ALMACAL IX: multi-band ALMA survey for dusty star-forming galaxies and the resolved fractions of the cosmic infrared background

Jianhang Chen, R.J. Ivison, Martin A. Zwaan, Ian Smail, Anne Klitsch, Céline Péroux, Gergö Popping, Andrew D. Biggs, Roland Szakacs, Aleksandra Hamanowicz, et al.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **518,** 1378–1397 (2023) https://doi.org/10.1093/mnras/stac2989
Advance Access publication 2022 October 21

ALMACAL IX: Multiband ALMA survey for dusty star-forming galaxies and the resolved fractions of the cosmic infrared background

Jianhang Chen [1]★ R. J. Ivison [1]★ Martin A. Zwaan,[1]★ Ian Smail [2] Anne Klitsch [3] Céline Péroux [1,4] Gergö Popping [1] Andrew D. Biggs [5] Roland Szakacs,[1] Aleksandra Hamanowicz [6] and Claudia Lagos[7,8,9]

[1]*European Southern Observatory (ESO), Karl-Schwarzschild-Strasse 2, D-85748 Garching, Germany*
[2]*Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[3]*DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200 Copenhagen, Denmark*
[4]*Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, F-13388 Marseille, France*
[5]*UK Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[6]*Space Telescope Science Institute, 3700 San Martin Dr., Baltimore, MD 21218, USA*
[7]*International Centre for Radio Astronomy Research (ICRAR), M468, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia*
[8]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia*
[9]*Cosmic Dawn Center (DAWN)*

Accepted 2022 October 14. Received 2022 October 7; in original form 2022 August 5

ABSTRACT

Wide, deep, blind continuum surveys at submillimetre/millimetre (submm/mm) wavelengths are required to provide a full inventory of the dusty, distant Universe. However, conducting such surveys to the necessary depth, with sub-arcsec angular resolution, is prohibitively time-consuming, even for the most advanced submm/mm telescopes. Here, we report the most recent results from the ALMACAL project, which exploits the 'free' calibration data from the Atacama Large Millimetre/submillimetre Array (ALMA) to map the lines of sight towards and beyond the ALMA calibrators. ALMACAL has now covered 1001 calibrators, with a total sky coverage around 0.3 deg^2, distributed across the sky accessible from the Atacama desert, and has accumulated more than 1000 h of integration. The depth reached by combining multiple visits to each field makes ALMACAL capable of searching for faint, dusty, star-forming galaxies (DSFGs), with detections at multiple frequencies to constrain the emission mechanism. Based on the most up-to-date ALMACAL data base, we report the detection of 186 DSFGs with flux densities down to $S_{870\,\mu\mathrm{m}} \sim 0.2$ mJy, comparable with existing ALMA large surveys but less susceptible to cosmic variance. We report the number counts at five wavelengths between 870 μm and 3 mm, in ALMA bands 3, 4, 5, 6, and 7, providing a benchmark for models of galaxy formation and evolution. By integrating the observed number counts and the best-fitting functions, we also present the resolved fraction of the cosmic infrared background (CIB) and the CIB spectral shape. Combining existing surveys, ALMA has currently resolved about half of the CIB in the submm/mm regime.

Key words: galaxies: high-redshift – cosmic background radiation – submillimetre: galaxies.

1 INTRODUCTION

The cosmic infrared background (CIB) – covering mid-infrared to submillimetre/millimetre (submm/mm) wavelengths – comprises an important component of the energy emitted by galaxies over the history of the Universe (Fixsen et al. 1998; Hauser et al. 1998). Along with the optical background, it represents the energy produced by the formation and evolution of galaxies, and all related processes, across all cosmic time. Since the first measurements of the CIB, a primary goal in astrophysics has been to identify the sources responsible for it, and thus to understand the lifetime energy budget of the Universe (e.g. Cooray & Sheth 2002; Lagache, Puget & Dole 2005; Dole et al. 2006; Planck Collaboration XVI 2014).

Several decades after the discovery of the CIB, various galaxy populations are known to make significant contributions. Among them, strongly star-forming galaxies with total infrared luminosities larger than 10^{11} L$_\odot$ (Sanders & Mirabel 1996) are thought to be the main contributors, with galaxies at different redshifts contribute to different parts of the CIB. In the mid- and far-infrared (-IR), the CIB largely comprises luminous and ultraluminous IR galaxies (LIRGs and ULIRGs) at $z < 1.5$ (Elbaz et al. 2002). In the submm/mm wavelength regime, the CIB is dominated by dusty, star-forming galaxies (DSFGs) at higher redshifts ($z \gtrsim 1.5$). In the submm/mm bands, dimming due to increasing redshift is compensated by stronger dust emission as the observing frequency traces rest frequencies progressively closer to the peak of a typical dust spectral energy distribution (SED). This makes the submm/mm wavelength regime particularly well suited to the detection of dust-rich galaxies at high redshift (e.g. Blain & Longair 1993; Smail et al. 2002).

★ E-mail: cjhastro@outlook.com (JC); Rob.Ivison@eso.org (RJI); mzwaan@eso.org (MAZ)

The first submm-selected galaxies (SMGs, the brightest DSFGs) were found by continuum surveys at 850 μm with the 15-m James Clerk Maxwell Telescope (e.g. Smail, Ivison & Blain 1997; Barger et al. 1998; Hughes et al. 1998; Eales et al. 1999), which ushered in the era of submm cosmology. Since then, SMGs have been studied extensively at different wavelengths, using radio, X-ray, and optical telescopes to trace their properties (e.g. Ivison et al. 2002; Alexander et al. 2005; Chapman et al. 2005, respectively). New instruments and associated surveys have increased the family of SMGs selected at the original wavelength (e.g. Chen et al. 2016; Geach et al. 2017), at longer wavelengths ($\lambda > 1$ mm, e.g. Scott et al. 2012; Magnelli et al. 2019), as well as at shorter wavelengths, e.g. with the *Herschel Space Observatory* ($\lambda < 500$ μm, e.g. Eales et al. 2010; Oliver et al. 2010).

A significant advance came with the advent of the Atacama Large Mm/submm Array (ALMA), which was immediately able to pinpoint and even spatially resolve many of the earlier bright, single-dish-selected SMGs (e.g. Hodge et al. 2013; Ikarashi et al. 2015; Simpson et al. 2015b; Stach et al. 2019). Some especially bright lensed sources proved to be ideal targets for detailed studies of their internal structures and physical states (e.g. Geach et al. 2018; Rizzo et al. 2020; Dye et al. 2022), while other bright SMGs were resolved into multiple sources, leading to the discovery of protoclusters of extreme DSFGs (e.g. Ivison et al. 2013; Miller et al. 2018; Oteo et al. 2018).

ALMA also offers the sensitivity and spatial resolution necessary to push below the confusion limit imposed by single-dish imaging, allowing the discovery of much fainter DSFGs. As happened with the single-dish telescopes that came before, two methods have been used to undertake surveys for DSFGs: long integrations in conventional 'deep fields' (e.g. Walter et al. 2016; Dunlop et al. 2017; Umehata et al. 2018) and the targeting of lensing clusters to probe smaller areas where the gravitational magnification is high (e.g. González-López et al. 2017; Laporte et al. 2021), enabling the discovery of faint DSFGs in the epoch of reionization (e.g. Laporte et al. 2017; Tamura et al. 2019; Fudamoto et al. 2021).

To create a full inventory of dusty galaxies, we need deep, wide, and blind surveys. Before ALMA, single-dish telescopes had already covered 10s to 100s of square degrees of the sky and found many thousands of bright FIR/submm/mm sources. However, the diffraction limit prevented us from delving below the intensely star-forming ULIRG regime, aside from a small number of strongly lensed systems. It was also necessary to make model-based corrections to the counts to deal with issues associated with source blending, and a significant portion of bright SMGs found by single-dish telescopes were later resolved into multiple sources by interferometers (e.g. Ivison et al. 2007; Wang et al. 2011; Karim et al. 2013). Using ALMA, it has been possible to extend the submm/mm detection limit down to the sub-mJy level. Even with the small number of antennas available in Cycle 0, ALMA pushed down to $S_{870\,\mu m} \sim 0.4$ mJy in the Extended *Chandra* Deep Field South (Hodge et al. 2013; Karim et al. 2013). Since then, extensive follow-up campaigns have been carried out for the bright SMGs found in earlier single-dish surveys (e.g. Weiß et al. 2013; Brisbin et al. 2017; Miettinen et al. 2017; Cowie et al. 2018; Stach et al. 2019; Simpson et al. 2020). Meanwhile, blind ALMA surveys have been continuously enlarging the survey area and/or improving the detection limit (e.g. Hatsukade et al. 2016, 2018; Walter et al. 2016; Dunlop et al. 2017; Umehata et al. 2018). Very recently, the ALMA Spectroscopic Survey in the Hubble Ultra Deep Field (ASPECS) Large Programme (Walter et al. 2016) has achieved a sensitivity of 10 μJy beam^{-1} (González-López et al. 2019, 2020) and although ALMA is not optimized for large sky coverage, the GOODS-ALMA project has observed the 72 arcmin^2 Great Observatories Origins Deep Survey South field (GOODS-South) in two different array configurations (Franco et al. 2018; Gómez-Guijarro et al. 2022). Considerable time has also been invested in the well-known legacy fields, which have also inspired dedicated data mining projects (e.g. Ono et al. 2014; Zavala et al. 2018; Liu et al. 2019).

All these ALMA surveys have helped us to constrain the number counts and the properties of DSFGs. Number counts – the projected galaxy surface density as a function of flux density – represent the most basic measurement we can glean from such observations. They therefore provide a simple test of the validity of models of galaxy formation and evolution. Indeed, both semi-analytic models (e.g. Somerville et al. 2012; Lacey et al. 2016; Lagos et al. 2020) and post-processing models in hydrodynamic simulations (e.g. Shimizu, Yoshida & Okamoto 2012; Cowley et al. 2019; McAlpine et al. 2019; Lovell et al. 2021) have struggled to reproduce the number counts of DSFGs. In addition, accurate and unbiased number counts are essential to resolve the CIB and determine the contributions of different galaxy populations to the total CIB. However, due the very high demand for ALMA time it has become difficult to go much deeper or wider, and it has proved even more difficult to justify covering the same areas at multiple submm/mm wavelengths.

ALMACAL is a novel submm/mm survey that exploits the 'free' data that ALMA must collect to ensure its observations can be processed to make spectra, images, and cubes with accurate positions, polarizations and flux densities. With ALMACAL, we survey the immediate vicinity of each calibrator, which are typically blazars at $z < 1$ (Bonato et al. 2018; Klitsch et al. 2019). Since ≈20 per cent of all ALMA observing time is spent on calibration, ALMACAL is already competitive with the widest and deepest submm/mm surveys. It does come with an obvious disadvantage: the lack of ancillary data at depths comparable with classic deep fields, such as the Cosmic Evolution Survey (COSMOS) or GOODS. However, it also comes with several major advantages. Number counts of ALMACAL are largely immune to cosmic variance, since it covers a great many pointings scattered across the observable sky, and the presence of an in-beam calibrator allows for perfectly calibrated submm/mm/radio imaging with a very high dynamic range. The blind detections share the same spatial and spectral set-ups as the science targets, whatever they may have been, which makes it possible to search for spectral lines (in absorption or emission) and to study morphologies at the very highest spatial resolution, up to 20 mas (see Oteo et al. 2017). Meanwhile, these fields will be visited repeatedly by future submm/mm and radio interferometers, which will keep improving the sensitivities in these fields at different wavelengths.

In our earliest attempt at mining the ALMA calibration data, Oteo et al. (2016) described the first search for DSFGs in ALMACAL. With multiband data available for most of the fields, they derived dual-band number counts, in bands 6 (1.2 mm) and 7 (870 μm), using the data collected before 2015. Since then, both the number of calibrators and their on-source integration times have grown by more than an order of magnitude. Using the new data, Klitsch et al. (2020) reported the first number counts in ALMA band 8, at 650 μm. Besides number counts, searches for molecular absorption or emission lines have put constrains on the evolution of molecular gas density over cosmic time (Klitsch et al. 2019; Hamanowicz et al. 2022). In this work, we extend the number counts of Oteo et al. (2016) and Klitsch et al. (2020) to the most up-to-date ALMACAL data set, covering ALMA bands 3, 4, 5, 6, and 7, at wavelengths from 3 mm down to 870 μm, respectively.

The paper is structured as follows: in Section 2 we introduce the ALMACAL project and its results to date; Section 3 details

our data analysis, including source detection, corrections for flux boosting, sample completeness, effective area, source classification, and selection bias. Section 4 presents the final source catalogue and the number counts in five different ALMA bands. Section 5 contains relevant discussion and comparison between our number counts with literature and model predictions. The various fractions of the CIB resolved by ALMA are also presented in Section 5. We summarize our key results in Section 6. Throughout this paper, we follow the terminology used in the review by Hodge & da Cunha (2020). We use SMGs to denote the classical submm-selected galaxies with $S_{870\,\mu m} \geq 1.0$ mJy and use DSFGs to include all the dusty, star-forming galaxies we can detect via ALMA continuum observations. In this paper, we assume a ΛCDM cosmology with $H_0 = 67.7$ and $\Omega_m = 0.31$ (Planck Collaboration VI 2020).

2 ALMACAL

ALMACAL[1] aims to exploit all ALMA calibrator scans for science (Zwaan et al. 2022). Running now for more than ten years, ALMACAL has already accumulated more than 1000 h of data. As of 2020 March, ALMACAL includes 1001 calibrators. About 97 per cent of them are classified as blazars (Bonato et al. 2018), which are quasars – active galactic nuclei – whose jets are oriented very close to our line of sight, such that relativistic beaming makes them extremely bright. The calibrators are spread all over the sky and used to calibrate science targets local to them. Depending on their brightness, compactness, and flux stability, they can be used to calibrate bandpass, gain (complex amplitude and phase), flux density, and polarization. In a typical ALMA scheduling block (SB), two or more calibrators will be observed, along with the science targets, so calibrators share the same instrumental configuration as the science targets. The rich data buried in the calibration observations make them far more useful than their original intention.

One intriguing example is the calibrator named J1058+0133, one of the brightest blazars close to Cosmic Evolution Survey (COSMOS) field. Oteo et al. (2017) found two $z = 3.4$ SMGs behind this calibrator. ALMA has invested hundreds of hours in the COSMOS field, which makes J1058+0133 one of the most frequently visited calibrators, observed in total for around 150 h. With the rich range of configurations that ALMA has employed in the COSMOS field, Oteo et al. (2017) were able to create a high signal-to-noise multiband image of the two SMGs, self-calibrated on the time-scale of 1 s using the in-beam calibrator, J1058+0133, with 20 mas spatial resolution. Additionally, multiple CO lines have been identified in this system, yielding a remarkable spectral line energy distribution, making this amongst the most extraordinary data sets gathered for any SMG.

A detailed description of the ALMACAL data retrieval and calibration pipeline was given in the first ALMACAL paper, Oteo et al. (2016), which we summarize briefly here. All the data reduction is carried out using the Common Astronomy Software Application (CASA – McMullin et al. 2007). First, the calibration data are requested from the ALMA science archive. Next, we apply the standard calibration pipeline, following the *ScriptForPI.py* Python script for each project and splitting out the calibrator data. For calibrators that lack flux density calibration – mostly the bandpass calibrators – their flux densities are recovered from the internal flux tables in the data delivery package. After that, two cycles of self-calibration are performed in the multifrequency synthesis (mfs) image. The first cycle is focused on phase-only solutions; the second cycle corrects the amplitude and phase together. Next, a point-source model of the calibrator is removed from the *uv* data by the CASA internal tool UVMODELFIT to make science-ready data. Finally, the fully calibrated measurements are re-binned to a channel width of 15.6 MHz. After these pre-processing steps, the scans contain the fully calibrated data, with the central calibrator subtracted out.

