



**HAL**  
open science

## **A3COSMOS & A3GOODSS: Continuum Source Catalogues and Multi-band Number Counts**

Sylvia Adscheid, Benjamin Magnelli, Daizhong Liu, Frank Bertoldi, Ivan Delvecchio, Carlotta Gruppioni, Eva Schinnerer, Alberto Traina, Matthieu Béthermin, Athanasia Gkogkou

► **To cite this version:**

Sylvia Adscheid, Benjamin Magnelli, Daizhong Liu, Frank Bertoldi, Ivan Delvecchio, et al.. A3COSMOS & A3GOODSS: Continuum Source Catalogues and Multi-band Number Counts. Astronomy and Astrophysics - A&A, 2024, 685, pp.A1. 10.1051/0004-6361/202348407. hal-04523172

**HAL Id: hal-04523172**

**<https://hal.science/hal-04523172v1>**

Submitted on 16 May 2024







**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS: Continuum source catalogues and multi-band number counts<sup>★</sup>

Sylvia Adscheid<sup>1</sup>, Benjamin Magnelli<sup>2</sup>, Daizhong Liu<sup>3</sup>, Frank Bertoldi<sup>1</sup>, Ivan Delvecchio<sup>4</sup>, Carlotta Gruppioni<sup>5</sup>, Eva Schinnerer<sup>6</sup>, Alberto Traina<sup>5,7</sup>, Matthieu Béthermin<sup>8,9</sup>, and Athanasia Gkogkou<sup>8</sup>

<sup>1</sup> Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany  
e-mail: [sadscheid@astro.uni-bonn.de](mailto:sadscheid@astro.uni-bonn.de)

<sup>2</sup> Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191 Gif-sur-Yvette, France

<sup>3</sup> Max-Planck-Institut für extraterrestrische Physik, Gießenbachstraße 1, 85748 Garching b. München, Germany

<sup>4</sup> INAF – Osservatorio Astronomico di Brera 28, 20121, Milano, Italy and Via Bianchi 46, 23807 Merate, Italy

<sup>5</sup> Dipartimento di Fisica e Astronomia (DIFA), Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy

<sup>6</sup> Max Planck Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

<sup>7</sup> INAF – Osservatorio di Astrofisica e Scienza dello Spazio (OAS), Via Gobetti 101, 40129 Bologna, Italy

<sup>8</sup> Aix-Marseille Univ., CNRS, CNES, LAM, Marseille, France

<sup>9</sup> Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France

Received 27 October 2023 / Accepted 25 January 2024

## ABSTRACT

**Context.** Galaxy submillimetre number counts are a fundamental measurement in our understanding of galaxy evolution models. Most early measurements are obtained via single-dish telescopes with substantial source confusion, whereas recent interferometric observations are limited to small areas.

**Aims.** We used a large database of ALMA continuum observations to accurately measure galaxy number counts in multiple (sub)millimetre bands, thus bridging the flux density range between single-dish surveys and deep interferometric studies.

**Methods.** We continued the Automated Mining of the ALMA Archive in the COSMOS Field project (A<sup>3</sup>COSMOS) and extended it with observations from the GOODS-South field (A<sup>3</sup>GOODSS). The database consists of ~4000 pipeline-processed continuum images from the public ALMA archive, yielding 2050 unique detected sources, including sources with and without a known optical counterpart. To infer galaxy number counts, we constructed a method to reduce the observational bias inherent to targeted pointings that dominate the database. This method comprises a combination of image selection, masking, and source weighting. The effective area was calculated by accounting for inhomogeneous wavelengths, sensitivities, and resolutions and for the spatial overlap between images. We tested and calibrated our method with simulations.

**Results.** We derived the number counts in a consistent and homogeneous way in four different ALMA bands covering a relatively large area. The results are consistent with number counts retrieved from the literature within the uncertainties. In Band 7, at the depth of the inferred number counts, ~40% of the cosmic infrared background is resolved into discrete sources. This fraction, however, decreases with increasing wavelength, reaching ~4% in Band 3. Finally, we used the number counts to test models of dusty galaxy evolution, and find a good agreement within the uncertainties.

**Conclusions.** By continuing the A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS archival effort, we obtained the deepest archive-based (sub)millimetre number counts measured to date over such a wide area. This database proves to be a valuable resource that, thanks to its substantial size, can be used for statistical analyses after having applied certain conservative restrictions.

**Key words.** galaxies: abundances – galaxies: high-redshift – submillimeter: galaxies

## 1. Introduction

A central aspect in the study of galaxy evolution is the understanding of the cosmic star formation history. The star formation rate (SFR) of galaxies is commonly traced by emission in the rest-frame ultraviolet (UV) and/or far-infrared (far-IR). The presence of short-lived massive stars emitting strongly at UV wavelengths indicates recent star formation activity; part of this UV radiation is absorbed by dust and re-emitted in thermal IR. Galaxy studies in UV to IR wavelengths have shown

that the cosmic SFR density was much higher in the past (~10× higher at its peak at  $z \approx 2$ ), with most of star formation activity being obscured by dust (~80%; see [Madau & Dickinson 2014](#), for a review). Dusty star forming galaxies (DSFGs) contribute significantly to the cosmic SFR density; they are bright in the submillimetre regime and faint or even undetected in rest-frame UV to optical (e.g. [Casey et al. 2014](#); [Wang et al. 2019](#); [Talia et al. 2021](#)). The number and flux density distribution of DSFGs results from the complex history of gas accretion, star formation, and dust production and their underlying physical mechanisms. The number density of galaxies above a given flux density threshold per unit effective area, commonly referred to as number count, is therefore a fairly simple

<sup>★</sup> Data products are available at the CDS via anonymous ftp to [cdsarc.cds.unistra.fr](https://cdsarc.cds.unistra.fr) (130.79.128.5) or via <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/685/A1>

measure, but a very useful tool for constraining galaxy evolution models. A plethora of single-dish surveys at wavelengths  $\geq 450 \mu\text{m}$  have been conducted over the last two decades to reveal the nature and properties of DSFGs, for example the SCUBA Lens Survey (Smail et al. 2002), the LESS survey (Weiß et al. 2009), the Lockman Hole North 1.2 mm survey (Lindner et al. 2011), six 1.1 mm AzTEC blank-fields (Scott et al. 2012), the SCUBA-2 COSMOS survey (Casey et al. 2013), the ACT Southern survey (Marsden et al. 2014), the GISMO 2 mm survey (Magnelli et al. 2019), S2COSMOS (Simpson et al. 2019), STUDIES (Wang et al. 2017; Dudzevičiūtė et al. 2021), and the N2CLS survey (Bing et al. 2023). These single-dish surveys have already set valuable constraints on the (sub)millimetre number counts. However, due to their large beam sizes, they are limited by source confusion to the brightest DSFGs. This leaves a large part of the population of DSFGs and the lower flux density regimes of models largely unconstrained (e.g. Béthermin et al. 2012, 2017; Casey et al. 2018a; Popping et al. 2020).

The Atacama Large Millimeter/submillimeter Array (ALMA) can help to alleviate these issues, being one of the highest-sensitivity instruments currently operating in the (sub)millimetre regime; it achieves resolutions that far exceed the capabilities of single-dish telescopes at these wavelengths due to its interferometric nature. However, due to its small field of view, deep blind surveys with ALMA are very time-consuming and hardly viable for very extended fields. Though some blind surveys have been performed with ALMA, such as ASAGAO (Hatsukade et al. 2018), ASPECS-LP (González-López et al. 2019, 2020), MORA (Zavala et al. 2021; Casey et al. 2021), and the GOODS-ALMA survey (Franco et al. 2018; Gómez-Guijarro et al. 2022), they are small in size compared to single-dish surveys (e.g.  $\sim 70 \text{ arcmin}^2$  for GOODS-ALMA vs.  $\sim 2 \text{ deg}^2$  for S2COSMOS). Instead, ALMA projects are more often designed as follow-ups to galaxy samples from larger single-dish programmes (e.g. Karim et al. 2013; Stach et al. 2019; Simpson et al. 2020). Therefore, they are focussed on the brighter, more starburst galaxies, neglecting the bulk of the DSFG population.

The aim of the A<sup>3</sup>COSMOS project<sup>1</sup> (Liu et al. 2019a,b) is to aggregate and homogeneously process all public ALMA archival data in the  $\sim 2 \text{ deg}^2$  Cosmic Evolution Survey (COSMOS; Scoville et al. 2007) in order to provide to the community homogeneously processed images<sup>2</sup> and catalogues of prior-based and blind (sub)millimetre photometry and coherently derived galaxy properties. Among other blind deep fields (e.g. GOODS-North/South, UDS, EGS), COSMOS has the largest HST/ACS contiguous coverage (Scoville et al. 2007) and the largest JWST/NIRCam ( $0.54 \text{ deg}^2$ ) and MIRI ( $0.19 \text{ deg}^2$ ) coverage (COSMOS-Web; Casey et al. 2023). Due to this rich multi-wavelength coverage and the legacy status of COSMOS, the ALMA coverage in COSMOS, and therefore the A<sup>3</sup>COSMOS database, is extensive and continuously growing. The A<sup>3</sup>COSMOS database is naturally dominated by single pointings and their respective targets, but these pointings are usually far more extended than the angular size of their targets, which also yields an abundance of serendipitous detections. This facilitates the study of galaxy number counts utilising the large sky coverage of A<sup>3</sup>COSMOS.

In this paper, we present the latest data version of A<sup>3</sup>COSMOS, for which we included another well-studied extragalactic field: the Great Observatories Origins Deep Survey Southern field (GOODS-S; Dickinson et al. 2003). Although the areal size of this field is smaller than COSMOS, originally covering  $\sim 160 \text{ arcmin}^2$ , it offers rich ancillary data from a number of deep survey programmes, for instance hosting the *Hubble* Ultra Deep Field (HUDF; Beckwith et al. 2006). The ancillary data available in GOODS-S is generally deeper than in COSMOS, which facilitates the study of fainter and higher-redshift galaxies. We then use the combined database to homogeneously derive galaxy number counts in several ALMA bands.

The structure of this paper is as follows. In Sect. 2 we describe the updated A<sup>3</sup>COSMOS database and additions made to the processing pipeline, including the GOODS-S field (i.e. A<sup>3</sup>GOODSS). In Sect. 3 we use the combined database to infer number counts in multiple ALMA bands. We describe the calculation of the effective area and corrections made to reduce the observational biases. These corrections are then tested using simulations. The results are discussed in Sect. 4, and summarised in Sect. 5.

In the following we assume a flat  $\Lambda$ CDM cosmology with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_M = 0.3$ , and a Chabrier (2003) initial mass function.

## 2. The A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS database

The goal of A<sup>3</sup>COSMOS is to assemble a large sample of galaxies detected at (sub)millimetre wavelengths and probing a wide redshift range ( $z \approx 0-6$ ). This is done by using a pipeline to retrieve publicly available observational data from the ALMA archive and homogeneously processing them into continuum images from which sources are then extracted. Galaxy properties (e.g. SFR, stellar mass, dust luminosity) are inferred by fitting the spectral energy distributions (SEDs) of the extracted galaxies, making use of both the extracted ALMA flux densities and ancillary data (see Sects. 2.2 and 2.3).

For the new data version (20220606), we extended the database to ALMA observations in GOODS-S. This not only improves the statistical properties of our sample by including more galaxies, but it also offers the opportunity to expand our sample to fainter objects due to the availability of deeper ancillary photometric and redshift information, even though the areal size of the GOODS-S field is much smaller than that of COSMOS.

### 2.1. The A<sup>3</sup>COSMOS pipeline

For a full in-detail description of the entire pipeline of data retrieval, processing, source extraction and SED fitting, we refer the reader to the original work by Liu et al. (2019a). Here, we give a brief summary of the most important steps. The photometric data used in the pipeline, as well as any updates and additions to those with respect to the previous data version, are listed separately in Sects. 2.2 and 2.3.

The archive query and download were done via the Python package *astroquery* (Ginsburg et al. 2019). Calibration and creation of continuum images from the raw data was done using the Common Astronomy Software Application (CASA; McMullin et al. 2007) in the recommended version for each individual ALMA project. For calibration, the *scriptForPI.py* provided by the ALMA Observatory was used. Imaging was performed in automatic pipeline mode, using Briggs-weighting with

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

<sup>2</sup> <https://irsa.ipac.caltech.edu/data/COSMOS/images/a3cosmos/>

`robust=2.0`, and choosing `specmode="cont"`. The continuum images were masked outside of a primary beam attenuation of 0.2. Source extraction from the produced images was then done in two separate ways: “blind” extraction was performed using the Python Blob Detector and Source Finder (PyBDSF; Mohan & Rafferty 2015); “prior” extraction was performed with our A<sup>3</sup>COSMOS prior extraction pipeline (see Sect. 2.4 in Liu et al. 2019a) that iteratively executes GALFIT (Peng et al. 2002, 2010) using positional priors of known sources from a previously assembled master catalogue. Sources are fit simultaneously in GALFIT to allow for a dissection of sources potentially blended in blind extraction. Fortunately, due to the high resolution of most ALMA observations, source blending plays only a minor role, affecting only ~2.3% of the sources in our blind catalogue, most of which are the targets of their respective observation. These blended sources are indicated by a flag in our blind source catalogue. We note that there are no blended sources in the source sample of our number counts analysis in Sect. 3.

The master catalogue used for prior extraction is a compilation of multiple multi-wavelength catalogues covering the COSMOS and GOODS-S fields (see Sects. 2.2 and 2.3, respectively), spatially cross-matched with a matching radius of 1 arcsec to ensure that all entries are unique sources. This radius corresponds to a false-match probability of 13.3% between the COSMOS2015 catalogue from Laigle et al. (2016) and other catalogues (see Sect. 2.3 in Liu et al. 2019a). To avoid a high contamination with spurious detections in both source extraction methods, a minimum peak signal-to-noise ( $S/N_{\text{peak}}$ ) threshold was applied. This threshold was set to 5.40 for PyBDSF (blind) and 4.35 for GALFIT (prior) extraction, as to limit the cumulative spurious fraction to  $\lesssim 8\%$  and  $\lesssim 12\%$  (see Sects. 2.8 and 3.2 in Liu et al. 2019a). Sources with  $S/N_{\text{peak}}$  below these respective thresholds were discarded for further analysis (but kept in our released photometry catalogues)<sup>3</sup>. The blind and prior source catalogues were then spatially cross-matched, with a matching radius of 1 arcsec, to create a combined sample of unique sources, retaining the information from both prior and blind extraction for sources retrieved through both methods. The ALMA extracted continuum flux densities of the sources in this combined sample were then combined with the available ancillary photometry and photometric and/or spectroscopic redshift information of their respective priors (naturally, except for sources unique to the blind catalogue).

To assess the reliability of the master catalogue priors as counterparts for the extracted ALMA sources (which we call counterpart association or “CPA”), we defined a number of parameters, such as total flux density signal-to-noise, spatial separation and source extension, and measured them at the prior and extracted source coordinates in both the ALMA images and different optical and near-infrared counterpart images (listed in Sects. 2.2 and 2.3). Based on these parameters, counterparts were classified as reliable or unreliable using a combination of visual inspection and machine learning. This process is described in detail in Sect. 4.2 of Liu et al. (2019a) and summarised in Appendix A. Sources with unreliable counterparts, called “CPA discarded”, were excluded from further analysis (~9% in total, see Table 1).

<sup>3</sup> These catalogues also provide information on the respective image RMS noise and beam size of each (non-)detection, which users can utilise to infer upper limits on the flux densities of sources below the detection threshold in the prior catalogue (see also Sect. 3.2.2).

To infer physical galaxy properties from the sample of ALMA sources with reliable optical-to-near-IR counterparts, we used the MAGPHYS SED fitting package with its high-redshift library update (da Cunha et al. 2008, 2015), which is applicable to both low- and high-redshift galaxies. This package finds the best fit SED by comparing the given photometry with a comprehensive library of SED models via a reduced  $\chi^2$  approach. If several redshifts were available in our master catalogue for one ALMA source, the one yielding the lowest  $\chi^2$  was chosen. In the process, the potential contamination of the ALMA continuum flux densities from line emission was assessed and subtracted using the information on the spectral set-up of the ALMA observations and empirical IR-to-line luminosity correlations for the brightest (sub)millimetre lines, that is, [C I] (Liu et al. 2015), [C II] (De Looze et al. 2011), [N II] (Zhao et al. 2016), and CO (Sargent et al. 2014; Liu et al. 2015). Since MAGPHYS does not include an AGN component, strong mid-IR contribution from AGNs can cause an overestimation of inferred stellar masses and SFRs. Liu et al. (2019a) showed that by running MAGPHYS iteratively, once with and once without considering any 24  $\mu\text{m}$  flux density information, and choosing the fit yielding the lowest  $\chi^2$ , the inferred stellar masses are in a good agreement with results from Delvecchio et al. (2017) obtained from SED fitting including an AGN component (see Sect. 4.6 in Liu et al. 2019a). Therefore, MAGPHYS was run twice, one time disregarding any potential mid-IR flux density information from the *Spitzer*/MIPS 24  $\mu\text{m}$  filter, and the better fit was chosen. The flux densities at all wavelengths of the extracted ALMA sources ( $S_{\text{OBS}}$ ) were compared to those inferred from the best fit SEDs ( $S_{\text{SED}}$ ) by plotting the distribution of the observed-to-predicted flux density ratios for all sources. Fitting a 1D Gaussian function to this distribution, we found a mean of  $\mu = 1$  and a scatter of  $\sigma \approx 0.05$  dex. Then, sources were discarded if their observed-to-predicted flux density ratio at any wavelength differed by more than  $5\sigma$  (~0.25 dex) from a ratio of one (called “SED outliers”, see Sect. 4.4 in Liu et al. 2019a, also Traina et al. 2024). This deviation likely indicates a mismatch between the measured ALMA flux densities and the ancillary photometry, hence implying a chance association of sources (~7% of sources, see Table 1). We note that sources can be classified simultaneously as both CPA discarded and SED outliers, which, however, does not apply to all cases. We also investigated the use of the distribution in  $(S_{\text{OBS}} - S_{\text{SED}})/\sigma_{\text{OBS}}$  as a measure of fit accuracy, which incorporates the uncertainty  $\sigma_{\text{OBS}}$  on the observed flux density. However, in addition to the uncertainty on the observed photometry, the quality of our fits is also determined by other factors, such as uncertainties from the ancillary photometry, fluctuations in ALMA as well as ancillary zero-point calibration, and MAGPHYS fitting limitations. Through a visual assessment, we found the simple flux ratio to overall yield the best balance between all those factors and thus to be the best method to identify obvious outliers.

From the best-fit SED, we obtained galaxy properties, such as stellar mass, dust mass, SFRs averaged over the last 100 Myr of their star-formation history, dust temperature, and total infrared luminosity. From the total infrared luminosity of these galaxies, we also inferred empirically derived SFRs using the calibration from Kennicutt (1998) and a Chabrier (2003) initial mass function. Finally, molecular gas masses were inferred using the rest-frame 850  $\mu\text{m}$  flux density of these galaxies and the Rayleigh–Jeans continuum method applying the calibration of Hughes et al. (2017). This calibration was chosen in accordance with Liu et al. (2019b), as it uses a luminosity-dependent conversion factor, as opposed to the commonly used calibration



from Scoville et al. (2017) which uses a single conversion factor. We note, however, that the Hughes et al. (2017) calibration yields systematically lower molecular gas masses (by  $\sim 0.1$  dex) than the one from Scoville et al. (2017), as found in a direct comparison carried out by Liu et al. (2019b, see their Sect. 3).

We note that some ALMA projects contain several coverages of the same pointing or mosaic. While combining the individual images in the  $uv$ -plane would result in a single deeper continuum image, we chose to instead pipeline-process all images homogeneously and thus separately. Hence, one source can have several flux density estimates at the same wavelength. This has no importance for our SED fit, as both information are mathematically combined via the  $\chi^2$ . But it also means that sources lying slightly below the noise thresholds in those images could possibly have been detected in the combined image. This is a current limitation of the pipeline, which we plan to improve in future versions. Fortunately, for some wide and/or deep surveys in the COSMOS and GOODS-S fields, the already combined continuum maps are publicly available. Performing source extraction on these maps (instead of the individual coverages from our pipeline) allows us to circumvent the depth limitation, at least in these specific surveys, and thus extend our galaxy number counts to lower flux densities. This is detailed in Sect. 2.4.

