = -2.48_{-0.68}^{+0.88}$  and  $-0.08_{-0.96}^{+1.41}$  for the mean

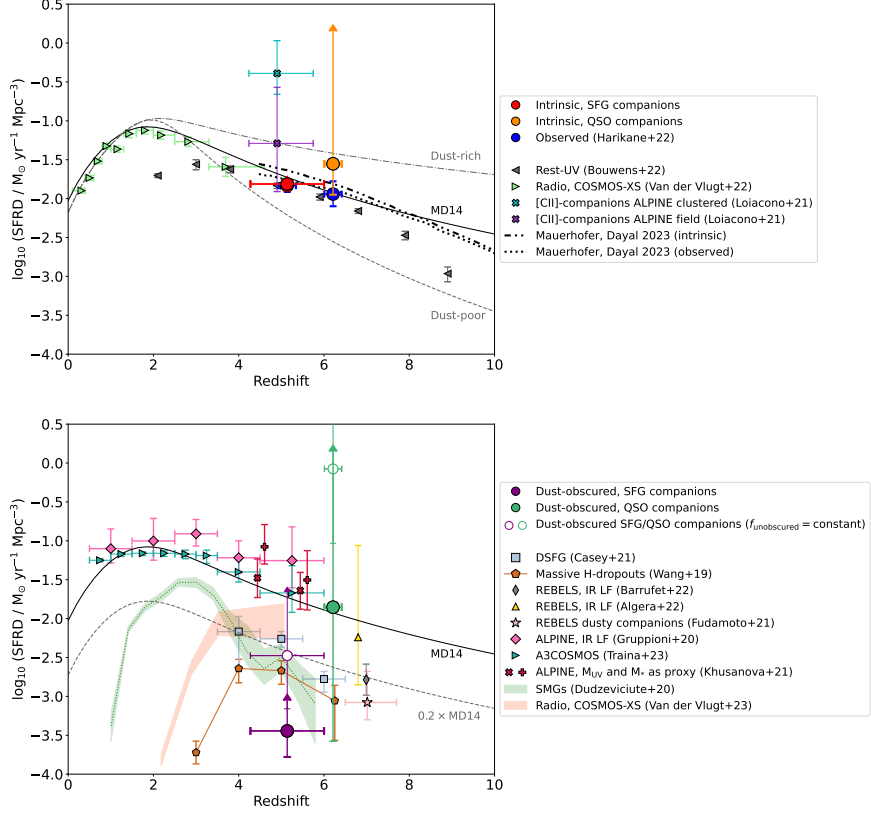
obscured fraction from the SFG and QSO companion sample, respectively. These results are presented in the lower panel of Fig. 2.10 with the empty purple and green markers.

As mentioned before, we find no relation between the obscured fraction and the total SFR. However, from the IRX- $M_*$  relation, one might reasonably expect a trend between the obscuration and the total SFR of a source. We therefore look into how a potential relation between obscuration and SFR would change our results. For this exercise we adopt the mass dependence found by Bowler et al. (2024) in the IRX- $M_*$  relation they derive based on ALPINE and REBELS galaxies. In Fig. 2.7 we can see that the relation from Bowler et al. (2017) indeed shows a similar slope between obscuration and SFR as the stacked results from ALPINE (B  thermin et al. 2020) and REBELS (Algera et al. 2023). To illustrate the impact such a dependence can have on our results, we introduce a small change in our procedure of correcting for the impact of dust obscuration by forward modeling the observed distribution of obscured fraction of the SFG companions. The change we make is to scale up or down the fractions according to the SFR of sources in the intrinsic LF using the same slope as Bowler et al. (2022) find for the IRX- $M_*$  relation. In the case of the obscuration being dependent on SFR we not only vary  $\phi_*$ , but make sure to also vary  $\beta_{\text{DP}}$ . Adopting the Bowler et al. (2024) dependence on SFR, we obtain a  $\log_{10}(\text{SFRD}/M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3})$  that is 0.37 dex smaller than our fiducial result assuming no dependence.

## 2.4 Discussion

### 2.4.1 Comparisons with earlier selections of [CII]-emitting Companion Galaxies

In this work we search for [CII]-emitting sources in the neighborhoods of luminous galaxies and quasars. Neighbours are identified through the detection of  $\text{SNR} \geq 6.2$  [CII] lines in the same ALMA data cubes as the target sources and with the requirement that the velocity difference with the primary target is  $\Delta v \leq 2000 \text{ km s}^{-1}$ . Our  $\text{SNR} \geq 6.2$  is based on a purity of  $> 95$  percent (Schouws et al. 2024, in prep.) for serendipitous detections with this SNR. Therefore, the probability that selected peaks originate from noise is very low, i.e.,  $< 5\%$ . We have examined the ALMA data cubes corresponding to 34 [CII]-detected target sources and found 2 neighbour candidates: J1208-0200C1 and J2228+0152C1. We expanded our sample with the neighbour galaxies from Loiacono et al. (2021), Venemans et al. (2020), Fudamoto et al. (2021) and REBELS-39 (Schouws et al. in prep.). We adopt the selection criteria of  $\Delta v \leq 500 \text{ km s}^{-1}$  to create a sample of galaxies that are most likely to be at  $z \gtrsim 5$ . However, we cannot



**Figure 2.10:** Constraints on the SFR density as a function of redshift. *Top panel:* The red and orange solid circles show the SFR density we estimate from our intrinsic UV luminosity functions based on dust obscuration from companions around SF galaxies and QSOs, respectively, integrated to  $-17$  mag. We only infer the intrinsic LF and dust corrections over the SFR range directly sampled by our work and therefore the dust corrected SFRDs we derive are lower limits. The blue solid circles show the SFR density when we integrate the observed UV LF from Harikane et al. (2022) to  $-17$  mag. The solid black line indicates the fiducial SFR density evolution from Madau & Dickinson (2014). Meanwhile, the dashed/dotted lines correspond to the dust-rich and dust-poor models from Casey et al. (2018a) and intrinsic and dust-attenuated models from Mauerhofer & Dayal (2023). Furthermore, we show the constraints on the rest-UV SFRD from Bouwens et al. (2022a) (*grey triangles*), the SFRD from radio-observations from van der Vlugt et al. (2022) (*green triangles*) and SFRDs from the [CII]-LF from the ALPINE companions from Loiacono et al. (2021) (*aqua and purple crosses*). *Lower panel:* The purple and green solid circles show the contribution of dust-obscured sources to the SFR density, based on the dust obscuration from SF galaxy and QSO companions, respectively. The empty purple and green circles show the dust-obscured SFRD when we assume a constant mean obscuration instead of using the distribution of obscured fractions. Other estimates of the contribution of obscured sources to the SFRD are shown, including those of Casey et al. (2021) (*light blue triangles*), Wang et al. (2019) (*orange pentagons*), Barrufet et al. (2023) (*gray thin diamond*), Algera et al. (2023) (*yellow triangle*), Fudamoto et al. (2021) (*pink star*), Gruppioni et al. (2020) (*pink diamonds*), Traina et al. (2024) (*aqua squares*), Khusanova et al. (2021) (*red cross and plus sign*), Dudzevičiūtė et al. (2020) (*green shaded region*) and van der Vlugt et al. (2023) (*orange shaded region*).

exclude the possibility that some neighbours are not [CII] emitters, but e.g. mid-J CO lines at a lower redshift. While we have verified that for the sources with available multi-wavelength observations there is no emission bluewards of the Lyman-alpha transition, fully excluding the low redshift possibility would require the detection of additional emission lines or spectral features in sources with spectroscopy. We have estimated in Sec. 2.3.3 the impact of excluding a possible lower redshift contaminant on the mean obscured fraction and found the result to still be within the errors on the mean obscured fraction. Multi-wavelength observations are also needed to robustly constrain the stellar masses of the sample, in order to more closely tie our results to the stellar mass of sources and not just their SFRs through [CII].