This specially designed pipeline takes full advantage of the bright, central source to do self-calibration, avoids possible effects due to variability, and it is straightforward to combine calibrated data from many different projects. The dynamic range of images made for calibrators exceeds 10^4, where the background sources or calibrator jets are typically 100 times weaker than the calibrators themselves. After removing the central, bright calibrator from each data set, we can also easily identify the observations that are suitable for combination. At the moment, we simply reject any problematic data (see Section 3.1 for details). Due to the immense size of the data set, it is not realistic to recalibrate all the problematic observations manually.

ALMA has invested significant observing time in various deep cosmology fields (see a summary in Hodge & da Cunha 2020). ASPECS is currently the deepest blind survey, covering 2.9 arcmin^2 with an r.m.s. sensitivity that reaches $\sigma_{1.2\,mm} \approx 9.3$ μJy beam^{-1} (González-López et al. 2020). Meanwhile, GOODS-ALMA represents the largest blind survey, with sky coverage of 72.4 arcmin^2, with a much shallower sensitivity, $\sigma_{1.2\,mm} \approx 70$ μJy beam^{-1} (Gómez-Guijarro et al. 2022). ALMA follow-up of bright sources from the single-dish surveys is the most efficient way to probe the very brightest SMGs and such an approach has been employed extensively in the existing legacy fields (e.g. Weiß et al. 2013; Stach et al. 2019; Simpson et al. 2020). ALMACAL manages to combine the best depth and sky coverage – see Section 3.6. In Fig. 1, we show the on-source time for all our calibrators in the different ALMA receiver bands. Compared with Oteo et al. (2016) in bands 6 and 7, the total on-source time and the number of available calibrators have both increased dramatically. Ultimately, ALMACAL will go deeper and wider than any ALMA blind survey, unless a significant strategic investment is made.

3 OBSERVATIONS AND ANALYSIS

We detail the steps we have taken regarding data selection and imaging in Section 3.1, then present the methods and steps taken to construct the source catalogue in Sections 3.2 and 3.3. After that, we derive the correction functions for our survey, including those for flux deboosting (Section 3.4), completeness (Section 3.5), effective area (Section 3.6), and effective wavelength (Section 3.7). Finally, we classify our sources in Section 3.8 and discuss selection biases in Section 3.9.

3.1 Data selection and imaging

A rose has its thorns. The bright calibrator at the centre of the field can be helpful for self-calibration but at times it brings problems. First, the calibrator is not always a point source; blazars have jets, and they are not always (or have not always been) oriented along the line of sight; they can be slightly resolved, or amplitude and phase errors can remain after the calibration steps. All these issues can introduce residuals after point-source subtraction of the central calibrator. To avoid residuals corrupting the final, combined image, we applied two imaging cycles.

[1] https://almacal.wordpress.com

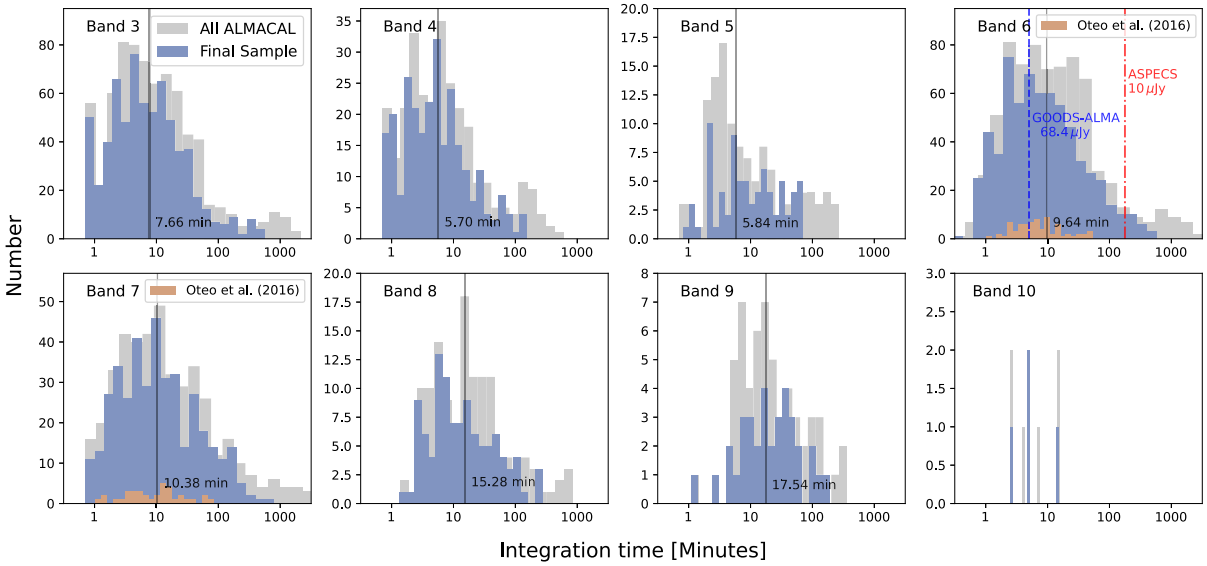

Figure 1. Statistics of on-source time of ALMACAL observations from ALMA bands 3 to 10. For every panel, the abscissa is the accumulated total on-source time in units of minutes. The ordinate is the number of fields in every time-step. We show the total data available in each band and the data used in this work. Some of them have been dropped because of calibration errors or strong residuals after calibrator removal (see Section 3.1). In bands 6 and 7, we also show the statistics for the data used by Oteo et al. (2016), illustrating the immense increase since that time. Both the number of fields and the sensitivity in other bands have increased, making it possible to expand the number counts to a wider flux density range and to multiple ALMA bands. In band 6, the two vertical lines show the sensitivity reached by the ASPECS (González-López et al. 2020) and GOODS-ALMA surveys (Franco et al. 2018; Gómez-Guijarro et al. 2022). ALMACAL provides a good compromise between their sensitivity and sky coverage.

In the first imaging cycle, we image the visibilities of every single data set after point-source removal, then we visually inspect all the images, discarding the ones with incorrect calibration or strong residuals. This step alone led to the loss of around half of our data, but it ensured the best quality of the final combined images. Having removed the poor quality data, we re-scale the weighting of different observations with the CASA task, STATWT, then combine the observations of each field using the task, CONCAT. All the images were made using TCLEAN in CASA version 5.7.0 in 'mfs' mode, which combines all the frequency channels into a single continuum image.

The second imaging cycle use the combined, re-scaled visibilities to create science-ready images. One of the major goals of ALMACAL is to search for DSFGs in the calibrator fields. From previous ALMA follow-up of known, bright DSFGs, we expect that they have compact dust morphologies, with a median size $\lessapprox$0.5 arcsec (FWHM; Ikarashi et al. 2015; Simpson et al. 2015a; Gullberg et al. 2019). To optimize the sensitivity to compact sources, we adopted a natural weighting scheme and the Hogbom (Högbom 1974) deconvolution algorithm to clean our images. In addition, to avoid resolving DSFGs we further tapered the visibilities with Gaussian kernels of 0.3 and 0.6 arcsec. We show the final resolution versus sensitivity in Fig. 2, our *uv* tapers ensure the images have at at least 0.3 arcsec and 0.6 arcsec spatial resolution, respectively. These two tapered images were used for source detection, and to quantify the fraction of missing flux density. During the deconvolution, we first estimated the sensitivity of the final combined visibilities using the function SENSITIVITY in CASA *Analysis Utilities*[2] and used this as the threshold in TCLEAN. In every major cleaning cycle, the built-in AUTO-MULTITHRESH algorithm in TCLEAN was used to search for secure emission and determine the clean regions in the residual maps. We cleaned the image within an area defined as 1.8 × the FWHM of the primary beam. After cleaning, IMPBCOR was used to correct the primary beam response. In the following steps, as a convenience, the image without the primary beam correction was used for source detection, while the primary-beam-corrected image was used for flux density measurements.

3.2 Source detection

Following our two earlier number counts papers, Oteo et al. (2016) and Klitsch et al. (2020), we performed the source detection using SEXTRACTOR as it can deal easily with complex source structures (Bertin & Arnouts 1996). The search for detections is made in the images before primary beam correction, corresponding to a signal-to-noise image. We detected residual signals from the calibrator in some fields, so we masked the central part of each map to a radius of 2 arcsec. This radius is 2 × larger than the masking used in band 8 (Klitsch et al. 2020) but the loss is small for the larger fields of view (FoV) at the longer wavelengths explored here. To improve the reliability of the source detection in ALMACAL, we applied the source-finding algorithm in the two tapered images simultaneously. We search sources with SNRs higher than 3σ and then only accept detections with peak SNR higher than 5σ in at least one of the two tapered images.[3] Our adoption of a 5σ final detection cut prevents us from finding very faint sources, but ensure our detected sources are robust. In Fig. 3, we show the source detections at different wavelengths in two example fields.

[2]https://casaguides.nrao.edu/index.php/Analysis_Utilities

[3]Most of our detections (>90 per cent) have been found in both of the two tapered images. Those detections found in just one of the tapered images, with a counterpart in another ALMA band, are also classified as robust detections; those without counterparts in any other ALMA band are marked as uncertain. The variation of the final number counts that include these uncertain sources are discussed in Online Fig. D1.

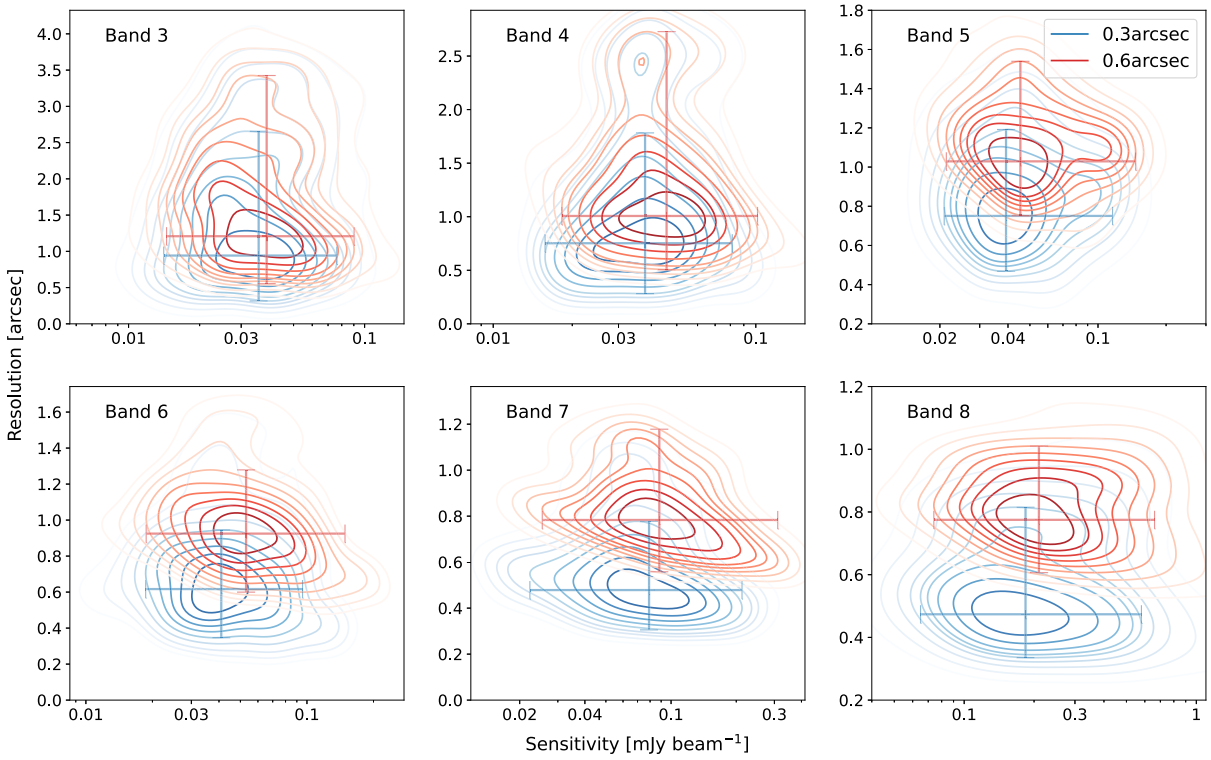

Figure 2. Density distribution of sensitivity versus resolution for all the combined images in each ALMA band. We show the distributions for each set of tapered images in different colours. The lowest contours include 90 per cent of the data points and decrease in steps of 10 per cent. Our two *uv* tapers ensure the images have at least 0.3 arcsec and 0.6 arcsec spatial resolution, respectively. In the final images, most of the sources should remain unresolved.

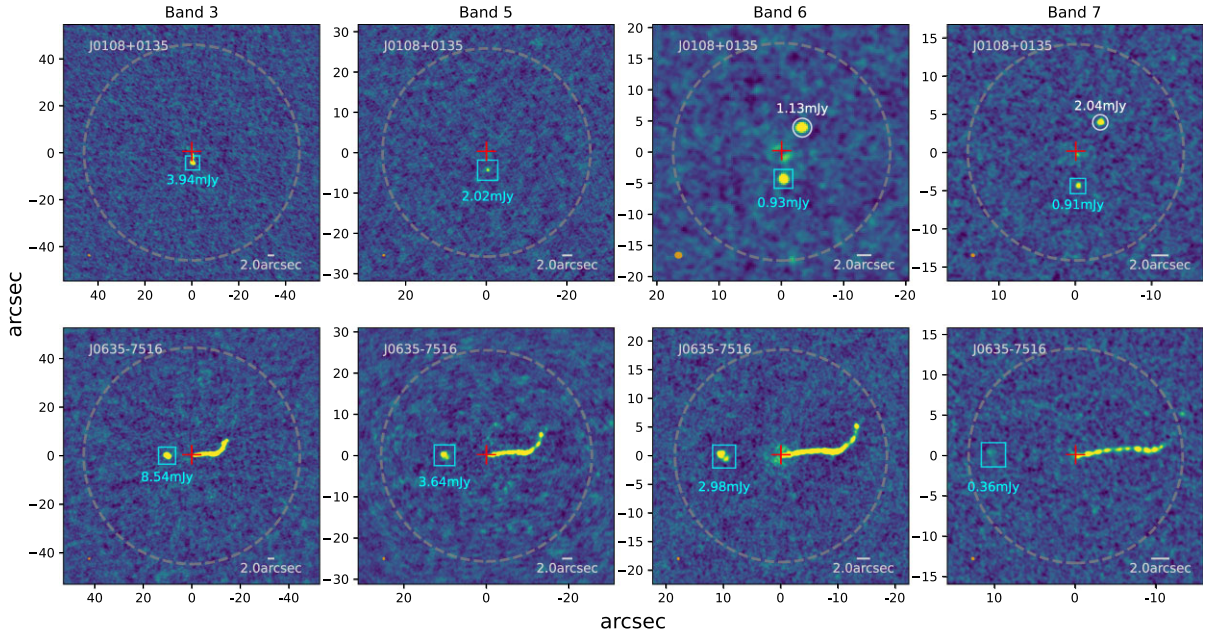

Figure 3. Two examples of detections in ALMACAL in different bands. The wavelength is decreasing from left to right. In every image, the central bright calibrator has been removed – its position is marked with a red cross – and all the available observations have been combined. The dashed grey circle is the FoV adopted in our work, which is 1.8 × FWHM of the respective primary beam. In the top row, the images in ALMA bands 4, 5, 6, and 7 towards the field J0108+0135 are shown, where one SMG (marked with a white circle) and one jet (marked with a cyan square) have been found. The SMG has higher flux densities at higher frequencies, while, in contrast, the flux density of the radio jet decreases. In the bottom row, the images from the field J0635−7516 are shown, where two-sided extended radio jets have been discovered. We have classified each detection based on its spectral index and morphology.

We searched for sources within 1.8 × the FWHM of the primary beam, which corresponds to 15 per cent of the peak sensitivity. This represents a reasonable trade-off, yielding a large total effective area whilst maintaining reasonable sensitivity. The final positions of all the detections are defined by their signal-to-noise centroid (Bertin & Arnouts 1996).