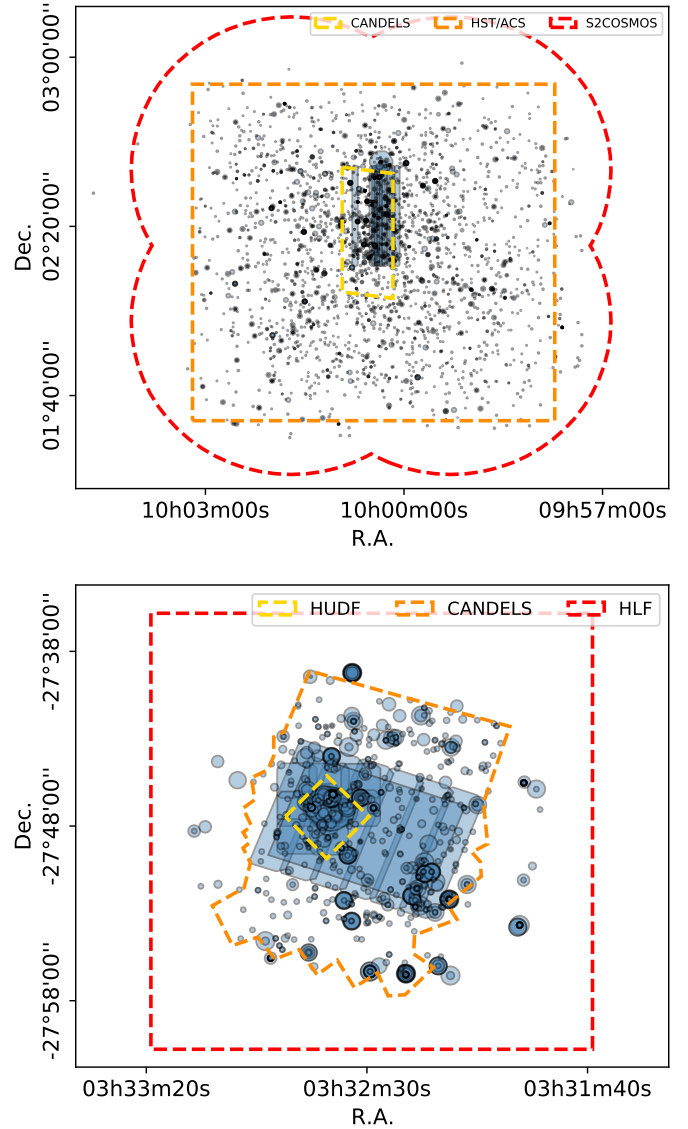
In the following, we describe the update of the database, the changes in ancillary data with respect to the previous data release, and introduce the new extension to the GOODS-S field, A<sup>3</sup>GOODSS.

## 2.2. COSMOS

We downloaded and pipeline-processed all ALMA archive data that became public up until June 6, 2022, as opposed to January 2, 2018 (version number 20180102; Liu et al. 2019a). The new images were added to the database, extending it from  $\sim 1500$  continuum images in Liu et al. (2019a) to  $\sim 3300$  from 173 individual ALMA projects. The spatial coverage of all produced continuum images is displayed in the top panel of Fig. 1. We also show the coverage in each individual ALMA band in Fig. B.1. Most pointings are located within the area of the COSMOS HST/ACS field (Koekemoer et al. 2007), covering it with nearly uniform density. We notice the existence of large mosaics from two blind surveys: MORA in Band 4 (Zavala et al. 2021; Casey et al. 2021, vertical stripes) and a hexagonal field in Band 3 (Keating et al. 2020).

The master catalogue was also updated. The individual prior catalogues used to assemble the previous master catalogue are listed in Table 2 of Liu et al. (2019a). We replaced the COSMOS2015 catalogue (Laigle et al. 2016) by the newer version COSMOS2020 (Weaver et al. 2022). Two catalogues are available for COSMOS2020: the CLASSIC catalogue and the FARMER catalogue, which use two different photometry methods. Since the two catalogues differ slightly in their spatial sky coverage, we included both. Additionally, two more prior catalogues were included: the *Spitzer*/IRAC catalogue from Ashby et al. (2018) and the *Spitzer*/24  $\mu$ m catalogue from Le Floc'h et al. (2009), which respectively add another 85 466 and 804 priors not contained in any of the other catalogues. With this updated master catalogue, we were able to identify a counterpart for  $\sim 93\%$  of our extracted ALMA sources in COSMOS (see Table 1), which is a notable improvement over the previous data version ( $\sim 74\%$  in Liu et al. 2019a).

The ancillary photometry of each ALMA source was taken from each of the catalogues that constitute the master catalogue, complemented by additional deblended mid- to far-IR



**Fig. 1.** Spatial sky coverage of all successfully imaged ALMA archival data available for COSMOS (top panel) and GOODS-S (bottom panel) as of June 6, 2022. The blue shaded shapes are outlines of single pointings and blind surveys; the darker shades of blue indicate an overlap of multiple images. For single pointings, the circle size corresponds to the area where the primary beam attenuation is  $>0.5$ . The outlines of blind surveys have been traced manually. For easier visibility the outlines of surveys are only drawn once, even if multiple coverages of the same survey exist in our database. The dashed coloured lines indicate the approximate outlines of different ancillary survey fields. In COSMOS: CANDELS (yellow; Grogin et al. 2011), COSMOS HST/ACS imaging (orange; Koekemoer et al. 2007), and S2COSMOS (red; Simpson et al. 2019). In GOODS-S: *Hubble* Ultra Deep Field (yellow; Beckwith et al. 2006), CANDELS (orange; as in Guo et al. 2013), and *Hubble* Legacy Fields (red; Illingworth et al. 2016).

photometry from Jin et al. (2018). Photometric redshifts were taken from the COSMOS2020 catalogue (Weaver et al. 2022; Salvato et al. 2011; Davidzon et al. 2017; Delvecchio et al. 2017; Jin et al. 2018). Spectroscopic redshifts were adopted from the compilation by M. Salvato as listed in Sect. 2 of Liu et al. (2019a), as well as supplementary spectroscopic redshifts from Riechers et al. (2014), Capak et al. (2015), Smolčić et al. (2015), Brisbin et al. (2017), Lee et al. (2017), and Pavesi et al. (2018).

**Table 1.** Properties of the A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS databases.

Database	COSMOS	GOODS-S
ALMA projects	173	74
Continuum images	3232	723
Blind survey images	14	49
Spatial coverage [arcmin <sup>2</sup> ]		
Band 3	427.7 (205.6)	101.0 (77.4)
Band 4	224.8 (205.2)	44.5 (21.2)
Band 5	2.6 (1.1)	1.9 (1.0)
Band 6	289.3 (127.9)	90.7 (86.8)
Band 7	183.3 (81.0)	38.5 (17.8)
Band 8	3.3 (1.4)	0.1 (0.05)
Band 9	0.08 (0.04)	0.5 (0.2)
Combined	896.3 (519.6)	149.6 (115.9)
Photometry catalogues		
Blind catalogue	2204	491
Prior catalogue	2540	568
Unique sources		
Total	1756	294
Blind-only	131	79
With prior photometry	1625	215
With prior redshift	1514	185
SED outliers	86	37
CPA discarded	117	60
Robust Galaxy Catalogue	1335	127

**Notes.** Listed by field are the number of ALMA projects, pipeline-produced continuum images, and blind surveys; spatial coverage (taking overlapping areas into account) per ALMA band considering the area where the primary beam attenuation is  $>0.2$  (in brackets:  $>0.5$ , as displayed in Fig. 1); number of valid detections in the blind ( $S/N_{\text{peak}} > 5.4$ ) and prior ( $S/N_{\text{peak}} > 4.35$ ) catalogues; number of unique sources with prior photometry available and with prior photometric or spectroscopic redshift information (both required for SED fitting), sources discarded due to disagreement with the SED fit (SED outliers), sources discarded due to unreliable counterpart association (CPA discarded), and final number of sources in our Robust Galaxy Catalogue including galaxy properties inferred from SED fitting.

For our CPA (see Sect. 2.1 and Appendix A), we made use of the following counterpart images: HST/ACS *i*-band from COSMOS (Koekemoer et al. 2007; Massey et al. 2010), *Spitzer*/IRAC at 3.6  $\mu\text{m}$  from the Cosmic Dawn survey (Moneti et al. 2022), VISTA/VIRCAM  $K_s$ -band from the Ultra-VISTA survey (release version DR4; McCracken et al. 2012) and VLA at 3 GHz from the VLA-COSMOS 3 GHz Large Project (Smolčić et al. 2017).

### 2.3. GOODS-South

We downloaded all public ALMA archive data available in the GOODS-S field at the same time as COSMOS, that is, on June 06, 2022. The search was centred on the position RA = 53.125° and Dec = -27.806°, with a radius of 10 arcmin. Data calibration and continuum imaging was performed using the pipeline as described in Sect. 2.1, yielding  $\sim 700$  images from 74 ALMA projects. The total spatial coverage is shown in the bottom panel of Fig. 1, and the individual coverage in each ALMA band in Fig. B.2. The pointings are mostly located inside the CANDELS field (Guo et al. 2013), but with the area within the

HUDF being sampled far more densely than the rest of the field. A number of blind surveys are available, most importantly the GOODS-ALMA survey in Band 6 (Franco et al. 2018; Gómez-Guijarro et al. 2022), the ASAGAO survey in Band 6 (Hatsukade et al. 2018), the ASPECS-LP survey in Bands 3 and 6 (Decarli et al. 2019; González-López et al. 2019, 2020) and a wide rectangular field in Band 3 (PI: R. Decarli).

We assembled a new master catalogue for the GOODS-S field and its surrounding area, using positional priors and photometry in UV to IR wavelengths from the CANDELS GOODS-South catalogue from Guo et al. (2013), the 3D-HST/CANDELS programmes in GOODS-S from Skelton et al. (2014), the S-CANDELS survey in CDFS (Ashby et al. 2015), the ZFOURGE survey in CDFS (Straatman et al. 2016) and in radio wavelength from the VLA 1.4 GHz survey in E-CDFS by Miller et al. (2008). The final master catalogue contains 82 519 priors over an area of  $\sim 1300$  arcmin<sup>2</sup>, with the majority of priors concentrated in a region of  $\sim 650$  arcmin<sup>2</sup> roughly centred on the HUDF. We also included additional far-IR photometry from the PEP-GOODS-*Herschel* programme (Magnelli et al. 2013).

We adopted photometric redshifts from Skelton et al. (2014), Straatman et al. (2016), Croom et al. (2001), Wuyts et al. (2008), and Momcheva et al. (2016). Spectroscopic redshifts were taken from Le Fèvre et al. (2004, 2013), Mignoli et al. (2005), Ravikumar et al. (2007), Vanzella et al. (2008), Balestra et al. (2010), Kriek et al. (2015), Morris et al. (2015), Tasca et al. (2017), Aravena et al. (2019), and Urrutia et al. (2019).

For our CPA, we adopted counterpart images from: *Spitzer*/IRAC at 3.6  $\mu\text{m}$  and 4.5  $\mu\text{m}$  from the SEDS survey (Ashby et al. 2013), HST/ACS *F775W*-band from the GOODS survey (Giavalisco et al. 2004), HST/WFC3 *F160W*-band from the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011; Skelton et al. 2014), and CFHT/WIRCam  $K_s$ -band from the TENIS survey (Hsieh et al. 2012).

### 2.4. Adding combined survey mosaics

Combined continuum maps are available for the following ALMA surveys: GOODS-ALMA (priv. comm.; at 1.1 mm; Gómez-Guijarro et al. 2022), ASPECS-LP<sup>4</sup> (at 1.2 mm and 3 mm; González-López et al. 2019, 2020), and MORA<sup>5</sup> (at 2.1 mm; Casey et al. 2021). We replaced our pipeline-imaged maps of these surveys by these publicly available combined maps to allow for a deeper detection limit (see Sect. 2.1). However, unlike our pipeline-processed images, these combined maps were not “cleaned”, hence their point spread function (PSF) differed from an ideal 2D Gaussian function (dirty vs. clean beam). This means that while PyBDSF could be used to detect sources in these maps, it could not be used to accurately measure their flux densities. We therefore had to resort to another approach.

For GOODS-ALMA and ASPECS-LP, the PSFs are available alongside the continuum maps, which allows for an aperture photometry approach: integrated flux densities were measured within circular apertures placed at the positions of source candidates (i.e. sources for which PyBDSF yields a  $S/N_{\text{peak}} \geq 5.4$ ) and normalised by the PSF. We verified that the so-extracted flux densities were in agreement with measurements from the original works of Gómez-Guijarro et al. (2022) and González-López et al. (2019, 2020). Naturally, this approach

<sup>4</sup> <https://almascience.nrao.edu/alldata/lp/ASPECS/>

<sup>5</sup> <https://www.as.utexas.edu/~cmcasey/alma2mmcosmos/>

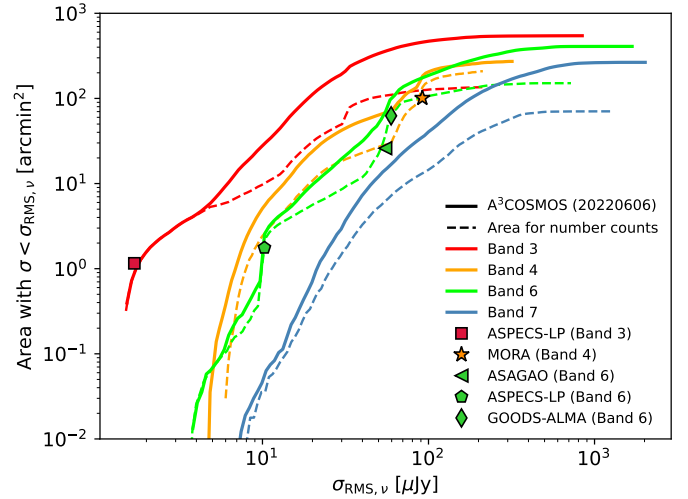
does not yield any information about the size of the sources, information that is therefore absent from our blind catalogue. While we do not provide modelled source size information in our blind source catalogue, an estimate of the angular extent of sources in GOODS-ALMA and ASPECS-LP can, however, still be deduced from the ratio of their peak to integrated flux densities. We used this approach for the computation of our number counts, which requires source size information (see Sect. 3.2). We note, however, that sources are mainly considered as point source at the coarse angular resolution of the ASPECS-LP survey ( $1.5 \times 1.1$  arcsec and  $1.8 \times 1.5$  arcsec at 1.2 mm and 3 mm, respectively) and only marginally resolved at the intermediate angular resolution of the GOODS-ALMA survey ( $0.7 \times 0.5$  arcsec). This is consistent with the finding from the original works of González-López et al. (2019, 2020) and Gómez-Guijarro et al. (2022).

For the MORA map, the PSF is not publicly available, preventing us from measuring flux densities via our aperture photometry approach. Fortunately, for this survey the flux densities obtained by PyBDSF assuming a 2D Gaussian PSF were in very good agreement with those from the original work of Casey et al. (2021). We only noted a slight overestimation by 10% of our PyBDSF flux densities, which we simply corrected by multiplying our flux densities by a factor 0.9. Finally, in line with Casey et al. (2021) and with the very coarse angular resolution of this survey ( $1.8 \times 1.4$  arcsec), we treated these detections as point sources.

## 2.5. Summary of A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS data release 20220606

Data products from this new data version of the A<sup>3</sup>COSMOS project, containing both the A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS databases, are publicly available at the CDS (for brevity, in the following we refer to A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS jointly as “A<sup>3</sup>COSMOS database”). This includes the blind and prior photometry catalogues, which contain all individual ALMA detections, and the Robust Galaxy Catalogue, which contains our final sample of unique galaxies (i.e. without CPA discarded and SED outlier sources) with their corresponding galaxy properties. For the description of the columns in those catalogues we refer to Tables 4 and 5 of Liu et al. (2019a). The properties of our final database and source catalogues are listed in Table 1. The ALMA bands with the most observations available are 3, 4, 6, and 7, each with a total spatial coverage of over 200 arcmin<sup>2</sup> combining both COSMOS and GOODS-S. Band 6 profits especially from the addition of the A<sup>3</sup>GOODSS database, expanding the total sky area covered in this band by ~30% compared to the coverage in COSMOS, due to the availability of large blind surveys. In Bands 3, 4, and 7 the total coverage is increased by ~20–24% compared to COSMOS. Bands 5, 8, and 9 are much less commonly used and contribute only a small number of single pointings each.

The areal coverage as a function of depth (i.e.  $1\sigma$  pixel RMS noise) is shown in Fig. 2. Single-dish surveys routinely cover areas of several hundred square arcminutes down to typical sensitivities of the order of ~1 mJy. At this depth, our database covers ~200–500 arcmin<sup>2</sup> depending on the ALMA band, approaching the regime of single-dish coverage but without their inherent source confusion problem. Deep interferometric surveys achieve much better sensitivities (~0.01–0.1 mJy depending on wavelength), but cover only small fields up to a few tens of square arcminutes. Several of these deep surveys are included in our database, plus a number



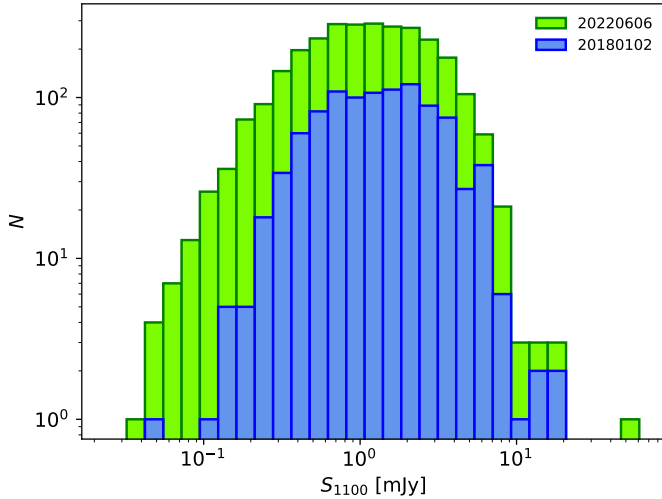
**Fig. 2.** Cumulative areal coverage of the A<sup>3</sup>COSMOS database, version 20220606, as a function of  $1\sigma_{\text{RMS}}$  sensitivity for ALMA Bands 3, 4, 6, and 7. The noise of all images in each band is normalised to the central wavelength of the band, and spatial overlaps of images are accounted for. The solid lines indicate an entire database; the image areas are considered out to a primary beam attenuation of 0.2. The dashed lines are images selected for the computation of the number counts after masking out the phase centres and cropping the edges at a primary beam attenuation of 0.3 (see Sect. 3.1). The symbols (see legend) represent the area and normalised depth of multiple ALMA blind surveys: ASPECS-LP (González-López et al. 2019, 2020), MORA (Zavala et al. 2021; Casey et al. 2021), ASAGAO (Hatsukade et al. 2018), and GOODS-ALMA (Gómez-Guijarro et al. 2022).

of even deeper single pointings. Finally, the large number of single pointings of varying sensitivity yield a continuous increase in area towards shallower depths, thus bridging the regimes of deep interferometric and large single-dish surveys.

With the new data version we more than double the number of detections compared to the last release. Figure 3 shows the normalised flux density distribution of unique ALMA detections in Bands 6 and 7 for the old release 20180102 and our new data version combining both A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS sources. The largest increase in source numbers is in the intermediate to lower flux density range (i.e. <2 mJy). The intermediate flux density regime (i.e. 0.2–2 mJy) benefits in particular from a large increase, in accordance with the fact that it is the most densely sampled regime in the first place. The lower flux density regime (i.e. <0.2 mJy), however, has the largest percentage increase with approximately an order of magnitude compared to the previous release.

