The ALMA Frontier Fields Survey: IV. Lensing-corrected 1.1 mm number counts in Abell 2744, MACS J0416.1-2403 and MACS J1149.5+2223

Muñoz Arancibia, A.M.; González-López, J.; Ibar, E.; Bauer, F.E.; Carrasco, M.; Laporte, N.; ... ; Zitrin, A.

The ALMA Frontier Fields Survey

IV: Lensing-corrected 1.1 mm number counts in Abell 2744, MACSJ0416.1-2403 and MACSJ1149.5+2223

A. M. Muñoz Arancibia[1], J. González-López[2,3], E. Ibar[1], F. E. Bauer[2,4,5], M. Carrasco[6], N. Laporte[7], T. Anguita[8,4], M. Aravena[3], F. Barrientos[2], R. J. Bouwens[9], R. Demarco[10], L. Infante[2,11], R. Kneissl[12,13], N. Nagar[10], N. Padilla[2], C. Romero-Cañizales[14,3], P. Troncoso[15,2], and A. Zitrin[16]

[1] Instituto de Física y Astronomía, Universidad de Valparaíso, Av. Gran Bretaña 1111, Valparaíso, Chile
e-mail: alejandra.munozar@uv.cl
[2] Instituto de Astrofísica y Centro de Astroingeniería, Facultad de Física, Pontificia Universidad Católica de Chile, Casilla 306, Santiago 22, Chile
[3] Núcleo de Astronomía de la Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Av. Ejército 441, Santiago, Chile
[4] Millennium Institute of Astrophysics, Chile
[5] Space Science Institute, 4750 Walnut Street, Suite 205, Boulder, CO 80301, USA
[6] Zentrum für Astronomie, Institut für Theoretische Astrophysik, Philosophenweg 12, 69120 Heidelberg, Germany
[7] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[8] Departamento de Ciencias Físicas, Universidad Andres Bello, Av. República 252, Santiago, Chile
[9] Leiden Observatory, Leiden University, 2300 RA Leiden, The Netherlands
[10] Department of Astronomy, Universidad de Concepción, Casilla 160-C, Concepción, Chile
[11] Carnegie Institution for Science, Las Campanas Observatory, Casilla 601, Colina El Pino S/N, La Serena, Chile
[12] Joint ALMA Observatory, Alonso de Córdova 3107, Vitacura, Santiago, Chile
[13] European Southern Observatory, Alonso de Córdova 3107, Vitacura, Casilla 19001, Santiago, Chile
[14] Chinese Academy of Sciences South America Center for Astronomy, National Astronomical Observatories, CAS, Beijing 100101, China
[15] Universidad Autónoma de Chile, Chile. Av. Pedro de Valdivia 425, Santiago, Chile
[16] Physics Department, Ben-Gurion University of the Negev, P.O. Box 653, Be'er-sheva 8410501, Israel

ABSTRACT

Context. Characterizing the number counts of faint (i.e., sub-mJy and especially sub-$100\,\mu$Jy), dusty star-forming galaxies is currently a challenge even for deep, high-resolution observations in the FIR-to-mm regime. They are predicted to account for approximately half of the total extragalactic background light at those wavelengths. Searching for dusty star-forming galaxies behind massive galaxy clusters benefits from strong lensing, enhancing their measured emission while increasing spatial resolution. Derived number counts depend, however, on mass reconstruction models that properly constrain these clusters.

Aims. We aim to estimate the 1.1 mm number counts along the line of sight of three galaxy clusters, Abell 2744, MACSJ0416.1-2403, and MACSJ1149.5+2223, which are part of the ALMA Frontier Fields Survey. We have performed detailed simulations to correct these counts for lensing effects, probing down to the sub-mJy flux density level.

Methods. We created a source catalog based on ALMA 1.1 mm continuum detections. We used several publicly available lensing models for the galaxy clusters to derive the intrinsic flux densities of these sources. We performed Monte Carlo simulations of the number counts for a detailed treatment of the uncertainties in the magnifications and adopted source redshifts.

Results. We estimate lensing-corrected number counts at 1.1 mm using source detections down to S/N = 4.5. In each cluster field, we find an overall agreement among the number counts derived for the different lens models, despite their systematic variations regarding source magnifications and effective areas. Combining all cluster fields, our number counts span ∼ 2.5 dex in demagnified flux density, from several mJy down to tens of μJy. Both our differential and cumulative number counts are consistent with recent estimates from deep ALMA observations at a 3σ level. Below ≈ 0.1 mJy, however, our cumulative counts are lower by ≈ 1 dex, suggesting a flattening in the number counts.

Conclusions. We derive 1.1 mm number counts around three well-studied galaxy clusters following a statistical approach. In our deepest ALMA mosaic, we estimate number counts for intrinsic flux densities ≈ 4 times fainter than the rms level. This highlights the potential of probing the sub-$10\,\mu$Jy population in larger samples of galaxy cluster fields with deeper ALMA observations.

Key words. gravitational lensing: strong - galaxies: high-redshift - submillimeter: galaxies

1. Introduction

Observations at far-infrared (FIR) to millimeter (mm) wavelengths have revealed a population of dusty star-forming galaxies (DSFGs, see Casey et al. 2014 and references therein). The detection of these sources benefits from the negative *k*-correction in their spectral energy distribution (SED), which keeps their measured flux density in the FIR-to-mm roughly constant up to redshift $z \approx 6 - 10$ (Blain et al. 2002). Bright sources were first

detected using single-dish telescopes (e.g., Smail et al. 1997; Hughes et al. 1998). After exhaustive identification efforts and spectroscopic campaigns, they were found to lie at high redshift with a peak at $z \sim 2 - 2.5$ (e.g., Chapman et al. 2005; Greve et al. 2005; Pope et al. 2006; Younger et al. 2007).

The surface density of DSFGs detected at different wavelengths is quantified through galaxy number counts (e.g., Blain et al. 1999). The bright end of the galaxy distribution has been extensively probed with single-dish telescopes (e.g., Coppin et al. 2006; Weiß et al. 2009). Recent interferometric follow-up observations of bright sources ($\gtrsim 5$ mJy at $870\,\mu$m) have resolved some of them into multiple components (Smolčić et al. 2012; Karim et al. 2013; Hodge et al. 2013). Fainter DSFGs comprise the bulk of the star formation activity at high redshifts. It has been estimated that sources having $\simeq 0.1 - 1$ mJy at 1.2 mm account for $\gtrsim 50\%$ of the total extragalactic background light (EBL) at mm wavelengths (e.g., Ono et al. 2014; Carniani et al. 2015; Aravena et al. 2016; Fujimoto et al. 2016; Hatsukade et al. 2016; Hsu et al. 2016; Oteo et al. 2016). Better constraints await a complete census of fainter galaxies at these wavelengths in order to fully understand the various contributions to the EBL. Importantly, measuring the source brightness at several FIR-to-mm bands helps to disentangle how the rest-frame FIR spectra vary among galaxy populations; this serves as a key constraint for models of galaxy formation and evolution (e.g., Hayward et al. 2013; Cowley et al. 2015; Muñoz Arancibia et al. 2015).

Faint flux densities can be probed in two ways, namely 1) performing deeper, high-resolution observations (compared to current confusion-limited single-dish data), or 2) using strong gravitational lensing by massive galaxy clusters (Hezaveh & Holder 2011). The high sensitivity of the Atacama Large Millimeter/submillimeter Array (ALMA) recently allowed the possibility to probe and characterize the faint end of the unlensed sub-mm population (Ono et al. 2014; Carniani et al. 2015; Oteo et al. 2016; Hatsukade et al. 2016; Aravena et al. 2016; Dunlop et al. 2017). On the other hand, the lensing power enhances the measured flux density of background sources and alleviates the effects of confusion (Blain et al. 1999). Some of the very first single-dish detections were done in galaxy cluster fields (Smail et al. 1997). Number counts from single-dish detected sources behind galaxy clusters have successfully probed flux densities down to the sub-mJy level albeit with a statistical approach, since counterparts are not firmly known (e.g., Knudsen et al. 2008; Zemcov et al. 2010; Johansson et al. 2011; Hsu et al. 2016). Combining both approaches can maximize their benefits. For instance, Fujimoto et al. (2016) derived 1.2 mm number counts down to a flux density of ~ 0.02 mJy ($\gtrsim 4\sigma$), using proprietary and archival deep ALMA data that included 66 blank fields and one lensed galaxy cluster field.