3.3 Flux density

We measured the flux densities of our detections using aperture photometry. Aperture photometry was undertaken using the PYTHON package, PHOTUTILS, with elliptical apertures. The size of the aperture is important: selecting too large or too small an aperture introduces uncertainties into the total flux density. Since most of our sources are unresolved, we use the synthesized beam as the shape of our aperture. Based on our testing, an aperture size that is twice the synthesized beam with an appropriate aperture correction delivers the most reliable flux density. The aperture correction assumes an unresolved point source. To better capture all the flux for our detections, we use the 0.6 arcsec tapered images as the primary images to measure their flux densities. The flux difference between the two *uv* tapered images are small, less than 10 per cent, supporting our assumption that most of emissions from our detections remain unresolved.

We also report the flux densities from 2D Gaussian fitting. The Gaussian fitting was achieved by ASTROPY.MODELING with the GAUSSIAN2D model and the *LevMarLSQFitter* fitting algorithm. Before applying the fitting algorithm, a cut-out is taken from the original image, measuring 5 × the major axis of the synthesized beam. During fitting, the amplitude, axial ratio, positional angle could change freely. The centre of the Gaussian function is allowed to shift by ±0.1 arcsec relative to the photometric centre. The flux density was calculated by integrating the best-fitting Gaussian.

The sensitivity of every ALMACAL field is not uniform, decreasing outwards from the phase centre. This means sources need to be brighter at larger radii to be detectable. Because the primary-beam correction scales up the flux density and noise by the same factor, the peak SNRs of detections remain the same. In the following two sections, we discuss flux boosting and sampling completeness as a function of peak SNR, which is convenient to apply to all our fields.

3.4 Flux deboosting

The well-known effect that we must consider when measuring accurate flux densities is the so-called 'flux boosting' of faint sources (Hogg & Turner 1998). It has two origins (e.g. Coppin et al. 2006; Casey, Narayanan & Cooray 2014). First, we are more likely to detect faint sources that have been scattered towards higher flux densities by random noise (Eddington flux boosting), an effect that can be mitigated to some extent by choosing a high SNR detection threshold. Secondly, where there is a rapid increase in the number of sources as we delve fainter, i.e. when the source counts are steep, then the resolution element of a telescope may include additional sources that are individually fainter than the detection threshold. The latter issue has been common for single-dish submm/mm telescopes, because of their typically large diffraction limit, $\gtrsim$10 arcsec, and because the source counts of DSFGs at $S_{850\,\mu\mathrm{m}} \gtrsim 6$ mJy are steep (Simpson et al. 2020).

However, the blending effects should be very much less significant for deep, high-resolution interferometric observations, such as those obtained using ALMA. The spatial resolution of ALMA is considerably higher than that of single-dish telescopes, which makes the confusion noise significantly lower. At the same time, the number counts of DSFGs have been found to flatten at fainter flux densities (e.g. Stach et al. 2018; González-López et al. 2020). The combination of these two differences with historic work ensures that blending effects are smaller for interferometer-based work. Nevertheless, we have tested the degree of flux boosting affecting our detections via bespoke Monte Carlo simulations in all our fields.

To preserve the noise characteristics of our fields, as well as any undetected faint sources, we injected artificial sources directly into our observed images. Before we begin, we clean the image and run our source-finding algorithm to search for detections, masking any before passing the image to the simulation. For each cycle of the simulation, we generated 20 artificial sources for each field. The artificial sources were randomly assigned flux densities such that their SNRs lay between 2 and 20 σ. Their positions were also randomly assigned, within the 1.8 × FWHM area of the primary beam. Next, we removed all close pairs (mutual distance <3 arcsec) of artificial sources to avoid unnatural source blending. After this pruning, the artificial sources were convolved with the synthesized beam and added to the image. This process was repeated 1000 × for each field, to gather a statistically significant sample of simulated sources. We then repeat all the steps for different source sizes. We model the shape of the injected sources with 2D Gaussians, with the intrinsic size varying across a grid of five different sizes (point sources with FWHM = 0.1, 0.2, 0.3, 0.6 arcsec). Finally, we searched these images for detections in the same way as we do for our sources (see Sections 3.2 and 3.3), using the same methods to measure their flux densities.

We also tested the injection of artificial sources into the *uv* visibilities. Similar to the image-plane simulation, we also randomly generate 20 sources per cycle per field. Then, ALMA task FT was used to convert the point sources into visibilities and ALMA task UVSUB is used to add them into the visibilities. The two simulations give consistent results, but conducting simulations in the visibilities is time-consuming, especially considering that we need simulations for different fields, so we chose to conduct all the simulations in the image plane for all the ALMACAL fields.

Fig. 4 shows the effects of flux boosting as a function of SNR. We only show the results for our fiducial source size of 0.2 arcsec. The flux boosting also varies for different source sizes, but the flux boosting converges to a single boosting function above $\mathrm{SNR}_{\mathrm{peak}} \geq 5$ (see Online Appendix A). Moreover, size measurements are problematic at low SNRs, which complicates any size-based correction. As a compromise, for simplicity, we focus on the flux boosting for sources with a fixed size (FWHM = 0.2 arcsec) in our analysis.

The flux boosting effect is slightly different for different flux-measuring methods. We found a smaller boosting factor for aperture photometry than with the 2D Gaussian fitting. For aperture photometry, we did not find noticeable boosting for bright sources (SNR > 5σ); while for 2D Gaussian fitting, the boosting effects are still not negligible for sources with SNR up to 10. This boosting factor is consistent with Oteo et al. (2016), whose results were based on 2D Gaussian fitting with CASA task, IMFIT. The key difference between aperture photometry and Gaussian fitting, as deployed here, is that the latter allows the centroid to move slightly (±0.1 arcsec) and has the freedom to vary the source shape, such that it will suffer more from Eddington flux boosting. However, Gaussian fitting is more robust than aperture photometry in crowded regions. In our final catalogue, both of the flux densities are reported and corrected by their own deboosting function. In the following analysis, we use the flux density from aperture photometry, but changing to Gaussian fitted flux density does not change our results significantly.

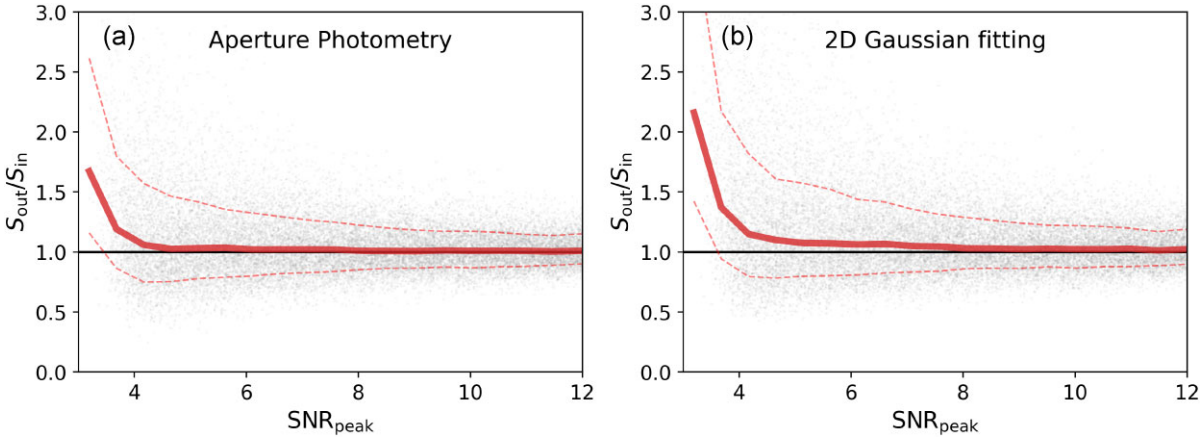

Figure 4. Flux boosting as a function of peak SNR. The boosting effects have been estimated from the recovered flux density of artificial sources (0.2 arcsec FWHM) randomly injected in the observed image. The abscissa is the peak SNR of the injected source, and the ordinate is the ratio between the measured flux density, S_{out}, and the injected flux density, S_{in}. The two methods used for flux measurements are shown in the two panels. Panel (a) is used aperture photometry; panel (b) is used 2D Gaussian fitting. The red solid line is the median value in different SNR bins, the red dashed lines enclose 68 per cent of the points in each bin. Our simulations indicate that different flux measuring methods will lead to slightly different deboosting functions, where the Gaussian fitting tends to capture more positive noise during the fitting. In our final catalogue, we report the two measurements corrected by their own deboosting functions.

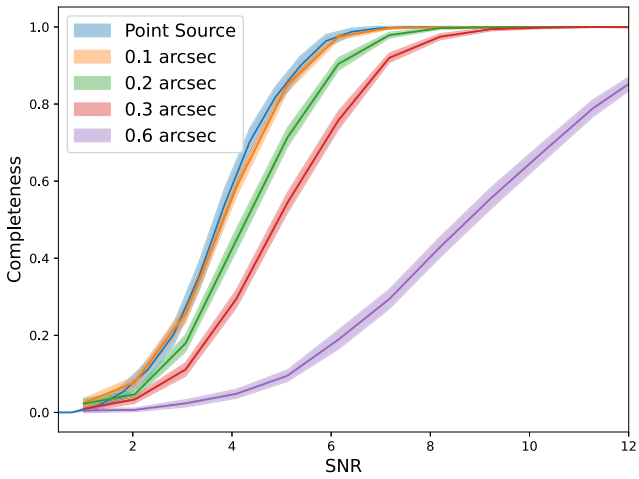

Figure 5. Sample completeness as a function of peak SNR for various source sizes (in FWHM). The larger the source size, the lower the completeness at a given SNR. However, source size with SNR < 10 cannot be measured robustly. Based on size distribution of DSFGs from current surveys at similar sensitivity, we adopted 0.2 arcsec as the fiducial source size to correct the sample completeness. We caution that our survey is less sensitive to very extended sources (FWHM > 0.6 arcsec).

3.5 Completeness and reliability

We use the same simulation as discussed in Section 3.4 to derive the completeness correction for our catalogue. We apply the source-finding procedures to these simulated images, using the methods described in Section 3.2. An injected source was marked as recovered if it was matched by a detection within the synthesized beam, and was otherwise marked as missed. If a source was detected without a corresponding injected source, it was marked as a false detection.

As shown in Fig. 5, the resulting completeness varies for different source sizes. Size-based corrections have been discussed in the literature (e.g. Béthermin et al. 2020), but the intrinsic source size cannot be well constrained for sources with SNR < 10 (Simpson et al. 2015a). We therefore adopt the completeness correction derived from our fiducial source size (FWHM = 0.2 arcsec). Besides a robust size measurement, the intrinsic size distribution of DSFGs is needed to fully correct the completeness, which requires future observational effort (e.g. Gullberg et al. 2019; Smail et al. 2021). We did not find many extended sources in the existing surveys with sensitivities and resolutions similar to ALMACAL. Therefore, we stick to the fiducial completeness correction but caution that this is a limitation of our analysis.

With a 5σ detection threshold, we found no spurious detections in our simulations. The spurious fraction goes down to zero around SNR $\sim$ 4.7 in most of our fields, hence the adopted 5σ threshold will give us a clean and robust sample.

3.6 Effective area

Unlike surveys made via mosaics of uniform ALMA pointings, the effective area of ALMACAL changes with sensitivity. First, different fields have different total integration times, which leads us to different sensitivities. On top of this, the sensitivity decreases with increasing distance from the phase centre, as per the primary beam response. To quantify this effect, we measured the effective area as a function of peak flux density. For every field, the effective area for a particular flux density is dictated by the radius at which its SNR drops to 5σ. The maximum effective radius is limited by the 1.8 $\times$ FWHM for source detection. Both the SNR threshold and the maximum radius are the same as those used for source detection. We performed the same calculation for all of the available fields to determine the total effective area at different peak flux densities.

Fig. 6 shows the total effective area as a function of peak flux density for the different ALMA bands. Due to the larger FoV at longer wavelengths, and its popularity with ALMA users, ALMA band 3 has the largest effective area, close to 700 arcmin^2 at a flux density of 0.1 mJy. It is followed by bands 4 and 6. The ALMA FoV in band 6 is much smaller than in band 4, but band 6 has more observations $-$39 per cent of all the ALMACAL observing time. Bands 5 and 7 have similar effective areas at flux densities larger than 0.3 mJy, but band 7 goes much deeper. Besides ALMACAL, we also plot the sky coverage of other pointed and blank-field surveys (Franco et al. 2018; Stach et al. 2019; González-López et al. 2020; Gómez-Guijarro et al. 2022). By comparison, ALMACAL offers a good balance between sensitivity and effective area, and will continue to be an essential complement to blind surveys.

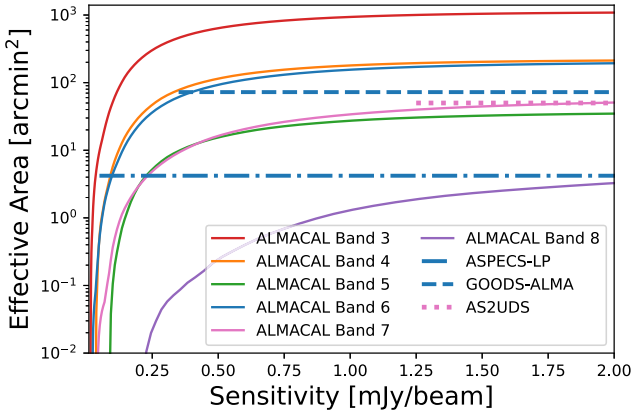

Figure 6. Effective area as a function of sensitivity (5σ) for different ALMA bands. The sensitivity reached in each field is different, and decreases with increasing distance to the phase centre within each pointing. These effects lead us to the final effective area changing with sensitivity. In addition to ALMACAL, we also show the sensitivities and effective areas of ASPECS (González-López et al. 2020) and GOODS-ALMA (Gómez-Guijarro et al. 2022) conducted in ALMA band 6, and the AS2UDS survey (Stach et al. 2019) in ALMA band 7. Compared with existing dedicated ALMA surveys, ALMACAL offers a good balance between sensitivity and effective area and will continue to be an essential complement to existing or ongoing blind surveys.

Klitsch et al. (2020) estimated the cosmic variance of band 8 footprints from ALMACAL following the methodology described in Driver & Robotham (2010). They found the cosmic variance is less than 5 per cent level in ALMACAL band 8 observations. All the bands with longer wavelengths have at least 10 × larger sky coverage, which should suffer from less than 1 per cent cosmic variance.

3.7 Effective wavelength

The spectral coverage of ALMACAL is much more complex than that of previous ALMA surveys where the observations were typically carried out at a fixed frequency. ALMA calibrators share the same configurations as their science targets, such that the same calibrator observed by different projects will potentially have different spectral configurations. Fig. 7 shows the frequency coverage of ALMACAL from ALMA band 3 to 10. The vertical lines are the exposure-time-averaged mean frequency for each band. For example, the observations in band 6 are spread across the whole band. The exposure-weighted mean frequency in band 6 is 237.95 GHz (1.26 mm), which is slightly lower than the most commonly used 250 GHz (1.2 mm). By comparison, the coverage in band 7 is simpler, with most of the observations undertaken at around 340 GHz. Its mean frequency is 338.08 GHz (887 μm). The effects of these differences need to be corrected before a comparison is made with literature results, as discussed in Section 4.2.

3.8 Source classification

Sources detected by ALMACAL are typically either thermal in nature, e.g. continuum emission from dust in star-forming galaxies, or non-thermal in nature, e.g. synchrotron emission from radio jets associated with the calibrator or other radio sources in the field. Thanks to the multiband coverage of ALMACAL, these emission mechanisms can be separated by their spectral index, α, where $S_\nu \propto \nu^\alpha$. For a typical DSFG at $z \sim 2$, we are probing the Rayleigh–Jeans tail of emission from warm dust with $\alpha \approx +3.5$ (Ivison et al. 2010; Swinbank et al. 2010), so the emission becomes brighter at higher frequencies. For radio AGN and jets, on the other hand, the flux density typically declines at higher frequencies or stays flat, such that $\alpha \lesssim 0$. We categorized a source as a DSFG if its multiband flux densities are consistent with a dust SED, and as synchrotron source if its spectral index instead betrays synchrotron emission (see also Klitsch et al. 2020).