One important property of our sample that can indicate that we are observing [CII] is the equivalent width of the candidate [CII] emission lines. Our sample (with dust detections) has a high median EW of  $\approx 1.1\mu\text{m} \approx 2000 \text{ km s}^{-1}$ . High EW ( $> 1000 \text{ km s}^{-1}$ ) are also found for [CII]-emission from e.g. Lyman break galaxies (Capak et al. 2015) at  $z \sim 5 - 6$  and QSO host galaxies at  $z > 6.6$  (Venemans et al. 2016) and local starbursts (Sargsyan et al. 2014). In contrast to this, the median EW of CO at  $z \sim 1 - 2.5$  is typically lower (Decarli et al. 2016), e.g.,  $\approx 1000 \text{ km s}^{-1}$ .

As previously stated we find two neighbour galaxies from our own examination of 34 archival ALMA observations. Out of the 34 observations analysed in this work, 18 of these are observations of quasars. The 2 neighbour galaxies are found to accompany quasars, therefore 11 percent of the examined quasars have a neighbour galaxy. In comparison, Venemans et al. (2020) found 13 quasar host galaxies with companions that would qualify as a neighbour in 27 quasar observations, almost 50 percent. In Fig. 2.1 we see that the targets of Venemans et al. (2020) have higher [CII] luminosities compared to our sample of archival observations. Likely one of the most important factors in the lower prevalence of companion around the bright galaxy samples we consider vs. that seen in the Venemans et al. (2020) are the halo masses of the targets considered. Given their brightness and rarity, the bright QSOs studied by Venemans et al. (2020) almost certainly have much higher halo masses and therefore much more significant clustering (Mo & White 1996) than most of the sources we consider. While halo mass is likely the largest factor, it is also possible that feedback from the QSO could contribute to the prevalence of nearby companions by boosting their SFRs, as Zana et al. (2022) and Ferrara et al. (2023) discuss. Moreover, the Venemans et al. (2020) data is relatively deep, with quasars that have on source observation times larger than an hour, likely resulting in more significant serendipitous detections.

Another possibility for the low amount of companions that are found could be due to the resolution of the observations. We restricted our search to archival observations with an angular resolution of  $\gtrsim 0.2 \text{ arcsec}$ . We have



found the [CII]-emitting companions J1208-0200C1 and J2228+0152C1 that have observations with resolutions of  $\sim 0.4$  arcsec. The majority (21) of the 34 observations used to search for companion galaxies have resolutions of  $\gtrsim 0.4$  arcsec. It is however possible that we over resolve the [CII]-emission in those observations we consider that have higher spatial resolution and thus we miss companions there. As we do not obtain the intrinsic UV LF from the number of companion galaxies that are found, but rather based on the obscured fraction of the companions, we do not expect this to be an issue for our analysis. However, it could be an issue if there is a trend between the obscuration of a galaxy and source size, for example if some extended sources that are missed are generally less dust-obscured. If that would be the case we might overestimate the obscuration if we have missed some more extended companion galaxies.

One last contributing factor to why we do not find as many companions as in previous work around QSOs might be due to us only considering archival observations of sources without published companions in our search. Given the lesser frequency with which researchers report null results, we might be focusing on sources where companion searches have already been done, but the results are not deemed interesting enough to publish.

### 2.4.2 Obscured star formation derived from [CII]-Selected Samples

We estimate the obscured star formation based on the FIR emission we infer from the (dust) continuum of the galaxies in our sample. When we convert continuum fluxes to IR luminosities it is necessary to rely on assumptions on the dust temperature and emissivity as we cannot constrain these parameters of the dust SED. While there is literature that suggests dust might be hotter at higher redshift, there is also a large scatter in dust temperatures observed (e.g. Schreiber et al. 2018; Viero et al. 2022; Sommovigo et al. 2022a). As IR luminosity scales with the dust temperature and dust emissivity ( $L_{\text{IR}} \propto M_{\text{dust}} T_{\text{dust}}^{\beta_{\text{em}}+4}$ ) this translates into a larger uncertainty on the dust-obscured star formation. We include a systemic uncertainty of a factor of 3, to account for the dust temperature varying by 10K and  $\beta_{\text{em}}$  between 1.5-2.5. If the dust in galaxies in our sample were systematically colder than our fiducial assumptions, the IR luminosities and obscured SFR would decrease. Conversely, if the dust were, systematically hotter, this would increase the obscured SF.

An important challenge for the present inferences is the very small numbers of galaxies in our samples, particularly at lower SFRs. Clearly, larger samples of companion sources are required to extend the range of SFRs where the characteristics of companion galaxies can be well sampled and to improve the statistical confidence on these results. Yet, even with the

present small selection of companion sources, we can see that dust-obscured star formation contributes quite substantially to the SFR in individual star-forming galaxies at  $z > 4$ .

As we have seen in Sec. 2.3.5, whether the obscured fraction is characterised by the distribution of obscured fractions or a mean obscured fraction can have a significant impact on the resulting SFRD. The use of the mean obscured fraction is a simplified approach that would very likely lead to an overestimate of the dust-obscured SFRD, due to the omission of the impact of the modest fraction of sources where the dust extinction is substantially less. Therefore, we prefer to characterise the obscured fraction with the distribution of obscured fractions and it is even more important to obtain larger samples in order to better sample the distribution of obscured fractions.

### 2.4.3 The use of [CII] as a SFR tracer

Having constructed a sample of [CII]-selected companion galaxies, we aimed to use this selection to characterise the relation between the [CII] SFR of a galaxy and the fraction of the SFR that is directly visible in the rest-UV. The advantage of performing this characterization on a [CII]-selected sample is that we would not expect the results to suffer from the same selection biases that would impact the same characterization performed on galaxies identified with a UV selection, as is the case for the ALPINE and REBELS galaxies themselves.

We find the total SFR as obtained from the UV and IR of our galaxies and the  $L_{[\text{CII}]}$  of the galaxies to generally follow the relation from De Looze et al. (2014). This is consistent with e.g. results from the high-resolution SERRA simulations that finds  $z \sim 8$  galaxies to lie on the standard [CII]-SFR relation (Pallottini et al. 2022). However, there is also literature questioning the use of [CII] as a tracer of SFR at high redshift (e.g. Liang et al. 2024). Liang et al. (2024) find with the use of the FIRE simulations that early Epoch of Reionization galaxies could have a [CII] deficit due to their low metallicities. If such a [CII]-deficit is present in our galaxy sample, this would imply less obscured star formation rates at a certain [CII] SFR. It may furthermore suggest that we may overestimate the obscured SFR if our sample is significantly deficient in [CII]. However, [CII] deficits are mostly observed in very bright and compact sources such as ULIRGs ( $L_{\text{IR}} \geq 10^{12} L_{\odot}$ ), SMGs and QSOs. 3 out of the 18 galaxies in our sample have IR luminosities corresponding to ULIRGs, however these galaxies do not show a greater [CII] deficit than the other galaxies in the sample. As we do not have estimates on the gas properties (e.g. metallicities) of the galaxies in our sample we cannot use the  $L_{[\text{CII}]}$ -SFR relation as described in Liang et al. (2024).