In this work, we derive 1.1 mm number counts using dedicated ALMA observations (González-López et al. 2017, hereafter Paper I) and recent publicly available lensing models. We exploit ALMA's unique capabilities to search for sources behind three well-studied galaxy clusters, which are part of the Frontier Fields survey (FFs, Lotz et al. 2017). This is a legacy project combining the power of gravitational lensing of massive clusters with extremely deep multiband HST and Spitzer imaging of six strong-lensing clusters and adjacent parallel fields. With the help of several detailed mass models for each galaxy cluster, we can harness the magnification power of these clusters to recover the intrinsic (i.e., "delensed") emission from background galaxies. In turn this may allow us to probe fainter flux densities when compared to observations from blank field surveys. Combining several cluster fields also helps to reduce the impact of cosmic

variance, that is, the field-to-field variation found in the volume density of sources due to large scale structure (Trenti & Stiavelli 2008).

This paper is organized as follows. §2 briefly describes the observational 1.1 mm data and public lensing models used in this work. §3 details the methodology used to derive the number counts, including a careful treatment of the uncertainties in magnification for a given lens model, source position and adopted redshift. §4 presents our derived demagnified 1.1 mm counts and places them in context compared to recent estimates from deep ALMA observations. §5 summarizes our main findings. Throughout this paper, we adopt a flat ΛCDM cosmology with parameters $H_0 = 70\,\mathrm{km\,s^{-1}Mpc^{-1}}$, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$, in order to match the cosmology for which the lens models were produced.

2. Data

2.1. Observations with ALMA

2.1.1. High-significance detections

The sources used in this work are drawn from the individual ALMA 1.1 mm detections in three of the galaxy clusters that comprise the FF survey, namely, Abell 2744 ($z = 0.308$), MACSJ0416.1-2403 ($z = 0.396$), and MACSJ1149.5+2223 ($z = 0.543$), hereafter A2744, MACSJ0416, and MACSJ1149, respectively. They were observed as part of the ALMA Frontier Fields Survey (cycle 2 project #2013.1.00999.S, PI: F. Bauer). Paper I introduces the 1.1 mm mosaic images, data reduction and analysis for these galaxy clusters. Each field covers an observed area of ~ 4.6 arcmin^2, and thus sum to a total image-plane area of ~ 14 arcmin^2. This corresponds to ~ 3 times the area of the Hubble Ultra Deep Field (HUDF, Dunlop et al. 2017) and ~ 14 times the initial ALMA Spectroscopic Survey in the HUDF (ASPECS, Walter et al. 2016; Aravena et al. 2016). With natural weighting, our continuum data reach rms depths of $\sim 55 - 71\,\mu$Jy beam^{-1} and have synthesized beam sizes between $\sim 0\farcs5 - 1\farcs5$. A2744 was partially observed in a more extended configuration compared to the other cluster fields, leading to a longer mean projected baseline. As a result, A2744 achieves the highest resolution among these fields (see Paper I for details).

For each cluster field, we take into account the primary beam (PB) correction. The source extraction is done within the region where PB > 0.5, that is, where the PB sensitivity is at least 50% of the peak sensitivity. Sources are detected by searching for pixels with signal-to-noise ratio (S/N) ≥ 5, which are then grouped as individual sources using the DBSCAN python algorithm (Pedregosa et al. 2012). In the following, we refer to the source S/N as the ratio of the peak intensity and the background rms. We note that depending on the spatial PB correction, sources having the same S/N may have different PB-corrected peak intensities. Unless noted, in the following we refer to source flux densities and peak intensities using PB-corrected values.

At S/N ≥ 5, we detect seven sources in A2744, four in MACSJ0416 and one in MACSJ1149. Since some sources appear to be resolved, we measure their integrated flux densities performing two-dimensional elliptical Gaussian fits in the *uv*-plane using the UVMCMCFIT python algorithm (Bussmann et al. 2016). These fits also deliver the centroid position and size parameters for each source. Before applying lensing corrections, detected sources have peak intensities in the range $\sim 0.33 - 1.43$ mJy beam^{-1}, integrated flux densities in the range $\sim 0.41 - 2.82$ mJy, effective radii in the range $\lesssim 0\farcs05 - 0\farcs37$

and axial ratios in the range ~ 0.17 − 0.66. All of these sources have near-infrared (NIR) detected counterparts (based on deep HST F160W imaging). None of them are members of a FF cluster. We refer the reader to Paper I for more details regarding the source extraction procedure, the choice of the uv-plane for estimating integrated source flux densities and sizes, and the search for NIR counterparts.

2.1.2. Going to fainter flux densities: $4.5 \leq$ S/N < 5

We push below the S/N ≥ 5 threshold of the 12 detections already reported in Paper I (and reintroduced in §2.1.1) in order to extract more information from the maps contributing to the number counts. We decide to include all sources having S/N ≥ 4.5 in the natural-weighted mosaics, being extracted through the same procedure as high-significance detections. This adds four sources to A2744, one to MACSJ0416, and two to MACSJ1149. Although the fraction of spurious sources increases for all fields as we move to lower S/N values, we can statistically correct the counts for this effect.

Table 1 lists these low-significance detections, together with the high-significance detections from Paper I. Peak intensities of $4.5 \leq$ S/N < 5 sources range from ~ 0.25 to ~ 0.52 mJy beam^{-1}. Since a two-dimensional Gaussian fit in the uv-plane gives a highly uncertain measure of the integrated source flux density at low S/N, we use the peak intensity of the detections for estimating of the integrated flux densities as follows. For all our low-significance sources, we adopt as their observed effective radius and axial ratio the median values found for the high-significance sources, namely, $r_{\mathrm{eff,obs}} = 0\farcs23$ and $q_{\mathrm{obs}} = 0.58$ (see Paper I). Assuming this source size is consistent with Fujimoto et al. (2017) values. From source injection simulations (see §3.1), we find a typical ratio between the peak and integrated flux density for these size parameters of 0.85, 0.96, and 0.96 in A2744, MACSJ0416, and MACSJ1149, respectively. Scaling the peak intensities by these ratios, the integrated flux densities of the $4.5 \leq$ S/N < 5 detections range from ~ 0.30 to ~ 0.55 mJy. For estimating the centroid coordinates of each source, we take the S/N ≥ 4.5 pixel that established the detection, plus all surrounding pixels having S/N ≥ 4. We collect the coordinates of these pixels, obtain the median right ascension and declination among all of them and set these median values as estimates of the source centroid coordinates.

Including these detections, our final catalog is comprised by 19 sources. We highlight that none of the $4.5 \leq$ S/N < 5 sources are part of the lists of lensed galaxies used by the lens modeling teams, therefore they do not influence to the lens models.

2.2. Source redshifts

In a galaxy cluster field, the observed magnification by gravitational lensing of a background source varies with both its relative position and redshift. Since we have accurate positions and deep HST imaging, we thus consider available spectroscopic and photometric redshift information.

Laporte et al. (2017, hereafter Paper II) determine photometric redshifts for all our S/N > 5 detections via SED fitting, finding a mean redshift of $z = 1.99 \pm 0.27$. Five of these high-significance sources (A2744-ID01, A2744-ID02, MACSJ0416-ID01, MACSJ0416-ID02, and MACSJ1149-ID03) have spectroscopic redshifts from the GLASS survey (Treu et al. 2015), which are consistent with the photometric redshifts found. We refer the reader to Paper II for more details regarding the mul-

tiwavelength data used, photometry estimates and SED-fitting procedure.