We show the SED from examples of confirmed sources in Fig. 8. To obtain the dust SEDs of sources covered by different ALMA bands, we extend the flux measurement to all the available bands whenever a detection is confirmed. The aperture size and aperture correction are calculated using the same way we do for true detections. We show all the measurements that have reliable flux densities (SNR $\geq$ 3) in at least three ALMA bands in Fig. 8. In general, DSFGs have distinguishable SEDs from radio sources. The median α in our sample is $\alpha = 3.2 \pm 0.3$ and $\alpha = -0.9 \pm 0.1$ for DSFGs and Synchrotron, respectively.

We commonly encounter the radio jets of our blazar calibrators, which are not always perfectly aligned with the line of sight. Most are relatively compact, beginning close to the calibrator; they tend to be much more prominent in bands 3 and 4 and are easily identified by their positions and elongated morphologies. However, extended radio jets have been found in about 10 per cent of the fields. These jets could be resolved into several blobs and morphologically they resemble point sources at higher frequencies. With detections in multiple bands, these jets can be easily excluded via their spectral indices and their morphological alignment (see the second example in Fig. 3).

The classification of an ALMACAL source is uncertain if the source is only covered by one ALMA band. In such cases, we first we cross-matched with radio images from the Monitoring Of Jets in Active galactic nuclei with VLBA Experiments (MOJAVE),[4] Faint Images of the Radio Sky at Twenty-cm (FIRST – Becker, White & Helfand 1995), the NRAO VLA Sky Survey (NVSS – Condon et al. 1998), and the NRAO VLA Archive Survey (NVAS).[5] If the source has been detected by the radio surveys given their typical depth, it is classified as synchrotron. The additional radio images are crucial for sources found in band 3 due to the generally wider FoV and higher sensitivity. Examples of our synergetic classification, combining data from ALMA and the VLA, are illustrated in Online Fig. B2. In band 3, we are able to confirm 14 per cent more of the total sources with the help of radio surveys.

For those unclassified sources lacking radio coverage, because of their small numbers (<20 per cent), we did not find a large difference in number counts by including these uncertain sources, so in the main context we focus on reliable SMGs that are confirmed by at least two ALMA bands or one ALMA band with the radio images available. Details of the flux density distribution, radial distance distribution, and the contribution to the number counts of the unclassified sources can be found in Online Figs B1 and D1. Source classifications are available from our on-line version of the catalogue.

3.9 Selection bias

ALMACAL is not a truly blind survey. The pre-selection of the calibrator fields can bias our survey if the central calibrator has influenced the detection of nearby sources. Since roughly 97 per cent of the calibrators are blazars, the blazar activity may have been

[4] https://www.physics.purdue.edu/astro/MOJAVE/
[5] https://www.vla.nrao.edu/astro/nvas/

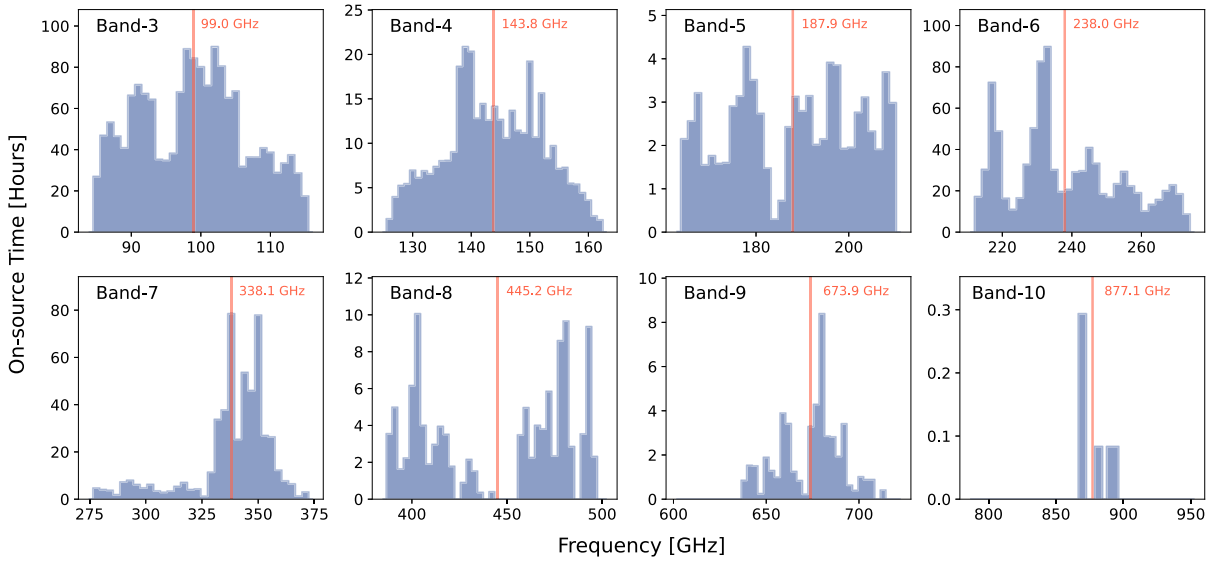

Figure 7. ALMACAL frequency coverage in the different ALMA bands. The abscissa is the frequency coverage in each ALMA band. The ordinate is the total integration time for the selected observations at each frequency. For each band, the vertical red line is the time-weighted mean frequency. In contrast to existing ALMA blind surveys, the frequency coverage of ALMACAL is much wider. To make meaningful comparisons with literature results, number counts have been rescaled to the characteristic frequencies based on the SED of a modified blackbody (see Section 4.2 for more detail).

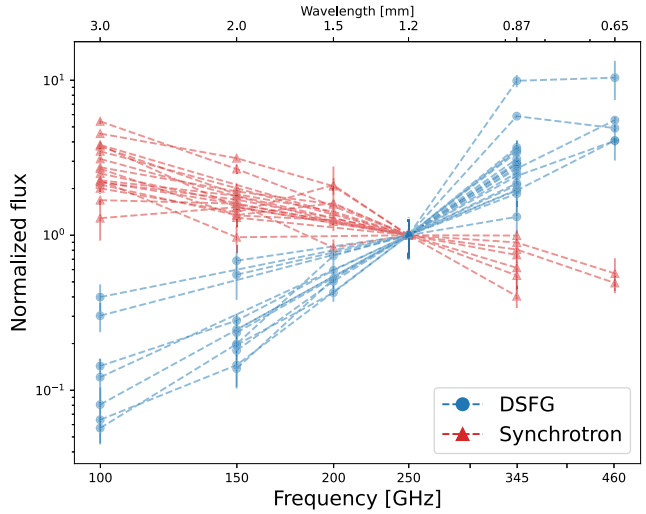

Figure 8. The submm/mm SED of a few confirmed DSFGs and Synchrotron sources. Only the sources detected in band 6 plus at least two more bands are shown here. All the SEDs have been normalized by their band 6 flux densities. We see that DSFGs have positive spectral indices with their fluxes rising at higher frequencies, while the Synchrotron sources typically have negative spectral indices with their fluxes decreasing at higher frequencies.

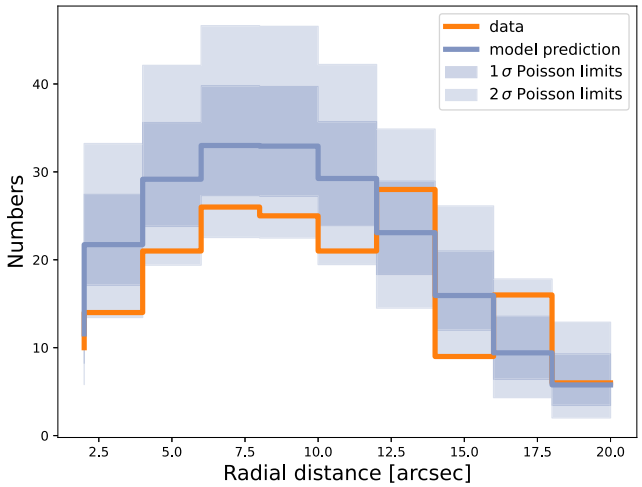

Figure 9. Radial distribution of DSFGs detected in band 6 in the ALMACAL survey. The abscissa is the radial distance of the detections from their field centre. The orange histogram is the radial distribution of the confirmed DSFGs. The blue histogram is the predicted DSFGs distribution based on the expected number counts from ASPECS González-López et al. (2020). The blue and light blue shadows are the 1σ and 2σ Poisson variations of the model predictions. Small differences between ALMACAL and ASPECS are expected. However, the radial distribution of our DSFGs is generally consistent with the expected of randomly distributed DSFGs, which indicates that the thermal sources are not clustered around the blazars.

triggered by an interaction, perhaps also giving rise to a nearby starburst. In addition, the blazar host galaxy may act as a gravitational lens. However, a blazar is bright mainly due to the beaming effect, which does not always traces the most massive galaxies and dark matter haloes. The redshift distribution of the ALMACAL blazars peaks at $0.5 < z < 1.0$ (Bonato et al. 2018), well short of the typical redshifts to DSFGs (Brisbin et al. 2017; Aravena et al. 2020; Dudzeviciute et al. 2020). Moreover, we note that a systematic search for galaxy overdensities around three distant quasars with known companions failed to find associated DSFGs (Meyer et al. 2022).

To further quantify the possible biases that might be introduced by the central bright sources, in Fig. 9 we compare the radial distribution of all our DSFGs with the prediction of the ALMA blind survey, ASPECS. We used the best-fitting number counts from ASPECS as the count model and then created a mock survey in the manner of ALMACAL. We show the radial distribution from band 6 where the number of detections is the highest. If the same density of sources detected by ASPECS is recovered here, scattered randomly throughout our fields, we would have unbiased DSFG positions, regardless of the central blazar. We made mock observations by using the same observational set-ups as ALMACAL, covering the same effective

Table 1. Statistical summary of ALMACAL detections.

ALMA band	Ref. λ (mm)	Sky coverage (arcmin^2)	On-source time (h)	⟨ν⟩ (GHz)	Number of detections	DSFGs	Synchrotron	Unclassified sources
3	3.00	817	250.6	99.0	63	8	44	12
4	2.00	157	53.1	143.75	54	20	23	11
5	1.50	27	20.8	200	21	10	9	2
6	1.20	149	344.7	250	228	132	54	42
7	0.87	45	275.2	345	130	93	22	15
8	0.65	5	52.7	460	17	13	3	1
9	0.45	1	20.9	666	2	2	0	0
10	0.35	0.1	0.5	857	0	0	0	0
total	–	1201.1	1018.5	–	371	186	102	83

area at different flux densities, and the uncertainty in the radial distribution is given by the Poisson error. Fig. 9 shows the 1 σ and 2 σ Poisson confidence levels indicating that the radial distribution of our detections is broadly consistent with ASPECS. In particular, we do not see any excess of detections near the central calibrators, indicating there is little evidence for clustering of the DSFGs around the blazars (see also Online Appendix C and Online Fig. C1 for the difference in number counts between inner and outer parts of all the fields.).

4 RESULTS

In this section, we present the source catalogue and the number counts at different wavelengths.

4.1 Source catalogue

Our classification of sources is carried out by the procedures presented in Section 3.8 and the numbers of different detections are summarized in Table 1. In total, we have detected 371 sources, including 186 secure DSFGs with multiband confirmations of their purity and spectral indices. The largest number of detections is in ALMA band 6, where we have detected 228 sources, including 132 DSFGs. This is followed by band 7, where 93 DSFGs from 130 detections have been confirmed. The detection rate in band 4 is also promising, with 20 secure DSFGs detected. Band 5 has ten DSFGs, which enables us to derive the number counts at 1.5 mm for the first time. In band 3, sky coverage is significantly larger than higher frequency bands and they are typically deeper, which makes the radio images much more important to classify the detections. Luckily, around half of the ALMACAL band 3 footprints have been covered by VLA archival surveys. Combining ALMACAL with the VLA archive, we have confirmed eight DSFGs in band 3. Band 9 has only two detections found in 1 arcmin^2 – too few to meaningfully constrain the number counts. For ALMA band 8, we adopt the results of Klitsch et al. (2020), which already benefited from the updated version of ALMACAL.

As we can see here, the number of detections provided by ALMACAL is competitive with dedicated ALMA cosmological surveys, offering a complementary way to refine the counts. We also report the detected numbers of synchrotron sources in Table 1. In this paper, we mainly focus on the number counts of DSFGs.

4.2 DSFG number counts

One of the major goals of ALMACAL is to constrain the number counts of DSFGs. The multiband coverage of the calibrators makes it possible to constrain the number counts at multiple wavelengths.

We derive the number counts following the formula used in Oteo et al. (2016). For every individual detection, its contribution to the total cumulative number counts is:

$$N_i(S_i) = \frac{1 - f_{\mathrm{sp}(S_i)}}{C(S_i) \cdot A(S_i)}, \tag{1}$$

where S_i is the flux density of the detection; $f_{\mathrm{sp}}(S_i)$ is the fraction of spurious sources at S_i, which is equal to unity with our 5σ detection threshold; $A(S_i)$ and $C(S_i)$ are the effective area and completeness at S_i, respectively.

The final number counts are calculated in flux density bins, S_j, evenly spaced in log scale. For each bin, the cumulative number counts is defined as:

$$N_j(>S_j) = \sum_i^{n_i} N_i(S_i), \tag{2}$$

where n_i is the number of N_i that satisfies $S_i \leq S_j$. The differential number counts for each flux density bin is given by:

$$\frac{\mathrm{d}N_j}{\mathrm{d}S} = \frac{\sum_i^{n_{ij}} N_i}{S_{j+1} - S_j}, \tag{3}$$

where n_{ij} is the number of N_i that meets $S_j < S_i \leq S_{j+1}$.

The uncertainties in the number counts are calculated via Monte Carlo simulations. First, the flux density of every detection was randomly sampled according to its 1σ uncertainty. Then, the whole DSFG sample is grouped into the different flux density bins, S_j. The flux density bins are constructed to include at least three detections. Their cumulative and differential number counts are calculated using equations (2) and (3), respectively. We ran this simulation 1000 times to determine the scatter of the number counts. Finally, Poisson errors were added to the uncertainties of number count according to the number of detections in each bin. We discuss the number counts that include the unclassified sources in Online Fig. D1.

To compare with literature results and theoretical models, we rescaled the flux densities of all our detections to reference wavelengths in each ALMA band. The reference wavelengths used in this work are summarized in Table 1. The rescaling factor is determined by the flux density ratio at different wavelengths assuming the composite SED from the AS2UDS survey at a reference redshift of $z = 2.0$ (Dudzeviciute et al. 2020), which also has a power-law index $\alpha \sim 3.2$. During the scaling, we add 20 per cent of error to the number counts to reflect the variation of the composite SED (Dudzeviciute et al. 2020). The wavelength rescaling does not provide a perfect correction for every detection, because each will have a different, unknown redshift, but since the correction is small (∼10 per cent), it should be reasonable in a statistical sense. The number counts before and after rescaling are displayed in Online Fig. D1.

Table 2. Cumulative and differential number counts in various ALMA bands.