One further caveat to the above argument concerns the origin of the [CII] emission. [CII] emission is believed to originate mainly from photodis-

sociation regions at high redshift (Hollenbach & Tielens 1999; Wolfire et al. 2022). However, it is unclear what fraction of [CII] emission originates from the neutral gas. Croxall et al. (2017) find that in local star forming galaxies 60% to 80% of the [CII] emission is from neutral gas. This neutral gas is heated by the photo-electric effect on small dust grains and cooled via [CII] emission. Therefore, since an unknown fraction of the [CII]-emission in high redshift galaxies could be due to photoelectric heating, it could subsequently bias us to slightly dustier galaxies. We note that this method is still preferred over UV-selection since the [CII] line is not affected by dust attenuation and we find our results to be in agreement with other selection methods seen in Sec. 2.4.5.2.

The use of [CII] as a SFR tracer could also be questioned if the [CII] and IR or UV show significant offsets, as the emission would probe different environments. We do not observe such large offsets, but the resolution of the [CII] observations used in this work is limited and therefore we cannot conclude whether these offsets are present. Lastly, some of the [CII] emission might correspond to the halo emission of a galaxy. With high-resolution modeling Schimek et al. (2024) find that up to 10% of the emission of [CII] is from the circumgalactic medium (CGM). If a significant percentage of [CII] emission does indeed result from the CGM in contrast to the disks of galaxies, this would decrease the reliability of [CII] as a SFR tracer.

#### 2.4.4 Characterising the obscured fraction of SFR from [CII] emitters found in overdense regions

One important uncertainty is connected to our overall methodology employed here. Due to the use of companion galaxies we are biased to overdense regions. We emphasise that we do not use the companion galaxies to derive the UV LF results directly as we base our results on the LF function from Harikane et al. (2022). We only use the properties of the companion galaxies to infer the dust obscuration at  $z \gtrsim 5$  to circumvent the bias that one introduces by UV-selection. For our method we assume that the [CII]-selected companion galaxies used to compute the dust obscuration and the UV-selected galaxies used to compute the LF are drawn from the same distribution of star-forming galaxies. Moreover, we assume that this population of star-forming galaxies dominate the UV LF (especially at the bright end) and is the main contributor to the SFRD at high redshift. Although we are not biasing our LF to overdense regions, we might ask ourselves if companion galaxies identified in overdense regions around massive galaxies or QSOs are more likely to be dust-obscured at a given SFR than galaxies in lower density environments. While posing this question, it is important to keep in mind that the richest environments contribute an increasingly large fraction of the SFR density at high redshift, with  $\sim 30\%$  of the SFR density at  $z \sim 6$  expected to originate from only  $\sim 5\%$  of the cosmic volume

(Chiang et al. 2017).

In theory, it is possible that galaxies in overdense regions are more dusty given that outflows and starbursts can enrich the IGM with metals. This has indeed been shown to be the case in studies of simulated galaxies (e.g. Graziani et al. 2020; Di Cesare et al. 2023). The *Astraeus* framework (Hutter et al. 2021) that fully couples galaxy formation and reionization has been used to show that radiative feedback preferentially quenches star formation more in low-mass galaxies ( $M_{\text{halo}} < 10^{9.5} M_{\odot}$ ) in overdense environments, affecting their star formation rates (Legrand et al. 2023). However, the metal and dust contents of early galaxies in the framework are primarily determined by a combination of their stellar masses and star formation rates (Ucci et al. 2023). For example, low-mass halos with high star formation rates show extremely low metallicities due to outflows removing their dust and metal contents. Moreover, the *FLARES* simulations find that both SFR and metallicity depend primarily on the stellar mass of sources and show no strong dependence on the environment (Wilkins et al. 2023). Additionally, Champagne et al. (2018) find no evidence for an overdensity of continuum sources around  $z > 6$  quasars, in contrast to the overdensity of [CII] emitters, but concede that deeper observations may be needed to arrive at more robust conclusions.

Interestingly, we observe the companion galaxies around QSOs to be more dust-obscured than companion galaxies with similar SFRs around SF galaxies. The inferred mean obscured fractions imply that companion galaxies around QSOs have SFRs that are on average 93% obscured, whereas companion galaxies around SF galaxies have less obscured SFR, at 63%. The suggestion that QSO environments might be more dusty is not new, as this has been used to explain the lack of overdensities around QSOs in the rest-UV (e.g. García-Vergara et al. 2019). Generally, the QSO observations have higher spatial resolution than the observations of star-forming galaxies. This could mean that we are biased to more compact systems and as mentioned before we could overestimate the obscuration if there is a trend between the obscuration of a galaxy and the source size. To understand if QSO environments are statistically more dusty and what the underlying physical scenarios are, we need larger and deeper samples. Moreover, we would benefit from even higher resolution data to understand whether an interaction between QSO and the companions is evident and could be the cause of increased dust obscuration. Larger samples could further help us to correct for a potential bias due to overdensity, e.g. if we could quantify obscuration as function of overdensity factor.

The SF-galaxy sample shows some low-attenuation outliers as we can clearly see in Fig. 2.7. The number of galaxies would be too small to conclude whether this is a statistically different population of galaxies. However, regardless of the size of sample we do not observe these objects to be special (e.g. they are not at larger velocity differences than the rest of our

sample). As these outliers are non-detected in the continuum we cannot infer whether the IR and rest-UV emission in these objects are spatially offset ([CII] and the rest-UV are spatially coincident) and if this could be responsible for the low-attenuation observed in these objects.

### 2.4.5 Evolution of the SFRD

In Figure 2.10, we show the SFRD we derive by integrating our intrinsic UV LF to  $-17$  mag (*red and orange circles*), as well as the contribution from dust-obscured star formation alone (*purple and green circles*) and the SFRD without any dust correction (*blue circles*). When integrating to  $-17$  mag, we make use of the dust-uncorrected UV LF from Harikane et al. (2022) at the fainter magnitudes we do not probe in this work. Therefore, we show our SFRDs as lower limits, as we have assumed zero attenuation for galaxies with lower SFRs. Besides the relation found by Madau & Dickinson (2014) plotted with the solid black line, we also show the dust-poor and dust-rich models from Casey et al. (2018a). The dust-poor model indicates the SFRD where dusty star-forming galaxies (DSFGs) contribute  $< 10\%$  to the SFRD at  $z > 4$ , whereas the dust-rich model shows the SFRDs where DSFGs contribute  $\sim 90\%$  of the SFRD at  $z > 4$ .

#### 2.4.5.1 Comparison of total SFRD

The top panel of the figure shows that the SFRD from the observed UV LF from Harikane et al. (2022) also agrees with the rest-UV SFRD (integrated to  $-17$  mag) from Bouwens et al. (2022a). The intrinsic SFRD from the dust-obscured SF galaxy companions agrees with the total SFRD inferred from radio (3 GHz) observations (van der Vlugt et al. 2022), which is not surprising given the fact that dust obscuration would also have no impact on radio probes. The high SFRD inferred on basis of QSO companions is suggestive of a potentially much higher contribution from dust-obscured star formation at  $z > 5$ , but there might be environmental biases which impact this result and cause it to be higher than in other probes.