We search for counterparts to our $4.5 \leq$ S/N < 5 sources in several public catalogs reporting photometric and spectroscopic redshift estimates, including: photometric redshifts estimated by the CLASH team (Postman et al. 2012; Molino et al. 2017) and the ASTRODEEP survey (Castellano et al. 2016; Di Criscienzo et al. 2017); catalogs of spectroscopic redshifts by Owers et al. (2011), Ebeling et al. (2014), Jauzac et al. (2016), Kawamata et al. (2016), Treu et al. (2016), Mahler et al. (2018), the GLASS survey (Hoag et al. 2016), and the CLASH survey using VIMOS (Balestra et al. 2016) and MUSE (Grillo et al. 2016; Caminha et al. 2017) at VLT; and redshift estimates for Herschel detections (Rawle et al. 2016). We find that only MACSJ0416-ID05 has a counterpart within $\approx 0\farcs3$ with a secure spectroscopic redshift $z = 1.849$. This was measured from NIR spectra as part of the GLASS survey, confirmed by fitting the continuum grism spectra to SED templates. This galaxy also has extensive multiwavelength broadband data from ASTRODEEP and CLASH.

For the remaining $4.5 \leq$ S/N < 5 sources, all galaxies in the aforementioned catalogs having reliable redshift estimates are beyond $\approx 1''$ from ALMA peak positions. In a few cases, these are contaminated by foreground sources. This makes identification of likely faint NIR emission particularly challenging, thus it is hard to gauge the veracity of these sources.

The choice of source redshifts is as follows. First, we use the spectroscopic redshifts for the five S/N > 5 and one $4.5 \leq$ S/N < 5 detections, respectively. These are presented in Table 1. For the remaining S/N > 5 sources, we use the photometric redshift probability distributions obtained in Paper II. In the aforementioned table, best fit values and 1σ errors from these distributions are presented for reference.

For sources lacking any redshift information (i.e., all but one $4.5 \leq$ S/N < 5 sources), we assume a Gaussian redshift distribution centered at $z = 2$ with standard deviation 0.5. This assumption is supported by the mean photometric redshift found in Paper II for the S/N > 5 sources and by results from the literature found in blind mm detections reaching the sub-mJy level (e.g., Aravena et al. 2016; Dunlop et al. 2017). It is also consistent within $\approx 1\sigma$ with the median redshift of dusty galaxies at 1.1 mm predicted by Béthermin et al. (2015b) using an empirical model, both including and not including strongly-lensed sources, for our chosen S/N threshold (assuming point sources).

2.3. Lensing models

A massive object (e.g., a galaxy cluster) deforms the space-time in its vicinity, acting as a gravitational lens (see Kneib & Natarajan 2011 for a review). In cluster fields, the light from background sources is deflected and magnified. Magnification estimates at each source position are essential for obtaining lensing-corrected flux densities and thus, the number counts. For this, we make use of gravitational lensing models produced by independent teams. Detailed explanations for the models (and their several versions) provided by each team can be found in the readme files publicly available in the FF website[1]. In the following, different model versions from a given team are treated as separated models.

Each modeling team uses their own choice of assumptions and methods. Lensing mass inversion techniques include parametric, free-form (or "non-parametric") and hybrid (i.e., a mixture of both) models. Parametric models, as the name suggests,

---

[1] http://archive.stsci.edu/prepds/frontier/lensmodels/

Table 1. Continuum detections at S/N ≥ 4.5.

ID	RA$_{J2000}$ [hh:mm:ss.ss]	Dec$_{J2000}$ [±dd:mm:ss.ss]	S/N	$S_{1.1\,mm,peak}$ [mJy beam^{-1}]	$S_{1.1\,mm,uv\text{-}fit}$ [mJy]	z
A2744-ID01[a]	00:14:19.80	-30:23:07.66	25.9	1.433 ± 0.056	1.570 ± 0.073	2.9[c]
A2744-ID02[a]	00:14:18.25	-30:24:47.47	14.4	1.292 ± 0.091	2.816 ± 0.229	2.482[c]
A2744-ID03[a]	00:14:20.40	-30:22:54.42	13.9	0.798 ± 0.058	1.589 ± 0.125	$2.52^{+0.23}_{-0.45}$[d]
A2744-ID04[a]	00:14:17.58	-30:23:00.56	13.8	0.932 ± 0.068	1.009 ± 0.074	$1.02^{+0.32}_{-0.09}$[d]
A2744-ID05[a]	00:14:19.12	-30:22:42.20	7.7	0.655 ± 0.086	1.113 ± 0.135	$2.01^{+0.69}_{-0.16}$[d]
A2744-ID06[a]	00:14:17.28	-30:22:58.60	6.5	0.574 ± 0.089	1.283 ± 0.241	$2.08^{+0.13}_{-0.08}$[d]
A2744-ID07[a]	00:14:22.10	-30:22:49.67	6.2	0.455 ± 0.074	0.539 ± 0.082	$1.85^{+0.16}_{-0.14}$[d]
A2744-ID08[b]	00:14:24.73	-30:24:34.20	4.8	0.270 ± 0.056	. . .	. . .
A2744-ID09[b]	00:14:21.23	-30:23:28.70	4.7	0.256 ± 0.055	. . .	. . .
A2744-ID10[b]	00:14:17.72	-30:23:02.25	4.5	0.286 ± 0.063	. . .	. . .
A2744-ID11[b]	00:14:22.63	-30:23:30.45	4.5	0.253 ± 0.056	. . .	. . .
MACSJ0416-ID01[a]	04:16:10.79	-24:04:47.53	15.4	1.010 ± 0.066	1.319 ± 0.103	2.086[c]
MACSJ0416-ID02[a]	04:16:06.96	-24:03:59.96	6.8	0.406 ± 0.062	0.574 ± 0.132	1.953[c]
MACSJ0416-ID03[a]	04:16:08.81	-24:05:22.58	5.8	0.389 ± 0.067	0.411 ± 0.072	$1.29^{+0.11}_{-0.39}$[d]
MACSJ0416-ID04[a]	04:16:11.67	-24:04:19.44	5.1	0.333 ± 0.066	0.478 ± 0.166	$2.27^{+0.17}_{-0.61}$[d]
MACSJ0416-ID05[b]	04:16:10.52	-24:05:04.77	4.6	0.302 ± 0.066	. . .	1.849[e]
MACSJ1149-ID01[a]	11:49:36.09	+22:24:24.60	5.9	0.442 ± 0.074	0.579 ± 0.134	1.46[c]
MACSJ1149-ID02[b]	11:49:40.32	+22:24:42.00	4.6	0.524 ± 0.113	. . .	. . .
MACSJ1149-ID03[b]	11:49:35.41	+22:23:38.60	4.5	0.326 ± 0.072	. . .	. . .

Notes. Column 1: Source ID. Columns 2, 3: Centroid J2000 position of ID. Column 4: Signal-to-noise of the detection. Column 5: PB-corrected peak intensity and 1σ error. Column 6: PB-corrected integrated flux density and 1σ error from *uv* fitting. Column 7: Source redshift. [a] High-significance (S/N ≥ 5) detections. Already reported in Paper I. [b] Low-significance (4.5 ≤ S/N < 5) detections. Instead of performing a *uv* fitting, we estimate the integrated flux density using the peak intensity and assuming given source size parameters (see §2.1.2). Since all but one of them lack clear counterparts (partly due to contamination) and spectroscopic redshifts, nor were they included in Paper II study, we assume a Gaussian redshift distribution with mean $z = 2$ and $\sigma = 0.5$ for these sources. [c] Spectroscopic redshift from GLASS, already noted in Paper II. [d] Photometric redshift found in Paper II. Best fit value and 1σ error from SED fitting are presented here only for reference, as we use the full probability distribution found for each photometric redshift. [e] Spectroscopic redshift from GLASS.

Table 2. Lensing models considered in this work.