Cumulative number counts					Differential number counts				
S (mJy)	N	$N(>S)$ (deg^{-2})	$\delta N(>S)_{\rm lower}$	$\delta N(>S)_{\rm upper}$	S (mJy)	N	dN/dS (mJy^{-1}deg^{-2})	δ (dN/dS)$_{\rm lower}$	δ(dN/dS)$_{\rm upper}$
				Band 3 (3 mm)					
0.07	7	380	170	230	0.17	4	1990	1130	1690
0.26	3	20	10	20	–	–	–	–	–
				Band 4 (2 mm)					
0.10	21	5310	1850	2040	0.14	4	54440	33170	47700
0.18	17	1550	450	530	0.25	7	9050	3880	5260
0.33	10	410	130	180	0.46	7	1480	560	810
0.60	3	70	40	70	–	–	–	–	–
				Band 5 (1.5 mm)					
0.41	10	2320	750	1020	0.72	7	4080	1570	2250
1.03	3	370	200	360	–	–	–	–	–
				Band 6 (1.2 mm)					
0.10	132	14350	2090	2090	0.13	7	136600	57430	78580
0.16	125	7470	880	880	0.20	7	26900	11620	15700
0.25	118	5300	530	530	0.33	18	14820	3760	4620
0.40	100	3410	350	350	0.53	30	7990	1500	1780
0.65	70	1780	210	210	0.85	28	2720	510	620
1.05	42	890	130	130	1.37	22	960	200	250
1.69	20	390	80	100	2.21	17	390	90	120
2.73	3	50	20	50	–	–	–	–	–
				Band 7 (870 μm)					
0.19	92	21900	3250	3250	0.25	8	73860	29940	39660
0.32	84	12950	1640	1640	0.42	15	28600	8210	10180
0.52	69	7390	930	930	0.69	16	10230	2670	3350
0.86	53	4120	560	560	1.15	14	2780	740	960
1.43	39	2660	420	420	1.90	14	1280	340	440
2.36	25	1540	300	370	3.14	12	530	150	200
3.91	13	770	210	270	5.19	9	220	70	100
6.47	4	230	110	180	–	–	–	–	–

Table 2 reports our cumulative and differential number counts and they are plotted in Fig. 10, which shows the number counts in the different ALMA bands. This is the first time that we have number counts calculated consistently across several bands. In the following section, we discuss comparisons with the literature, and with several different model predictions, and the implications of these number counts in the context of the underlying galaxy populations and cosmic infrared background.

5 DISCUSSION

5.1 Comparison with the literature

We compare our number counts reported in Table 2 only with those from other interferometric studies – all of them from ALMA – which are relatively free from source blending. In the following subsections, we will discuss number counts in each band separately, comparing ALMACAL with the literature.

5.1.1 870 μm

Number counts around 870 μm – ALMA band 7 – were the first to be established in the field of submm cosmology, where useful data have been collected since the first single-dish survey (Smail, Ivison & Blain 1997; Barger et al. 1998; Hughes et al. 1998; Blain et al. 1999). In Fig. 11 we compare the number counts from ALMACAL with the results from previous ALMA surveys at 870 μm. Karim et al. (2013) published the first ALMA interferometric number counts at 870 μm, following up the bright single-dish sources which had been detected in the Extended *Chandra* Deep Field South using LABOCA on the 12-m APEX telescope (LESS: Weiß et al. 2009). After that, Simpson et al. (2015b), Stach et al. (2018), and Simpson et al. (2020) targeted a larger sample of similarly bright SCUBA-2 sources with ALMA. More recently, Béthermin et al. (2020) probed fainter flux densities using serendipitous detections from the ALMA Large Programme to INvestigate [C II] at Early times (ALPINE) survey. Note that this latter survey is expected to be more biased than ALMACAL, since they targeted known high-redshift sources. With ALMACAL we are going even deeper, detecting sources down to 0.2 mJy at the 5σ level, whilst remaining able to detect relatively bright sources due to the large number of fields covered. As Fig. 11 show at the faint end, our number counts follow the general trend seen in earlier work and are consistent with the ALPINE survey (Béthermin et al. 2020). At the bright end, our results match well with the follow-up ALMA surveys of known, bright submm sources (Karim et al. 2013; Stach et al. 2018; Simpson et al. 2020).

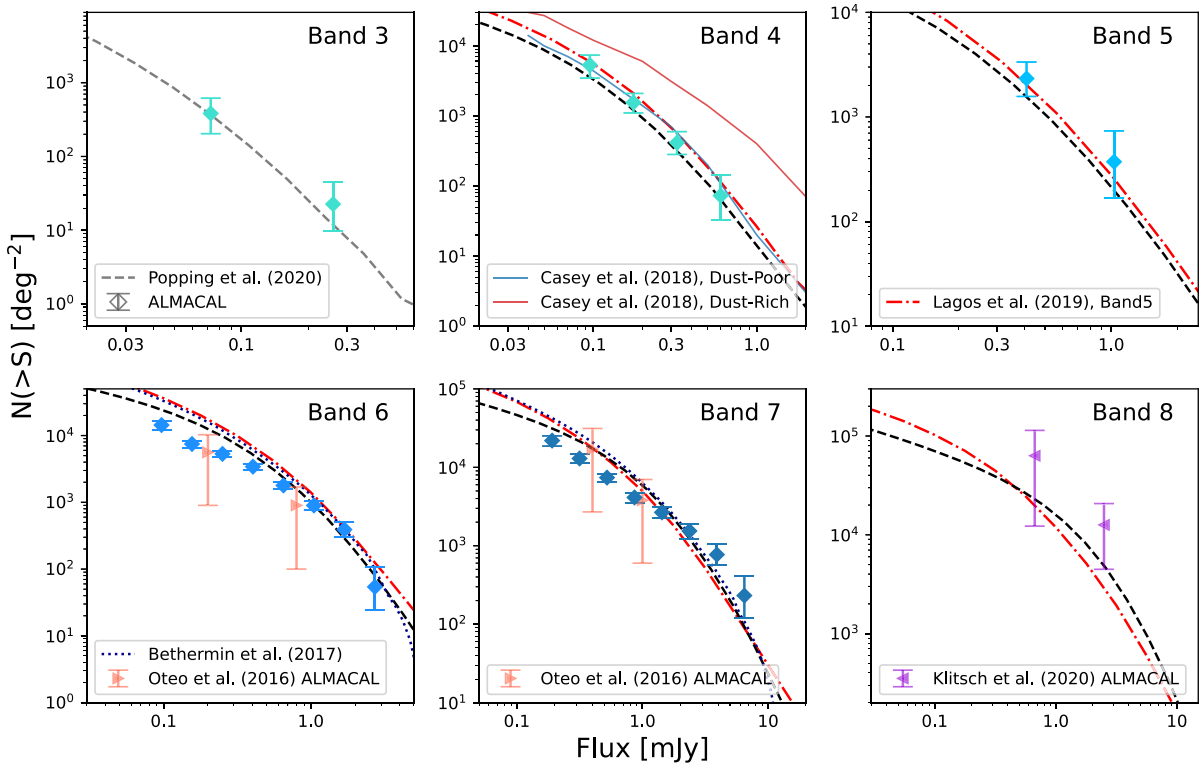

Figure 10. Multiband number counts from ALMACAL. The number counts reported in this work are displayed in each panel. In bands 6 and 7, we also show the first ALMACAL number counts by Oteo et al. (2016); in band 8, we adopt the results from Klitsch et al. (2020). We also show model predictions of the number counts at various wavelengths, including the semi-analytical model from Lagos et al. (2020) and the semi-empirical models from Popping et al. (2020) and Béthermin et al. (2017). A successful model should be able to explain the number counts at the different wavelengths simultaneously. The most recent semi-analytical and semi-empirical models are broadly consistent with the number counts of ALMACAL at bright fluxes, but show apparent differences at fainter limits in band 6 and 7 (see also Figs 11 and 12), which need to be confirmed by the future surveys. Therefore, multiwavelength number counts are a powerful tool to validate those models.

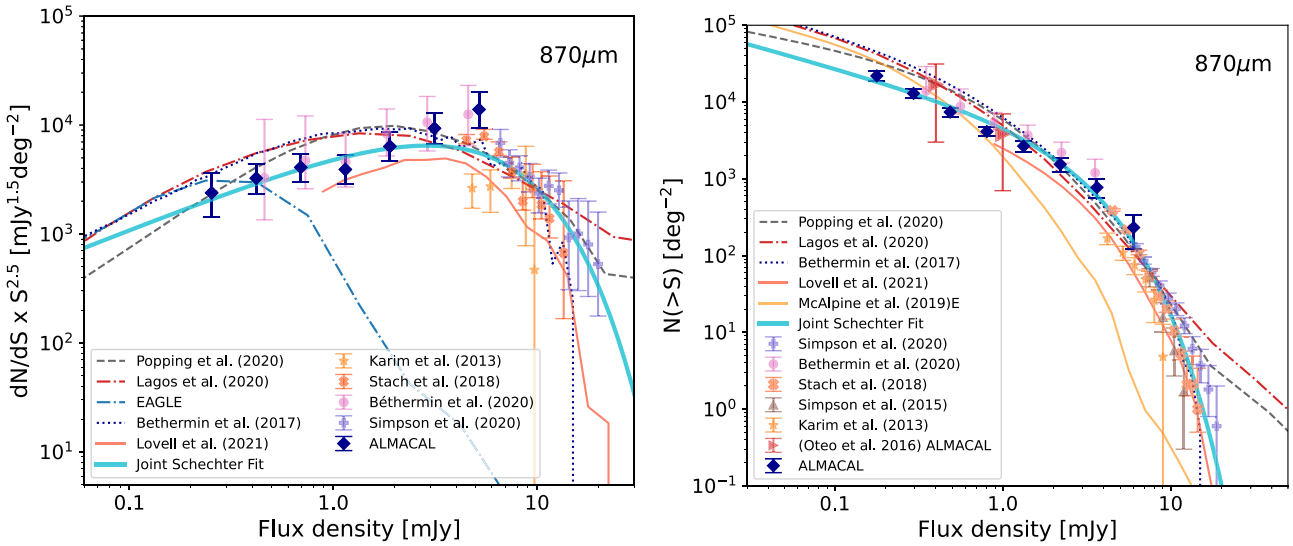

Figure 11. Differential number counts (left-hand panel) and cumulative number counts (right-hand panel) at 870 μm in band 7. The differential counts have been normalized by $S^{2.5}$ to reduce the dynamic range. We show the new ALMACAL results and the previous ALMACAL band 7 number counts from Oteo et al. (2016). We also plot the number counts from surveys that used ALMA (Karim et al. 2013; Simpson et al. 2015b, 2020; Stach et al. 2018; Béthermin et al. 2020). Model predictions from Fig. 10 are included here, with the same line styles. The predictions based on EAGLE (Camps et al. 2018; McAlpine et al. 2019) and SIMBA (Lovell et al. 2021) cosmological hydrodynamic simulations are also included as a comparison. The radiative transfer post-processing is only applied to SIMBA galaxies with SFR $> 20\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$, which can only give complete number counts above $S_{870\,\mu\mathrm{m}} > 1$ mJy. The number counts from ALMACAL are consistent with previous results in a wide flux density range and represent the deepest survey available at the moment. The best joint Schechter fits are shown in each plot.

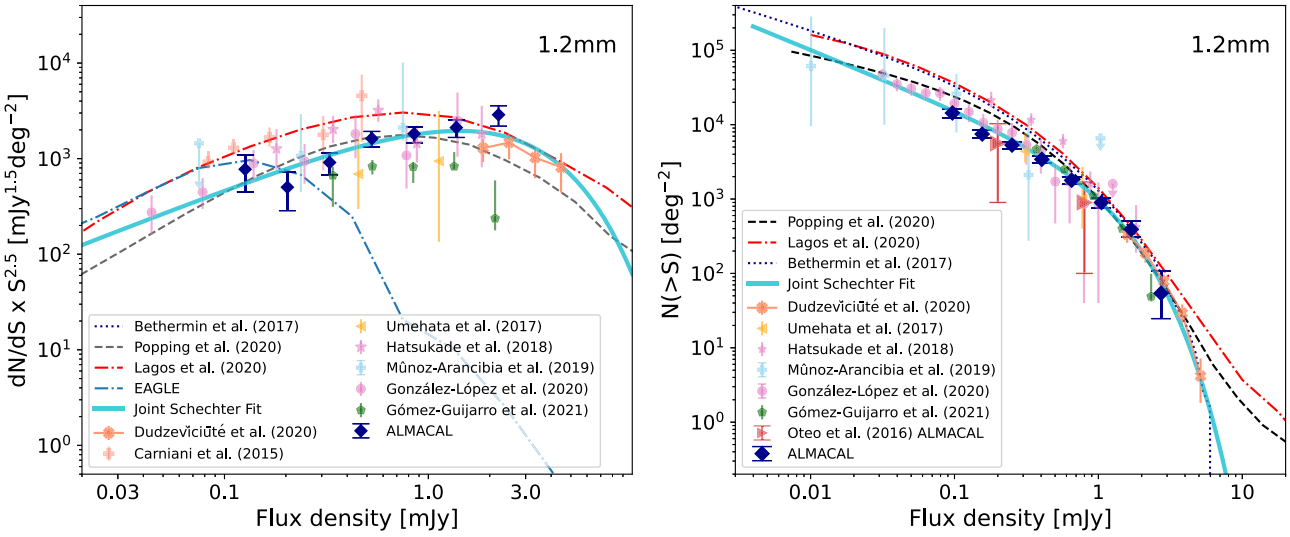

Figure 12. Differential number counts (left-hand panel) and cumulative number counts (right-hand panel) at 1.2 mm. We show the new ALMACAL results and the previous ALMACAL band 6 number counts from Oteo et al. (2016). Interferometric results reported by Umehata et al. (2018), Hatsukade et al. (2018), Muñoz Arancibia et al. (2018) are also shown. From Umehata et al. (2018), only the results from the field are shown. From Muñoz Arancibia et al. (2018), only the combined results are shown (see the updated results from the corrigendum Muñoz Arancibia et al. 2019). We also show the predicted SED-scaled number counts from AS2UDS (Dudzeviciute et al. 2020) to provide constraint in the higher flux density range. Model predictions from Fig. 10 are also included here. In band 6, ALMACAL overlaps with GOODS-ALMA and empirical AS2UDS prediction at the brighter end and is consistent with ASPECS at the fainter end. The best joint Schechter fits are shown in each plot.

5.1.2 1.2 mm

Number counts around 1.2 mm – ALMA band 6 – have also been well studied for many years. It started with single-dish surveys, such as the early single-dish surveys at the IRAM 30-m telescope using the Max-Planck Millimeter Bolometer array (MAMBO) (e.g. Greve et al. 2004), the Bolocam at the 10-m Caltech Submillimeter Observatory (e.g. Laurent et al. 2005), and AzTEC at the 15-m James Clerk Maxwell Telescope and the 10-m Atacama Submillimetre Telescope Experiment (ASTE) (e.g. Scott et al. 2012). The first deep interferometric observations were made by Hatsukade et al. (2013) towards the Subaru/*XMM–Newton* Deep Survey field. Since then, band 6 number counts have been extended by mining existing deep fields in the ALMA archive (Ono et al. 2014; Carniani et al. 2015; Fujimoto et al. 2016). More recently, dedicated ALMA blind surveys, ASPECS (González-López et al. 2019, 2020) and the ALMA Frontier Fields survey (González-López et al. 2017; Muñoz Arancibia et al. 2018) have greatly improved the depth, while GOODS-ALMA (Franco et al. 2018; Gómez-Guijarro et al. 2022) has increased the sky coverage substantially, up to 72-arcmin^2. All of these surveys are plotted in Fig. 12. ALMACAL offers a good balance between survey depth and sky coverage. In the flux density range between 0.1 and 1.0 mJy, most of the number counts are consistent within 2σ. At the brighter end, ALMACAL is consistent with GOODS-ALMA and represents the widest survey available.

At 1.2 mm, discrepancies have been reported in number counts both at the fainter end (<0.1 mJy) and at the brighter end (>1.0 mJy). At the fainter end, Fujimoto et al. (2016) reported number counts that appear systematically high relative to other surveys, possibly due to the use of a relatively low SNR detection threshold (SNR = 3σ) and the bias introduced using the targeted fields. At these flux limits, we find that ALMACAL is consistent with the results from ASPECS (González-López et al. 2020) and the ALMA Frontier Fields Survey (González-López et al. 2017; Muñoz Arancibia et al. 2019). At the bright end, there are no direct number counts measurements from ALMA for sources brighter than $S_{1.2\mathrm{mm}} > 2$ mJy. To better constrain the 1.2 mm bright-end number counts, we adopted the SED-scaled number counts from AS2UDS (Stach et al. 2018, 2019). Dudzeviciute et al. (2020) derived SED fits for all SMGs within the $\approx$1 deg^2 AS2UDS, providing the opportunity to derive the expected counts in other wavelengths. We thus predict 1.2 mm flux densities for each SMG based on its best-fitting SED and then derive the number counts after correcting for the original survey completeness down to $S_{1.2\,\mathrm{mm}} \approx 3.6$ mJy (see also Stach et al. 2018). The AS2UDS prediction at 1.2 mm is overplotted in Fig. 12. ALMACAL is also consistent with the AS2UDS predictions in the overlapping regime. In the future, a larger blind ALMA survey, or follow-up of the bright sources from wide field 1.2-mm single-dish surveys, is needed to verify our results and confirm the number counts at the brightest flux densities.