Finally, we note that the total number of unique sources in the A<sup>3</sup>GOODSS catalogue is significantly lower than the number in the A<sup>3</sup>COSMOS catalogue, but this difference is understandable when considering the difference in spatial coverage (i.e. 149.6 arcmin<sup>2</sup> vs. 896.3 arcmin<sup>2</sup>). In fact, we verified that the number density of ALMA detections in the GOODS-S field per unit area at a given flux density (i.e. the number counts, see Sect. 3) matches that in the COSMOS field. However, the fraction of unique sources entering our final Robust Galaxy Catalogue is lower in GOODS-S (~43%) than in COSMOS (~76%). There are two reasons for this difference. Firstly, in GOODS-S there is a higher fraction of sources without any optical/IR counterpart. This can be explained by deeper ALMA observations in GOODS-S and by the presence of more deep blind surveys in





**Fig. 3.** Flux density distribution of detected sources in ALMA Bands 6 and 7 in the A<sup>3</sup>COSMOS database versions 20180102 and 20220606 (including A<sup>3</sup>GOODSS). Flux densities are normalised to 1100  $\mu$ m assuming that they probe the Rayleigh–Jeans tail of the dust continuum, with a dust spectral index  $\beta = 1.8$  (Scoville et al. 2014).

this database, which are tailored to detect sources without previously known optical counterparts. And secondly, GOODS-S also has a higher fraction of discarded sources due to unreliable counterparts, identified either through our dedicated CPA approach (i.e. CPA discarded) or revealed by bad SED fits (i.e. SED outliers). The ancillary data available in GOODS-S is generally deeper than in COSMOS, which increases the surface density of potential counterparts and thus the ambiguity of our CPA.

### 3. Selection bias correction towards accurate number counts

Galaxy number counts are a useful measurement to constrain galaxy evolution models and the longer wavelength regimes covered by ALMA are especially valuable to constrain the population of high-redshift DFSGs in these models (e.g. Casey et al. 2018a,b). Number counts within different ALMA bands have been established in the past (see e.g. Karim et al. 2013; Oteo et al. 2016; Stach et al. 2018; Hatsukade et al. 2018; González-López et al. 2019, 2020; Simpson et al. 2020; Béthermin et al. 2020; Zavala et al. 2021; Gómez-Guijarro et al. 2022; Chen et al. 2023). Ideally, number counts are obtained from blind surveys, which are characterised by a large contiguous area and the absence of selection biases. Although containing a number of blind surveys, our A<sup>3</sup>COSMOS database is an intrinsically inhomogeneous accumulation of observations from a large number of projects with different scientific objectives, tailored to the particular targets and requirements set by each individual PI. This circumstance negates both advantages inherent to blind surveys and introduces two major challenges: removing selection biases and determining the correct effective sky area. The first challenge is associated with the fact that most of the images in the A<sup>3</sup>COSMOS database are targeted on specific sources. Single pointings are targeted observations as a rule and therefore usually have a source at or near their phase centre. If these targets are counted, this results in a massive overestimation of the source surface density. The second challenge is associated with the inhomogeneous nature of the database. To determine

the effective area of this “survey”, the differences in wavelength, depth, and resolution between observations as well as their spatial overlap need to be properly taken into account.

There have been past attempts to infer number counts from single-pointing-based data with ALMA (Oteo et al. 2016; Zavala et al. 2018; Béthermin et al. 2020; Chen et al. 2023). Those studies, however, either deal with few projects or use calibrator observations, so they can reliably exclude their targets manually. With our large data volume, it is unfeasible to do a manual selection. Therefore, we have to define an algorithm.

In this section, we describe the procedure by which we attempt to derive number counts from the fundamentally heterogeneous A<sup>3</sup>COSMOS database (see also Traina et al. 2024, for a similar method). The procedure is schematically outlined in Fig. 4 and consists of two main parts: the selection criteria applied to correct for observational biases and achieve a “blinding” of the data, and the calculation of the effective area.

We chose the blind source catalogue as the basis for our number counts analysis in order to also include sources without a known optical counterpart. Sources can be listed in this catalogue multiple times if they are covered by multiple pointings or surveys. We therefore first reduced the catalogue to only one detection per source and ALMA band. In cases of multiple detections for one source in a given band, we only considered the detection with the highest significance (i.e. highest  $S/N_{\text{peak}}$ ).

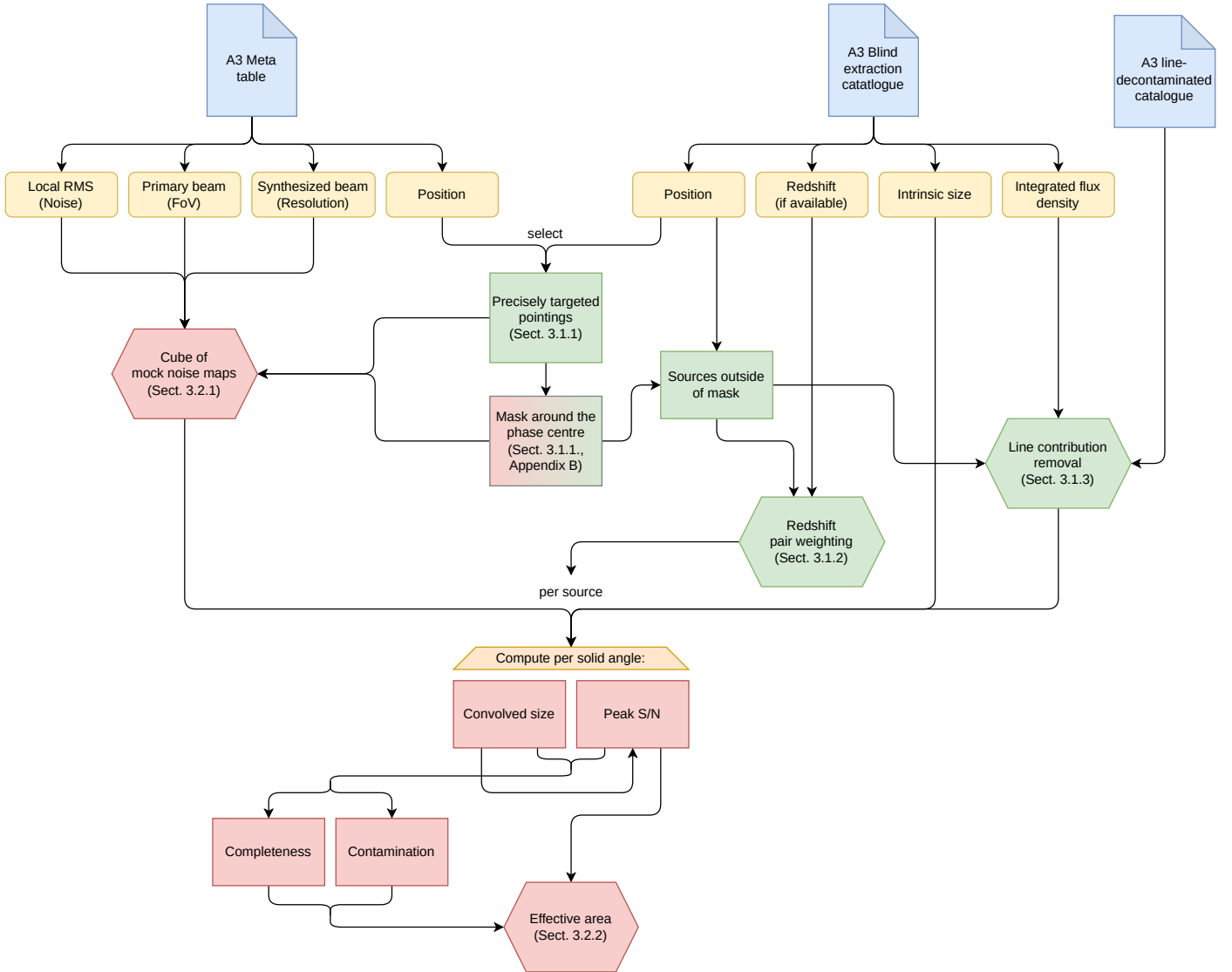
#### 3.1. “Blinding”: Source selection

The removal of selection biases from a database in which the scientific objectives of each observation were selected by individual PIs with all types of galaxies is challenging but vital in attempting to measure the underlying source surface density. A form of selection and exclusion has to be made to the detected sources in order to correct for those biases. The most important factor is the identification and exclusion of observational targets. Additionally, sources physically associated with those targets due to clustering can cause an artificial excess in the number counts. While clustering is an integral part of the distribution of sources in the sky, intentional observation of clusters (e.g. Umehata et al. 2017) and/or bright sources, which have an increased probability of being in a cluster environment, can result in artificially high source counts (e.g. Gruppioni et al. 2020). In the following sections, the selection criteria to minimise those biases are described and tested via simulations.

##### 3.1.1. Offset and masking

Identifying the target of single pointings without prior knowledge on the observational goals of the respective projects poses a big challenge. Single pointings are naturally aimed at the position of a known source, the target is thus expected to be exactly at the phase centre of the respective pointing. However, the known prior position can be offset from the source position when observed with ALMA, for example, if it is a follow-up of submillimetre galaxies selected from single-dish surveys which are characterised by large positional uncertainties ( $\geq 3$  arcsec; Hodge et al. 2013). From the A<sup>3</sup>COSMOS images alone, observational targets offset from the phase centre are indistinguishable from serendipitous sources that happen to be in the field of view, especially if a source is offset only by a few arcseconds from the phase centre. As there are  $\sim 230$  unique projects and  $\sim 4000$  images in the entire database, it is impossible to examine individually every image considering the respective project in order





**Fig. 4.** Schematic of the process for retrieving continuum number counts from the inhomogeneous A<sup>3</sup>COSMOS database. The red panels are steps in the computation of the effective area, as described in Sect. 3.2; the green panels are steps to correct for selection biases, as described in Sect. 3.1. This process is repeated once for each ALMA band.

to determine which source/s was/were the intended target/s in each individual case.

Instead of making uncertain assumptions about which sources can or cannot be considered targets, we chose a more conservative approach (see also Traina et al. 2024): we limited our analysis to what we refer to in the following as “precisely targeted” single pointings that have a source right at their phase centre (i.e. the extracted position is within a radius of 1 arcsec around the phase centre). Since the probability of a source being serendipitously detected within this small area is low (of the order of  $\sim 0.1\%$ ), we have a degree of certainty to say that these sources are actual targets and should be excluded from our number counts analysis. To exclude these targets from our analysis as well as further potential contaminants close to them due to clustering, we masked a central circular area around the phase centre of each precisely targeted pointing. All sources inside the mask were removed from our analysis and the corresponding area was excluded from the calculation of the effective area (Sect. 3.2.1). The radius of this mask was chosen to be the smallest radius beyond which the number counts do not significantly change

when the radius is further increased. In Bands 3, 4, and 6 the mask radius was set to 1 arcsec, in Bands 7–4 arcsec. The choice of the mask was also tested through our simulations described in Sect. 3.3 (see Appendix D for details).

It would be intuitive to assume that masking out all (usually bright) target sources would in turn introduce a negative bias to the bright end of our number counts. However, the sky distribution of galaxies follows a Poisson distribution (not considering clustering) and thus the probability of detecting a bright source in a given line of sight (LOS) is independent of the presence of another bright source in its vicinity. Therefore, bright sources are still detected serendipitously and the bright end of the number counts can still be constrained. The only limitation to this is clustering, which we accounted for by considering the redshift information of targets and serendipitous sources (see Sect. 3.1.2). Our simulations (see Sect. 3.3) show that using this approach, even when targeting only the brightest sources in a given field and masking them out, it does not introduce a negative bias on the bright end of the number counts.

In blind surveys, no specific selection or masking is needed since there are no dedicated targets. Thus, every source was treated as serendipitous.

### 3.1.2. Redshift pairs

While a central mask can exclude the most proximate cluster neighbours, the typical angular scale of galaxy clusters (e.g.  $\sim 240$  arcsec for a cluster at  $z = 2$  with a 2 Mpc diameter) exceeds the primary beam size of typical ALMA pointings (e.g.  $\sim 60$  arcsec in Band 3). It is therefore necessary to identify among our serendipitous sources (i.e. outside the central mask) those which are physically associated with the target-source (i.e. at the phase centre) by more than purely angular proximity.

To this end, we made use of the availability of prior redshift information for most detected sources within our A<sup>3</sup>COSMOS database. The expected maximum redshift difference between two galaxies located within the same cluster is of the order of  $\Delta z \approx 10^{-3}$  (for a cluster at  $z \lesssim 6$  with a 2 Mpc diameter). However, measured redshifts (especially when measured photometrically) are subject to uncertainties that often exceed this order of magnitude. It is therefore not possible to make absolute statements about two sources being associated in the same cluster, and instead we have to assess the likeliness of such an association. To do so, we used the integral of the overlap between the redshift probability distribution function (PDF) of a serendipitous source and its corresponding target-source as an estimate of the likelihood  $P_{\text{pair}}$  of both sources having the same redshift and thus being located within the same galaxy cluster. For sources with only a photometric redshift estimates we used as PDFs a Gaussian function centred on their redshifts with a width of  $0.06(1+z)$ , that is, the median photometric redshift uncertainty in the COSMOS2020 catalogues (Weaver et al. 2022). For sources with a spectroscopic redshift we used a PDF width of  $0.001(1+z)$  (e.g. Fernández-Soto et al. 2001). The contribution of each serendipitous source to the number counts was then weighted by its likelihood to have the same redshift as the target-source by multiplying it with a factor  $1 - P_{\text{pair},i}$ , where  $P_{\text{pair}}$  ranges from 0, for a very unlikely physical association between the serendipitous source and the target of the pointing, to 1 for a very likely association.

Naturally, we cannot apply this weighting to a fraction of source pairs ( $\sim 20\%$ ) for which no redshift information is available (for either the target or the serendipitous source). In that case,  $P_{\text{pair}}$  was simply set to a value of 0, with the exception of one source known to be in a cluster with the observational target, where we set  $P_{\text{pair}} = 1$  (CRLE, Pavesi et al. 2018).

### 3.1.3. Line contamination

There is a small probability that the continuum flux densities of some of sources selected for our number counts may be contaminated by emission from bright lines that happen to fall within the frequency window of the observation. To account for this, we made use of the fact that our A<sup>3</sup>COSMOS pipeline already includes a routine that calculates the contribution of bright emission lines to the continuum flux densities (see Sect. 2.1). However, as this requires an SED fit it can only be performed for sources that have prior counterparts and redshift information. This is the case for almost all serendipitous sources in Bands 3 and 7, and for  $\sim 60\%$  of the serendipitous sources in Bands 4 and 6. We find a significant contribution of line emission to the continuum (i.e.  $\geq 10\%$  of the measured integrated flux density) for

one source in Band 3 and for  $\sim 25\%$  and  $\sim 10\%$  of the sources in Bands 4 and 6, respectively. No significant line contribution is found in Band 7. Despite this rather low degree of line contamination, we decided to use the line-decontaminated total flux densities to calculate the effective area associated with each galaxy (see Sect. 3.2). Additionally, we excluded sources from our number counts analysis if their line-decontaminated peak flux density fell below our minimum threshold of  $S/N_{\text{peak}} = 5.4$ . This applied only to one source in Band 3, two sources in Band 4, and two sources in Band 6. We note that although we chose to correct for line contamination for the sake of thoroughness, this does not significantly affect the results of our analysis as the differences are well within the uncertainties.

### 3.2. Effective area

Number counts can be calculated by inversely summing the effective area of all sources above a given flux density  $S$ :

$$N(>S) = \sum_i \frac{1}{A_{\text{eff}}(S_i, \theta_i)}, \quad (1)$$

When introducing pair-weighting (see Sect. 3.1.2), this equation becomes:

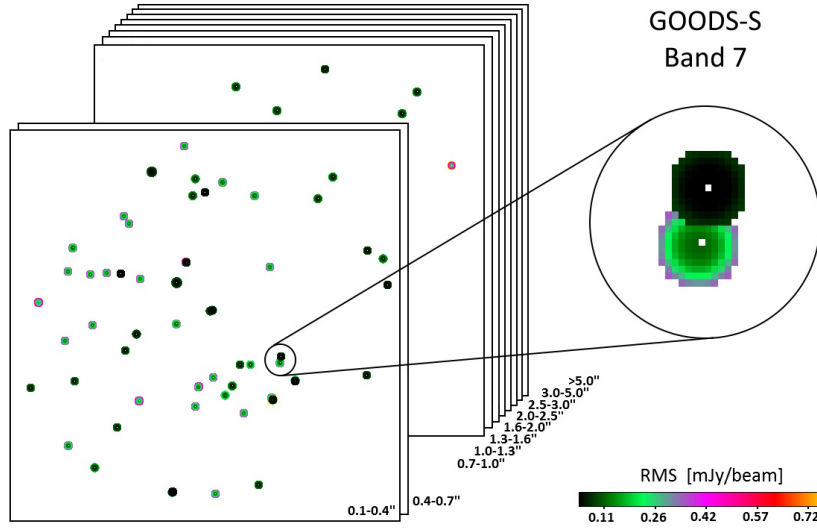
$$N(>S) = \sum_i \frac{1 - P_{\text{pair},i}}{A_{\text{eff}}(S_i, \theta_i)}. \quad (2)$$

Here, the effective area of a source,  $A_{\text{eff}}$ , is given by

$$A_{\text{eff}}(S_i, \theta_i) = \iint_{\text{Survey}} \frac{C_{\text{compl.}}(S_i, \theta_i, x, y)}{1 - C_{\text{contam.}}(S_i, \theta_i, x, y)} D(S_i, \theta_i, x, y) d\Omega, \quad (3)$$

and depends on the detectability  $D$ , the completeness  $C_{\text{compl.}}$  and the contamination  $C_{\text{contam.}}$ <sup>6</sup>. A source is considered detectable if its  $S/N_{\text{peak}}$ , which depends on the total source flux density  $S_i$  and the beam-convolved source size  $\theta_i$ , at this particular position is greater than the blind extraction threshold of 5.4 (see Sect. 2.1), therefore  $D = 1$  where  $S/N_{\text{peak}} > 5.4$  and  $D = 0$  otherwise. The completeness (i.e. the recovery rate of real sources) and contamination (i.e. the likelihood of a source being a spurious detection) both depend on  $S/N_{\text{peak}}$  and  $\theta_i$  and range from 0 to 1. For any given source, the detectability, completeness and contamination thus vary largely between different observations, depending on wavelength, sensitivity and resolution of the respective images. A two-step approach was thus chosen in order to incorporate the effect of detectability, completeness and contamination on the calculation of the effective area of a given source: we first created noise coverage maps of the COSMOS and GOODS-S fields for each ALMA band into which we wrote the positions and sensitivities of each ALMA image, split into bins of different spatial resolution and prioritising low over higher RMS noise in areas of spatial overlap. We then used these maps to determine the detectability, completeness and contamination for each source in each solid angle of the COSMOS and GOODS-S fields and thereby calculate its associated effective area.

<sup>6</sup> The contamination is an equivalent measure to the commonly used ‘‘purity’’, which corresponds to  $1 - C_{\text{contam.}}$ . Here, we use contamination, as it is directly obtained from the spurious fraction (see Sect. 3.2.2).



**Fig. 5.** Illustration of the noise cube for the GOODS-S field in ALMA Band 7 (see Sect. 3.2.1). Each cube consists of ten layers for different angular resolutions, as indicated in the bottom right of each layer. Each pixel represents a solid angle of  $2 \times 2$  arcsec. The colour of each pixel indicates the lowest RMS noise available between all pointings covering this solid angle in that particular band and at this angular resolution. The enlarged cutout shows how, as a result, pointings with lower RMS overwrite those with higher RMS in cases of spatial overlap between pointings with similar angular resolution.

### 3.2.1. Noise maps

In order to assess the detectability of a source of a given flux density and size in each solid angle of the COSMOS and GOODS fields, we need to take into account both the RMS and spatial resolution of all images available in our database. While the detectability of point-like sources depends only on the image sensitivity, an extended source is more easily detectable in a low angular resolution image than in a high angular resolution image with the same sensitivity per beam.