Model	References
Caminha v4[a]	Caminha et al. (2017)
CATS v4, v4.1	Jullo & Kneib (2009); Richard et al. (2014); Jauzac et al. (2014, 2015, 2016)
Diego v4, v4.1	Diego et al. (2005, 2007, 2015)
GLAFIC v4[b]	Oguri (2010); Kawamata et al. (2016, 2018)
Keeton v4	Keeton (2010); Ammons et al. (2014); McCully et al. (2014)
Sharon v4	Jullo et al. (2007); Johnson et al. (2014)
Williams v4	Liesenborgs et al. (2006, 2007); Sebesta et al. (2016)

Notes. All these models cover the region where our ALMA sources lie. [a] Only available for MACSJ0416. [b] Only available for A2744 and MACSJ0416.

assume that the mass distribution can be represented by a super-position of analytical functions that depend on a limited number of free parameters. In most cases, these models are guided by the distribution of cluster members and their luminosities. Free-form models do not use this assumption, but find the solution directly from the multiple-image constraints (as a result, their resolution is often lower). Parametric models include Caminha (Caminha et al. 2017), CATS (Jullo & Kneib 2009; Richard et al. 2014; Jauzac et al. 2014, 2015, 2016), GLAFIC (Oguri 2010; Kawamata et al. 2016, 2018), Keeton (Keeton 2010; Ammons et al. 2014; McCully et al. 2014), and Sharon (Jullo et al. 2007; Johnson et al. 2014). Williams (Liesenborgs et al. 2006, 2007; Sebesta et al. 2016) is a free-form model, while Diego (Diego et al. 2005, 2007, 2015) is hybrid. Brief descriptions of these

and more models can be found in Coe et al. (2015) and Priewe et al. (2017).

Table 2 lists the models considered in this work. These models are constrained by input archival observations (both from HST and ground based), redshifts and multiple image identifications. The reliability of these constraints has been collectively assigned by all teams, ranking each constraint as Gold, Silver or Bronze (see Priewe et al. 2017). Model versions v3 and newer are based on FF observations, with v4 and newer models using a considerably larger set of arcs and spectroscopic redshifts compared to previous versions. Model versions v4 and v4.1 vary in the set of constraints chosen by each team, with v4 models using only the most reliable constraints (i.e., images from the Gold

sample[2]). For details regarding the selection of constraints and their reliability, we refer to the readme files publicly available for each lens model. We attempt to use the best data to date, so for all cluster fields we consider only v4 or newer models.

Lens models are comprised of maps of the normalized mass surface density κ and shear γ of the galaxy cluster, assuming a redshift $z = \infty$ background. The deflection field $\overrightarrow{\alpha}$ around the lensing object can be estimated from κ as

$$\nabla \cdot \overrightarrow{\alpha} = 2\kappa \tag{1}$$

(Coe et al. 2008). These maps are scaled to the source-plane z of interest as

$$\kappa(z) = \kappa \frac{D_{LS}}{D_S}, \quad \gamma(z) = \gamma \frac{D_{LS}}{D_S}, \quad \overrightarrow{\alpha}(z) = \overrightarrow{\alpha} \frac{D_{LS}}{D_S}, \tag{2}$$

where the angular diameter distances D_S and D_{LS} are computed from source to observer and source to lens respectively. The magnification map at a given source-plane z is obtained as (see Coe et al. 2015)

$$\mu(z) = \frac{1}{|(1 - \kappa(z))^2 - \gamma(z)^2|}. \tag{3}$$

For each release, teams provide a lens model coined as "best", plus a range of individual reconstructions (hereafter the "range" maps) that sample the range of uncertainties, that is, there is one κ and γ map for each realization. The field of view and angular resolution adopted for presenting the maps, as well as the number of realizations provided, may vary across teams and model versions.

We use the full set of mass reconstructions for estimating uncertainties in both source magnifications and effective source-plane areas in a given lens model. These in turn are propagated to the number counts as explained in §3. In order to use the models, "range" maps for κ and γ are reprojected to the size and resolution of the ALMA maps using a first order interpolation. Based on these, we obtain magnification maps for a given source redshift using Eq. 3, and deflect these maps (together with the PB-corrected rms maps) to the source plane using the deflection fields; if several pixels in the image plane are deflected to only one in the source plane, only the image-plane pixel with the highest magnification is kept and assigned to the source-plane pixel (following Coe et al. 2015). This is needed as effective areas are measured in the source plane. However, we adopt redshift probability distributions for most of the detections (see §2.2), and therefore need to create source-plane maps for several redshift values. In order to make our Monte Carlo simulations faster at this step, we precompute source-plane maps for a fixed grid in redshift, using steps of $\Delta z = 0.2$ in the range $z_{min} = 0.4$ to $z_{max} = 4$ for A2744 and MACSJ0416 ($z_{min} = 0.6$ for MACSJ1149, given the higher cluster redshift). It is safe to consider only this redshift range since it contains all the adopted

---

[2] Note, however, that the choice of constraints for v4 models is not completely homogeneous across teams. For instance, Sharon included also few Silver and Bronze images in regions where the number of Gold images is small. Similarly, teams that released v4.1 versions added particular lower-ranked constraints: CATS added Silver images plus some very (photometrically) convincing candidates, while Diego added the full Silver and Bronze sets.

spectroscopic redshifts; also, all our photometric redshift distributions have zero values at $z \geq 4$, and this limit is at 4σ from the mean redshift assumed for sources lacking redshift information.

When sampling the source redshift distributions across the Monte Carlo simulations, we also find (for each random z) the two closest values used in our set of precomputed source-plane maps, and use them for interpolating the effective source-plane areas at a given demagnified peak intensity. It is safe to use this approximation even for sources having spectroscopic redshifts, as we check that the predictions using their two closest redshift bins have no significant variation for most detections.

All v4 and v4.1 lens models cover the region where our detections lie. However, a fraction of the region where the ALMA maps have PB > 0.5 is not fully covered by the GLAFIC v4 model ($\sim$ 0.4% for A2744) and the Williams v4 model ($\sim$ 2%, 13%, and 5% for A2744, MACSJ0416, and MACSJ1149, respectively). In these cases, we impose $\mu = 1$ for the missing pixels in magnification maps, as their closest pixels have $\mu \approx 1$. In total, we adopt for use eight, nine, and seven lens models for A2744, MACSJ0416, and MACSJ1149, respectively (see Table 2). Since not all modelers provide deflection field maps for all realizations, we use Eq. 1 to compute these maps from the κ maps provided for the "range" models.

3. Methodology

We compute demagnified number counts at 1.1 mm. This requires recovering the demagnified (i.e., source-plane) integrated flux density S_{demag} for each source, which is obtained as

$$S_{demag} = \frac{S_{obs}}{\mu}. \tag{4}$$

Here, S_{obs} corresponds to the measured (i.e., image-plane) integrated flux density and μ the source magnification (see §3.3). We obtain the differential number counts at the j-th flux density bin as

$$\frac{dN}{d \log(S)} = \frac{1}{\Delta \log(S)} \sum_i^n X_i, \tag{5}$$

where we sum the individual contribution X_i to these counts by the sources that have demagnified flux densities within that bin. Similarly, we compute the cumulative number counts for the k-th flux limit as

$$N(> S_k) = \sum_i^m X_i, \tag{6}$$

where we sum over the sources having $S_{demag,i} \geq S_k$. In these two expressions, we estimate the contribution by the i-th source as

$$X_i = \frac{1 - p_{false,i}}{C_i A_{eff,i}}. \tag{7}$$

Here, C_i is a completeness correction (see §3.1) and $p_{false,i}$ the fraction of spurious sources (see §3.2). $A_{eff,i}$ corresponds to the effective area where that source can be detected (see §3.5), depending on the source redshift and lens model that is adopted.

A detailed treatment for all these quantities is described in this section. Throughout our number count analysis, we consider

ALMA detections down to S/N = 4.5. This S/N threshold is chosen as an appropriate balance between the correction factors that are related to the source detection, even when the fraction of spurious sources is not exactly the same among cluster fields (see §3.2).