As we have made use of the sample of companion galaxies identified in ALPINE from Loiacono et al. (2021) we want to see if our results agree with the SFRD based on the [CII] luminosity function from that work. The [CII] LF is constructed from the companion galaxies in the ALPINE observations. Therefore the LF is biased by the use of overdense regions, we can see this clearly as the clustered SFRD is much higher than expected from previous results. We aim to arrive at more representative results by basing our results on the wide-area rest-UV LF results of Harikane et al. (2022) and then correcting for the impact of dust using our obscured fraction determinations derived using our companion sample. The SFRD computed

**Table 2.4:** Summary of our key results for galaxies at  $z \sim 5$ -6.

<b>Unobscured SFR</b>	
Mean unobscured fraction	$f_{\text{unobscured,SFG}} = 0.33^{+0.16}_{-0.14}$
$\frac{\text{SFR}_{\text{UV}}}{\text{SFR}_{\text{UV}} + \text{SFR}_{\text{IR}}} = (1 + A \times \text{SFR}_{\text{[CII]}}^B)^{-1}, B = 0$	$A_{\text{SFG}} = 1.6^{+1.7}_{-0.8}$
	$f_{\text{unobscured,QSO}} = 0.07^{+0.05}_{-0.05}$
	$A_{\text{QSO}} = 11.6^{+15.6}_{-5.4}$
<b>Star formation rate densities (lower limits)</b>	
	$\log_{10}(\text{SFRD}_{\text{SFG}} / M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3})$
Intrinsic (Including obscured, $M_{\text{UV}} < -17$ )	$-1.81^{+0.09}_{-0.08}$
Observed (Based on rest-UV, $M_{\text{UV}} < -17$ )	$-1.83^{+0.08}_{-0.08}$
Dust-obscured ( $M_{\text{UV}} < -17$ )	$-3.44^{+0.47}_{-0.33}$
	$\log_{10}(\text{SFRD}_{\text{QSO}} / M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3})$
	$-1.55^{+1.78}_{-0.40}$
	$-1.94^{+0.16}_{-0.16}$
	$-1.85^{+2.08}_{-1.72}$



from the field galaxy found in the ALPINE data by Loiacono et al. (2021) is more in line with previous literature and our own SFRD estimates.

Also shown in the top panel of Figure 2.10 are the results of the DELPHI semi-analytic model by Mauerhofer & Dayal (2023) integrated to  $-17$  mag. In this model the maximum star-formation efficiency and fraction of supernova energy coupled to gas are baselined to galaxies from ALPINE and REBELS. Their intrinsic UV LF indeed implies more SFRD than the dust-attenuated (or observed) LF.

### 2.4.5.2 Dust-obscured SFRD

It is interesting to compare the present estimates of the dust-obscured star formation from the [CII]-selected companions with estimates from previous studies as shown in the lower panel of Figure 2.10, including those by Casey et al. (2021), Wang et al. (2019), Barrufet et al. (2023), Algera et al. (2023), Fudamoto et al. (2021), Gruppioni et al. (2020), Dudzevičiūtė et al. (2020), Traina et al. (2024), Khusanova et al. (2021) and van der Vlugt et al. (2023). In comparing to previous work in the literature, it is important that we pay close attention to the SFR limits that various studies probe. We compute the SFRD from integrating our obtained intrinsic luminosity function from  $0.20 M_{\odot} \text{ yr}^{-1}$  ( $-17$  mag) to  $\approx 150$  and  $\approx 320 M_{\odot} \text{ yr}^{-1}$  for the SF galaxy and QSO companion sample, respectively. Important to note is that for the magnitudes fainter than those that are sampled by the SFG and QSO companions, we assume no dust-obscuration, and therefore our SFRD estimates are lower limits.

Starting with the dust-obscured SFRD from the Casey et al. (2021) 2mm probe down to  $\sim 2 \times 10^{12} L_{\odot}$  ( $200 M_{\odot} \text{ yr}^{-1}$ ), we find a lower and higher dust-obscured SFRD we derive from our SF galaxy and QSO companion sample, respectively. Moving towards slightly deeper probes of obscured star formation in the  $z > 4$  universe, we consider the search for sub-millimetre galaxies (SMGs) provided by Dudzevičiūtė et al. (2020) to  $S_{870} \geq 1$  mJy, corresponding to  $\approx 100 M_{\odot} \text{ yr}^{-1}$  and the ( $M_{*} > 10^{10.3} M_{\odot}$ ) H-dropouts from Wang et al. (2019). Again the dust-obscured SFRD from the SFG/QSO companion are smaller/greater than the SFRD from these SMGs. We find a much higher dust-obscured SFRD using the obscured fraction of QSO companions, but this may be due to galaxies near bright QSOs being subject to more dust obscuration. We note however that the uncertainty on the intrinsic and observed SFRD based on the distribution of obscured fractions of the QSO companions are large.

Slightly deeper yet, in Fudamoto et al. (2021), two companion sources in the REBELS data are described with obscured star formation rates of  $40 - 70 M_{\odot} \text{ yr}^{-1}$ . Other results for the dust-obscured SFRD from REBELS obtained by Algera et al. (2023) and Barrufet et al. (2023) arrive at similar SFR density determinations to Fudamoto et al. (2021). These SFRD

estimates are at slightly higher redshift,  $z \sim 7$ , and they are higher than the obscured SFRD from our SF galaxy companion sample, while they are slightly lower than from our inferences using our QSO companion sample.

Lastly, results from the dust-obscured SFRD from the ALPINE targets by Gruppioni et al. (2020) show higher SFRDs than in other studies. This may be due to their SFRD results being integrated down to  $\simeq 10^{-2} M_{\odot} \text{ yr}^{-1}$ . Results based on the ALPINE survey by Khusanova et al. (2021) are also shown and indicate a relation between  $L_{\text{IR}}$  and both the UV magnitudes and stellar mass. The values shown in Fig. 2.10 are the obscured SFRD they infer extrapolating their sample results (obtained to  $-20$  mag) to  $-17$  mag. Also shown is the contribution of optically dark radio sources in COSMOS-XS as found by van der Vlugt et al. (2023). At  $z \sim 5$  they find radio sources down to SFRs of  $\approx 500 M_{\odot} \text{ yr}^{-1}$  and integrate over the whole LF to obtain the dust-obscured SFRD, that agrees with our SFRD based on the QSO galaxy companion sample. Results from Traina et al. (2024) from the A<sup>3</sup>COSMOS survey that utilises archival ALMA data on the COSMOS field to conduct an unbiased probe, are integrated down to the same SFRs as the results from ALPINE and show comparable SFRDs. Our dust-obscured SFRD from the QSO companions agrees with these results, and our results imply that a higher SFRD than observed from UV-selected galaxies with a small dust correction (as shown by the relation from Madau & Dickinson 2014) is probable.

The corrected SFRD estimates from this work demonstrate again that dust-obscured star formation likely contributes meaningfully to the total SFRD of the Universe at  $z \gtrsim 5$ . Moreover, we see that SFRD estimates derived based on UV-selected samples may under predict the SFRD due to biases associated with a UV selection. Massive H-dropouts and SMGs contribute significantly to the obscured SFRD and their densities agree with the average dust-obscured SFRD we obtain from the SFG and QSO companion sample together.

For context, we also show the dust-obscured SFRD we obtain if we assume that all sources show the same mean obscured fraction as shown in Fig. 2.7 and do not utilise the full distribution to perform the correction (as in our fiducial calculations). With larger samples, we should be able to better constrain the full distribution of the obscured fraction and its dependence on the total SFR and improve constraints on the dust-obscured SFRD at  $z \gtrsim 5$ .

## 2.5 Conclusions

Here we present an entirely new methodology for estimating the contribution from obscured star formation in the high-redshift Universe. Essential to our methodology is the selection of sources using the [CII] line, which has

been shown to be an excellent tracer of the total SFR in galaxies across cosmic time (De Looze et al. 2014; Schaerer et al. 2020; Bouwens et al. 2022b; Carniani et al. 2020) and not to depend sensitively on the extent to which SFR is obscured. Selection of sources via [CII] line emission can be performed in the same ALMA data sets as used for the characterization of bright galaxies and QSOs using the same line, but key is to construct these samples from sources found at similar redshift as the (UV-selected) target sources. This method of identifying dusty, far-IR bright galaxies is significantly cheaper than the use of blank field surveys as the ALMA FOV is relatively small. Given that the selected galaxies will be identified using their [CII]-derived SFRs alone, we can make use of rest-UV observations (where available) to determine the fraction of their SFR which is obscured.