To account for the effect of both RMS and angular resolution, we chose the following approach: For each ALMA band, we set up empty maps of the COSMOS and GOODS-S fields with a pixel-scale of 2 arcsec. Each map has ten layers corresponding to different angular resolutions, which are: 0.1–0.4'', 0.4–0.7'', 0.7–1.0'', 1.0–1.3'', 1.3–1.6'', 1.6–2.0'', 2.0–2.5'', 2.5–3.0'', 3.0–5.0'' and  $>5''$ . These bins were chosen so that the spatial overlap of ALMA pointings within each layer is minimised. The RMS of all available images from the database ( $\sigma_v$ ) was normalised to the central frequency of its respective band ( $\nu_{v,\text{ref}}$ ) under the assumption of being in the regime of the Rayleigh-Jeans tail of the dust continuum emission and choosing a dust spectral index  $\beta = 1.8$  (e.g. [Scoville et al. 2014](#)):

$$\sigma_{v,\text{ref}} = \sigma_v \cdot \left( \frac{\nu_{\text{ref}}}{\nu} \right)^{3.8}. \quad (4)$$

These normalised RMS maps were then corrected for the primary beam attenuation, causing the noise to increase with distance from the phase centre. Areas beyond a primary beam response of 0.3 were disregarded as well as a central radius of 1 or 4 arcsec around the phase centre of single pointings (Sects. 3.2.2, 3.1 and Appendix D). The renormalised primary-beam corrected RMS maps were then written into the layer corresponding to their angular resolution of the respective field maps at the positions corresponding to their sky coordinates. In case of spatial overlap between pointings of similar angular resolution, lower RMS was favoured over higher RMS. In addition to this so-created “noise cube”, the resolution properties of all images (that is, major and minor axis as well as position angle of

the synthesised beam) were stored in equivalent map cubes. As a result, we have for each ALMA band four ten-layered map cubes per field, one storing the normalised noise and three storing the corresponding resolution information. These cubes are the basis on which the effective area for each source was calculated. An example of such a noise cube for Band 7 in the GOODS-S field is shown in Fig. 5. This example illustrates the primary-beam corrected noise profiles, the masking, and the treatment of spatial overlap between two pointings of similar angular resolution in our calculation of the effective area.

### 3.2.2. Computing the effective area

The sources chosen for our number counts analysis are blindly detected sources with  $S/N_{\text{peak}} \geq 5.4$ . Since the edges of images are more likely to be contaminated with spurious sources due to a higher and not well-behaved noise, we only considered the area of images where the primary beam attenuation is  $>0.3$  and used only sources found therein. Blind source extraction provides the integrated flux density  $S_{\text{total}}$  and peak flux density  $S_{\text{peak}}$  of each source. From this, the circularised intrinsic source size,  $\theta_{\text{intr.}}$ , was backwards-inferred by dividing  $S_{\text{total}}$  by  $S_{\text{peak}}$  and performing a deconvolution with the circularised size of the synthesised beam of the image in which the source was originally detected,  $\theta_{\text{beam}}$ :

$$\theta_{\text{intr.}} = \theta_{\text{beam}} \cdot \sqrt{\frac{S_{\text{total}}}{S_{\text{peak}}} - 1}. \quad (5)$$

We chose to backwards-infer the intrinsic source size from the flux density measurements, rather than from the PyBDSF 2D Gaussian modelling, as it allowed us to consistently predict the measured  $S/N_{\text{peak}}$  of each source in their respective original image. While in most cases this also works with the source size information provided by PyBDSF, there are also rare cases in which it is not possible, for example when a source is fitted by PyBDSF with two or more Gaussian profiles. Sources extracted from low angular resolution images are typically retrieved as unresolved (i.e.  $\theta_{\text{intr.}} = 0$ ). To account for the fact that such



sources could be resolved in images of higher angular resolution, we assigned them a typical intrinsic source size of radius  $r = 0.1 \pm 0.05$  arcsec, following the findings of Gómez-Guijarro et al. (2022).

We used  $\theta_{\text{intr.}}$  and  $S_{\text{total}}$  to compute the expected detectability of each source in a given LOS of the COSMOS and GOODS-S fields, using

$$S/N_{\text{peak,LOS}} = \frac{S_{\text{total},\nu} \cdot \left(\frac{\nu_{\text{ref}}}{\nu}\right)^{3.8}}{\frac{\theta_{\text{intr.}}^2 + \theta_{\text{beam,LOS}}^2}{\theta_{\text{beam,LOS}}^2} \cdot \text{RMS}_{\text{LOS},\nu_{\text{ref}}}}, \quad (6)$$

where  $\nu$  is the observed frequency,  $\nu_{\text{ref}}$  is the central frequency of the band,  $\theta_{\text{beam,LOS}}$  is the circularised beam axis of the most sensitive observation in that angular resolution layer and  $\text{RMS}_{\text{LOS}}$  is the corresponding RMS value. If the expected peak signal-to-noise was above the minimum threshold of 5.4, the source was considered detectable at that angular resolution and within the solid angle intersecting the area  $A_{\text{LOS}} = 4 \text{ arcsec}^2$  (i.e. the previously defined pixel-scale of 2 arcsec of our noise cubes).

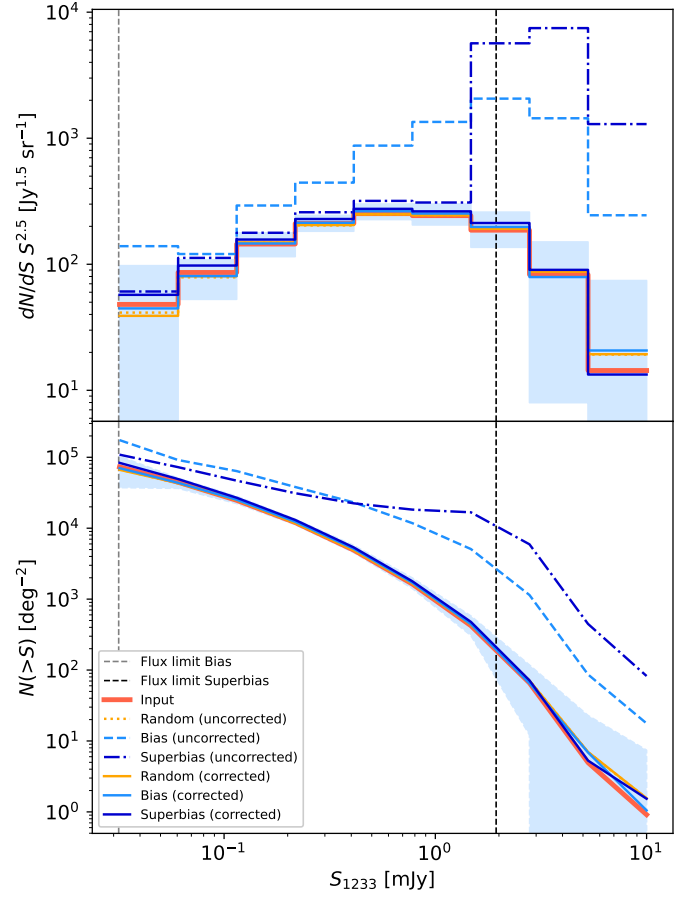
Using  $S/N_{\text{peak,LOS}}$  and  $\theta_{\text{conv,LOS}} = \sqrt{\theta_{\text{intr.}}^2 + \theta_{\text{beam,LOS}}^2}$  the correction factors for completeness  $C_{\text{compl,LOS}}$  and contamination  $C_{\text{contam,LOS}}$  were determined. The completeness correction for a given  $S/N_{\text{peak,LOS}}$  and  $\theta_{\text{conv,LOS}}$  was adopted from Liu et al. (2019a) who inferred these corrections using Monte Carlo simulations in which artificial sources were inserted into residual images and recovered with the same source extraction method that was used to build the A<sup>3</sup>COSMOS blind catalogue. As a rule, the completeness increases with  $S/N_{\text{peak}}$  and is slightly higher for less extended sources. Above our detection threshold of  $S/N_{\text{peak}} > 5.4$ , the correction is small as the completeness is always  $\gtrsim 70\%$ . The contamination correction was determined from the fraction of spurious sources. This spurious fraction was obtained by dividing the number of detections in the A<sup>3</sup>COSMOS negative ALMA images (i.e. original A<sup>3</sup>COSMOS images multiplied by  $-1$ , so any detection is by definition spurious) by the number of detections in the positive (i.e. original) A<sup>3</sup>COSMOS images (i.e. consisting of both real and spurious sources). This spurious fraction decreases with both signal-to-noise and source size. Again, for most cases the correction is small, that is, on average  $\sim 20\%$  at  $S/N_{\text{peak}} = 5.4\text{--}6.0$  and  $< 15\%$  for  $S/N_{\text{peak}} > 6.0$ . The corrections for completeness and contamination were directly applied to the area associated with a given solid angle:

$$A_{\text{eff,LOS}} = \frac{A_{\text{LOS}} \cdot C_{\text{compl,LOS}}}{1 - C_{\text{contam,LOS}}}. \quad (7)$$

With this approach, the effective area for a given source in a given angular resolution layer is the sum of  $A_{\text{eff,LOS}}$  over all lines of sight in the COSMOS and GOODS-S noise cubes. Finally, for sources detectable in a given solid angle in more than one angular resolution layer, we counted the layer with the highest value of  $A_{\text{eff,LOS}}$ . Naturally, this calculation of the effective area was made independently for each band (i.e. to infer the number counts in a given band, we only account for the source detected in this band and considering the noise cubes of this band).

### 3.3. Simulations

In order to test the applicability of our selection criteria, and to some degree also our effective area calculation, to accurately retrieve number counts from a biased database such as ours, we



**Fig. 6.** Simulated differential (top) and cumulative (bottom) number counts for single pointings in ALMA Band 6 based on the SIDES simulated sky catalogue (Béthermin et al. 2017; Gkogkou et al. 2023), the pointing distribution of A<sup>3</sup>COSMOS, and our applied selection criteria. The coloured lines show the number counts for different simulated biases (see Sect. 3.3). Dark orange indicates the input number counts as inferred from the complete SIDES catalogue. Light orange is for the pointings randomly distributed on the sky (Random). Dark blue is for the pointings that target all the brightest sources (Superbias). Light blue is for the pointings that reflect the target source brightness distribution in A<sup>3</sup>COSMOS (Bias). The dashed, dash-dotted, and dotted lines are number counts inferred from the complete pointings; the solid lines show the number counts when applying a central mask of 1 arcsec and weighting by redshift proximity. The shaded region shows the  $1\sigma$  dispersion between all realisations of the corrected Bias line, but is approximately equal for all corrected lines. The vertical black and grey dashed lines indicate the flux density of the faintest target source in the Superbias and Bias cases, respectively. We note that the uncorrected Random line is not visible in the lower panel, as it fully agrees with the corrected Random line.

used simulations to recreate a biased sample of observations and infer number counts with our previously described method.

As basis for these simulations, we used the Simulated Infrared Dusty Extragalactic Sky (SIDES) from Gkogkou et al. (2023, see also Béthermin et al. 2017). SIDES offers 117 unique patches of simulated sky of  $1 \text{ deg}^2$  each, accounting for clustering and including information on redshift and continuum flux density in all relevant ALMA bands. Since the COSMOS and GOODS-S fields together cover an approximate area of  $2 \text{ deg}^2$ , we took 116 of those patches and combined them into independent 58 realisations of a  $2 \text{ deg}^2$  field.

In each of these 58 fields (i.e. per realisation), we made mock observations by placing in them mock ALMA pointings imitating the set of precisely targeted pointings used in our real number counts analysis. We then retrieved number counts from these mock observations using the SIDES catalogue to compute the  $S/N_{\text{peak}}$  of the mock sources located inside these mock pointings. For a detailed description of these mock observations and the number counts retrieval see Appendix C. We used three modes of simulations differing in the placement of their mock pointings (i.e. choice of mock target sources) to simulate different observational biases: “Superbias”, “Bias”, and “Random”. The Superbias mode models the extreme case in which only the brightest sources in a given field have been targeted by the ALMA observers. The Bias mode represents a more realistic case as the mock targets are chosen based on the flux density distribution of the real observational targets in our pointings. Finally, the Random mode serves as a sanity check, as the number counts are measured on a set of mock pointings not targeting any mock sources, distributed randomly instead.

In all three modes, we used two approaches to retrieve the mock number counts: the “uncorrected” number counts are retrieved accounting for the complete areal coverage of the mock pointings and all sources within; the “corrected” number counts are inferred after applying the same corrections as in the real case (i.e. masking and weighting). Per realisation, the number counts were computed separately for all bias modes, in both the uncorrected and corrected cases. The final recovered mock number counts are the mean over all 58 realisations. Variation from one realisation to another is characterised through the dispersion over these 58 realisations and corresponds to the cosmic variance affecting our estimates. We verified that this cosmic variance is dominated by pure Poisson error at the areal coverage of our database.

The Band 6 mock number counts inferred using these simulations and with a central masking radius of 1 arcsec radius are shown in Fig. 6. These are compared to the “Input” number counts which correspond to the number counts inferred using the complete input SIDES catalogue. Because the cosmic variance is similar for all the different modes and number counts retrieval approaches, we show for clarity the dispersion only for the corrected Bias line. The Random uncorrected and corrected number counts are in good agreement with the Input number counts. This is expected, as in the absence of observational bias this set of mock pointings behaves effectively as a blind survey, only affected by cosmic variance. Applying additional masking has no further impact in that case other than slightly decreasing their effective area and thus increasing the cosmic variance. The Random mode validates our measurement of the effective area. The uncorrected Bias and Superbias number counts show an expected large excess over the real distribution, up to two orders of magnitude for the bright end in Superbias mode and still more than one order of magnitude in Bias mode. The excess in the uncorrected Superbias mode is predominantly at the bright end, since all existing sources above a certain flux density limit are targeted and counted. This limit is indicated by a vertical line in Fig. 6. Below this limit, however, there is still a notable, albeit much smaller, systematic excess (by  $\sim 25\%$ ), which is due to clustering, as massive and bright galaxies at  $z > 1$  are more likely to be in an over-dense environment (in SIDES and reality) and therefore have a higher number of less bright neighbours. The excess of the uncorrected Bias number counts is spread over the entire flux density range, albeit largest at the bright end. This bias probably reflects the actual situation in the A<sup>3</sup>COSMOS database, in which bright sources are often favoured as observa-

tional targets, but fainter sources are also being investigated. The effect of clustering here is not as obvious, because the target bias dominates over the whole flux density range. These uncorrected number counts demonstrate that it is indispensable to apply a form of masking in attempting to extract the underlying real number counts from these biased observations.

In the corrected Bias and Superbias modes, there is a clear improvement compared to the respective uncorrected number counts. The remaining systematics affect the number counts at a level which is an order of magnitude lower than the uncertainties associated with the cosmic variance (see Appendix D). We notice that these uncertainties increase towards the bright end due to the lower number of sources present in this flux density range, and towards the faint end due to the decrease in effective areal coverage.

The good agreement of the corrected Superbias number counts with the Input number counts also shows that, as previously mentioned, masking out targets does not prevent constraining the bright end of the number counts, even if all brightest sources are targeted. Since sources are Poisson distributed, bright sources, while masked out as targets, are also occasionally detected as serendipitous sources in other pointings and thus still enter our number counts analysis.

We conclude from these simulations that with a central mask and redshift pair weighting applied, an ensemble of single pointings is able to yield accurate (with a mean difference of no more than  $\sim 10\%$ ) number counts.

## 4. Results on the number counts

Having demonstrated that our selection bias correction method can successfully recover the number counts in simulations (see Sect. 3.3), we apply the same technique to our real data and infer the number counts. Table 2 gives an overview of the total number of single pointings available per band in our database compared to the number that remains after excluding pointings without a detection at their phase centre. The table also lists the number of precisely targeted pointings that have a serendipitous detection outside of their central mask. Figure 2 shows as dashed lines the areal coverage of these precisely targeted pointings and mosaics, after cropping the edges and masking out the phase centres. The total area available in Band 3 is reduced by  $\sim 75\%$ , in Band 7 by  $\sim 70\%$  and in Band 6 by  $\sim 63\%$ . Band 4 loses only  $\sim 24\%$  of total available area due to the fact that one large, albeit rather shallow, blind survey dominates the areal coverage in this band. Unfortunately, in Bands 5, 8, and 9, we do not have sufficient statistics in terms of number of precisely targeted pointings and serendipitous detections to obtain meaningful constraints on the number counts. We do not consider these bands in the rest of our analysis.

### 4.1. Single pointing and blind mosaic number counts

Using the method validated by our simulations, we inferred the number counts by combining precisely targeted pointings and blind surveys. The results are listed in Table 3 and shown in Fig. 7 as solid lines. The number counts are shown in both their differential ( $dN/dS$ ) and cumulative ( $N(>S)$ ) forms, together with number counts from the literature colour-coded by the wavelength of the corresponding ALMA band. Errors, shown as shaded regions, are the quadratic combination of the Poisson errors (using approximations from Gehrels 1986) with errors due to flux density uncertainties and errors due to uncertainties on the completeness and contamination corrections. Since our

**Table 2.** Number of images available to infer the number counts before and after applying our selection criteria.

ALMA Band	Blind survey images	Single pointings	Precisely targeted pointings	With serendipitous detection
Band 3	5	413	68	7
Band 4	2	158	75	16
Band 5	0	12	0	0
Band 6	24	1319	361	60
Band 7	0	1926	734	94
Band 8	0	42	12	0
Band 9	0	19	6	3

**Notes.** For each ALMA band the total number of single pointings and blind survey images is given (the latter not affected by selection bias). The next columns list the number of single pointings that contain a detected source at a distance  $<1$  arcsec from the phase centre (called the precisely targeted pointings) and among those the number of pointings that contain at least one serendipitous detection outside of the central mask of 1 arcsec radius (Band 7: 4 arcsec).

**Table 3.** Differential number counts inferred from the A<sup>3</sup>COSMOS database in ALMA Bands 3 (2998  $\mu\text{m}$ ), 4 (2082  $\mu\text{m}$ ), 6 (1233  $\mu\text{m}$ ), and 7 (925  $\mu\text{m}$ ).

$S_{2998}$ range [mJy]	$dN/dS S^{2.5}$ [Jy <sup>1.5</sup> sr <sup>-1</sup> ]	$S_{2082}$ range [mJy]	$dN/dS S^{2.5}$ [Jy <sup>1.5</sup> sr <sup>-1</sup> ]	$S_{1233}$ range [mJy]	$dN/dS S^{2.5}$ [Jy <sup>1.5</sup> sr <sup>-1</sup> ]	$S_{925}$ range [mJy]	$dN/dS S^{2.5}$ [Jy <sup>1.5</sup> sr <sup>-1</sup> ]
0.018–0.034	$0.81^{+1.87}_{-0.71}$	0.08–0.14	$39^{+80}_{-31}$	0.05–0.09	$17^{+19}_{-9}$	0.20–0.37	$449^{+1040}_{-374}$
0.034–0.064	$<0.75$	0.14–0.24	$21^{+30}_{-10}$	0.09–0.18	$114^{+66}_{-25}$	0.37–0.67	$206^{+168}_{-68}$
0.064–0.120	$2.41^{+3.28}_{-1.36}$	0.24–0.42	$15^{+29}_{-9}$	0.18–0.35	$87^{+70}_{-25}$	0.67–1.23	$458^{+299}_{-117}$
0.120–0.225	$1.45^{+2.57}_{-0.98}$	0.42–0.72	$16^{+14}_{-5}$	0.35–0.67	$174^{+108}_{-38}$	1.23–2.25	$739^{+428}_{-155}$
0.225–0.422	$<1.18$	0.72–1.26	$14^{+16}_{-6}$	0.67–1.28	$178^{+110}_{-35}$	2.25–4.12	$658^{+532}_{-178}$
		1.26–2.18	$<8$	1.28–2.47	$123^{+89}_{-30}$	4.12–7.54	$700^{+812}_{-256}$
				2.47–4.75	$144^{+149}_{-51}$	7.54–13.8	$<310$
				4.75–9.13	$<72$		

simulations indicate that the cosmic variance is dominated by pure Poisson error (see Sect. 3.3), no additional cosmic variance uncertainty was added to our error bars. The cumulative counts are shown at the measured total flux densities of each source in a band. The dashed horizontal lines indicate upper limits for the number counts above the flux density of the brightest source in a given band.