5.1.3 2 mm

There are relatively few published studies of number counts at longer wavelengths $\lambda \gg 1$ mm. The first single-dish 2 mm survey was conducted by the Goddard IRAM Superconducting Millimeter Observer (GISMO) at the IRAM 30-m telescope (Staguhn et al. 2014; Magnelli et al. 2019). The number counts were later confirmed by the Mapping Obscuration to Reionization with ALMA (MORA) survey (Zavala et al. 2021). We revisit the 2 mm number counts with the detections from ALMACAL. In Fig. 13, we plot the new measurements from ALMACAL, as well as the data from Magnelli et al. (2019) and Zavala et al. (2021). Our new measurements are consistent with the two previous surveys at high flux densities and go 4 × deeper, down to $S_{2\,\mathrm{mm}} \sim 0.1$ mJy. We also predict the 2 mm number counts based on the AS2UDS sample, following the same method we used for 1.2 mm in Section 5.1.2. The AS2UDS predictions are shown in Fig. 13, which matches well with MORA, suggesting that there is no substantial new population of sources appearing at $\lambda \gg 1$ mm.

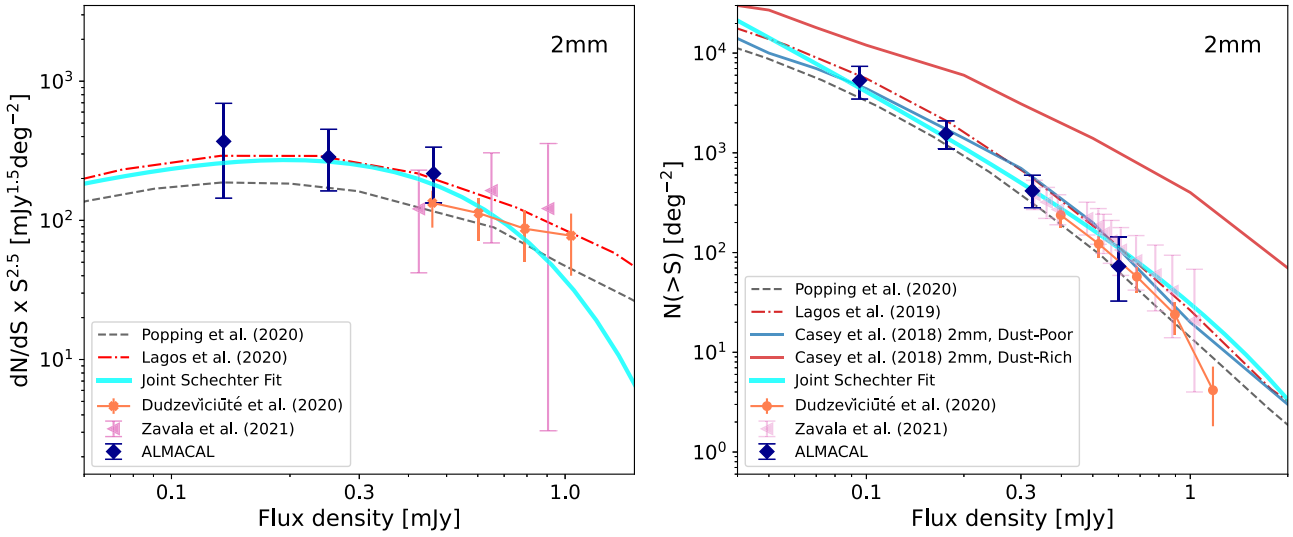

Figure 13. Differential number counts (left-hand panel) and cumulative number counts (right-hand panel) at 2 mm. We show the results from ALMACAL and the Mapping Obscuration to Reionization with ALMA (MORA) survey (Zavala et al. 2021). We also include the SED predicted 2 mm number counts from AS2UDS survey (Dudzeviciute et al. 2020). The two different models proposed by (Casey et al. 2018) are both included. Existing data support the dust-poor model and are consistent with the model predictions from Popping et al. (2020) and Lagos et al. (2020). The best joint Schechter fits are shown in each plot. ALMACAL is consistent with Zavala et al. (2021) and empirical SED-scaled AS2UDS counts at the bright end and present the deepest survey at 2 mm.

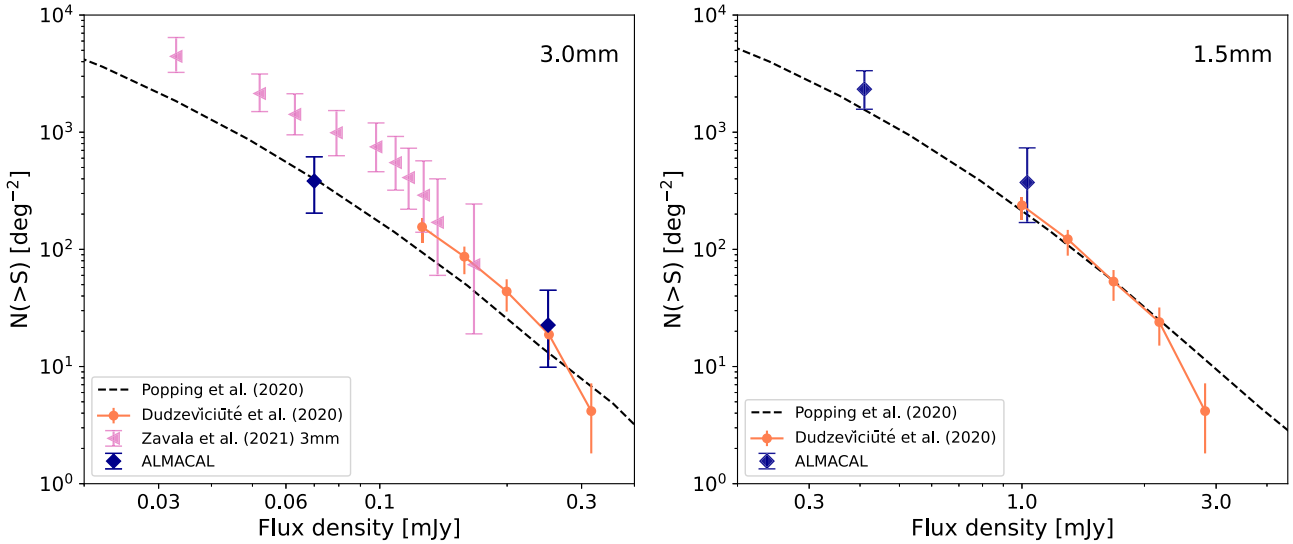

Figure 14. Cumulative number counts at 3 and 1.5 mm. We also include semi-empirical model predictions from Popping et al. (2020) and the SED-based prediction from AS2UDS survey (Dudzeviciute et al. 2020). At 1.5 mm, ALMACAL is the first survey to constrain the number counts of DSFGs. At 3.0 mm, we also include the measurements from Zavala et al. (2021), which is based on ALMA archival data. Number counts from ALMACAL overlap with Zavala et al. (2021) and empirical SED-scaled AS2UDS predictions at brighter end, but are systematically lower than Zavala et al. (2021) at the faint end.

5.1.4 1.5 and 3 mm

ALMACAL is the first survey to constrain the number counts of DSFGs at 1.5 mm, with detections down to $S_{1.5\,\mathrm{mm}} \sim 0.4$ mJy. At 3 mm, Zavala et al. (2018) provided constraints on the DSFGs number counts using the ALMA Science Archive towards regions around targeted sources in three extragalactic legacy fields: COSMOS, CDF-S, and the UDS. These results were later updated in Zavala et al. (2021) after removing three spurious detections. We compare our cumulative number counts with those of Zavala et al. (2021) in Fig. 14, along with the SED-based predictions from AS2UDS and various theoretical model predictions (Lagos et al. 2020; Popping et al. 2020). Our counts agree well with the those predicted from AS2UDS and Zavala et al. (2021) at the bright end. However, they are systematically lower than Zavala et al. at the faint end, which we attribute to their counts being biased high due to associations and potentially source blending in the original target selection.

5.1.5 Joint fitting

Combining different surveys offers a way to mitigate some of the biases that are particular to each individual survey. We have therefore perform a joint fit of the count data in ALMA bands where data is sufficient: 870 μm, 1.2 mm, and 2 mm. In band 7 (870 μm), we include the data from Karim et al. (2013), Simpson et al. (2015b),

Table 3. Best-fitting models for cumulative and differential number counts.

Cumulative number counts							
	Schechter			Double power law			
	α	S_0 mJy	N_0 deg^{-2}	α_1	α_2	S_0 mJy	N_0 deg^{-2}
B4 (2 mm)	$-1.7^{+0.3}_{-0.3}$	$1.0^{+0.6}_{-0.3}$	$0.09^{+0.01}_{-0.01}$	0.0	$2.1^{+0.2}_{-0.2}$	1.0	$0.04^{+0.01}_{-0.01}$
B6 (1.2 mm)	$-0.8^{+0.1}_{-0.1}$	$1.0^{+0.1}_{-0.1}$	$2.7^{+0.3}_{-0.3}$	$1.0^{+0.1}_{-0.1}$	$3.8^{+0.4}_{-0.3}$	$1.2^{+0.2}_{-0.2}$	$1.6^{+0.2}_{-0.2}$
B7 (870 μm)	$-0.6^{+0.1}_{-0.1}$	$2.2^{+0.1}_{-0.1}$	$9.6^{+0.9}_{-0.9}$	$1.0^{+0.1}_{-0.1}$	$4.7^{+0.2}_{-0.2}$	$4.5^{+0.4}_{-0.4}$	$3.6^{+0.7}_{-0.6}$
Differential number counts							
	Schechter			Double power law			
	α	S_0 mJy	N_0 mJy^{-1}deg^{-2}	α_1	α_2	S_0 mJy	N_0 mJy^{-1}deg^{-2}
B4 (2 mm)	-1.7	$0.2^{+0.4}_{-0.3}$	$6.0^{+0.9}_{-0.7}$	0.0	$3.1^{+0.8}_{-1.1}$	1.0	$0.1^{+0.2}_{-0.1}$
B6 (1.2 mm)	$-1.7^{+0.1}_{-0.1}$	$1.7^{+0.5}_{-0.3}$	$2.4^{+0.8}_{-0.9}$	$1.9^{+0.1}_{-0.2}$	$4.6^{+1.8}_{-1.0}$	$2.3^{+0.8}_{-0.9}$	$1.1^{+1.6}_{-0.5}$
B7 (870 μm)	$-1.7^{+0.2}_{-0.2}$	$3.9^{+0.7}_{-0.7}$	$2.2^{+1.4}_{-0.9}$	$2.0^{+0.2}_{-0.3}$	$4.7^{+0.9}_{-0.5}$	$5.6^{+2.2}_{-1.5}$	$0.8^{+1.6}_{-0.8}$
Joint fit							
	α/α_1	$/\alpha_2$	$S_{2.0\,\mathrm{mm}}$ mJy	$N_{2.0\,\mathrm{mm}}$ mJy^{-1}deg^{-2}	$S_{1.2\,\mathrm{mm}}$ mJy	$N_{1.2\,\mathrm{mm}}$ mJy^{-1}deg^{-2}	$S_{870\,\mu\mathrm{m}}$ mJy	$N_{870\,\mu\mathrm{m}}$ mJy^{-1}deg^{-2}
Schechter	$-1.7^{+0.1}_{-0.1}$		$0.3^{+0.1}_{-0.1}$	$5.0^{+6.9}_{-4.0}$	$1.9^{+0.5}_{-0.3}$	$1.8^{+0.8}_{-0.7}$	$3.9^{+0.5}_{-0.5}$	$2.1^{+0.7}_{-0.5}$
Double power-law	$1.9^{+0.1}_{-0.1}$	$4.2^{+0.3}_{-0.3}$	$0.3^{+0.1}_{-0.1}$	$3.7^{+6.9}_{-3.9}$	$2.2^{+0.5}_{-0.4}$	$1.1^{+0.7}_{-0.5}$	$4.3^{+0.7}_{-0.6}$	$1.3^{+0.6}_{-0.5}$

Note. Values without errors are fixed values during the fitting.

Stach et al. (2018), Béthermin et al. (2020), Simpson et al. (2020); in band 6 (1.2 mm), we include the data from Umehata et al. (2017), Muñoz Arancibia et al. (2018), González-López et al. (2020), and the SED-scaled number counts from AS2UDS (Stach et al. 2018); in band 4, we include the data from Zavala et al. (2021). We fit the differential and the cumulative number counts with the Schechter function (Schechter 1976) and double power law:

$$N(>S)\,\mathrm{d}S = N_0 \left(\frac{S}{S_0}\right)^{\alpha} \exp\left(-\frac{S}{S_0}\right) \mathrm{d}\left(\frac{S}{S_0}\right) \quad , \tag{4}$$

$$N(>S)\,\mathrm{d}S = N_0 \left[\left(\frac{S}{S_0}\right)^{\alpha_1} + \left(\frac{S}{S_0}\right)^{\alpha_2}\right]^{-1} \mathrm{d}\left(\frac{S}{S_0}\right) \quad , \tag{5}$$

where N_0 is the normalization factor, in units of deg^{-2} for the cumulative number counts and in units of mJy^{-1}deg^{-2} for the differential number counts. S_0 is the flux density at the turnover. The index, α, is the power-law index of the Schechter function in equation (4); while α_1 and α_2 are the slopes of the two independent power laws in equation (5). During the fitting, N_0, S_0 and α, or α_1 and α_2 are free parameters. The best-fitting models for the different bands are summarized in Table 3 and the best-fitting Schechter functions are plotted in Figs 11, 12, and 13. We adopt the best-fitting Schechter function as our fiducial model, because it provides the best fit to the flattening trend in the very deep 1.2-mm number counts (see discussion in González-López et al. 2020).

We constrain the free parameters using the maximum-likelihood minimization algorithm, MINIMIZE, from SCIPY.OPTIMIZE, and adopt the Markov chain Monte Carlo (MCMC) sampler, EMCEE, to derive the confidence levels. We treat different surveys equally, using their own reported errors, but increase the uncertainties if rescaling is applied (see more in Section 4.2). In band 4, the available data are not sufficient to constrain the models. We thus fix $\alpha = 1.7$ for the Schechter function and $\alpha_1 = 0$ and $S_0 = 1$ mJy for the double power law (see more discussion in Section 5.3).

From our best fits, we confirm the shallower trend seen in the faint flux density range at 870 μm (Béthermin et al. 2020) and 1.2 mm (González-López et al. 2020). Based on the joint Schechter function fitting for the differential number counts, the turnover flux density at 870 μm is $S_{870\,\mu\mathrm{m}} = 3.9^{+0.7}_{-0.7}$ mJy. At 1.2 mm, González-López et al. (2020) found that the number counts flattened below 0.1 mJy, and argued for a triple power law to fit the differential number counts, based on their $P(D)$ analysis. In our fitting, a Schechter function with a turnover at $S_{1.2\mathrm{mm}} = 1.7^{+0.5}_{-0.3}$ mJy with a power-law index of $\alpha = 1.7 \pm 0.1$ gives a reasonable fit to the existing data. However, as we mentioned earlier, considerable uncertainties exist both at the fainter and brighter end of the 1.2 mm number counts. At 2 mm, the constrained fitting gives a turnover flux $S_{2\mathrm{mm}} = 0.2^{+0.4}_{-0.3}$ mJy, which still suffers from large uncertainties.