In constructing [CII]-selected samples of companion galaxies for this purpose, we made use of selections of companion galaxies in the literature from Venemans et al. (2020), ALPINE (Loiacono et al. 2021), and from REBELS (Bouwens et al. 2022b; Fudamoto et al. 2021), but we also did a search for these sources in ALMA observations of  $z > 6$  galaxies from the archive, considering ALMA observations of 55 separate targets. Of the 34 sources where a [CII] line was found for the main target, two dust-obscured companion galaxies were found with the line search code MF3D: J1208-0200C1 and J2228+0152C1. Neither neighbour has a detection in the rest-UV, possibly due to obscuration by dust. A less obscured companion is found around REBELS-39, with a  $2\sigma$  detection in the rest-UV. Figure 2.1 shows the full range of redshifts and [CII] luminosity for targets we considered in constructing our samples.

Using these samples to characterise the fraction of SFR that is obscured, we only consider [CII]-selected companion galaxies where the available rest-UV observations are deep enough to detect sources in the rest-UV even if they are 85% obscured. Additionally, to maximise the probability that sources in these [CII]-selected samples are at  $z > 4$ , we only consider neighbor candidates whose systemic velocity difference with the target galaxies are  $\Delta v \leq 500 \text{ km s}^{-1}$ . On the basis of the rest-UV observations, we estimate two fractions of obscured SFR, one for galaxies found in the neighborhoods around SF galaxies and one for galaxies around QSOs. The companion sample near SF galaxies suggest  $f_{\text{obs,median}} = 63\%$  similar to those expected from UV-selected samples. The [CII]-selected galaxies around QSOs indicate  $f_{\text{obs,median}} = 93\%$ , suggesting that QSO environments are even more dust-enriched.

With the fraction of obscured SFR we derive, we used a forward modeling methodology to determine the intrinsic UV LF from the observed UV LF results by Harikane et al. (2022). By the intrinsic UV LF, we indicate the UV LF we would expect to obtain prior to the impact of dust on either the rest-UV brightness or selectability of sources. We integrate the intrinsic UV LF to  $-17 \text{ mag}$  to find  $\log_{10}(\text{SFRD}/M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}) \geq -1.81_{-0.08}^{+0.09}$  and

$-1.55_{-0.40}^{+1.78}$  applying the distribution of obscured fractions from companion galaxies nearby SF galaxies and QSOs, respectively (Table 2.4). It is possible the dust obscuration is underestimated due to the lack of constraints we have at low SFRs, although the QSO companion sample suggest that their environments may be dustier than the environment of SF galaxies. Using this new methodology we find that  $> 2_{-1}^{+4}\%$  and  $> 55_{-53}^{+44}\%$  of the SFRD is dust-obscured, based on the SF galaxy and QSO companion sample, respectively. If we integrate to  $-20$  mag we obtain  $\log_{10}(\text{SFRD}/M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}) = -2.25_{-0.09}^{+0.10}$  for the SF galaxy and  $-1.77_{-0.74}^{+1.99}$  for the QSO companion sample. This is  $7_{-4}^{+11}\%$  and  $84_{-74}^{+16}\%$  (or equivalently  $> 3\%$  and  $> 10\%$ ) of the total intrinsic SFRD integrated to  $-20$  mag for the SF galaxy and QSO companion samples, respectively. This result indicates that the effect of dust-obscuration is substantial at the bright-end of the luminosity function.

To constrain the obscured SFR relation more robustly, we need more sensitive probes for [CII] emission from companion galaxies and this will also provide us constraints to lower star formation rates. This is particularly important as we are not yet able to detect a dependence of the obscured fraction of SFR, but if the obscured fraction increases with SFR (as expected from most derived IRX- $M_*$  relations, e.g., Bowler et al. 2024), this would have a clear impact on the derived SFRD. Moreover, observations in more ALMA bands are needed to probe the full IR SED, to decrease uncertainties on the IR luminosities. Rest-UV observations are also needed for the majority of the companion galaxies that are found in the high redshift Universe, so that these sources can also be used to infer the dust obscuration. Future JWST observations (e.g. from ASPIRE; program ID 2078; PI: F. Wang) will be beneficial for our understanding of stellar masses of these companion galaxies and confirming the redshifts as found by [CII]. One way of making immediate progress on this front is to make use of the deeper observations being obtained around massive  $z > 4$  galaxies from CRISTAL (2021.1.00280.L, PI: Herrera-Camus, Rodrigo), an ALMA Large Program approved for execution in cycle 8 targeting 19 massive galaxies in the range  $z = 4-6$ . In addition, we will be making use of the ALMA observations from the  $\sim 49$ -hour Cycle-10 program CISTERN targeting 195 spectroscopically confirmed QSOs and bright galaxies. Among other open questions, the data should allow us to better determine if highly dust-obscured galaxies are substantially more common in QSO environments than they are around star-forming galaxies.

## Acknowledgements

The authors are very appreciative to our referee, Matthieu Bethermin, for the constructive feedback that significantly improved this paper. RJB acknowledges support from NWO grants 600.065.140.11.N211 (vrijcompeti-

tie) and TOP grant TOP1.16.057. JH acknowledges support from the ERC Consolidator Grant 101088676 (VOYAJ) and the VIDI research program with project number 639.042.611, which is (partly) financed by the Netherlands Organization for Scientific Research (NWO). MA acknowledges support from ANID basal grant FB210003. LB acknowledges support by ERC AdG grant 740246 (Cosmic-Gas). RB acknowledges support from an STFC Ernest Rutherford Fellowship [grant number ST/T003596/1]. PD acknowledges support from the NWO grant 016.VIDI.189.162 ("ODIN") and from the European Commission's and University of Groningen's CO-FUND Rosalind Franklin program. HI and HSBA acknowledge support from the NAOJ ALMA Scientific Research Grant Code 2021-19A. IDL acknowledges funding support from ERC starting grant 851622 DustOrigin. AF acknowledges support from the ERC Advanced Grant INTERSTELLAR H2020/740120. LG and RS acknowledge support from the Amaldi Research Center funded by the MIUR program "Dipartimento di Eccellenza" (CUP:B81I18001170001).

This paper makes use of the following ALMA data: ADS/JAO.ALMA#2017.1.00541.S, ADS/JAO.ALMA#2017.1.00428.L, ADS/JAO.ALMA#2019.1.01634.L. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

The project leading to this publication has received support from ORP, that is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004719 [ORP].

This research is based on observations made with the NASA/ESA Hubble Space Telescope obtained from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with program(s) 12184, 12974, 13645, 14185, 14876 and 15064.

The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID #2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing-Arizona Sky Survey (BASS; NOAO Prop. ID #2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; Prop. ID #2016A-0453; PI: Arjun Dey). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory (LBNL). The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. LBNL is managed by the Regents of the University of California under contract to the U.S. Department of Energy.

This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientifico e Tecnologico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenossische Technische Hochschule (ETH) Zurich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciencies de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig Maximilians Universitat Munchen and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories of China, the Chinese Academy of Sciences (the Strategic Priority Research Program "The Emergence of Cosmological Structures" Grant # XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. The BASS is also supported by the External Cooperation Program of Chinese Academy of Sciences (Grant # 114A11KYSB20160057), and Chinese National Natural Science Foundation (Grant # 12120101003, # 11433005).

The Legacy Survey team makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a



project of the Jet Propulsion Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration.