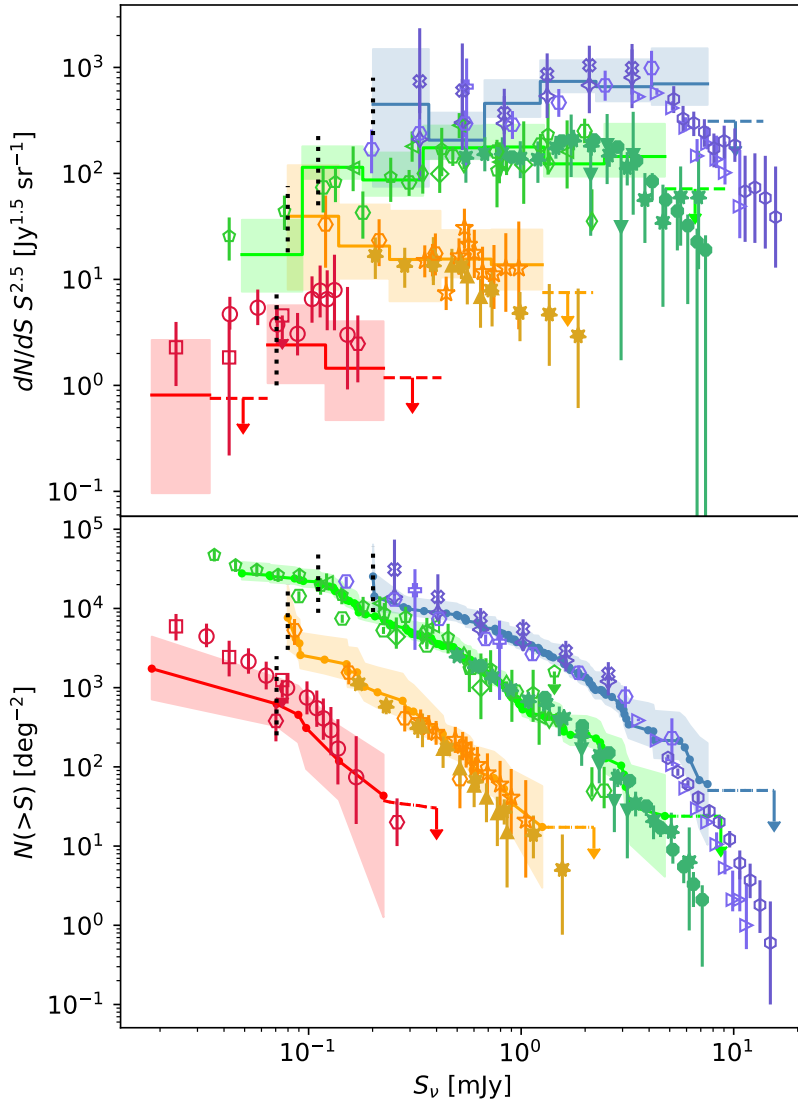
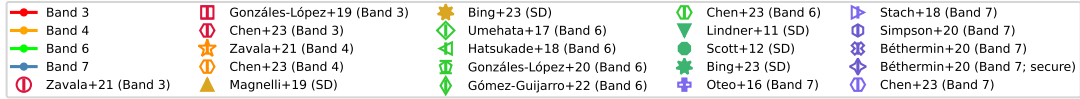
Significantly extended sources with low total flux densities are difficult to detect due to a low  $S/N_{\text{peak}}$  in images with a high or moderate angular resolution. Therefore, our database holds the risk of insufficiently detecting such sources, introducing an additional incompleteness effect, which lowers the accuracy of our number counts in lower flux density ranges. To see if this issue considerably affects our results, we defined the size of a significantly extended source (a circularised full width at half maximum of  $\sim 0.66$  arcsec, which corresponds to the 80th percentile of the distribution of circularised source radii in our blind catalogue) and determined the lowest total flux density at which a source of such size is detected in our source sample for each band. This limit is marked by the black vertical dotted lines in Fig. 7. Bands 3 and 6 are potentially affected in their low flux density bins  $\lesssim 0.1$  mJy.

#### 4.2. Comparison with previous ALMA number counts

In Band 3, we first compare our results to the ASPECS-LP ALMA 3 mm number counts (González-López et al. 2019) and find that they are by a factor  $\sim 3$  higher than ours. As the

combined map from this survey is contained in our database (see Sect. 2.4), this disagreement is somehow surprising. Therefore, we directly compared the sources extracted from this map with the six continuum detections listed in González-López et al. (2019). Only their brighter source (“CO1”) was recovered above our detection threshold of  $S/N_{\text{peak}} = 5.4$  and at a significantly lower flux density than listed in their work (18 vs. 33  $\mu\text{Jy}$ ). This source, as well as the other five, are also identified in González-López et al. (2019) as CO line emitters. This lead us to suspect that the continuum fluxes listed in and used in the number counts of González-López et al. (2019) were not measured on the publicly available “linefree” continuum map<sup>4</sup> but from a different version still containing line emission. To test this, we ran our prior-based GALFIT source extraction on the five frequency coverages of this survey, as downloaded from the ALMA archive and processed by our pipeline (see Sect. 2.1). Three out of the six continuum sources listed in González-López et al. (2019) were recovered at  $S/N_{\text{peak}} > 3$ : “CO1”, “CO2”, and “CO3”. The source “CO1” was found in four of these frequency coverages. In two of them, the GALFIT flux densities agree with our measurement on the “linefree” map, while in the other the flux density is significantly higher (i.e. 56 and 87  $\mu\text{m}$ ). The frequencies covered by these two latter tunings correspond to that of the CO line for this galaxy listed in González-López et al. (2019; CO 3–2 at  $z = 2.543$ ). The other two sources were recovered in only one frequency coverage, also corresponding to their respective CO lines (CO 3–2 at  $z = 2.696$





**Fig. 7.** Number counts inferred from the A<sup>3</sup>COSMOS database in four ALMA bands: red, Band 3 (2.6–3.6 mm); orange, Band 4 (1.8–2.4 mm); green, Band 6 (1.1–1.4 mm); blue, Band 7 (0.8–1.1 mm). Number counts were measured by combining blind surveys and single pointings containing a detected source within 1 arcsec radius around the phase centre and excluding a central area of 1 arcsec radius in Bands 3, 4, and 6, and 4 arcsec in Band 7. Top panel: Differential. Bottom panel: Cumulative. The number counts calculated in this work are shown as solid lines, with the shaded area indicating the uncertainty resulting from flux density, contamination, and completeness errors, quadratically combined with Poisson errors. The respective flux densities of the individual sources are shown as points in the cumulative panel. The vertical dotted lines indicate the flux density above which our number counts can be considered complete. Number counts from the literature are shown as symbols and are colour-coded to the respective ALMA band.

and CO 2–1 at  $z = 1.550$ ). These findings support our initial suspicion that the continuum flux densities of the sources listed in González-López et al. (2019) are based on a map that still contains line contribution. Therefore, our number counts can be considered as more accurate than those of González-López et al. (2019).

We then compare our Band 3 number counts with the results from Zavala et al. (2021) who, like us, used ALMA archival data (including the 3 mm ASPECS-LP) but where the targets and physically associated sources were masked out manually. At flux densities  $\geq 0.06$  mJy, our number counts are consistent within the

error bars with their results. However, at the lower end, their number counts are higher by a factor  $\sim 3$ . This can be explained by the inclusion of the 3 mm ASPECS-LP survey in the sample of Zavala et al. (2021). Like in our analysis, they recovered only one source from the ASPECS-LP map and noted the possible contamination by CO line emission of the flux density of this source. For this reason, they replaced the flux density of this source with the one given by González-López et al. (2019). However, as we previously argued, this flux density value is likely still contaminated by line emission. Fixing the continuum flux density of this source to our measured value would bring the

number counts of Zavala et al. (2021) into agreement, within the uncertainties, with our number counts at the low end.

Lastly, we compare our Band 3 number counts with the results from Chen et al. (2023), who used ALMA calibrator observations with masked-out targets. The results are consistent within the uncertainties.

In Band 4, we compare our number counts to the ALMA-based number counts from Zavala et al. (2021) and Chen et al. (2023) and the single-dish counts from Magnelli et al. (2019) and Bing et al. (2023), and find that they agree within the uncertainties. We reach a depth of  $\sim 0.08$  mJy, similar to the single-pointing-based number counts from Chen et al. (2023), which is 0.4 dex and 0.6 dex lower than the flux density limits of Bing et al. (2023) and Zavala et al. (2021). Utilising high-sensitivity single pointings allows for the exploration of much fainter sources than those probed by the relatively shallow blind 2 mm surveys used in Bing et al. (2023) and Zavala et al. (2021). We note that, despite still agreeing within the uncertainties, the number counts inferred here are systematically higher than the single-dish number counts from Magnelli et al. (2019) and Bing et al. (2023). This suggests that the resolution limitations inherent to single-dish instruments are not negligible at long wavelengths, even with the application of certain corrections.

In Band 6, we compare our results to number counts from the ALMA blind surveys from Hatsukade et al. (2018, ASAGAO survey), González-López et al. (2020, ASPECS-LP survey at 1.2 mm) and Gómez-Guijarro et al. (2022, GOODS-ALMA survey), to number counts from an ALMA protocluster field with removed cluster galaxies from Umehata et al. (2017), to the number counts from ALMA calibrator pointings from Chen et al. (2023), and to number counts from single-dish blind surveys from Lindner et al. (2011), Scott et al. (2012) and Bing et al. (2023). There is an overall very good agreement between our number counts and these previous studies. Our number counts bridge the flux density range between the deepest ALMA blind survey and the deepest single-dish surveys, reaching nearly as deep as ASPECS-LP but covering a wider area than any individual ALMA blind survey. In the bright end of Band 6, our number counts favour the results from the single-dish surveys over the ALMA-based results from Gómez-Guijarro et al. (2022), the latter likely being affected by large cosmic variance due to a small areal coverage. In the faint end, our number counts agree with those from González-López et al. (2020), although we do not reach quite the same depth as ASPECS-LP. This difference in depth is attributed to the higher  $S/N_{\text{peak}}$  source extraction threshold chosen in our work ( $S/N_{\text{peak}} = 5.4$  instead of 3.3 in González-López et al. 2020).

Band 7 is the only band in which only single pointings and no blind surveys are available in our database. We compare our results with ALMA-based number counts from Oteo et al. (2016), Béthermin et al. (2020), and Chen et al. (2023) who all used single pointings with masked-out targets, and from Stach et al. (2018) and Simpson et al. (2020), which are based on single-dish follow-up studies. The number counts are largely consistent with the literature within the error bars. Our number counts are systematically a bit lower than the ones from Béthermin et al. (2020) that are based on their full sample, only masking a circular region of radius 1 arcsec around the phase centre of the ALPINE pointings. However, our number counts are in almost perfect agreement with those inferred from their “secure” sample which is furthermore limited to sources at  $z < 4$  with an optical or near-IR counterpart. Even though this “secure” sample is meant to provide a lower limit, the resulting number counts are not systematically lower than ours. Unfortunately,

**Table 4.** Fraction of the CIB resolved into individual sources in our A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS databases.

ALMA Band	$S_{\text{total,ALMA}}$ [Jy deg <sup>-2</sup> ]	$S_{\text{total,COBE}}$ [Jy deg <sup>-2</sup> ]	Percentage resolved
Band 3 <sup>(*)</sup>	$0.08^{+0.02}_{-0.02}$	$2.2 \pm 1.1$	$4^{+2}_{-2}$
Band 4	$1.0^{+0.6}_{-0.3}$	$5.5 \pm 2.6$	$18^{+14}_{-10}$
Band 6	$6.6^{+0.5}_{-0.3}$	$19.1 \pm 8.2$	$34^{+15}_{-15}$
Band 7	$14.5^{+5.6}_{-2.1}$	$36.3 \pm 14.8$	$40^{+22}_{-17}$

**Notes.** Total flux density per unit area inferred from A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS number counts and CIB emission measured with COBE (Fixsen et al. 1998) for ALMA Bands 3 (2998  $\mu\text{m}$ ), 4 (2082  $\mu\text{m}$ ), 6 (1233  $\mu\text{m}$ ), and 7 (925  $\mu\text{m}$ ). The last column gives the fraction of the CIB that is resolved into discrete sources using our number counts. <sup>(\*)</sup>Band 3 estimates for COBE are an extrapolation as they lie outside of the data range constrained in Fixsen et al. (1998).

the A<sup>3</sup>COSMOS database does not yet probe the bright end in Band 7 as far as the single-dish follow-up studies, because we do not probe large enough areas to serendipitously detect those bright sources.

For sanity check, we also report the number counts separately for blind surveys and single pointings (see Appendix E). We find that both results are consistent within the uncertainties in their common flux density ranges. This confirms the practical applicability of our blinding method for single pointings.

#### 4.3. Contribution to the cosmic infrared background

Using our number counts, we can calculate the total flux density per unit area as emitted by these sources with  $S/N_{\text{peak}} \geq 5.4$  for the ALMA Bands 3, 4, 6, and 7. This is done by summing over the redshift-weighted (see Sect. 3.1.2) total flux densities of all sources in each band divided by the respective effective area:

$$S_{\text{total,ALMA}} = \sum_i \frac{S_i \cdot (1 - P_{\text{pair},i})}{A_{\text{eff}}(S_i, \theta_i)}. \quad (8)$$

This allows us to determine how much of the cosmic infrared background (CIB) is resolved into discrete sources at each wavelength. As a reference for the intensity of the CIB, we took the analytical fit by Fixsen et al. (1998) to the far-IR background light as measured by COBE/FIRAS observations. This relation was fitted on the wavelength range between 125  $\mu\text{m}$  and 2000  $\mu\text{m}$ , which covers our Bands 4, 6, and 7. Even though Band 3 at  $\sim 3$  mm is outside of this range, we still used the Fixsen et al. (1998) fit as a comparative value, although we emphasise that this is an extrapolation. The total flux densities per unit area emitted by all sources detected in a given band ( $S_{\text{total,ALMA}}$ ) as well as the ones inferred from the COBE data ( $S_{\text{total,COBE}}$ ) are listed in Table 4, in addition to the resulting fraction of the CIB that is thus resolved into discrete ALMA sources.

The percentage of the CIB resolved decreases notably with the increase in wavelength. While in Band 7, we resolve  $\sim 40\%$  of the CIB, in Band 6 it is  $\sim 34\%$  and in Band 4 only  $\sim 18\%$ . From our extrapolation in Band 3, the resolved fraction is even below 10%. We note, however, that in all cases the uncertainties are large, not least due to the already large uncertainties on the Fixsen et al. (1998) fit. This decline with longer wavelengths is probably attributed to the fact that ALMA is mainly probing the Rayleigh–Jeans tail of the dust emission of DSFGs, which gets

progressively fainter going further into the millimetre regime. Although the  $\sim 2\text{--}3$  mm wavelength regime serves as a filter for very high-redshift SFGs (Casey et al. 2018a,b, 2021), detecting the faint emission of these DSFGs in the millimetre still requires very long integration times, even with the good sensitivity of ALMA. Therefore, at these long wavelengths, we predominantly detect the brighter end of the population of DSFGs, missing out on most of the fainter ones which constitute the bulk of the total emission. Our sample is thus not representative of the majority of this population, so there is still a lot of information missing on their dust emissivity and general properties. Future observations with instruments of even higher sensitivity than ALMA at millimetre and longer wavelengths, for example the planned “Next Generation VLA”, will also aid in improving our understanding of this population by detecting more of these faint galaxies.

The fraction of the CIB resolved by ALMA into individual sources in Band 7 is higher than typical single-dish surveys achieve at similar wavelengths ( $\sim 20\text{--}30\%$ , e.g. Eales et al. 1999; Coppin et al. 2006), but consistent within the errors with results from Smail et al. (2002) where they resolve 60% using lensed sources. In Band 6, our resolved fraction is higher than single-dish results at 1.1 mm from Scott et al. (2010) and Hatsukade et al. (2011) who both resolve around 6–10%, but in agreement with Viero et al. (2013) who find  $45 \pm 8\%$  via a stacking analysis.

Not only is ALMA already largely outperforming single-dish observations in resolving the CIB into individual sources, but it is currently only limited by integration time and can thus potentially resolve even higher fractions by performing deeper observations. Single-dish telescopes provide precise measurements of the brighter populations of DSFGs, but are already mostly at their confusion limits. These limits are impossible to circumvent outside of using lensed sources.

#### 4.4. Comparison with number counts models

We plot our number counts in comparison to the models from Zavala et al. (2021), from the SPRITZ simulation of Bisigello et al. (2021), and from the SIDES simulation of Gkogkou et al. (2023). The Zavala et al. (2021) model is based on galaxy dust SEDs that were combined with an IR luminosity function (IRLF) with some free parameters to create mock sky maps matching the number counts at 1.2 mm, 2 mm, and 3 mm available at that time. The SPRITZ simulation (Bisigello et al. 2021) is based on a set of IRLFs for different galaxy populations and their corresponding semi-empirical SED templates from which physical galaxy properties and fluxes were derived to create simulated catalogues and images. The SIDES simulation (Gkogkou et al. 2023, see also Béthermin et al. 2017) was built from dark matter lightcones that were populated with galaxies based on a stellar mass function using an abundance matching technique. These galaxies were randomly split into star-forming and passive populations based on a stellar mass- and redshift-dependent probability, derived from the observed evolution of the star-forming fraction (Davidzon et al. 2017). They were subsequently assigned SFRs and SEDs based on their stellar mass and type (either main sequence galaxies or starbursts) to infer their observable flux densities in several filters. The model also considers lensing (de-)magnification and clustering properties. The comparison of the models with our measured number counts is shown in Fig. 8.

The predictions from the Zavala et al. (2021) model agree well with our measured number counts, thus their galaxy evolution model provides an overall good representation of the

(sub)millimetre galaxy population. This is not entirely surprising as this phenomenological model was developed to reproduce the number count from real 1.2 mm, 2 mm, and 3 mm data available at that time. Nevertheless, there are notable differences in Band 6 in the intermediate flux density range (i.e. 0.2–2.0 mJy) and Band 7 in the high flux density range ( $>4$  mJy). These differences are probably due to the limited number counts available at that time used for fitting the model in Zavala et al. (2021). In particular, in Band 6, the only constraints used were the ASPECS-LP number counts below  $\sim 0.4$  mJy. Higher flux densities in Band 6 or shorter wavelengths were not constrained with observational data. Despite these slight differences, the model of Zavala et al. (2021) provides a reasonable representation of the (sub)millimetre galaxy population. It is noteworthy that predictions at low flux density in Band 4, where no constraints were available at the time of creation of this model, are fully consistent with our measurements. This highlights the predictive power of this model.

The predicted number counts from SPRITZ are in good agreement with our measurements in Band 7. Although consistent with our number counts within the uncertainties in Band 6 and the faint end of Band 4, SPRITZ predicts significantly higher number counts (by a factor  $\sim 5\text{--}6$ ) in the bright end of Band 4 and over the full range of Band 3. The SPRITZ simulation and its underlying SED templates are derived from *Herschel*-based optical-to-IR (up to  $\sim 500$   $\mu\text{m}$ ; Gruppioni et al. 2013) data and optimised to make predictions in this specific regime. The (sub)millimetre regime, however, was not explicitly taken into account in the construction of the simulation and thus it is not well constrained at significantly longer wavelengths.