For the sake of simplicity, we assume that none of the ALMA continuum detections are multiply imaged over the S/N threshold. We verified this for all lens models using their "best" maps (see §2.3). For each ALMA detection, we create a set of image-plane mosaic pixels, comprised by its peak plus all the S/N ≥ 4 pixels surrounding it (hereafter Set 1). For a given lens model and redshift, we use the deflection fields for finding the spatial coordinates of Set 1 pixels in the source plane. We later search for all the image-plane pixels that are deflected from these source-plane coordinates. This new set includes Set 1 pixels (by construction) but may include new mosaic positions if the source-plane pixels are multiply imaged. If any of these new pixels belongs to any of the remaining ALMA detections, that is, matches another peak pixel or a S/N ≥ 4 pixel surrounding it, a detection is said to be multiply imaged (above the S/N threshold) in our mosaics.

Adopting the same redshift bins as when precomputing magnification maps, we find that none of our S/N ≥ 4.5 detections are a multiple image of another one in the catalog. Moreover, we check that none of the newly found image-plane positions have S/N ≥ 4. Therefore, if any of our detections both lies at one of the redshifts considered and is multiply imaged, then the predicted images could not be detected, unless a S/N < 4 criterion is used. We further assume that we have recovered the total 1.1 mm flux densities, within their respective errors.

3.1. Completeness

In presence of noisy data, number counts need to be corrected for the proportion of sources that were not detected, because their noise level shifted their peak S/N below our chosen threshold. We compute the completeness as a function of image-plane integrated flux density S_{obs} as follows. We draw 10^5 artificial image-plane sources from a uniform distribution in $\log(S_{obs})$ in the range $0.01 - 10$ mJy, a uniform distribution in scale radius $r_{eff,obs}$ in the range $0\rlap{.}''001 - 0\rlap{.}''5$ (sources with scale radii smaller than the pixel size are considered point sources) and a uniform distribution in axial ratio q_{obs} in the range $0 - 1$. The scale radius interval is chosen based on the image-plane scale radii and their 1σ errors found for our high-significance detections. One at a time, we inject these sources randomly in the PB-corrected mosaic. We later extract them and check if they meet our S/N ≥ 4.5 criterion. We restrict this source injection only to the PB > 0.5 region. We obtain the completeness C for each (injected) flux bin as the fraction of sources that have an (extracted) S/N ≥ 4.5 and are thus detected. We later calculate the completeness curves assuming extended sources in steps of $\Delta r_{eff,obs} = 0\rlap{.}''05$. The completeness corrections for all cluster fields are shown in Fig. 1. For point sources, a value of 50% is reached at image-plane flux densities of 0.27, 0.30, and 0.36 mJy for A2744, MACSJ0416, and MACSJ1149, respectively. However, the completeness drops to 24%, 35%, and 42% at the same flux densities for image-plane source sizes in the range $0\rlap{.}''20 - 0\rlap{.}''25$ (i.e., for the image-plane size assumed for our low-significance detections).

Since our source catalog is S/N limited, we note that measured source intensities may be systematically enhanced by noise fluctuations, such that they are boosted over the S/N threshold and thus bias the number counts. Correcting for this effect is known as flux deboosting (e.g., Hogg & Turner 1998; Weiß et al.
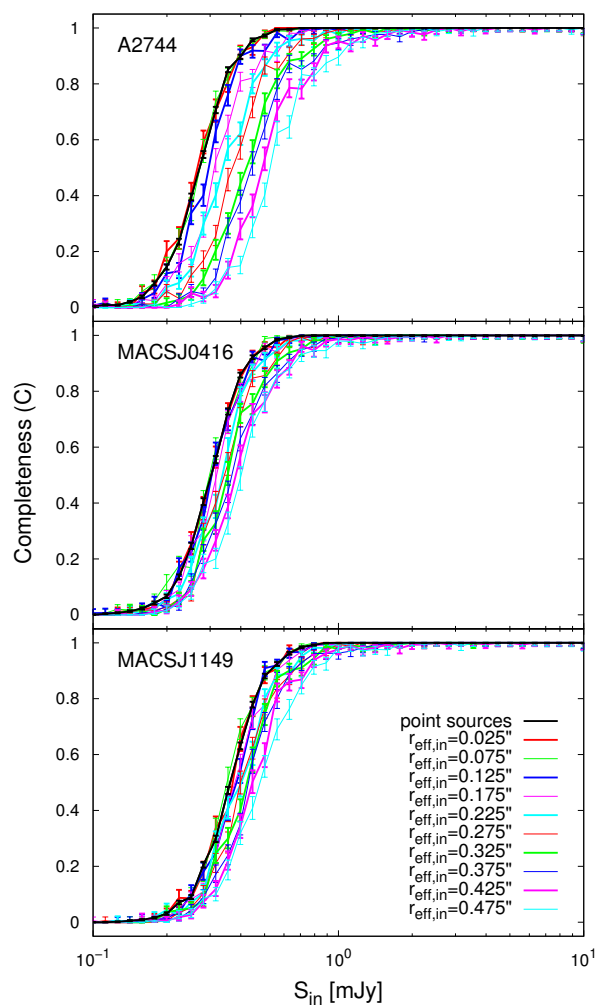
Fig. 1. Completeness correction C as a function of image-plane integrated flux density and separated in bins of image-plane scale radius. Error bars indicate binomial confidence intervals.

2009). Taking the source injection simulations used to estimate the completeness corrections, we select the simulated sources extracted down to S/N = 4.5 and compute the ratio between their extracted and injected flux densities. Figure 2 shows these ratios, together with the median values found as a function of S/N. At S/N = 4.5, we find that the noise boosts the flux densities by 8%, 6%, and 5% for A2744, MACSJ0416, and MACSJ1149, respectively. We use the median ratios found at each S/N for correcting both the observed peak intensities and integrated flux densities for our detections.

If the underlying distribution of source flux densities is steep, number counts derived in the image plane can be overestimated even more in the faint end due to noise fluctuations. This is known as the Eddington bias (Eddington 1913). Correcting the intrinsic number counts for this effect is not trivial, since it requires several assumptions to be made regarding the source properties and folding these through the various lens models. We choose to make no assumptions regarding the true underlying distribution of flux densities, supported by the low number density of ALMA sources in the FFs. However, we can obtain a rough estimate of the scope of any Eddington bias using a single "trial" lens model and assuming specific source flux density and redshift distributions. We choose to test this effect creating sets of 10^4 simulated sources drawn from the redshift and flux distri-
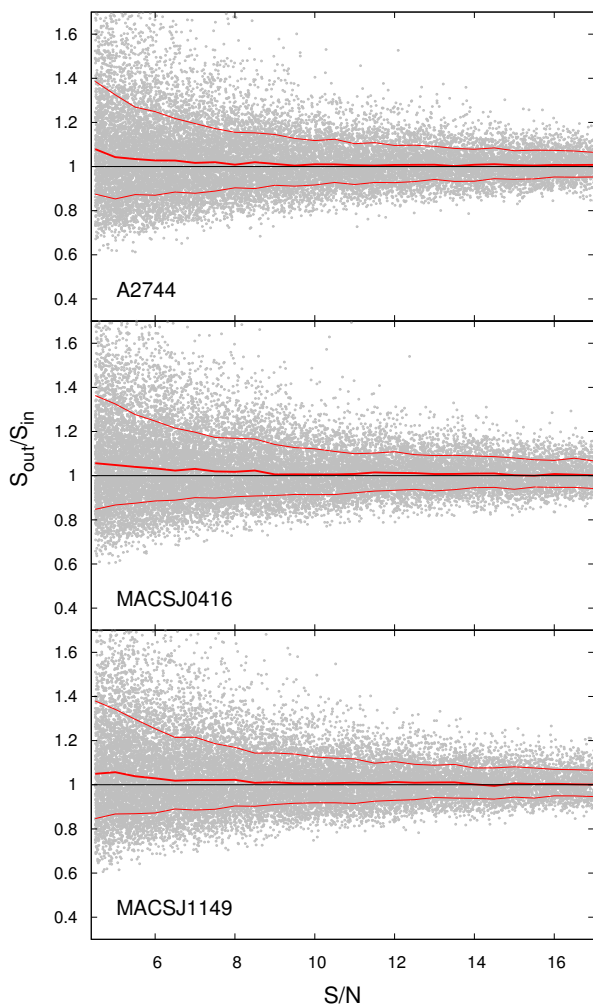
Fig. 2. Deboosting correction as a function of S/N. We display the ratio between the extracted and injected flux densities for our simulated sources as gray dots. Thick red lines correspond to median values while thin red lines indicate the 16th and 84th percentiles.