The Legacy Surveys imaging of the DESI footprint is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH1123, by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; and by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to NOAO.

## Data Availability

---

All ALMA observations used are publicly available in the ALMA Science Archive (<https://almascience.nrao.edu/aq/>). The rest-UV HST observations are obtained with the MAST portal (<https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html>). DECaLS observations are acquired from the Legacy Survey website (<https://www.legacysurvey.org/>). MF3D is publicly available on github (<https://github.com/pavesiricardo/MF3D>).

## 2.A Appendices

---

### 2.A.1 Low-redshift contamination

One potential uncertainty in this analysis is the nature of the serendipitous emission lines detected in the same data cubes as the primary  $z \geq 4$  ALMA targets. While the vast majority of these discovered lines very likely correspond to [CII], some of the discovered lines might originate from sources at lower redshift and in fact correspond to a CO transition. To estimate the probability of this, we use 27 target quasar host galaxies and 27 serendipitously identified galaxies of Venemans et al. (2020). We focus on the velocity difference between the serendipitous lines and the quasar host galaxy as given in Eq. 2.1. To estimate the number of low redshift contaminants, we look at the number of galaxies with  $\Delta v > 2000 \text{ km s}^{-1}$ , similarly to what is done in Venemans et al. (2020). Serendipitous lines with larger velocity differences are substantially less likely to be associated with the quasar host galaxy than lines with smaller velocity differences. At large velocity differences, serendipitously identified sources very likely correspond to the same type of sources found in random search fields, i.e., CO transitions from lower redshift galaxies (e.g. González-López et al. 2019). There are 8 serendipitous lines with  $\Delta v > 2000 \text{ km s}^{-1}$ , which we assume to be low redshift contaminants. The setup of the 27 quasars that are observed varied

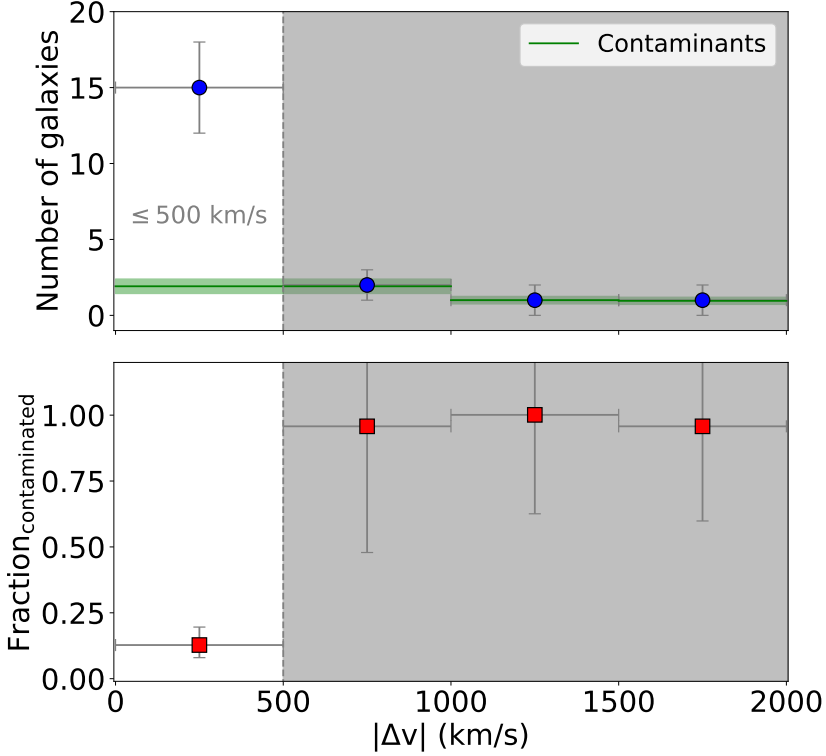
slightly from source to source, but on average two spectral windows overlapped in order to provide a coverage of 3.3 GHz around the quasar host galaxy, corresponding to  $3600 \text{ km s}^{-1}$  at  $z = 6$ . Two spectral windows are used at 15 GHz away in order to probe the continuum emission, so in total about  $7200 \text{ km s}^{-1}$  is covered per observation. Using the amount of contaminants at  $\Delta v > 2000 \text{ km s}^{-1}$  as a proxy for the low redshift sources we can expect per GHz, we estimate there would be 6 lower-redshift contaminants lying within  $\Delta v \leq 2000 \text{ km s}^{-1}$  of the main target. In Fig. 2.11 the upper panel shows the number of serendipitous lines as a function of the velocity difference. The green line shows the expected number of contaminants within one velocity bin. The lower panel shows the fraction of galaxies that are expected to be contaminants as a function of the velocity difference. The values plotted are the median values with 16th and 84th percentiles from bootstrapping 10000 times. We see that the fraction of contaminants is low ( $\approx 0.13$ ) for serendipitous lines with  $\Delta v \leq 500 \text{ km s}^{-1}$  and within that velocity limit we would expect to find about 2 contaminants for the sample of Venemans et al. (2020). At  $\Delta v > 500 \text{ km s}^{-1}$  the fraction increases significantly and would approach  $\approx 1.0$ . These numbers are an upper limit as some galaxies with  $\Delta v > 2000 \text{ km s}^{-1}$  could nevertheless be associated with the quasar host galaxy and be at  $z > 4$ . Choosing the velocity limit to be  $\Delta v \leq 500 \text{ km s}^{-1}$  is therefore a safe limit to assume for the serendipitously identified galaxies to be at high redshift as the fraction of expected contaminants is low.

## 2.A.2 Neighbour candidates

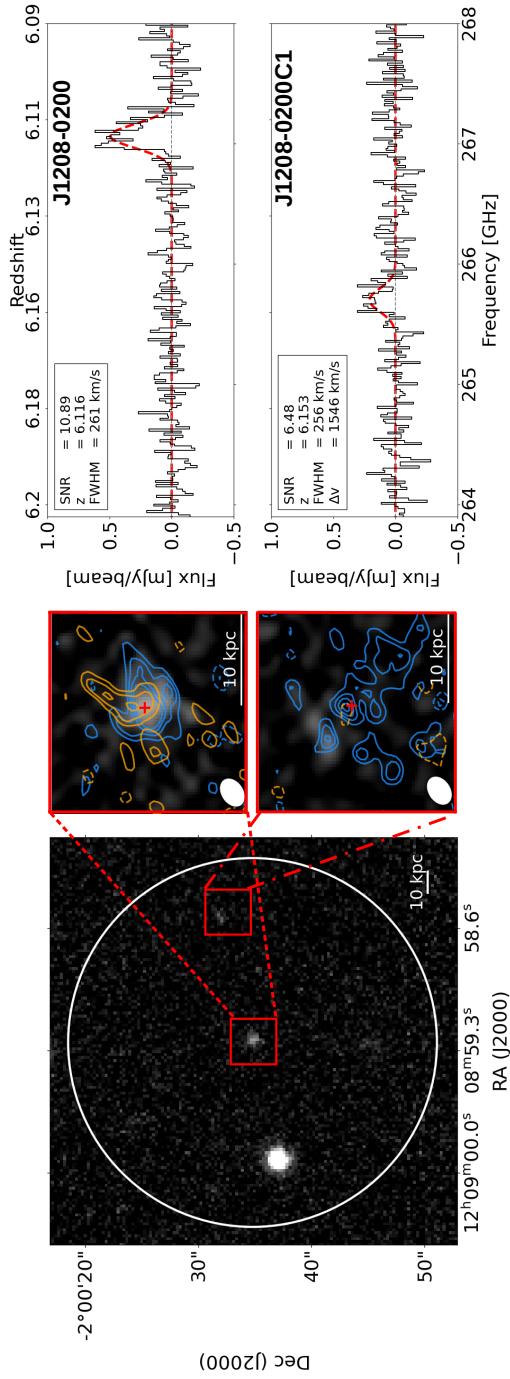
In this work we identify two neighbour candidates in archival ALMA data that had not been previously reported. In §2.3.1 we show one of our [CII] neighbours candidates: J2228+0152C1. Here we show the second candidate J1208-0200C1 in Figure 2.4. Similar to J2228+0152C1 J1208-0200C1 shows no detection in the available observations probing the rest-UV. Moreover, in Figure 2.13 we show the companion galaxy to REBELS-39, which shows a companion galaxy with a  $2\sigma$  detection in the rest-UV, in contrast to the dust-obscured galaxies from Fudamoto et al. (2021).