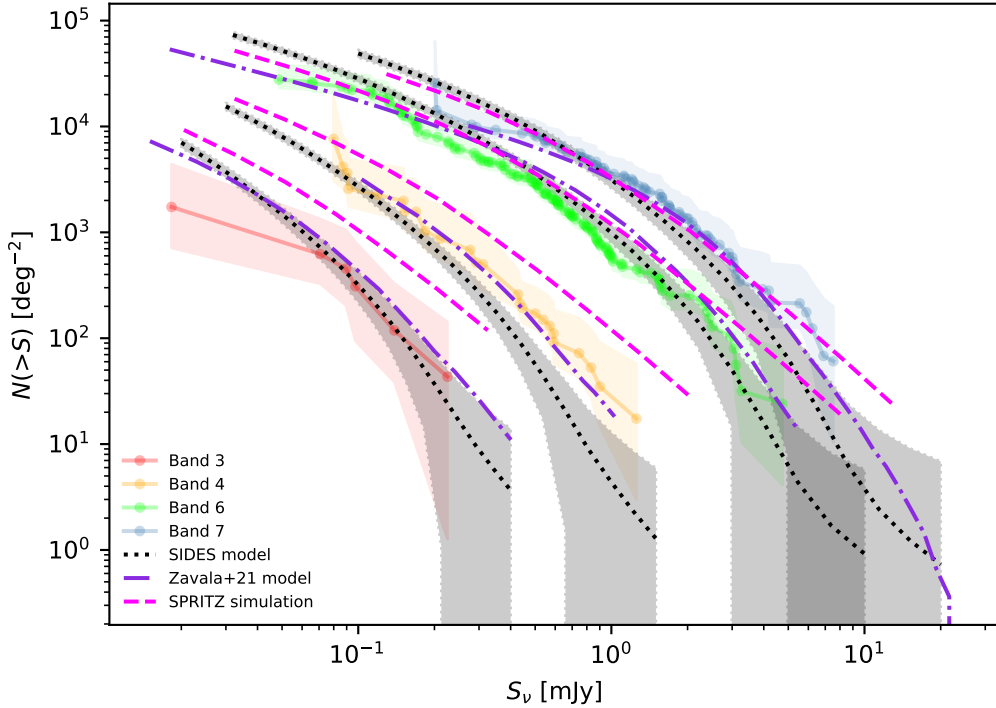
The SIDES model agrees with our measured number counts within the uncertainties. This is remarkable considering that SIDES was not explicitly fitted to any measured (sub)millimetre number counts. We notice, however, that the model tends to predict systematically lower number counts at the bright end of Bands 4 and 7. A slight under-prediction of SIDES at the bright end of the galaxy number counts was previously noted by Bing et al. (2023), who suggested a lack of either cold dust in the underlying SEDs or of galaxies with high SFRs.

## 5. Summary

In this work we presented the updated  $A^3\text{COSMOS}$  database, including the newly added  $A^3\text{GOODSS}$ , in data version 20220606. This more than doubles the size of the  $A^3\text{COSMOS}$  database compared to the previous release. With a robust sample of  $\sim 1400$  galaxies over a wide redshift range and with SED-inferred information such as SFR and gas or dust masses, it is a rich resource for studies on (dusty) galaxy evolution. This database, which includes the continuum images, blind and prior photometry catalogues, and value-added galaxy catalogues, is made public to the community and will receive regular updates.

We used the blind continuum detections from this database to infer in a homogeneous way, galaxy number counts across all major ALMA bands. To reduce the inhomogeneous observational bias inherent to the  $A^3\text{COSMOS}$  and  $A^3\text{GOODSS}$  databases, we used a conservative combination of selections and corrections: we restricted our analysis to blind surveys or single targeted pointings with a secure target at their phase centre, masked a circular area around the phase centre of these single pointings, and weighted any source detected outside of this mask by redshift proximity to the target. Finally, we used simulations based on the SIDES model (Béthermin et al. 2017; Gkogkou et al. 2023), to confirm the applicability of our method.





**Fig. 8.** Cumulative number counts as shown in Fig. 7 in comparison with the models from Zavala et al. (2021), Bisigello et al. (2021; i.e. the SPRITZ simulation), and Gkogkou et al. (2023; i.e. the SIDES model). The shaded regions of the SIDES curves indicate one standard deviation due to cosmic variance for an area corresponding to our coverage in each band.

Thanks to the sensitivity of ALMA and the large areal coverage offered by the combination of all ALMA archive projects in the COSMOS and GOODS-S fields, the number counts inferred here bridge the flux density gap between single-dish observations, which constrain only the bright end of the number counts, and pencil beam ALMA blind surveys, which constrain the faint end of the number counts (albeit with large uncertainties due to cosmic variance). In the flux density ranges in common, our number counts are overall in good agreement with those retrieved from the literature. Finally, down to the flux density limits of our number counts, about 40%, 34%, 18%, and 4% of the CIB is resolved into discrete sources in Bands 7, 6, 4, and 3, respectively. Thus, while the emission of DSFGs is well constrained at 0.9–1.3 mm wavelengths, it remains poorly understood at longer wavelengths. However, as ALMA observations are far less limited by confusion noise than single-dish facilities, future deeper observations with ALMA will undoubtedly allow us to resolve an ever larger fraction of the CIB at these wavelengths.

The homogeneous measurements of the number counts in ALMA Bands 3, 4, 6, and 7 over such broad flux density ranges offer unique constraints for galaxy evolution models. We thus compared our number counts with predictions from the models of Zavala et al. (2021), SPRITZ (Bisigello et al. 2021), and SIDES (B  thermin et al. 2017; Gkogkou et al. 2023). We find that our results are in very good agreement with the predictions from the model of Zavala et al. (2021). The SPRITZ model is consistent with our number counts in Bands 6 and 7, but its predictions are significantly higher than the observations at longer wavelengths (i.e. Bands 3 and 4). Our number counts are overall in good agreement with those predicted by the SIDES simulation, even though this one was not explicitly fitted to any measured (sub)millimetre number counts. However, this simulation tends to slightly underestimate the number of bright sources in Bands 4 and 7.

We have shown that it is possible to infer number counts from the heterogeneous target-driven A<sup>3</sup>COSMOS database. This method naturally has limitations; however, the effects on the number counts are negligible compared to the effects of cosmic variance. Our analysis therefore profits from the large areal coverage of the A<sup>3</sup>COSMOS database and the availability of blind surveys within it. Future updates will allow us to decrease the cosmic variance and further improve the accuracy of the measured number counts.

*Acknowledgements.* S.A. would like to thank Laura Bisigello and Jorge Zavala for providing the curves of their respective models. F.B. and S.A. gratefully acknowledge the Collaborative Research Center 1601 (SFB 1601 sub-project C2) funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 500700252. I.D. acknowledges support from INAF Minigrant “Harnessing the power of VLBA towards a census of AGN and star formation at high redshift”. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2011.0.00064.S, ADS/JAO.ALMA#2011.0.00097.S, ADS/JAO.ALMA#2011.0.00539.S, ADS/JAO.ALMA#2011.0.00742.S, ADS/JAO.ALMA#2012.1.00076.S, ADS/JAO.ALMA#2012.1.00173.S, ADS/JAO.ALMA#2012.1.00307.S, ADS/JAO.ALMA#2012.1.00323.S, ADS/JAO.ALMA#2012.1.00523.S, ADS/JAO.ALMA#2012.1.00536.S, ADS/JAO.ALMA#2012.1.00775.S, ADS/JAO.ALMA#2012.1.00869.S, ADS/JAO.ALMA#2012.1.00919.S, ADS/JAO.ALMA#2012.1.00952.S, ADS/JAO.ALMA#2012.1.00978.S, ADS/JAO.ALMA#2012.1.00983.S, ADS/JAO.ALMA#2013.1.00034.S, ADS/JAO.ALMA#2013.1.00092.S, ADS/JAO.ALMA#2013.1.00118.S, ADS/JAO.ALMA#2013.1.00139.S, ADS/JAO.ALMA#2013.1.00146.S, ADS/JAO.ALMA#2013.1.00151.S, ADS/JAO.ALMA#2013.1.00171.S, ADS/JAO.ALMA#2013.1.00205.S, ADS/JAO.ALMA#2013.1.00208.S, ADS/JAO.ALMA#2013.1.00250.S, ADS/JAO.ALMA#2013.1.00276.S, ADS/JAO.ALMA#2013.1.00470.S, ADS/JAO.ALMA#2013.1.00668.S, ADS/JAO.ALMA#2013.1.00718.S, ADS/JAO.ALMA#2013.1.00786.S, ADS/JAO.ALMA#2013.1.00815.S, ADS/JAO.ALMA#2013.1.00836.S, ADS/JAO.ALMA#2013.1.00884.S, ADS/JAO.ALMA#2013.1.00914.S, ADS/JAO.ALMA#2013.1.01258.S, ADS/JAO.ALMA#2013.1.01271.S, ADS/JAO.ALMA#2013.1.01292.S, ADS/JAO.ALMA#2015.1.00026.S, ADS/JAO.ALMA#2015.1.00039.S, ADS/JAO.ALMA#2015.1.00040.S, ADS/JAO.ALMA#2015.1.00055.S, ADS/JAO.ALMA#2015.1.00098.S, ADS/JAO.ALMA#2015.1.00122.S.

ADS/JAO.ALMA#2015.1.00137.S, ADS/JAO.ALMA#2015.1.00220.S, ADS/JAO.ALMA#2015.1.00260.S, ADS/JAO.ALMA#2015.1.00379.S, ADS/JAO.ALMA#2015.1.00540.S, ADS/JAO.ALMA#2015.1.00568.S, ADS/JAO.ALMA#2015.1.00704.S, ADS/JAO.ALMA#2015.1.00853.S, ADS/JAO.ALMA#2015.1.00862.S, ADS/JAO.ALMA#2015.1.00907.S, ADS/JAO.ALMA#2015.1.00948.S, ADS/JAO.ALMA#2015.1.01096.S, ADS/JAO.ALMA#2015.1.01111.S, ADS/JAO.ALMA#2015.1.01205.S, ADS/JAO.ALMA#2015.1.01362.S, ADS/JAO.ALMA#2015.1.01447.S, ADS/JAO.ALMA#2015.1.01590.S, ADS/JAO.ALMA#2015.A.00026.S, ADS/JAO.ALMA#2016.1.00048.S, ADS/JAO.ALMA#2016.1.00279.S, ADS/JAO.ALMA#2016.1.00330.S, ADS/JAO.ALMA#2016.1.00478.S, ADS/JAO.ALMA#2016.1.00564.S, ADS/JAO.ALMA#2016.1.00624.S, ADS/JAO.ALMA#2016.1.00721.S, ADS/JAO.ALMA#2016.1.00735.S, ADS/JAO.ALMA#2016.1.00790.S, ADS/JAO.ALMA#2016.1.00804.S, ADS/JAO.ALMA#2016.1.00954.S, ADS/JAO.ALMA#2016.1.00990.S, ADS/JAO.ALMA#2016.1.01012.S, ADS/JAO.ALMA#2016.1.01079.S, ADS/JAO.ALMA#2016.1.01184.S, ADS/JAO.ALMA#2016.1.01240.S, ADS/JAO.ALMA#2016.1.01426.S, ADS/JAO.ALMA#2016.1.01512.S, ADS/JAO.ALMA#2016.1.01559.S, ADS/JAO.ALMA#2017.1.00001.S, ADS/JAO.ALMA#2017.1.00127.S, ADS/JAO.ALMA#2017.1.00190.S, ADS/JAO.ALMA#2017.1.00300.S, ADS/JAO.ALMA#2017.1.00373.S, ADS/JAO.ALMA#2017.1.00428.L, ADS/JAO.ALMA#2017.1.00604.S, ADS/JAO.ALMA#2017.1.00856.S, ADS/JAO.ALMA#2017.1.01020.S, ADS/JAO.ALMA#2017.1.01099.S, ADS/JAO.ALMA#2017.1.01176.S, ADS/JAO.ALMA#2017.1.01259.S, ADS/JAO.ALMA#2017.1.01347.S, ADS/JAO.ALMA#2017.1.01359.S, ADS/JAO.ALMA#2017.1.01471.S, ADS/JAO.ALMA#2017.1.01618.S, ADS/JAO.ALMA#2017.1.01674.S, ADS/JAO.ALMA#2017.1.01713.S, ADS/JAO.ALMA#2017.A.00032.S, ADS/JAO.ALMA#2018.1.00081.S, ADS/JAO.ALMA#2018.1.00164.S, ADS/JAO.ALMA#2018.1.00222.S, ADS/JAO.ALMA#2018.1.00236.S, ADS/JAO.ALMA#2018.1.00329.S, ADS/JAO.ALMA#2018.1.00429.S, ADS/JAO.ALMA#2018.1.00543.S, ADS/JAO.ALMA#2018.1.00570.S, ADS/JAO.ALMA#2018.1.00681.S, ADS/JAO.ALMA#2018.1.00876.S, ADS/JAO.ALMA#2018.1.00938.S, ADS/JAO.ALMA#2018.1.01044.S, ADS/JAO.ALMA#2018.1.01128.S, ADS/JAO.ALMA#2018.1.01225.S, ADS/JAO.ALMA#2018.1.01359.S, ADS/JAO.ALMA#2018.1.01536.S, ADS/JAO.ALMA#2018.1.01594.S, ADS/JAO.ALMA#2018.1.01673.S, ADS/JAO.ALMA#2018.1.01824.S, ADS/JAO.ALMA#2018.1.01852.S, ADS/JAO.ALMA#2018.A.00022.S, ADS/JAO.ALMA#2019.1.00074.S, ADS/JAO.ALMA#2019.1.00207.S, ADS/JAO.ALMA#2019.1.00228.S, ADS/JAO.ALMA#2019.1.00299.S, ADS/JAO.ALMA#2019.1.00388.S, ADS/JAO.ALMA#2019.1.00543.S, ADS/JAO.ALMA#2019.1.00664.S, ADS/JAO.ALMA#2019.1.00821.S, ADS/JAO.ALMA#2019.1.00861.S, ADS/JAO.ALMA#2019.1.00870.S, ADS/JAO.ALMA#2019.1.00928.S, ADS/JAO.ALMA#2019.1.01074.S, ADS/JAO.ALMA#2019.1.01105.S, ADS/JAO.ALMA#2019.1.01171.S, ADS/JAO.ALMA#2019.1.01212.S, ADS/JAO.ALMA#2019.1.01379.S, ADS/JAO.ALMA#2019.1.01495.S, ADS/JAO.ALMA#2019.A.00009.S, ADS/JAO.ALMA#2016.1.00012.S, ADS/JAO.ALMA#2016.1.00171.S, ADS/JAO.ALMA#2016.1.00324.L, ADS/JAO.ALMA#2016.1.00463.S, ADS/JAO.ALMA#2016.1.00534.S, ADS/JAO.ALMA#2016.1.00567.S, ADS/JAO.ALMA#2016.1.00646.S, ADS/JAO.ALMA#2016.1.00726.S, ADS/JAO.ALMA#2016.1.00778.S, ADS/JAO.ALMA#2016.1.00798.S, ADS/JAO.ALMA#2016.1.00932.S, ADS/JAO.ALMA#2016.1.00967.S, ADS/JAO.ALMA#2016.1.01001.S, ADS/JAO.ALMA#2016.1.01040.S, ADS/JAO.ALMA#2016.1.01155.S, ADS/JAO.ALMA#2016.1.01208.S, ADS/JAO.ALMA#2016.1.01355.S, ADS/JAO.ALMA#2016.1.01454.S, ADS/JAO.ALMA#2016.1.01546.S, ADS/JAO.ALMA#2016.1.01604.S, ADS/JAO.ALMA#2017.1.00046.S, ADS/JAO.ALMA#2017.1.00138.S, ADS/JAO.ALMA#2017.1.00270.S, ADS/JAO.ALMA#2017.1.00326.S, ADS/JAO.ALMA#2017.1.00413.S, ADS/JAO.ALMA#2017.1.00486.S, ADS/JAO.ALMA#2017.1.00755.S, ADS/JAO.ALMA#2017.1.00893.S, ADS/JAO.ALMA#2017.1.01027.S, ADS/JAO.ALMA#2017.1.01163.S, ADS/JAO.ALMA#2017.1.01217.S, ADS/JAO.ALMA#2017.1.01276.S, ADS/JAO.ALMA#2017.1.01358.S, ADS/JAO.ALMA#2017.1.01451.S, ADS/JAO.ALMA#2017.1.01512.S, ADS/JAO.ALMA#2017.1.01659.S, ADS/JAO.ALMA#2017.1.01677.S, ADS/JAO.ALMA#2017.A.00013.S, ADS/JAO.ALMA#2017.A.00034.S, ADS/JAO.ALMA#2018.1.00085.S, ADS/JAO.ALMA#2018.1.00216.S, ADS/JAO.ALMA#2018.1.00231.S, ADS/JAO.ALMA#2018.1.00251.S, ADS/JAO.ALMA#2018.1.00348.S, ADS/JAO.ALMA#2018.1.00478.S, ADS/JAO.ALMA#2018.1.00567.S, ADS/JAO.ALMA#2018.1.00635.S, ADS/JAO.ALMA#2018.1.00874.S, ADS/JAO.ALMA#2018.1.00933.S, ADS/JAO.ALMA#2018.1.00992.S, ADS/JAO.ALMA#2018.1.01079.S, ADS/JAO.ALMA#2018.1.01136.S, ADS/JAO.ALMA#2018.1.01281.S, ADS/JAO.ALMA#2018.1.01521.S, ADS/JAO.ALMA#2018.1.01551.S, ADS/JAO.ALMA#2018.1.01605.S, ADS/JAO.ALMA#2018.1.01739.S, ADS/JAO.ALMA#2018.1.01841.S, ADS/JAO.ALMA#2018.1.01871.S, ADS/JAO.ALMA#2018.A.00037.S, ADS/JAO.ALMA#2019.1.00102.S, ADS/JAO.ALMA#2019.1.00151.S, ADS/JAO.ALMA#2019.1.00399.S, ADS/JAO.ALMA#2019.1.00477.S, ADS/JAO.ALMA#2019.1.00678.S, ADS/JAO.ALMA#2019.1.00909.S, ADS/JAO.ALMA#2019.1.01075.S, ADS/JAO.ALMA#2019.1.01142.S, ADS/JAO.ALMA#2019.1.01286.S, ADS/JAO.ALMA#2019.1.01491.S, ADS/JAO.ALMA#2019.1.01528.S, ADS/JAO.ALMA#2019.1.01600.S, ADS/JAO.ALMA#2019.1.01634.L, ADS/JAO.ALMA#2019.1.01722.S, ADS/JAO.ALMA#2019.2.00118.S, ADS/JAO.ALMA#2019.2.00246.S, ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

## References

Aravena, M., Decarli, R., González-López, J., et al. 2019, *ApJ*, **882**, 136  
 Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2013, *ApJ*, **769**, 80  
 Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2015, *ApJS*, **218**, 33  
 Ashby, M. L. N., Caputi, K. I., Cowley, W., et al. 2018, *ApJS*, **237**, 39  
 Balestra, I., Mainieri, V., Popesso, P., et al. 2010, *A&A*, **512**, A12  
 Beckwith, S. V. W., Stiuvelii, M., Koekemoer, A. M., et al. 2006, *AJ*, **132**, 1729  
 Béthermin, M., Daddi, E., Magdis, G., et al. 2012, *ApJ*, **757**, L23  
 Béthermin, M., Wu, H.-Y., Lagache, G., et al. 2017, *A&A*, **607**, A89  
 Béthermin, M., Fudamoto, Y., Ginolfi, M., et al. 2020, *A&A*, **643**, A2  
 Bing, L., Béthermin, M., Lagache, G., et al. 2023, *A&A*, **677**, A66  
 Bisigello, L., Gruppioni, C., Feltre, A., et al. 2021, *A&A*, **651**, A52  
 Brisbin, D., Miettinen, O., Aravena, M., et al. 2017, *A&A*, **608**, A15  
 Capak, P. L., Carilli, C., Jones, G., et al. 2015, *Nature*, **522**, 455  
 Casey, C. M., Chen, C.-C., Cowie, L. L., et al. 2013, *MNRAS*, **436**, 1919  
 Casey, C. M., Narayanan, D., & Cooray, A. 2014, *Phys. Rep.*, **541**, 45  
 Casey, C. M., Zavala, J. A., Spilker, J., et al. 2018a, *ApJ*, **862**, 77  
 Casey, C. M., Hodge, J., Zavala, J. A., et al. 2018b, *ApJ*, **862**, 78  
 Casey, C. M., Zavala, J. A., Manning, S. M., et al. 2021, *ApJ*, **923**, 215  
 Casey, C. M., Kartaltepe, J. S., Drakos, N. E., et al. 2023, *ApJ*, **954**, 31  
 Chabrier, G. 2003, *PASP*, **115**, 763  
 Chen, J., Ivison, R. J., Zwaan, M. A., et al. 2023, *MNRAS*, **518**, 1378  
 Coppin, K., Chapin, E. L., Mortier, A. M. J., et al. 2006, *MNRAS*, **372**, 1621  
 Croom, S. M., Warren, S. J., & Glazebrook, K. 2001, *MNRAS*, **328**, 150  
 da Cunha, E., Charlot, S., & Elbaz, D. 2008, *MNRAS*, **388**, 1595  
 da Cunha, E., Walter, F., Smail, I. R., et al. 2015, *ApJ*, **806**, 110  
 Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, *A&A*, **605**, A70  
 Decarli, R., Walter, F., González-López, J., et al. 2019, *ApJ*, **882**, 138  
 De Looze, I., Baes, M., Bendo, G. J., Cortese, L., & Fritz, J. 2011, *MNRAS*, **416**, 2712  
 Delvecchio, I., Smolčić, V., Zamorani, G., et al. 2017, *A&A*, **602**, A3  
 Dickinson, M., Giavalisco, M., & GOODS Team 2003, in *The Mass of Galaxies at Low and High Redshift*, eds. R. Bender, & A. Renzini (Springer-Verlag), 324  
 Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2021, *MNRAS*, **500**, 942  
 Eales, S., Lilly, S., Gear, W., et al. 1999, *ApJ*, **515**, 518  
 Fernández-Soto, A., Lanzetta, K. M., Chen, H.-W., Pascarelle, S. M., & Yahata, N. 2001, *ApJS*, **135**, 41  
 Fixsen, D. J., Dwek, E., Mather, J. C., Bennett, C. L., & Shafer, R. A. 1998, *ApJ*, **508**, 123  
 Franco, M., Elbaz, D., Béthermin, M., et al. 2018, *A&A*, **620**, A152  
 Gehrels, N. 1986, *ApJ*, **303**, 336  
 Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJ*, **600**, L93  
 Ginsburg, A., Sipőcz, B. M., Bressan, C. E., et al. 2019, *AJ*, **157**, 98  
 Gkogkou, A., Béthermin, M., Lagache, G., et al. 2023, *A&A*, **670**, A16  
 Gómez-Guijarro, C., Elbaz, D., Xiao, M., et al. 2022, *A&A*, **658**, A43  
 González-López, J., Decarli, R., Pavesi, R., et al. 2019, *ApJ*, **882**, 139  
 González-López, J., Novak, M., Decarli, R., et al. 2020, *ApJ*, **897**, 91  
 Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, **197**, 35  
 Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, *MNRAS*, **432**, 23  
 Gruppioni, C., Béthermin, M., Loiacono, F., et al. 2020, *A&A*, **643**, A8  
 Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, **207**, 24  
 Hatsukade, B., Kohno, K., Aretxaga, I., et al. 2011, *MNRAS*, **411**, 102  
 Hatsukade, B., Kohno, K., Yamaguchi, Y., et al. 2018, *PASP*, **70**, 105