Fig. 3. Fraction of spurious sources at a given S/N. We display curves for A2744, MACSJ0416, and MACSJ1149 in red, green, and blue, respectively. A vertical dotted line indicates our S/N threshold of 4.5.

field, we generate 300 simulated non-PB-corrected maps, having the same size and resolution as the true ALMA mosaics. Each fake map is comprised by pure Gaussian noise with mean zero and variance one (in S/N units), convolved with the ALMA synthesized beam and later renormalized by the standard deviation of the noise distribution (for preserving the initial variance). We extract sources from each simulated map just as done with the true maps (see Paper I). Since the effective number of independent beams is twice the value expected from Gaussian statistics (see Condon 1997; Condon et al. 1998), we double the number of sources detected in each noise map; doubling this number gives good agreement with the amount of sources found in the negative ALMA mosaics. We obtain the fraction of spurious sources at a given S/N, p_{false}, defined as the average ratio between the number of sources detected over that peak S/N in the true mosaic and in the simulated noise maps.

Figure 3 shows the fraction of spurious sources per S/N limit for the three clusters. At S/N ≥ 4.5 p_{false} is $\approx 20\% - 30\%$ among the cluster fields. Based on the source extraction on the 300 simulated noise maps, the average number of spurious sources at S/N ≥ 4.5 is 2.98 ± 2.37 (A2744), 0.90 ± 1.30 (MACSJ0416), and 0.81 ± 1.32 (MACSJ1149). This is consistent within 1σ with both the amount of spurious sources from the negative mosaics (five, one, and one, respectively) and the number of sources beyond $1''$ of an optical counterpart (four, zero, and two, respectively).

3.3. Source magnifications

Predicting how much is the source brightness amplified by the gravitational lensing effect is necessary for estimating the intrinsic emission from background sources. Lens models applying different techniques predict different values for that magnification.

The centroid pixel of each ALMA detection (see §2.1), together with the "range" maps (see §2.3), are used to calculate the magnification for each source. Indeed, we obtain the magnification distribution for a given source and lens model using the μ values found for the source centroid pixel in all the "range" maps. This choice implies neglecting the effects of differential magnification, and is done in order to simplify the calculations. This is safe as most detections lie far from critical lines (i.e., where magnification formally diverges), and thus magnifications do not have a strong variation across the image-plane extension of these sources. A few detections are found close to critical lines

bution at 1.1mm predicted by the SIDES galaxy formation model (Béthermin et al. 2017), assuming random source coordinates, and lensing them using the "best" CATS v4 model. We then inject and extract these sources in our ALMA mosaics down to S/N = 4.5, estimate the demagnified flux densities for these extracted sources and compute the ratio between output and input demagnified flux density as a function of S/N. At S/N = 4.5, we estimate flux enhancements by 15%, 11%, and 13% for A2744, MACSJ0416, and MACSJ1149, respectively. We find that within the uncertainties, these ratios are consistent with the deboosting corrections obtained in Fig. 2, which were computed assuming a flat distribution in $\log(S_{obs})$. We note that the counts predicted by the SIDES simulation agree with our demagnified counts at a 1σ level (see Fig. 13 and §4.2). However, the SIDES simulation predicts steeper counts at flux densities $0.01 - 0.1$ mJy compared to our median estimates. Therefore, we consider that it is safe to skip any additional Eddington bias correction in this work, including only the deboosting correction shown in Fig. 2.

3.2. Fraction of spurious sources

We compute the fraction of spurious sources (i.e., generated by noise) as a function of S/N as follows. For each galaxy cluster
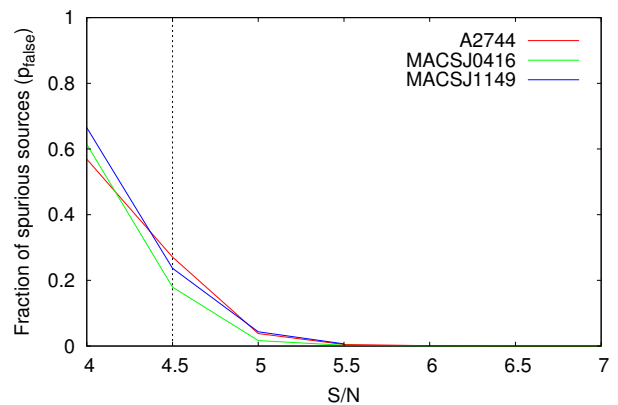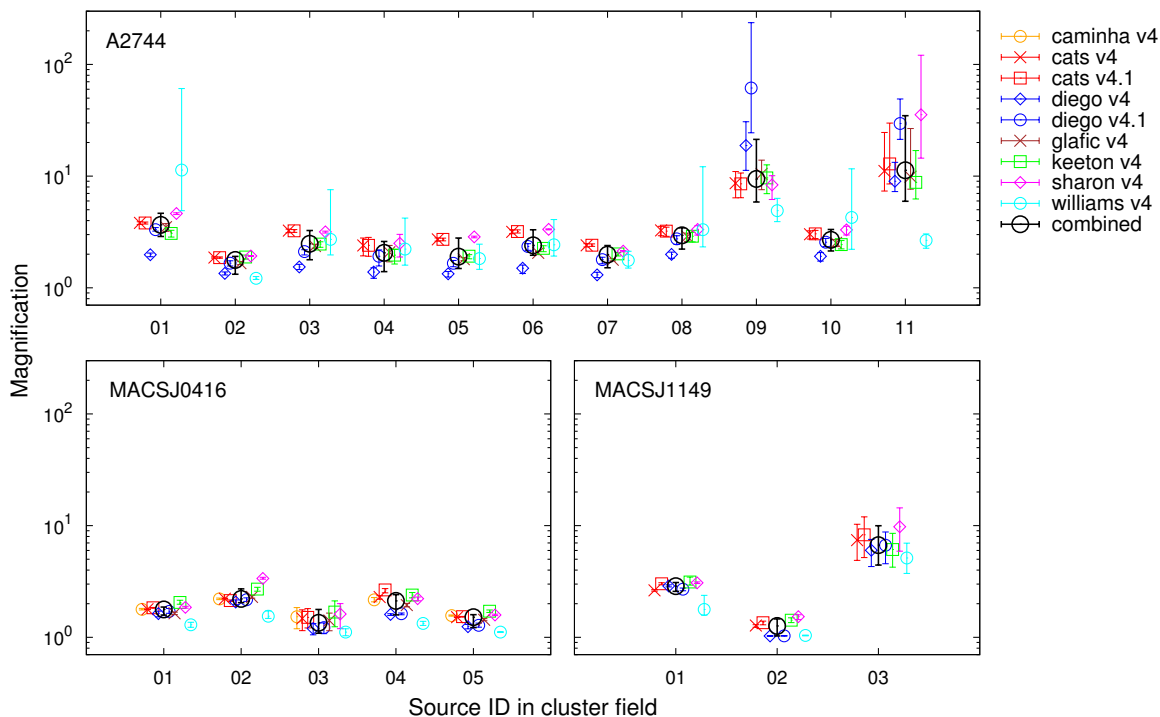
Fig. 4. Median magnification per source for the lens models listed in Table 2 (colored symbols), and also combining all models for each cluster field (large black circles). Error bars indicate the 16th and 84th percentiles (see §3.3). Values for each model have been offset around the source ID for clarity.

(A2744-ID09 and A2744-ID11), being as close as ≈ 1 synthesized beam away from them in a limited number of lens models and assumed redshifts. Unfortunately, these sources lack redshift information, making it difficult to constrain their source magnification (see Fig. 4). Notably, the predictive power of lens models is lower close to critical lines (see below), and thus these sources have large uncertainties in their magnifications.