5.2 Comparing with model predictions

Simulations and analytical models of galaxy formation and evolution must also explain the DSFG population. Depending on basic assumptions and underlining techniques, these modelling methods can be roughly divided into semi-analytic models (SAMs), semi-empirical models, and hydrodynamic (N-body) simulations.

SAMs are the most widely used tools (Cole et al. 2000). They start with the merging trees of dark matter halo from either cosmological N-body simulations or self-consistent Monte Carlo simulations and use analytical equations for baryonic processes. Their flexibility and relatively cheap computational cost make them powerful to explore large regions of parameter space. However, it was difficult to reproduce submm/mm number counts by incorporating standard feedback processes along with radiative transfer to balance the

FUV-to-NIR and the FIR-to-submm/mm emission of galaxies simultaneously across cosmic time (Baugh et al. 2005; Somerville et al. 2012; Lacey et al. 2016; Cowley et al. 2019). Lagos et al. (2019, 2020) adopted attenuation curves obtained from the radiative transfer analysis of hydrodynamical simulations (Trayford et al. 2020) to reproduce the panchromatic emission of galaxies. Their model is better able to match the multiwavelength DSFG number counts. We show predictions from Lagos et al. (2020) in Fig. 10. They are broadly consistent with ALMACAL in the various ALMA bands, but exhibit overprediction at the fainter end of bands 6 and 7.

Due to the increasing complexity of SAMs and the computational cost of fully 3D dust radiative transfer, observational empirical scaling relations have been coupled into SAMs – creating semi-empirical models – to simplify the calculations (Béthermin et al. 2017; Popping et al. 2020). Popping et al. (2020) provide a new framework to explain the number counts at 1.2 mm (González-López et al. 2020), modelling the submm/mm emission of galaxies with simple functions from full radiative transfer simulations, where the submm/mm flux density of a galaxy is a function of SFR and total dust mass (Hayward et al. 2011). The Popping et al. model is extended here to provide the predictions on the Rayleigh-Jeans tail of the submm SED by fitting the original 850-μm and 1.1-mm predictions with a grey body. Their predictions for different ALMA bands are shown in Fig. 10. They are slightly lower than the predictions from Lagos et al. (2020) at fainter flux densities and are more consistent with the observations but at the cost that there are no true ab initio predictions. Casey et al. (2018) also explored the number counts with different fractions of DSFGs at $z > 4$. They proposed two simple models: their dust-poor model has a steep slope in the luminosity function in the early Universe, while their dust-rich model has a much shallower slope and predicts many more DSFGs. The two models thus predict different number counts in the mm bands, with the dust-rich model having a higher number density of DSFGs. Our 2 mm number counts favour the dust-poor model, which is also closer to the prediction from Popping et al. (2020) and Lagos et al. (2020).

Hydrodynamic models, on the other hand, still struggle to reproduce the number counts of DSFGs and usually underpredict their numbers (Shimizu et al. 2012). With recent advances in computational resources, fully 3D radiative transfer (RT) has been integrated with modern cosmological simulations with different recipes for stellar and AGN feedback (Camps et al. 2018; McAlpine et al. 2019; Lovell et al. 2021). Here, we compare our results with EAGLE (Evolution and Assembly of GaLaxies and their Environment; Schaye et al. 2015) and SIMBA (Davé et al. 2019) cosmological simulations that have 3D radiative transfer implementations (Camps et al. 2018; Lovell et al. 2021). To compute the number counts of EAGLE, we first retrieve the publicly available observer-frame fluxes in ALMA bands 6 and 7 of all galaxies from $z = 10$ to $z = 0$ (see details about the flux calculation in Camps et al. 2018). Then, we use the fluxes and all the snapshots of the simulation to construct a light cone. For each redshift, we compute the projected sky area of 100 Mpc^2 and the redshift bin implied by a 100 Mpc comoving distance. We then calculate the number of galaxies per unit area, per unit redshift, at each flux bin contributed by each snapshot $N(z, S)$. Finally, We integrate $N(z, S)$ along redshift to obtain the cumulative number counts $N(S)$. For SIMBA, we adopt the results from Lovell et al. (2021), which reproduced the population of bright DSFGs but – due to the computational cost – their predicted number counts are only complete down to $S_{850\,\mu\mathrm{m}} \sim 1$ mJy. We show all the predictions in Figs 11 and 12.

Modelling the basic observations can help us to understand the relevant physical processes behind galaxy formation and evolution. Number counts are one of the most basic measurements from observations and play an essential role in validating models. By construction, semi-empirical models closely match the observed scaling relations, producing reasonable number counts at most of the observable flux densities, which can help design future surveys. SAMs start with basic physical assumptions, representing our understanding of the physics behind the observables. They can already give reasonable predictions for the flux densities covered by existing surveys. Future surveys with the flexibility offered by SAMs will continue to be a powerful tool to understand the fundamental physics behind the number counts. Hydrodynamic cosmological simulations remain the most computational expensive method. However, they also offer the most detailed description of the internal/external galactic structures and their environments. The discrepancy between the model predictions and the observations will continue to motivate more detailed sub-grid physics in cosmological simulations. Multiwavelength number counts – especially in the submm/mm range – offer additional constraints on the redshift distribution of DSFGs, thanks to the negative-*K*-correction (Baugh et al. 2005; Casey et al. 2014). Future, deep multiband DSFG survey will therefore continue to be a critical benchmark for modern galaxy evolution models.

5.3 The galaxy populations behind the number counts

It has long been argued that submm/mm surveys at different depths and different wavelengths probe different galaxy populations and redshifts (Chapman et al. 2005; Casey et al. 2014). This claim has been explored by semi-empirical models and semi-analytical models (Béthermin et al. 2015; Lagos et al. 2020; Popping et al. 2020). Béthermin et al. (2015) argued that shallower surveys at longer wavelengths yield DSFGs with a higher median redshift. Popping et al. (2020) emphasized that the contribution from galaxies at $z > 4$ to the number counts is small and the flattening of the number counts at $S_{1.2\,\mathrm{mm}} < 0.1$ is likely caused by the shallow faint-end luminosity functions of $z = 1$–2 DSFGs. Lagos et al. (2020) also suggested that the redshift increases with flux density for bright SMGs, but this trend flattens out for fainter DSFGs. In general, simulations reproduced the observational trend that reshift increases with flux density for bright SMGs (see also Simpson et al. 2020, see also Ivison et al. 2007; Brisbin et al. 2017; Simpson et al. 2017; Stach et al. 2019), but we still have limited observational constraints about the redshift distribution of fainter DSFGs.

Thanks to the unique wavelength coverage of ALMACAL, we have been able to constrain the multiband number counts for faint DSFGs simultaneously. Even though all the parameters are fitted independently, band 6 and 7 share remarkable similar power-law indices (see Table 3). If we were probing a similar galaxy population at different wavelengths and if their redshift distributions do not vary significantly, the conversion of number counts between different wavelengths could then be approximated by a simple colour difference. Inspired by this idea, we jointly fit the differential number counts in band 6 and band 7 by linking their power-law indices. In Fig. 15, we show the joint Schechter and double power-law fitting of ALMACAL measurements. In the flux density regime of ALMACAL, the differential number counts in band 6 and band 7 can be well fit with the same power-law index with $\alpha_1 = -1.9 \pm 0.1$, implying the number counts in band 6 and band 7 are dominated by the same galaxy population. A tentative fit with the same power-law index can also be applied to the fainter end number counts at 2 mm, but the statistical significance is limited. The joint Schechter fitting gives $\alpha = 1.7 \pm 0.1$, which is the same for independent fitting for

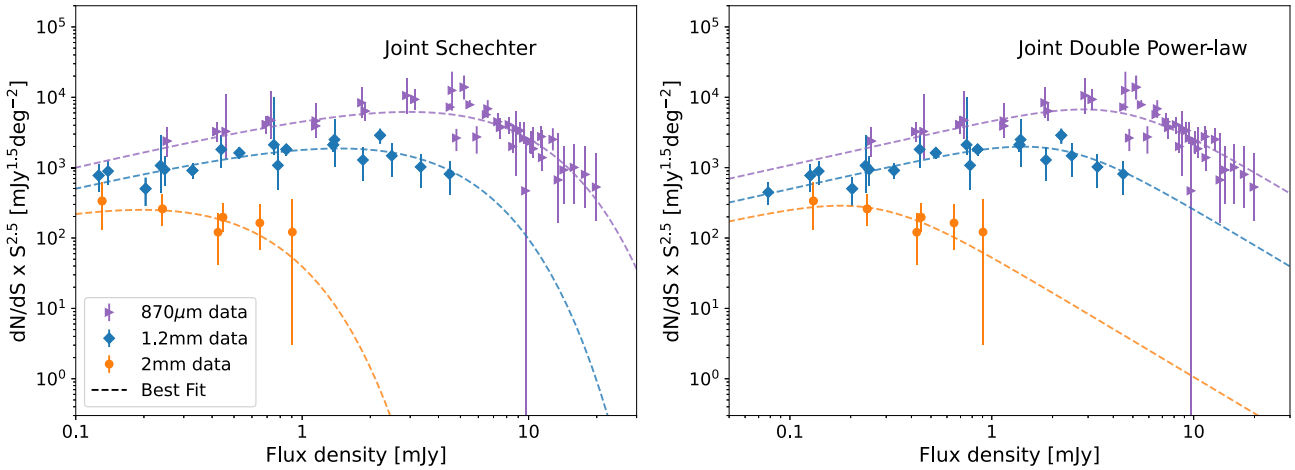

Figure 15. Joint fits combining differential number counts at 2 mm, 1.2 mm, and 870 μm. Left-hand panel: Joint Schechter fits for all the data mentioned in Section 5.1.5. During the fitting, the α index of the Schechter function is bound for all the three wavelengths. The best final fits give $\alpha = -1.7 \pm 0.1$. Right-hand panel: Joint double power-law fits. In the flux density regime of ALMACAL, all the measurements can be fitted with the same power-law index $\alpha = -1.9 \pm 0.1$, indicating the same or similar galaxy population behind the number counts.

band 6 and 7. To give more robust total CIB prediction in band 4, we fixed $\alpha = 1.7$ for the Schechter fitting of differential number counts.

5.4 Resolved cosmic infrared background

Two methods have been used to constrain the CIB: the direct measurement and the integrated intensity based on galaxy number counts. For direct measurements, the pioneering experiment were achieved by the Far Infrared Absolute Spectrophotometer (FIRAS) aboard *Cosmic Background Explorer* (*COBE*) (Puget et al. 1996; Fixsen et al. 1998) and were later continued by instruments including *ISOPHOT* (Juvela et al. 2009), *AKARI* (Matsuura et al. 2011), and *Planck*/HFI (Planck Collaboration XVI 2014) at different wavebands. A direct measurement requires accurate foreground modelling for the scattered light from solar interplanetary dust and the emission from the Milky Way, which usually needs additional assumptions or ancillary data for calibration (Fixsen et al. 1998; Odegard et al. 2019). On the other hand, if we have good knowledge of the galaxy number counts, we can also reconstruct the CIB by integrating the intensities from all the galaxies at different wavelengths. Due to the limiting sensitivity of the instruments used to derive number counts, extrapolations are normally needed to recover the undetected faint sources. Comparing the integrated energy density of galaxy number counts with direct measurements, we can quantify the energy budgets of different galaxy populations. The accurate comparison between the two can also provide us clues about the presence of any unidentified populations.

After nearly two decades of efforts, a major part of the CIB has been resolved into single sources (Casey et al. 2014). However, large uncertainties still exist in the submm/mm bands. First, the accuracy of direct CIB measurement in the original FIRAS spectrum is decreasing toward longer wavelengths, as one can see in Fig. 16. Secondly, the integration based on galaxy number counts also suffers from large uncertainties due to the poorly constrained slope at faint flux densities and also cosmic variance uncertainties. Because of this, different works report quite different fractions of resolved CIB, both at 870 μm (e.g. Smail et al. 2002; Chen et al. 2013; Oteo et al. 2016; Béthermin et al. 2020) and 1.2 mm (e.g. Fujimoto et al. 2016; González-López et al. 2020).

To make optimal use of the existing data and to minimize their biases, we choose to derive the resolved CIB using our joint best-fitting model. Two values are calculated: the resolved intensity and the extrapolated total intensity. The resolved intensity is the integration of the best-fitting differential number counts within the flux density range that has been covered by existing surveys. The extrapolated total intensity is based on the extrapolation of the best-fitting differential number counts to include both faint and bright sources that are beyond the detection limits of current surveys. It should be noted that the extrapolation can introduce large uncertainties if we adopt incorrect models. The deep ALMA blind surveys ASPECS and the ALMA Frontier Fields Survey at 1.2 mm (Muñoz Arancibia et al. 2018; González-López et al. 2020) suggested that the number counts flatten out at the very faint flux density end ($S_{1.2\,\mathrm{mm}} < 0.1$ mJy), implying that the counts can be best described by a Schechter function. We thus made the predictions based on our best-fitting Schechter functions, using the theoretical integration formula:

$$\int_0^\infty N(>S)\, S\, \mathrm{d}S = N_0 \times S_0 \times \Gamma(\alpha + 2). \tag{6}$$

The resolved CIB and the predictions at 2 mm, 1.2 mm, and 870 μm are summarized in Table 4.

In Fig. 16, we compare our results with the direct measurements from *COBE*/FIRAS and the most recent recalibrated *Planck*/HFI data (Odegard et al. 2019). Both the resolved intensities and the predictions from ALMA are shown.

Our results are consistent with the literature and our predictions are close to the direct measurements from FIRAS. At 870 μm, the resolved CIB for $S_{870\,\mu\mathrm{m}} > 0.2$ mJy is 18.3 ± 4.7 Jy deg^{-2}, which matches the resolved intensity of 16.4 ± 2.7 Jy deg^{-2} from the secure sample of Béthermin et al. (2020). Compared with direct measurements from FIRAS, the resolved intensity at 870 μm accounts for ∼44 per cent of the total CIB. At 1.2 mm, our resolved intensity for $S_{1.2\,\mathrm{mm}} > 45$ μJy is 11.2 ± 2.2 Jy deg^{-2} and the extrapolated DSFGs' contribution to the CIB is 14.1 ± 2.9 Jy deg^{-2}. Both of these values are slightly larger than the measurement of ASPECS (González-López et al. 2020), but closer to the direct measurements. At 2 mm, the resolved CIB for $S_{2\mathrm{mm}} > 0.1$ mJy is 2.4 ± 1.7 Jy deg^{-2}, which is close to half of the direct measurement. However, since large

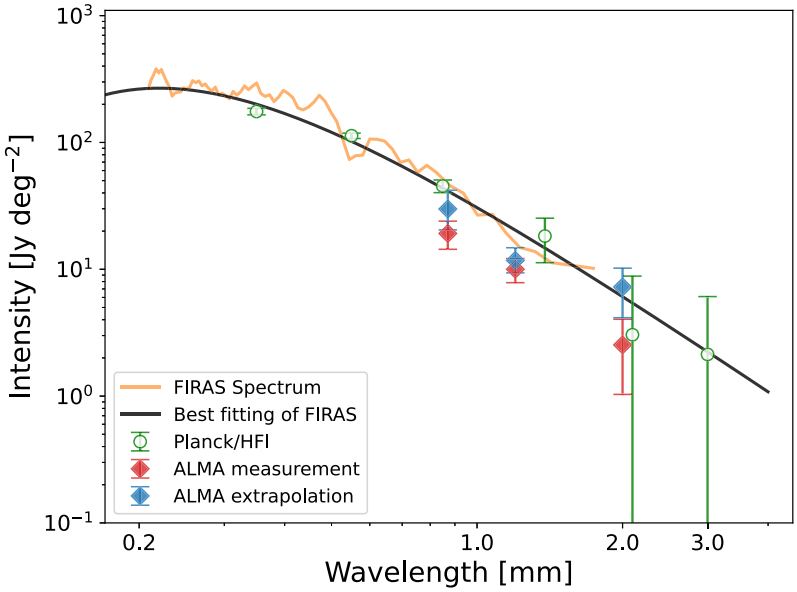

Figure 16. ALMA resolved cosmic infrared background. The resolved CIB at different ALMA bands is the integration of joint Schechter fitted differential number counts within the observed flux density ranges. The predicted CIB is the whole integration of the best-fitting differential number counts. The direct measurement from FIRAS on *COBE* is shown in orange. The black curve is the best-fitting FIRAS spectrum made by Fixsen et al. (1998). We also show the recalibrated *Planck*/HFI data (Odegard et al. 2019). Comparing with the direct measurements from FIRAS, ALMA currently resolved nearly one half of the CIB from 870 μm to 2 mm, but the fraction is highly uncertain due to the large uncertainties of the FIRAS spectrum at the submm/mm wavelengths.