- Hodge, J. A., Karim, A., Smail, I., et al. 2013, *ApJ*, 768, 91
- Hsieh, B.-C., Wang, W.-H., Hsieh, C.-C., et al. 2012, *ApJS*, 203, 23
- Hughes, T. M., Ibar, E., Villanueva, V., et al. 2017, *MNRAS*, 468, L103
- Illingworth, G., Magee, D., Bouwens, R., et al. 2016, arXiv e-prints [arXiv:1606.00841]
- Jin, S., Daddi, E., Liu, D., et al. 2018, *ApJ*, 864, 56
- Karim, A., Swinbank, A. M., Hodge, J. A., et al. 2013, *MNRAS*, 432, 2
- Keating, G. K., Marrone, D. P., Bower, G. C., & Keenan, R. P. 2020, *ApJ*, 901, 141
- Kennicutt, R. C., Jr. 1998, *ApJ*, 498, 541
- Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, *ApJS*, 172, 196
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36
- Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, *ApJS*, 218, 15
- Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, *ApJS*, 224, 24
- Lee, N., Sheth, K., Scott, K. S., et al. 2017, *MNRAS*, 471, 2124
- Le Fèvre, O., Vettolani, G., Paltani, S., et al. 2004, *A&A*, 428, 1043
- Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, *A&A*, 559, A14
- Le Flocc'h, E., Aussel, H., Ilbert, O., et al. 2009, *ApJ*, 703, 222
- Lindner, R. R., Baker, A. J., Omont, A., et al. 2011, *ApJ*, 737, 83
- Liu, D., Gao, Y., Isaak, K., et al. 2015, *ApJ*, 810, L14
- Liu, D., Lang, P., Magnelli, B., et al. 2019a, *ApJS*, 244, 40
- Liu, D., Schinnerer, E., Groves, B., et al. 2019b, *ApJ*, 887, 235
- Madau, P., & Dickinson, M. 2014, *ARA&A*, 52, 415
- Magnelli, B., Popesso, P., Berta, S., et al. 2013, *A&A*, 553, A132
- Magnelli, B., Karim, A., Staguhn, J., et al. 2019, *ApJ*, 877, 45
- Marsden, D., Gralla, M., Marriage, T. A., et al. 2014, *MNRAS*, 439, 1556
- Massey, R., Stoughton, C., Leauthaud, A., et al. 2010, *MNRAS*, 401, 371
- McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, *A&A*, 544, A156
- McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in *Astronomical Data Analysis Software and Systems XVI*, eds. R. A. Shaw, F. Hill, & D. J. Bell, *ASP Conf. Ser.*, 376, 127
- Mignoli, M., Cimatti, A., Zamorani, G., et al. 2005, *A&A*, 437, 883
- Miller, N. A., Fomalont, E. B., Kellermann, K. I., et al. 2008, *ApJS*, 179, 114
- Mohan, N., & Rafferty, D. 2015, *Astrophysics Source Code Library* [record ascl:1502.007]
- Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, *ApJS*, 225, 27
- Moneti, A., McCracken, H. J., Shuntov, M., et al. 2022, *A&A*, 658, A126
- Morris, A. M., Kocevski, D. D., Trump, J. R., et al. 2015, *AJ*, 149, 178
- Oteo, I., Zwaan, M. A., Ivison, R. J., Smail, I., & Biggs, A. D. 2016, *ApJ*, 822, 36
- Pavesi, R., Riechers, D. A., Sharon, C. E., et al. 2018, *ApJ*, 861, 43
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *J. Mach. Learn. Res.*, 12, 2825
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, *AJ*, 124, 266
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, *AJ*, 139, 2097
- Popping, G., Walter, F., Behroozi, P., et al. 2020, *ApJ*, 891, 135
- Ravikumar, C. D., Puech, M., Flores, H., et al. 2007, *A&A*, 465, 1099
- Riechers, D. A., Carilli, C. L., Capak, P. L., et al. 2014, *ApJ*, 796, 84
- Salvato, M., Ilbert, O., Hasinger, G., et al. 2011, *ApJ*, 742, 61
- Sargent, M. T., Daddi, E., Béthermin, M., et al. 2014, *ApJ*, 793, 19
- Scott, K. S., Yun, M. S., Wilson, G. W., et al. 2010, *MNRAS*, 405, 2260
- Scott, K. S., Wilson, G. W., Aretxaga, I., et al. 2012, *MNRAS*, 423, 575
- Scoville, N., Aussel, H., Brusa, M., et al. 2007, *ApJS*, 172, 1
- Scoville, N., Aussel, H., Sheth, K., et al. 2014, *ApJ*, 783, 84
- Scoville, N., Lee, N., Vanden Bout, P., et al. 2017, *ApJ*, 837, 150
- Simpson, J. M., Smail, I., Swinbank, A. M., et al. 2019, *ApJ*, 880, 43
- Simpson, J. M., Smail, I., Dudzevičiūtė, U., et al. 2020, *MNRAS*, 495, 3409
- Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, *ApJS*, 214, 24
- Smail, I., Ivison, R. J., Blain, A. W., & Kneib, J. P. 2002, *MNRAS*, 331, 495
- Smolčić, V., Karim, A., Miettinen, O., et al. 2015, *A&A*, 576, A127
- Smolčić, V., Novak, M., Bondi, M., et al. 2017, *A&A*, 602, A1
- Stach, S. M., Smail, I., Swinbank, A. M., et al. 2018, *ApJ*, 860, 161
- Stach, S. M., Dudzevičiūtė, U., Smail, I., et al. 2019, *MNRAS*, 487, 4648
- Straatman, C. M. S., Spitler, L. R., Quadri, R. F., et al. 2016, *ApJ*, 830, 51
- Talia, M., Cimatti, A., Giuliotti, M., et al. 2021, *ApJ*, 909, 23
- Tasca, L. A. M., Le Fèvre, O., Ribeiro, B., et al. 2017, *A&A*, 600, A110
- Traina, A., Gruppioni, C., Delvecchio, I., et al. 2024, *A&A*, 681, A118
- Umehata, H., Tamura, Y., Kohno, K., et al. 2017, *ApJ*, 835, 98
- Urrutia, T., Wisotzki, L., Kerutt, J., et al. 2019, *A&A*, 624, A141
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2008, *A&A*, 478, 83
- Viero, M. P., Moncelsi, L., Quadri, R. F., et al. 2013, *ApJ*, 779, 32
- Wang, W.-H., Lin, W.-C., Lim, C.-F., et al. 2017, *ApJ*, 850, 37
- Wang, T., Schreiber, C., Elbaz, D., et al. 2019, *Nature*, 572, 211
- Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, *ApJS*, 258, 11
- Weiß, A., Kovács, A., Coppin, K., et al. 2009, *ApJ*, 707, 1201
- Wuyts, S., Labbé, I., Förster Schreiber, N. M., et al. 2008, *ApJ*, 682, 985
- Zavala, J. A., Casey, C. M., da Cunha, E., et al. 2018, *ApJ*, 869, 71
- Zavala, J. A., Casey, C. M., Manning, S. M., et al. 2021, *ApJ*, 909, 165
- Zhao, Y., Lu, N., Xu, C. K., et al. 2016, *ApJ*, 819, 69



## Appendix A: Counterpart association

Here, we provide an overview of the counterpart association method and the machine learning (ML) model applied in the process. For a detailed description we refer the reader to the work of Liu et al. (2019a, their Sect. 4.2 and Appendix D).

To assess the accuracy of the association between priors and extracted ALMA sources, we made use of several ancillary images from optical and near-infrared instruments covering the COSMOS and GOODS-S fields, as listed at the end of Sects. 2.2 and 2.3, respectively. The following parameters were then measured for each ALMA source-counterpart pair:

- The angular separation between the positions of the ALMA source and the counterpart, normalised by the angular radius of the ALMA source (‘Sep.’);
- The signal-to-noise ratio of the total flux of the ALMA source ( $S/N_{\text{ALMA}}$ );
- The signal-to-noise ratio measured in the counterpart images at the positions of the ALMA source ( $S/N_{\text{S}}$ ) and at the reference position of the counterpart source ( $S/N_{\text{Ref.}}$ ), determined via an aperture photometry approach;
- The amount of extended emission in the counterpart images in-between the ALMA and counterpart source positions, determined by measuring the flux in a series of apertures placed between the respective positions.

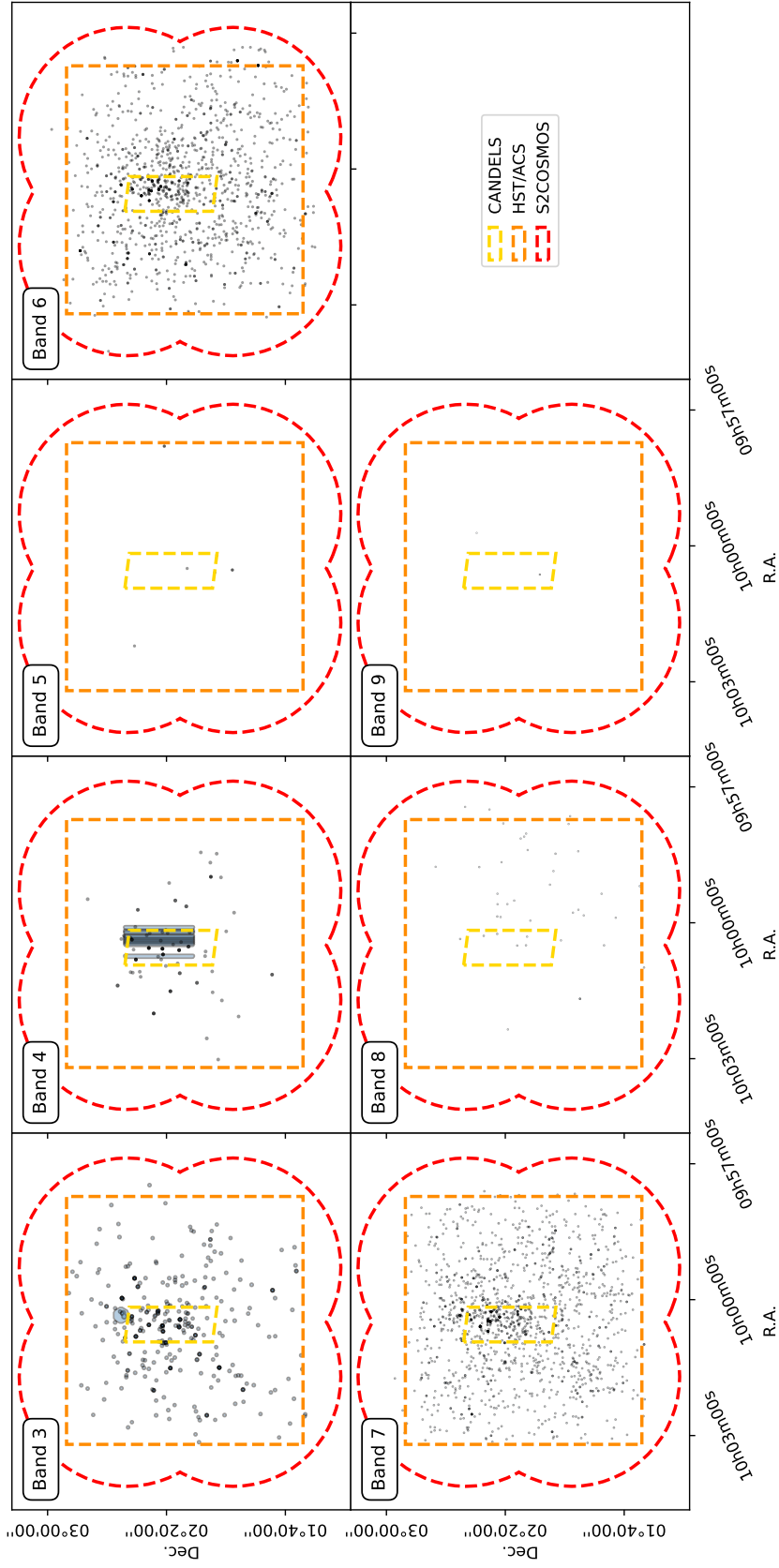
To assess the reliability of the association based on these parameters, Liu et al. (2019a) trained a supervised ML model using the Python package Scikit-learn (Pedregosa et al. 2011) with the Gaussian Process classifier. The training data for the model consisted of a A<sup>3</sup>COSMOS subsample of ~600 ran-

domly selected sources which had been manually classified as either robust or spurious counterpart. The accuracy of the trained model was then assessed by running it on a validation sample of another ~400 manually classified sources not contained in the training data. They found an agreement between the manual and ML classification for ~97% of the sources in the validation sample. The same accuracy was also obtained for the combined sample of ~1000 sources.

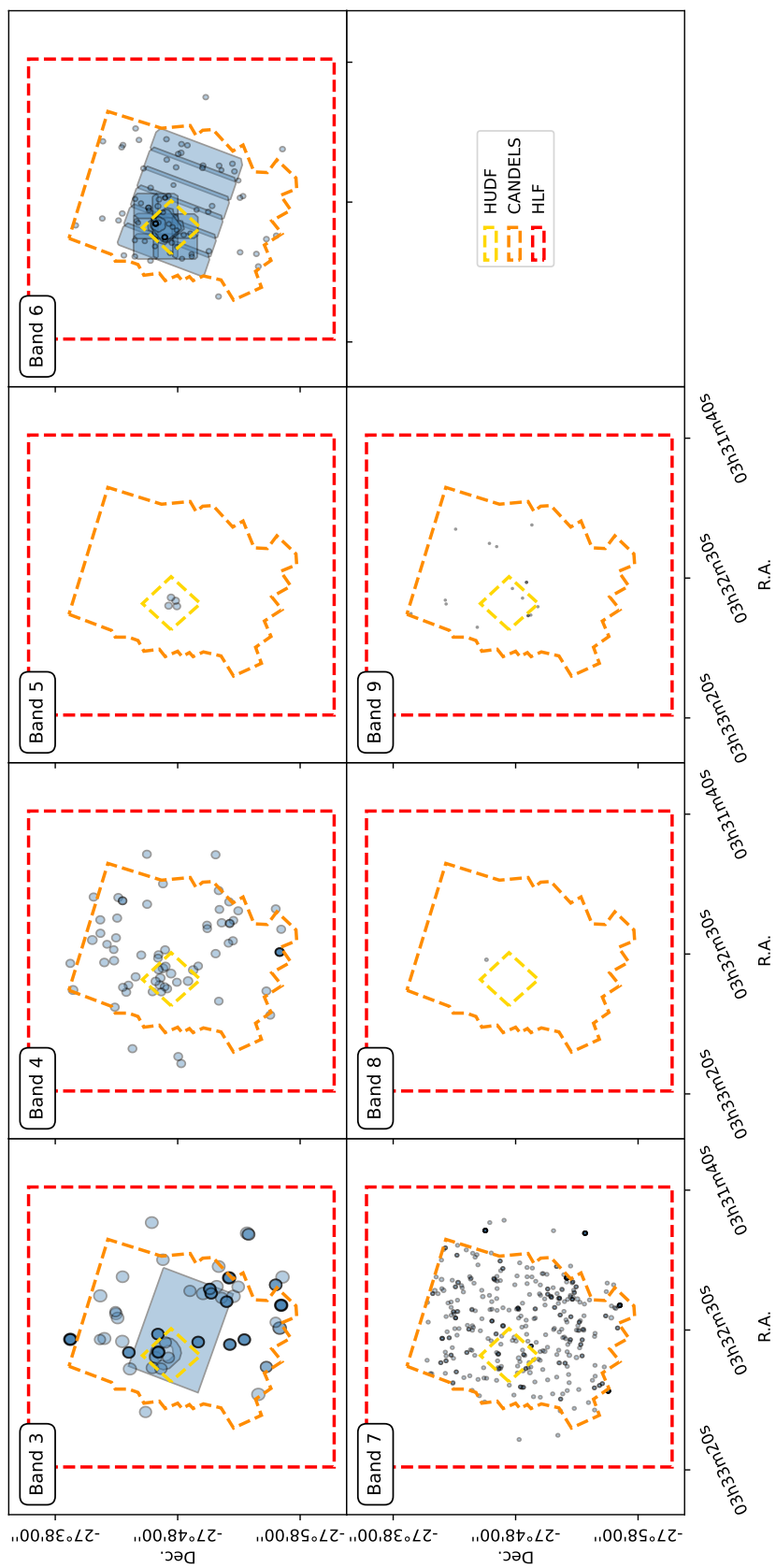
For our A<sup>3</sup>COSMOS database update, we applied this trained model to determine the reliability of our counterpart association for the newer COSMOS data, using the same counterpart images as Liu et al. (2019a), only replacing the *Spitzer*/IRAC 3.6  $\mu\text{m}$  image by a newer image from Moneti et al. (2022) (see Sect. 2.2). However, naturally, new counterpart images and most importantly different filters had to be used for GOODS-S, making the model not applicable to A<sup>3</sup>GOODSS. The number of sources in our current version of the A<sup>3</sup>GOODSS prior source catalogue (i.e. 568) approximately corresponds to the size of the initial training dataset of the ML model. We therefore classified all A<sup>3</sup>GOODSS sources manually, creating the training dataset for the ML model in preparation for the future updates of the A<sup>3</sup>GOODSS database.

## Appendix B: A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS sky coverage by band

Here, we show the spatial sky coverage of our A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS databases, as displayed cumulatively in Fig. 1, separately for each ALMA band. The values of the individual coverages are also listed in Table 1.



**Fig. B.1.** Spatial sky coverage of A<sup>3</sup>COSMOS as shown in the top panel of Fig. 1, and split into individual ALMA bands, as indicated in the top left of each panel. The dashed coloured lines indicate the approximate outlines of the CANDELS (yellow), COSMOS HST/ACS (orange), and S2COSMOS (red; Simpson et al. 2019) fields.