## 2.A.3 SExtractor

As explained in §2.2.4 we use SExtractor to measure the rest-UV fluxes of the sources imaged with HST and DECaLS. Unlike the ALPINE sources, we do not have sufficient data on these sources to use MOPHONGO. Table 2.5 shows the aperture fluxes and corresponding magnitudes of the Venemans et al. (2020) quasars and neighbour galaxies. The aperture fluxes of the neighbour sources found in this study are presented in Table 2.6.

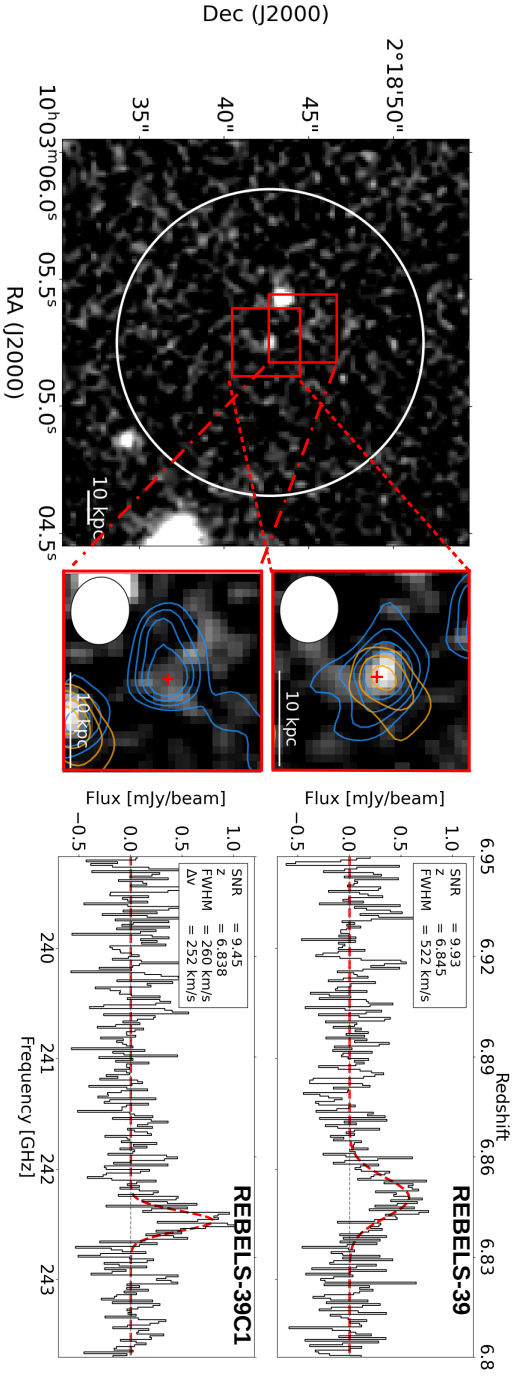


**Figure 2.11:** Top panel: the number of serendipitous lines identified by Venemans et al. (2020) as a function of the absolute velocity difference ( $\Delta v = \frac{z_{\text{comp}} - z_{\text{target}}}{1 + z_{\text{target}}}$ ). In green the number of expected contaminants within one velocity bin is plotted. Lower panel: the fraction of galaxies that are contaminants as function of the velocity difference. For galaxies with  $\Delta v \leq 500 \text{ km s}^{-1}$  the fraction of contaminants is the smallest ( $\approx 0.13$ ) and therefore the probability of them being high redshift sources is high.



**Figure 2.12:** Similar to Figure 2.4 from the main text but for the QSO target J1208-0200. Here we present a newly identified companion galaxy J1208-0200C1.





**Figure 2.13:** Similar to Figure 2.4 from the main text but for the REBELS target REBELS-39 and the background image is the J-filter from VISTA/VIRCAM. Here we present a newly identified companion galaxy REBELS-39C1.



**Table 2.5:** The rest-UV aperture flux and magnitudes of the Venemans et al. (2020) quasars and companion galaxies as measured by SExtractor. Whenever possible, various HST images probing the rest-UV light are stacked to increase the S/N.

Source	Flux <sub>aper</sub> (nJy)
J0100+2802Q	291383 ± 143
J0100+2802C1	236 ± 143
J0305-3150Q	17734 ± 15
J0305-3150C1	16 ± 15
J0842+1218Q	41163 ± 14
J0842+1218C1	19 ± 14
J0842+1218C2	122 ± 14
J1319+0950Q	42189 ± 23
J1319+0950C2	31 ± 23
P231-20Q	36992 ± 13
P231-20C1	-0.5 ± 13
J2054-0005Q	20576 ± 10
J2054-0005C1	7 ± 10
J2100-1715Q	16832 ± 26
J2100-1715C1	-24 ± 26

**Table 2.6:** The rest-UV aperture flux and magnitudes of the two quasars and neighbours from the ALMA Science Archive as obtained through SExtractor. We used the observations of the DECaLS survey in the  $z$ -filter.

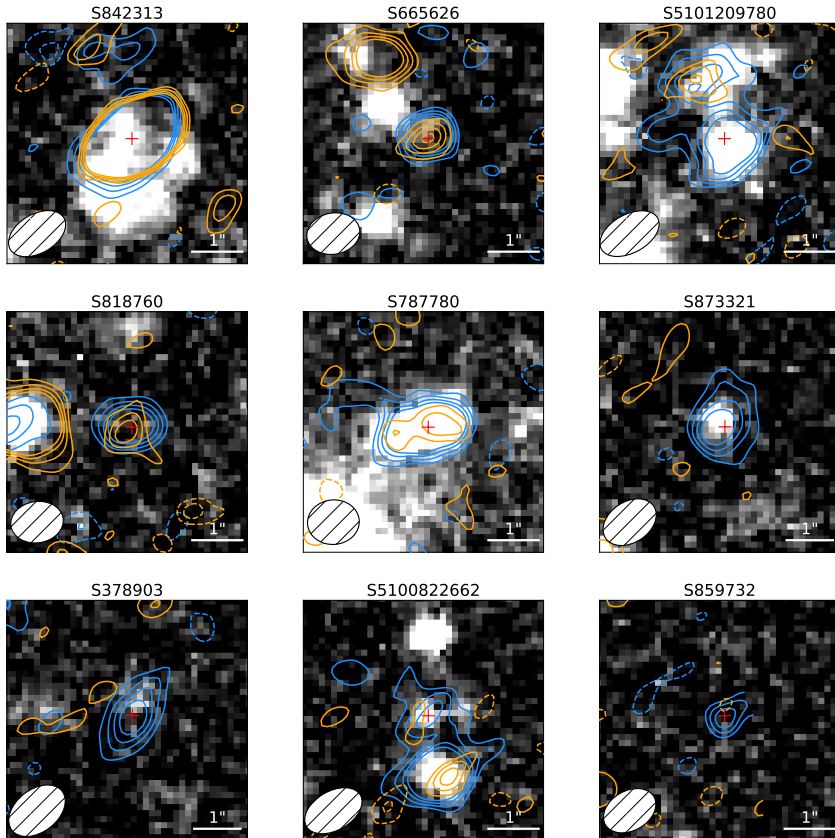
Source	Flux <sub>aper</sub> (nJy)
J1208-0200Q	1358 ± 80
J1208-0200C1	171 ± 80
J2228+0152Q	784 ± 84
J2228+0152C1	1 ± 84