Since we are adopting a non-unique redshift, we use a Monte Carlo approach with 1000 realizations. Each time, we draw a random z value from the source redshift probability distribution (see §2.2), choose randomly one of the "range" sets of κ and γ maps, and obtain the corresponding μ value using Eq. 3. If a sample z is lower than the cluster redshift (e.g., the photometric redshift distribution has a non-zero probability which extends below the cluster redshift), we assume $\mu = 1$ for the source (i.e., the source is not affected by lensing at that redshift), use its observed flux density and compute the corresponding effective area in the image plane (i.e., assuming all map pixels have $\mu = 1$). This happens only to sources A2744-ID03 and A2744-ID04 and at a very low rate (∼ 3% and < 1% of the realizations, respectively), thus the inclusion of photometric redshift tails below the cluster redshift has a negligible impact in our results.

The magnification distribution sampled for each source is then a combination of distributions obtained at the source position for several redshifts. From this sampling, we can compute a median magnification for each source and estimate uncertainties using the 16th and 84th percentiles (following Coe et al. 2015). This is shown in Fig. 4 for the models listed in Table 2, and also combining all models for each cluster field. Median (combined) magnification values for our sample range from 1.3 to 11.3.

In a given lens model, we find that sources having higher median magnifications have also larger dispersions. Some sources having median $\mu \gtrsim 10$ reach dispersions $\gtrsim 0.5$ dex, such

as sources A2744-ID09 in the Diego v4.1 model and A2744-ID11 in the Sharon v4 model. Magnification distributions are broad and asymmetrical for sources A2744-ID01, A2744-ID03, A2744-ID04, A2744-ID08, and A2744-ID10 in the Williams v4 model, although most of them have median $\mu < 10$. Sources in MACSJ0416 have very similar magnifications in all models, showing small individual dispersions.

Previous works have used the lens models publicly available in the FFs for quantifying systematic uncertainties in predicted magnifications, applying the lens models both to observations (e.g., Bouwens et al. 2017, Lotz et al. 2017, Priewe et al. 2017) and simulations (e.g., Johnson & Sharon 2016, Acebron et al. 2017, Meneghetti et al. 2017). Our trend of increasing dispersion with source magnification (see Fig. 4) is in line with results by Zitrin et al. (2015), Meneghetti et al. (2017) and Bouwens et al. (2017). Zitrin et al. (2015) presented a comprehensive lensing analysis of the complete CLASH cluster sample, examining several lens models produced by their team. They found that the systematic differences (relative to one of the models) increase rapidly with the magnification value. Meneghetti et al. (2017) made a detailed comparison of the mass reconstruction techniques applied by different teams using two simulated galaxy clusters, which resemble the depth and resolution of the FFs. They found that the largest uncertainties in lens models are close to cluster critical lines, with the predictive power of the lens models worsening at $\mu > 10$. For instance, they estimated that the accuracy in the magnifications predicted by some models degrades from ∼ 10% at $\mu = 3$ to ∼ 30% at $\mu = 10$. Bouwens et al. (2017) found similar results using a sample of 160 lensed, NIR-detected sources at $z \sim 6$ in the first four FFs. They constrained the faint end of the $z \sim 6$ ultraviolet luminosity function (UV LF), finding systematic variations in the LF of several orders of magnitude at $M_{UV,AB} = -12$ mag and fainter. They attributed this
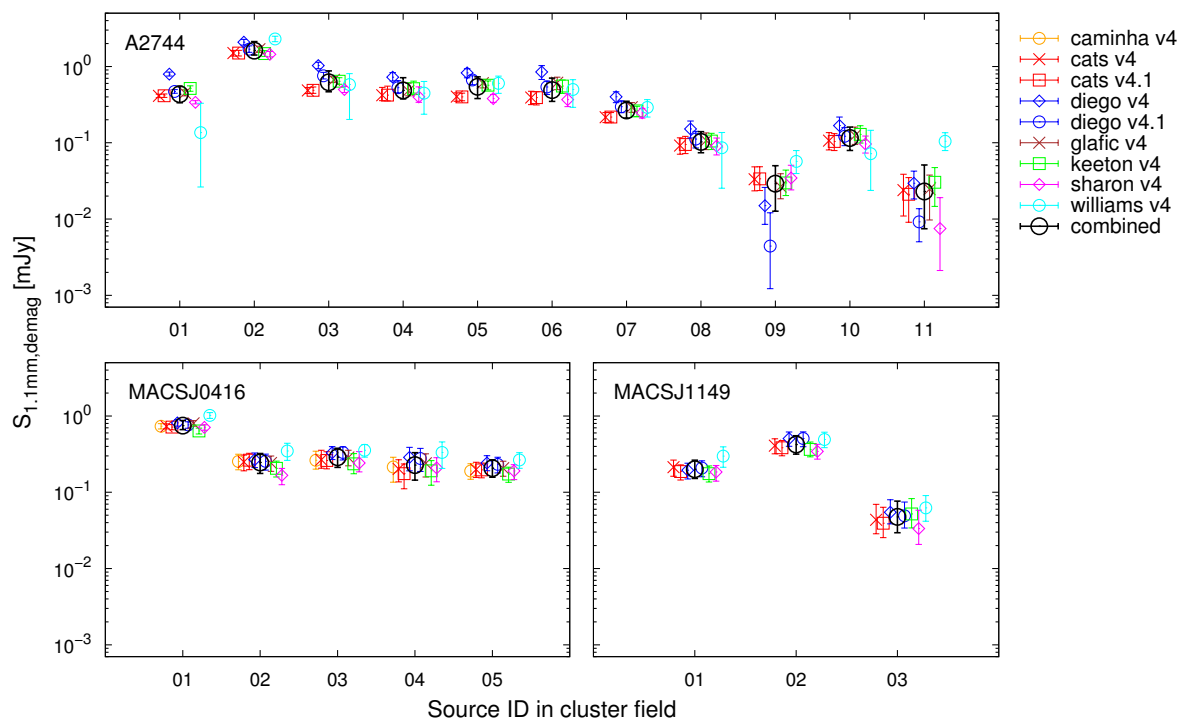
Fig. 5. Median demagnified integrated flux density per source for the lens models listed in Table 2 (colored symbols), and also combining all models for each cluster field (large black circles). Error bars indicate the 16th and 84th percentiles. Values for each model have been offset around the source ID for clarity.

to the large systematic uncertainties inherent at high magnifications, with models having a poor predictive power specially at $\mu > 30$.

Furthermore, Lotz et al. (2017) computed method-to-method standard deviations for the subset of models in A2744 and MACSJ0416 that kept using the same methodology across versions (i.e., for both pre- and post-FF data). They found no significant reduction in the magnification variations across methodologies, reporting median systematic uncertainties in magnification of $< 26\%$ and 15%, for v3 models in A2744 and MACSJ0416, respectively. However, Priewe et al. (2017) found a systematic uncertainty of 70% at $\mu \sim 40$, using the dispersion between v3 or newer lens models in those cluster fields for a $z = 9$ source plane. They argued that the discrepancies in the magnification predictions among models, which often exceed the statistical uncertainties reported by individual reconstructions, were driven by lensing degeneracies, that is, different mass distributions may reproduce the same observational constraints. Moreover, they found the Williams v3 model gives the largest magnification uncertainties at most sky locations in A2744. The broad magnification distributions that we find for some sources in A2744 in the Williams v4 model (see Fig. 4) are in line with these findings.