**Fig. B.2.** Spatial sky coverage of A<sup>3</sup>GOODSS as shown in the bottom panel of Fig. 1, and split into individual ALMA bands, as indicated in the top left of each panel. The dashed coloured lines indicate the approximate outlines of the Hubble Ultra Deep Field (yellow; Beckwith et al. 2006), CANDELS (orange; as in Guo et al. 2013), and Hubble Legacy Fields (red; Illingworth et al. 2016).



## Appendix C: Simulating observations in SIDES

In order to ‘observe’ mock sources in each of our 58 realisations of a  $2 \text{ deg}^2$  sky field made from SIDES, we created a set of mock pointings. For this, we used the observational properties (i.e. areal coverage and sensitivity profile) of each precisely targeted pointing used in our real number counts analysis. Since the SIDES simulation does not include source size information we, however, assumed an angular resolution of 1 arcsec (which corresponds to the mean angular resolution in our A<sup>3</sup>COSMOS database) for all mock pointings and that all mock sources are point sources at this resolution.

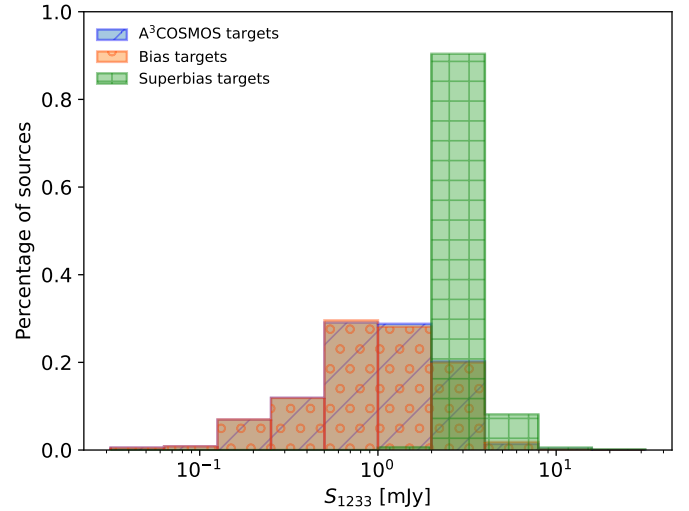
These mock pointings were then placed in each of the 58 fields, centred on mock target sources according to the respective observational bias mode (i.e. Bias, Superbias or Random, see Sect. 3.3). We selected sources from the SIDES catalogue in each respective field as mock targets based on their flux density (except the Random mode in which the pointings were distributed randomly). In the Superbias mode, all the brightest sources in a given field were chosen as targets. In the Bias mode, mock target sources were drawn randomly based on their flux density, with a probability according to the flux density distribution of the real target sources in all precisely targeted pointings in our database in the ALMA band in consideration. Figure C.1 shows the flux density distribution of our chosen mock targets in the Bias and Superbias modes, compared to that of the real A<sup>3</sup>COSMOS targets in Band 6. The Bias and real target distributions are almost identical by construction, with the majority of targets being in the intermediate to higher flux density range ( $\sim 0.5 - 4 \text{ mJy}$ ). In contrast, Superbias targets are exclusively bright sources ( $\geq 2 \text{ mJy}$ ). In both the Bias and Superbias modes, the choice of which mock pointing to assign to which mock target was made to ensure that every mock target is detectable in its assigned mock pointing (as it is the case in our set of real pointings).

The source extraction from these mock observations followed the same constraint as our real source extraction, that is, only sources with  $S/N_{\text{peak}} > 5.4$  were considered as detected, but assuming that above this threshold  $C_{\text{compl.}} = 1$  and  $C_{\text{contam.}} = 0$ . Here,  $S/N_{\text{peak}}$  is calculated from the sensitivity of a mock pointing at the position of the mock source covered and the flux density of that source as given in the SIDES catalogue (i.e.  $S_{\text{peak}} = S_{\text{total}}$ , as all mock sources are assumed to be point sources).

For the corrected mock number counts, a central circular region around the phase centre of each mock pointing was disregarded during source extraction. In addition, in order to allow for a weighting of mock sources by redshift proximity to their

**Table C.1.** Fractions of sources detected within precisely targeted A<sup>3</sup>COSMOS pointings that have a prior spectroscopic, photometric, or no prior redshift, split into sources within a radius of 1 arcsec around the phase centre (Targets) and sources further away from the phase centre than 1 arcsec (Serendipitous).

	Targets	Serendipitous
Spec. $z$	40.66%	9.23%
Photo. $z$	54.58%	73.85%
No prior $z$	4.76%	16.92%



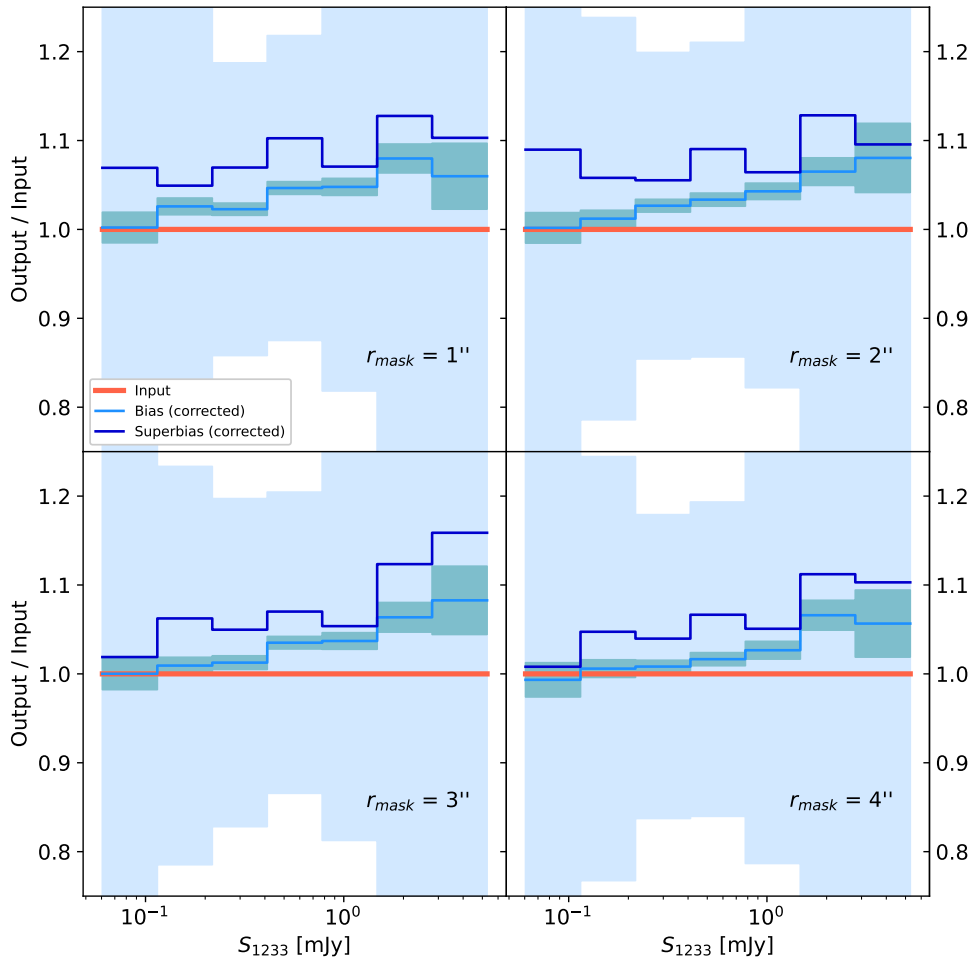
**Fig. C.1.** Band 6 flux density distribution of targets in the simulation compared to the real targets of precisely targeted pointings in the A<sup>3</sup>COSMOS database. Orange with circles indicates Bias mode; green with squares is Superbias mode; and blue with diagonal lines is for real targets.

respective mock targets (as in our real analysis), we assigned to each detected mock source a mock redshift uncertainty. For this, we randomly assigned each detected mock (serendipitous) source and each mock target to either a photometric, spectroscopic, or no redshift, based on the fractions of our real sources (see Table C.1). We then assigned to each mock source and mock target a redshift randomly picked within a Gaussian PDF centred on their original SIDES redshift,  $z$ , with a width of  $0.001(1+z)$  for spectroscopic or  $0.06(1+z)$  for photometric redshifts. These assigned redshifts were then used for weighting as described in Sect. 3.1.2.

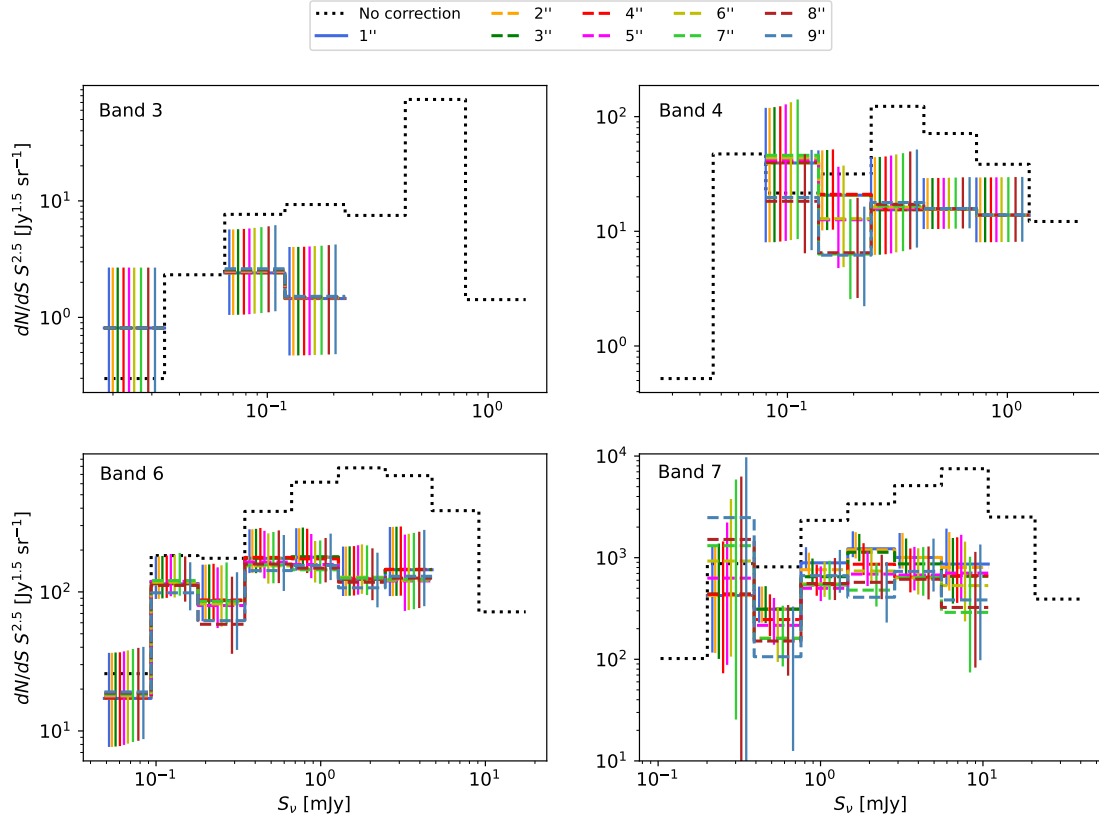
## Appendix D: Selecting the radius of the central mask

We wanted to determine the optimal size of the mask put on the centre of our precisely targeted pointings. To do so, we ran our simulation (see Sect. 3.3) with several mask radii of increasing size and compared the retrieved corrected number counts (‘Output’) to the underlying source distribution in SIDES (‘Input’). The results are shown in Fig. D.1 for Band 6 with radii from 1 arcsec to 4 arcsec. The lighter shaded region indicates the  $1\sigma$  scatter between realisations (i.e. the cosmic variance). The darker shaded region indicates the error on the mean over all realisations (i.e.  $\sigma/\sqrt{58}$ ). For clarity, the uncertainties are shown only for the Bias line. While we see a slight convergence, there is no significant overall improvement with increasing radius. The excess of the Bias over the Input number counts at the bright end remains consistently at  $\sim 6\%$ . The deviation, however, is far smaller than the cosmic variance in all flux density bins at any radius. Cosmic variance introduces an uncertainty of  $\sim 15\text{--}20\%$  in the intermediate flux density ranges and  $\sim 50\%$  and  $\sim 90\%$  at the faint and bright end, respectively. From a simulation viewpoint, increasing the radius of the mask above 1 arcsec therefore offers no additional benefit.

However, these conclusions are only meaningful if the clustering properties in the simulations are an accurate representation of reality. To test the applicability on our A<sup>3</sup>COSMOS database, we therefore investigated the influence of varying the mask radius on the inferred number counts. Figure D.2 shows the number counts inferred using our method using different mask radii up to 9 arcsec as coloured lines. In Bands 3, 4, and 6 there is no significant change in the number counts within the uncertainties when increasing the radius beyond 1 arcsec. In Band 7, however, there is a slight trend towards lower number counts with increasing radius, but with a converging behaviour. For  $r \gtrsim 4$  arcsec the number counts are consistent within their uncertainties in all flux density bins. This is most likely due to a higher contamination from close-by neighbours of the target source, are not caught as well in this band as in other bands by our redshift weighting method (see Sect. 3.1.2). Band 7 is the band in which the lowest fraction of serendipitous sources in close angular proximity (i.e.  $< 4$  arcsec) to the target have a redshift counterpart in our ancillary catalogues. This reduces our possibility to perform redshift weighting of these sources. Thus, while redshift weighting excludes most close neighbours due to clustering in Bands 3, 4, and 6, it is not sufficient in Band 7. We therefore had to choose a larger exclusion radius in this band. In brief, we chose for our central mask a radius of 1 arcsec in Bands 3, 4, and 6, and a radius of 4 arcsec for Band 7.



**Fig. D.1.** Simulated differential number counts in ALMA Band 6, retrieved with the corrected Bias and Superbias modes for four different central mask radii, divided by the underlying Input number counts. Shown here is the mean over ten run-throughs of the simulation (i.e. 580 realisations). The light shaded regions show the standard deviation due to cosmic variance; the dark shaded regions show the error on the mean (i.e.  $\sigma_{\text{cos.var.}}/\sqrt{580}$ ). The errors are of approximately the same size for the Bias and Superbias modes, but for clarity they are only shown for the Bias mode.



**Fig. D.2.** Differential continuum number counts computed from the A<sup>3</sup>COSMOS database in four ALMA bands (as labelled in the top left corner of each panel) for nine different radii of the mask around the phase centre of single pointings. Uncertainties are shown as vertical coloured lines in the respective bins. The dotted grey line shows the number counts using all data in the A<sup>3</sup>COSMOS database and applying no corrections. Error bars are not shown for the uncorrected number counts.

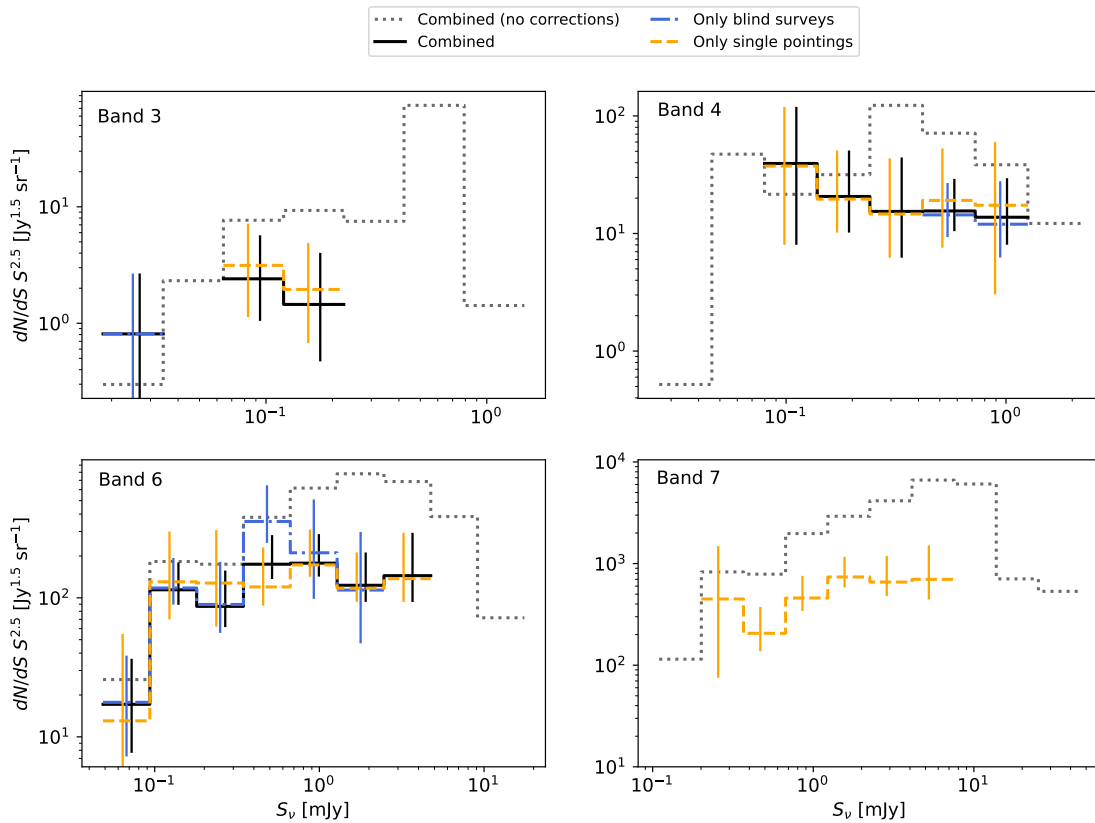
## Appendix E: Comparing blind surveys and single pointings

In order to check the validity of our selection, masking and weighting methods on single pointings, we computed the number counts from our A<sup>3</sup>COSMOS database for four different cases. The first is using the entire database, that is, all images and sources, applying no selection, masking or weighting. This way we see the full effect of the observational biases in the data. The second case is computing the number counts exclusively from blind surveys in our database and the sources detected therein. Since blind surveys do not have selection biases by design, no masking or weighting by redshift is necessary to infer accurate number counts. The third case is using only single pointings with a central target, applying our masking and weighting method, so we can see the extent to which our method constitutes an improvement. The fourth case is using blind surveys combined with the selected and masked single pointings, which represents our final results as shown in Fig. 7.

The comparison of these four cases is shown in Fig. E.1. In all bands, there is a significant improvement over the uncor-

rected number counts, especially towards higher flux densities, which is primarily due to brighter sources being targeted preferentially. In Bands 3, 4, and 6, the blind survey, single pointing, and combined number counts are consistent with each other within their range of uncertainties. This makes us confident of the validity of our selection and correction method for single pointings to retrieve accurate number counts. Blind survey and combined number counts are not shown for Band 7, as there are no blind surveys available for this band within the A<sup>3</sup>COSMOS and A<sup>3</sup>GOODSS databases. The only noteworthy difference that can be seen between the blind survey and the combined number counts appears in one flux density bin in Band 6 (i.e. 0.35 – 0.67 mJy). While still within the errors, the blind survey count in that bin is by a factor 2 higher than the combined. However, this disagreement can entirely be attributed to cosmic variance and small number statistics.

We conclude from this comparison that our selection and correction method for single pointings is valid, as it yields consistent results with blind surveys.



**Fig. E.1.** Differential continuum number counts computed from the A<sup>3</sup>COSMOS database in four ALMA bands (as labelled in the top left corner of each panel). The number counts are measured for four different sets of data. The grey dotted line indicates use of all images and blindly extracted sources contained within the entire database, applying no selection or correction; the blue dash-dotted line is for the use of images only from ALMA blind surveys and sources found therein; the orange dashed line is for the use of only precisely targeted single pointings, after masking out a central circular region of 1 arcsec (4 arcsec in Band 7) radius and weighting the contribution of each detected source outside the masked region by the proximity of its redshift to that of the respective target (see Sect. 3.1); the black solid line is for both blind surveys and single pointings with targets, after masking the phase centre and weighting sources by redshift. As the single pointing line overlaps with other lines in several bins, it has been slightly shifted to improve visibility. In Band 3, it is increased by 5%, in Band 4 and 6 it is lowered by 5%. Error bars are shown for each bin of all lines except the uncorrected case, in the colour of the respective line. No blind surveys exist in our database for Band 7, and therefore no number counts are shown for that and the combined case.