### 2.A.4 ALPINE companion galaxies

In this work we make use of companion sources identified around [CII]-emitting galaxies from ALPINE in constraining the obscured fraction of SFR. Given the lack of published coordinates for the companion sources identified by Loiacono et al. (2021), we conducted our own search for these sources with MF3D. The 9 sources we found are shown in Figure 2.14. The red crosses mark the positions of the peaks found by MF3D. In blue and orange we show the 2 to 5 $\sigma$  contours of the [CII] and continuum emission, respectively. We find no source with similar properties to S5100969402 with MF3D, but Loiacono et al. (2021) only report a fidelity of 0.51 for the source, so this is not surprising. We have excluded S848185 from our sample given this source is positioned near the edge of the ALMA cube. S510327576 is also excluded given the large differences in apparent velocities for the putative [CII] line ( $\Delta v > 2000 \text{ km s}^{-1}$ ), significantly decreasing our confidence that S510327576 actually corresponds to a  $z > 4$  galaxy (see Fig. 2.11 and Appendix 2.A.1). The positions and properties of the 9 serendipitously identified ALPINE galaxies used in this work are summarised in Table 2.7. MOPHONGO is used to measure the rest-UV fluxes as discussed in §2.2.4.

### 2.A.5 Archival ALMA Programs

A summary of all ALMA projects where we have explicitly reduced the ALMA observations and done our own independent search for high-redshift companion galaxies is provided in Table 2.8.



**Figure 2.14:** Companion sources identified in the neighborhood of target galaxies from ALPINE and following upon the analysis by Loiacono et al. (2021). The blue and orange contours show the  $2\sigma$ ,  $3\sigma$ ,  $4\sigma$ ,  $5\sigma$  [CII] and the continuum emission, respectively. All background images correspond to rest-UV  $z$ -band light seen with Hyper Suprime-Cam. In the lower left corners are the beam sizes of the ALMA observations.

**Table 2.7:** Characteristics we find for the [CII]-emitting companion galaxies from ALPINE (and which were previously reported by Loiacono et al. 2021).

Source	RA (ICRS)	Dec (ICRS)	$\nu_c$ (GHz)	FWHM (km s <sup>-1</sup> )	$F_{\text{[CII]}}$ (Jy km s <sup>-1</sup> )	$S_{\text{IR}}$ (mJy)	$M_{1500}$
S842313	10:00:54.42	+02:34:35.83	343.162 ± 0.011	845 ± 48	6.99 ± 0.1	8.517 ± 0.092	-21.63 ± 0.17
S665626	10:01:13.83	+02:18:40.42	340.791 ± 0.315	277 ± 22	1.14 ± 0.11	0.353 ± 0.074	> -20.59
S5101209780	10:01:33.45	+02:22:08.31	341.272 ± 0.442	388 ± 23	2.40 ± 0.28	0.065 ± 0.063	-22.88 ± 0.06
S818760	10:01:54.69	+02:32:31.30	341.465 ± 0.186	164 ± 17	0.62 ± 0.06	0.331 ± 0.052	> -20.88
S787780	09:59:56.85	+02:29:48.08	344.984 ± 0.302	263 ± 20	2.33 ± 0.15	0.365 ± 0.057	-23.38 ± 0.03
S873321	10:00:03.22	+02:37:37.46	308.763 ± 0.264	256 ± 27	2.47 ± 0.32	-0.007 ± 0.114	-21.22 ± 0.23
S378903	10:01:11.03	+01:52:07.88	295.905 ± 0.293	297 ± 37	0.64 ± 0.02	0.194 ± 0.070	> -20.60
S5100822662	09:58:57.94	+02:04:52.80	344.280 ± 0.179	156 ± 9	0.54 ± 0.06	0.153 ± 0.072	-20.25 ± 0.22
S859732	09:59:59.75	+02:36:22.62	343.145 ± 0.111	97 ± 13	1.00 ± 0.11	-0.195 ± 0.255	> -20.36



**Table 2.8:** A summary of all the ALMA programs and targets analysed in our search for [CII]-emitting companion galaxies.<sup>a</sup>

Project ID	ALMA source name	Angular resolution (")
2012.1.00536.S	MSDM_80+3 ×	0.60
	MSDM_71-5 ×	0.46
	MSDM_29.5_5 ×	0.45
2012.1.00676.S	CFHQSJ0055+0146 *	0.41
	CFHQSJ2229+1457 *	0.62
2012.1.00719.S	BDF-521 ×	0.39
	SDF-46975 ×	0.79
2012.1.00962.S	Abell383_z1	0.23
2013.1.00815.S	CLM1	0.40
	WMH13	0.88
2013.1.01241.S	A383	0.49
	MS0451 ×	0.88
2015.1.00091.S	RXJ1347_1216	0.45
2015.1.00122.S	CR7	0.26
2015.1.00834.S	WMH_5	0.21
2015.1.00997.S	SDSS_J231038.88+185519.7 *	0.75
	SDSS_J205406.49-000514.8 *	0.59
	SDSS_J012958.51-003539.7 *	0.31
	ULAS_J131911.29+095051.4 *	0.97
2015.1.01096.S	UDF-640-1417	0.69
2015.1.01105.S	COSMOS13679 ×	0.74
	COSMOS24108 ×	0.67
	NTTDF6345 ×	0.94
	UDS16291 ×	0.72
	UDS4812 ×	0.24
2015.1.01136.S	Abell383-iD	0.22
2015.1.01178.S	Abell_611 ×	0.90
2015.1.01406.S	A1689-zD1	0.23
2016.1.01240.S	20521	1.12
	2203 ×	1.18
	2313	0.67
2016.1.01423.S	J0859+0022 *	0.42
	J1152+0055 *	0.42
	J1202-0057 *	0.60
	J2216-0016 *	0.40
2017.1.00541.S	J1208-0200 *	0.41
	J2228+0152 *	0.41
2017.1.01451.S	VR7	0.42
	MASOSA	0.43
2017.A.00026.S	MACS1149-JD1	0.52
2018.1.00566.S	J0439 *	0.29
2018.1.00570.S	z7_PAR2_17941793 ×	0.43
2018.1.00781.S	system_D1_T1 ×	0.61
2019.1.01003.S	PJ308-SMG1 ×	0.64
	PJ308-SMG2 ×	0.64
	PJ308-SMG3 ×	0.64
	PJ308-SMG4 ×	0.64
	PJ308-SMG6 ×	0.64
	PJ308-SMG7 ×	0.64
	PJ308-SMG8 ×	0.64
	J0252m0503 *	0.24
2019.1.01025.S	J0525m2406 *	0.24
	J0923p0753 *	0.24
	J1007p2115 *	0.28
2019.1.01436.S	PSO_J083+11 *	0.27

<sup>a</sup> Quasar host galaxies targets are marked with a \*. [CII] non-detections of the target galaxies are marked with a ×. The angular resolution is defined as the beam of the observations along the minor axis. 34 of the 55 primary targets presented here are detected with [CII] and could be used to search for companion galaxies at similar redshift.