Table 4. ALMA resolved cosmic infrared background.

Wavelength mm	Resolved intensity Jy deg^{-2}	Extrapolated intensity Jy deg^{-2}	Resolved fraction[a] Observed	Resolved fraction[b] Extrapolation
2000	2.4 ± 1.7	7.0 ± 3.6	$0.40^{+0.23}_{-0.11}$	$0.36^{+0.27}_{-0.25}$
1200	11.2 ± 2.2	14.1 ± 2.9	$0.54^{+0.32}_{-0.15}$	$0.79^{+0.18}_{-0.15}$
870	18.3 ± 4.7	27.0 ± 7.8	$0.44^{+0.26}_{-0.12}$	$0.67^{+0.17}_{-0.17}$

Note. The resolved intensity is the integration within the flux density range that is covered by ALMA; [a]the resolved fraction compared with the FIRAS measurement; [b]the resolved fraction compared with the prediction of this work.

uncertainties also exist in the FIRAS measurement at this wavelength, the comparison of the calculation of the resolved fraction is not very accurate. More recently, Odegard et al. (2019) have improved the accuracy of the CIB direct measurement with Planck/HFI data, but the uncertainties at the longer wavelengths have not significantly improved. Therefore, to have a more precise comparison, we still need more accurate direct measurements of the total CIB at submm/mm wavelengths. Comparing our predictions based on extrapolations, we have resolved 36^{+27}_{-25}, 79^{+18}_{-15}, and 67^{+17}_{-17} per cent of the CIB at 2 mm, 1.2 mm, and 870 μm, respectively.

6 CONCLUSIONS

ALMACAL is a survey that is exploiting calibration data that are accumulating 'for free' with every scheduled ALMA observing project. Before the observatory shutdown due to Covid-19 in 2020 March, ALMACAL had accumulated more than 1000 h of observing time and covered 1001 calibrator fields. The sensitivity, total sky coverage, and wide frequency sampling make ALMACAL a promising data set to undertake blind surveys for DSFGs. Within the ALMACAL footprints, we detect 371 sources, including 186 DSFGs confirmed by their spectral indices. We report the number counts based on these DSFGs for ALMA band 3 to band 7 (wavelengths spanning from 3 mm to 870 μm), which are mostly in agreement with existing surveys at overlapping flux densities. In band 4 (2 mm) and 7 (870 μm), ALMACAL represents the deepest survey available to this date. In band 5 (1.5 mm), ALMACAL is the first survey to be able to constrain the number counts of DSFGs. Together with the previously reported band 8 results from Klitsch et al. (2020), we are now able to present number counts covering almost the entire wavelength range covered by ALMA, from 0.65 to 3 mm.

We compare our number counts with various model predictions. The semi-analytic models from Lagos et al. (2020) and semi-empirical models from Béthermin et al. (2017) and Popping et al. (2020) match the number counts in ALMA bands reasonably well. Hydrodynamic simulations, although finding it generally harder to reproduce the number counts of DSFGs, have shown promising results from the recent efforts (e.g. McAlpine et al. 2019; Lovell et al. 2021). As demonstrated in this work, multiwavelength number counts can be a powerful benchmark to validate galaxy formation and evolution models. Future surveys, including the ongoing ALMACAL, will continue to improve the DSFG number counts by covering a larger sample and a wider flux density range.

We also turn ALMACAL into a cosmological survey to constrain the energy budget from DSFGs to the cosmic infrared background. We provide joint fits for number counts of DSFGs by combining ALMACAL with literature ALMA surveys. Compared with the direct measurements from FIRAS/*COBE* (Fixsen et al. 1998) and

HFI/*Planck*, we report ALMA has directly resolved 40^{+23}_{-11}, 54^{+32}_{-15}, and 44^{+26}_{-12} per cent of CIB at 2 mm, 1.2 mm, and 870 μm, respectively. Due to the large uncertainties of direct measurements at submm/mm wavebands, we still suffer from large errors, thus demanding more accurate direct CIB measurements.

The large number of detections in ALMACAL suggests the promising opportunity to conduct a large submm/mm survey with ALMA calibration data. We only explore the continuum image of these calibrator fields, more treasures are still buried in this data base. In addition, ALMACAL is just a small part of the ALMA archive data base. How to make use of the enormous ALMA archive effectively is challenging, but also a promising way to find more hidden gold.

ACKNOWLEDGEMENTS

We greatly thank U. Dudzevičiūté for providing the SEDs of the AS2UDS sample and Christopher C. Lovell for providing the simulations results from SIMBA. This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2094–390783311. IRS acknowlege the support from STFC (ST/T000244/1). AK gratefully acknowledges support from the Independent Research Fund Denmark via grant number DFF 8021-00130. CL has received funding from the ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

This research made use of: `NumPy`, a fundamental package for scientific computing with Python; `Matplotlib`, a plotting library for Python; `Astropy`, a community-developed core Python package for astronomy; `IPython`, an interactive computing system for Python; `photutils`, an affiliated package of `AstroPy` to provide tools for detecting and performing photometry of astronomical sources (Bradley et al. 2020).

DATA AVAILABILITY

All the data used in the work is published/available from ALMA science archive https://almascience.eso.org/alma-data. The catalogue is available on https://almacal.wordpress.com and https://github.com/cjhang/almacal.

The code used for this paper is also available online: https://github.com/cjhang/almacal.

REFERENCES

Alexander D. M., Bauer F. E., Chapman S. C., Smail I., Blain A. W., Brandt W. N., Ivison R. J., 2005, ApJ, 632, 736
Aravena M. et al., 2020, ApJ, 901, 79
Barger A. J., Cowie L. L., Sanders D. B., Fulton E., Taniguchi Y., Sato Y., Kawara K., Okuda H., 1998, Nature, 394, 248
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Becker R. H., White R. L., Helfand D. J., 1995, ApJ, 450, 559
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Béthermin M., De Breuck C., Sargent M., Daddi E., 2015, A&A, 576, L9
Béthermin M. et al., 2017, A&A, 607, A89
Béthermin M. et al., 2020, A&A, 643, A2
Blain A. W., Longair M. S., 1993, MNRAS, 264, 509
Blain A. W., Kneib J. P., Ivison R. J., Smail I., 1999, ApJ, 512, L87
Bonato M. et al., 2018, MNRAS, 478, 1512
Bradley L. et al., 2020, astropy/photutils: 1.0.0. Available at: https://zenodo.org/record/4044744
Brisbin D. et al., 2017, A&A, 608, A15
Camps P. et al., 2018, ApJS, 234, 20
Carniani S. et al., 2015, A&A, 584, A78
Casey C. M., Narayanan D., Cooray A., 2014, Phys. Rep., 541, 45
Casey C. M. et al., 2018, ApJ, 862, 77
Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772
Chen C.-C., Cowie L. L., Barger A. J., Casey C. M., Lee N., Sanders D. B., Wang W.-H., Williams J. P., 2013, ApJ, 776, 131
Chen C.-C. et al., 2016, ApJ, 820, 82
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Coppin K. et al., 2006, MNRAS, 372, 1621
Cowie L. L., González-López J., Barger A. J., Bauer F. E., Hsu L. Y., Wang W. H., 2018, ApJ, 865, 106
Cowley W. I., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., Lagos C. d. P., 2019, MNRAS, 487, 3082
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Dole H. et al., 2006, A&A, 451, 417
Driver S. P., Robotham A. S. G., 2010, MNRAS, 407, 2131
Dudzeviciute U. et al., 2020, MNRAS, 494, 3828
Dunlop J. S. et al., 2017, MNRAS, 466, 861
Dye S. et al., 2022, MNRAS, 510, 3734
Eales S., Lilly S., Gear W., Dunne L., Bond J. R., Hammer F., Le Fèvre O., Crampton D., 1999, ApJ, 515, 518
Eales S. et al., 2010, PASP, 122, 499
Elbaz D., Cesarsky C. J., Chanial P., Aussel H., Franceschini A., Fadda D., Chary R. R., 2002, A&A, 384, 848
Fixsen D. J., Dwek E., Mather J. C., Bennett C. L., Shafer R. A., 1998, ApJ, 508, 123
Franco M. et al., 2018, A&A, 620, A152
Fudamoto Y. et al., 2021, Nature, 597, 489
Fujimoto S., Ouchi M., Ono Y., Shibuya T., Ishigaki M., Nagai H., Momose R., 2016, ApJS, 222, 1
Geach J. E. et al., 2017, MNRAS, 465, 1789
Geach J. E., Ivison R. J., Dye S., Oteo I., 2018, ApJ, 866, L12
Gómez-Guijarro C. et al., 2022, A&A, 658, A43
González-López J. et al., 2017, A&A, 608, A138
González-López J. et al., 2019, ApJ, 882, 139
González-López J. et al., 2020, ApJ, 897, 91
Greve T. R., Ivison R. J., Bertoldi F., Stevens J. A., Dunlop J. S., Lutz D., Carilli C. L., 2004, ApJ, 354, 779
Gullberg B. et al., 2019, MNRAS, 490, 4956
Hamanowicz A. et al., 2022, MNRAS, in press
Hatsukade B., Ohta K., Seko A., Yabe K., Akiyama M., 2013, ApJ, 769, L27
Hatsukade B. et al., 2016, PASJ, 68, 36
Hatsukade B. et al., 2018, PASJ, 70, 105
Hauser M. G. et al., 1998, ApJ, 508, 25
Hayward C. C., Kereš D., Jonsson P., Narayanan D., Cox T. J., Hernquist L., 2011, ApJ, 743, 159
Hodge J. A., da Cunha E., 2020, R. Soc. Open Sci., 7, 200556
Hodge J. A. et al., 2013, ApJ, 768, 91
Högbom J. A., 1974, A&AS, 15, 417
Hogg D. W., Turner E. L., 1998, PASP, 110, 727
Hughes D. H. et al., 1998, Nature, 394, 241
Ikarashi S. et al., 2015, ApJ, 810, 133
Ivison R. J. et al., 2002, MNRAS, 337, 1
Ivison R. J. et al., 2007, MNRAS, 380, 199
Ivison R. J. et al., 2010, A&A, 518, L35
Ivison R. J. et al., 2013, ApJ, 772, 137
Juvela M., Mattila K., Lemke D., Klaas U., Leinert C., Kiss C., 2009, A&A, 500, 763
Karim A. et al., 2013, MNRAS, 432, 2
Klitsch A. et al., 2019, MNRAS, 490, 1220
Klitsch A. et al., 2020, MNRAS, 495, 2332
Lacey C. G. et al., 2016, MNRAS, 462, 3854
Lagache G., Puget J.-L., Dole H., 2005, ARA&A, 43, 727

Lagos C. d. P. et al., 2019, MNRAS, 489, 4196
Lagos C. d. P., da Cunha E., Robotham A. S. G., Obreschkow D., Valentino F., Fujimoto S., Magdis G. E., Tobar R., 2020, MNRAS, 499, 1948
Laporte N. et al., 2017, ApJ, 837, L21
Laporte N. et al., 2021, MNRAS, 505, 4838
Laurent G. T. et al., 2005, ApJ, 623, 742
Liu D. et al., 2019, ApJS, 244, 40
Lovell C. C., Geach J. E., Davé R., Narayanan D., Li Q., 2021, MNRAS, 502, 772
McAlpine S. et al., 2019, MNRAS, 488, 2440
McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, ASP Conf. Ser. Vol. 376, Astronomical Data Analysis Software and Systems XVI. Astron. Soc. Pac., San Francisco, p. 127
Magnelli B. et al., 2019, ApJ, 877, 45
Matsuura S. et al., 2011, ApJ, 737, 2
Meyer R. A. et al., 2022, ApJ, 927, 141
Miettinen O. et al., 2017, A&A, 606, A17
Miller T. B. et al., 2018, Nature, 556, 469
Muñoz Arancibia A. M. et al., 2018, A&A, 620, A125
Muñoz Arancibia A. M. et al., 2019, A&A, 631, C2
Odegard N., Weiland J. L., Fixsen D. J., Chuss D. T., Dwek E., Kogut A., Switzer E. R., 2019, ApJ, 877, 40
Oliver S. J. et al., 2010, A&A, 518, L21
Ono Y., Ouchi M., Kurono Y., Momose R., 2014, ApJ, 795, 5
Oteo I., Zwaan M. A., Ivison R. J., Smail I., Biggs A. D., 2016, ApJ, 822, 36
Oteo I., Zwaan M. A., Ivison R. J., Smail I., Biggs A. D., 2017, ApJ, 837, 182
Oteo I. et al., 2018, ApJ, 856, 72
Planck Collaboration XVI, 2014, A&A, 571, A30
Planck Collaboration VI, 2020, A&A, 641, A6
Popping G. et al., 2020, ApJ, 891, 135
Puget J.-L., Abergel A., Bernard J.-P., Boulanger F., Burton W. B., Desert F.-X., Hartmann D., 1996, A&A, 308, L5
Rizzo F., Vegetti S., Powell D., Fraternali F., McKean J. P., Stacey H. R., White S. D. M., 2020, Nature, 584, 201
Sanders D. B., Mirabel I. F., 1996, ARA&A, 34, 749
Schaye J. et al., 2015, MNRAS, 446, 521
Schechter P., 1976, ApJ, 203, 297
Scott K. S. et al., 2012, MNRAS, 423, 575
Shimizu I., Yoshida N., Okamoto T., 2012, MNRAS, 427, 2866
Simpson J. M. et al., 2015a, ApJ, 799, 81
Simpson J. M. et al., 2015b, ApJ, 807, 128
Simpson J. M. et al., 2017, ApJ, 839, 58
Simpson J. M. et al., 2020, MNRAS, 495, 3409
Smail I., Ivison R. J., Blain A. W., 1997, ApJ, 490, L5
Smail I., Ivison R. J., Blain A. W., Kneib J. P., 2002, MNRAS, 331, 495
Smail I. et al., 2021, MNRAS, 502, 3426
Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, MNRAS, 423, 1992
Stach S. M. et al., 2018, ApJ, 860, 161
Stach S. M. et al., 2019, MNRAS, 487, 4648
Staguhn J. G. et al., 2014, ApJ, 790, 77
Swinbank A. M. et al., 2010, Nature, 464, 733
Tamura Y. et al., 2019, ApJ, 874, 27
Trayford J. W., Lagos C. d. P., Robotham A. S. G., Obreschkow D., 2020, MNRAS, 491, 3937
Umehata H. et al., 2017, ApJ, 835, 98
Umehata H. et al., 2018, PASJ, 70, 65
Walter F. et al., 2016, ApJ, 833, 67
Wang W.-H., Cowie L. L., Barger A. J., Williams J. P., 2011, ApJ, 726, L18
Weiß A. et al., 2009, ApJ, 707, 1201
Weiß A. et al., 2013, ApJ, 767, 88
Zavala J. A., Casey C. M., da Cunha E., Spilker J., Staguhn J., Hodge J., Drew P. M., 2018, ApJ, 869, 71
Zavala J. A. et al., 2021, ApJ, 909, 165
Zwaan M. et al., 2022, The Messenger , 186, 10

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Appendix A. Flux boosting for sources with different sizes.
Appendix B. Radial distance and flux densities of all the detections.
Appendix C. Number counts in the inner and outer parts of the fields.
Appendix D. Robustness of number counts.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.