3.4. Lensing-corrected source flux densities

Once the magnification distribution for each source is obtained, the demagnified integrated flux density is recovered using Eq. 4 for the different lens models. We do this by adopting a Gaussian distribution for S_{obs} with standard deviation given by its reported statistical error, and the distribution described in §3.3 for the magnification. Using both, we resample 1000 times the ratio given in Eq. 4 to obtain a distribution for S_{demag}.
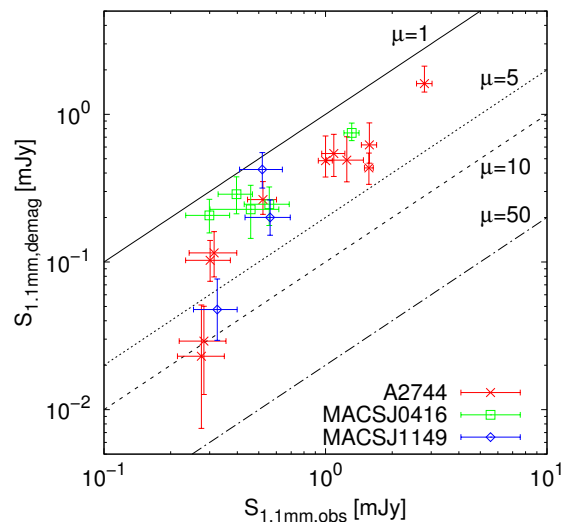

Fig. 6. Median demagnified integrated flux density as a function of observed integrated flux density for A2744 (red crosses), MACSJ0416 (green squares), and MACSJ1149 (blue diamonds). Median values are obtained combining all models for each cluster field. Error bars in demagnified fluxes correspond to the 16th and 84th percentiles while for observed fluxes are 1σ statistical uncertainties. As a reference, black lines indicate magnification values of one (solid), five (dotted), ten (dashed) and 50 (dot-dashed).

Figure 5 shows the median demagnified integrated flux density for each source, computed from both the distributions obtained for each model and joining all of them for each cluster field. Median (combined) lensing-corrected flux densities range
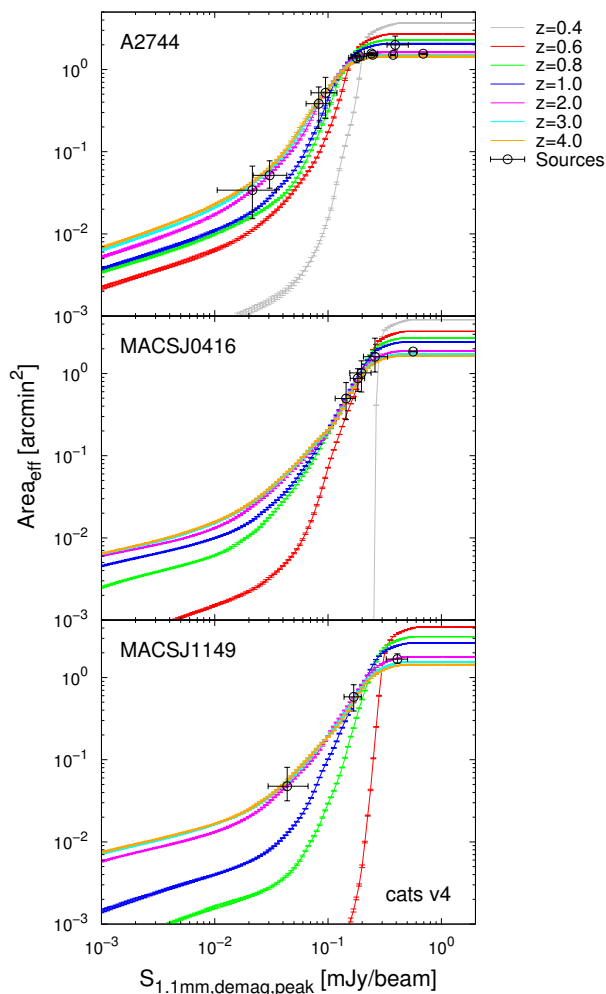
Fig. 7. Median effective area as a function of demagnified peak intensity at several redshifts as indicated in the key (colored lines) for the CATS v4 lens model. Values for our S/N ≥ 4.5 sources (black symbols) are shown for this model (corresponding to the red crosses in Fig. 8). Error bars indicate the 16th and 84th percentiles. For each curve, these are obtained using the "range" maps at the corresponding redshift, while for symbols they are computed as described in §3.5. At $z \geq 2$, areas do not differ significantly with redshift for this lens model.

from ∼ 0.02 to 1.62 mJy, with both the faintest and brightest sources in the sample being found around A2744. Naturally, sources having broad magnification distributions have also large uncertainties in their median S_{demag} values. Within the uncertainties, combined demagnified flux densities cover around 2.5 orders of magnitude.

At $S_{obs} \gtrsim 0.4$ mJy, we find a trend of brighter observed sources being also brighter intrinsically, while sources having lower observed flux densities tend to span ≈ 1.5 dex in demagnified flux. This is shown in Fig. 6. We also find that sources with the highest magnifications ($\mu \gtrsim 5$) are among the faintest ones both in observed and lensing-corrected flux ($S_{obs} \lesssim 0.4$ mJy and $S_{demag} \lesssim 0.06$ mJy, respectively).

3.5. Source effective areas

For computing counts, a key step is to estimate the effective area, A_{eff}, over which the source can be detected. That is, the angular area in the source plane where a map pixel having a given peak

intensity can be detected over a given S/N threshold. The effective area at a given demagnified peak intensity depends not only on the PB response, but also on the source redshift assumed and the lens model adopted.

At a given redshift, we estimate the effective area as a function of demagnified peak intensity $S_{demag,peak}$ (corrected for PB attenuation) as follows. We consider a PB-corrected rms map for each cluster. For each "range" map in a given lens model, we deflect both the PB-corrected rms and magnification maps to the source plane using the deflection fields (see §2.3). If several pixels in the image plane are deflected to only one in the source plane, only the image-plane pixel with the highest magnification is kept and assigned to the source-plane pixel (following Coe et al. 2015). The lensing-corrected rms level for each source-plane pixel, σ_{demag}, is then given by the ratio between its PB-corrected rms and magnification. At a given $S_{demag,peak}$, we collect all the source-plane pixels where $S_{demag,peak}/\sigma_{demag} \geq 4.5$. The effective area corresponds to the sum of areas of source-plane pixels meeting this criterion, each of them given by the ALMA mosaic resolution. We precompute A_{eff} vs $S_{demag,peak}$ curves for each of the redshifts used in our set of precomputed "range" magnification maps (see §2.3).

For each source, we used its full distribution of demagnified peak intensities to compute its effective area. We obtain the $S_{demag,peak}$ distribution as in §3.4, but using a Gaussian distribution for the image-plane peak intensity $S_{obs,peak}$ instead of S_{obs}. We perform a Monte Carlo simulation where we use the same number of realizations and follow the same approach for obtaining both random $S_{obs,peak}$ and z values as in §3.3. This time, however, we need to resample directly the set of "range" magnification maps, in such a way that the same magnification map is used for obtaining both $S_{demag,peak}$ and A_{eff}. This is required in order to have consistency between their values, since both depend on μ values (of the source centroid pixel and all PB > 0.5 pixels, respectively) in an individual "range" map.

This resampling is done using the "range" map identifiers, which are numbered from 0 to $N_{range} - 1$ (with N_{range} the number of "range" maps provided for each model). We draw a random "range" map identifier using a uniform distribution bounded by zero and $N_{range} - 1$. Using the "range" map corresponding to that identifier, we obtain the source magnification in the realization at the random z value. We then use Eq. 4 for computing the source demagnified peak intensity, and then use the two closest redshift bins in our precomputed set (see §2.3) for estimating the source effective area for that "range" map: first linearly interpolating precomputed A_{eff} vs $S_{demag,peak}$ curves in both redshifts bins, and later linearly interpolating the A_{eff} vs z trend within these redshift limits.

The mass reconstruction for each cluster and lens model predicts a distinct proportion between high-μ and low-μ pixels at a given redshift. This is the main driver shaping the slope of the A_{eff} vs $S_{demag,peak}$ curve. Finding small effective areas at low demagnified peak intensities is a natural consequence of having few regions in the maps with very high magnification. In general, the effective area increases steeply with peak intensity until some point where it reaches a plateau. In a given model, both the slope at low peak intensities and plateau level at high peak intensity depend on the modeled cluster field and adopted source redshift.

We illustrate this in Fig. 7 for the CATS v4 model. At $z = 2$, for instance, the largest effective areas found are $1.63^{+0.02}_{-0.02}$, $1.87^{+0.01}_{-0.02}$, and $1.79^{+0.02}_{-0.01}$ arcmin^2 for A2744, MACSJ0416, and MACSJ1149, respectively. They sum to a total effective area
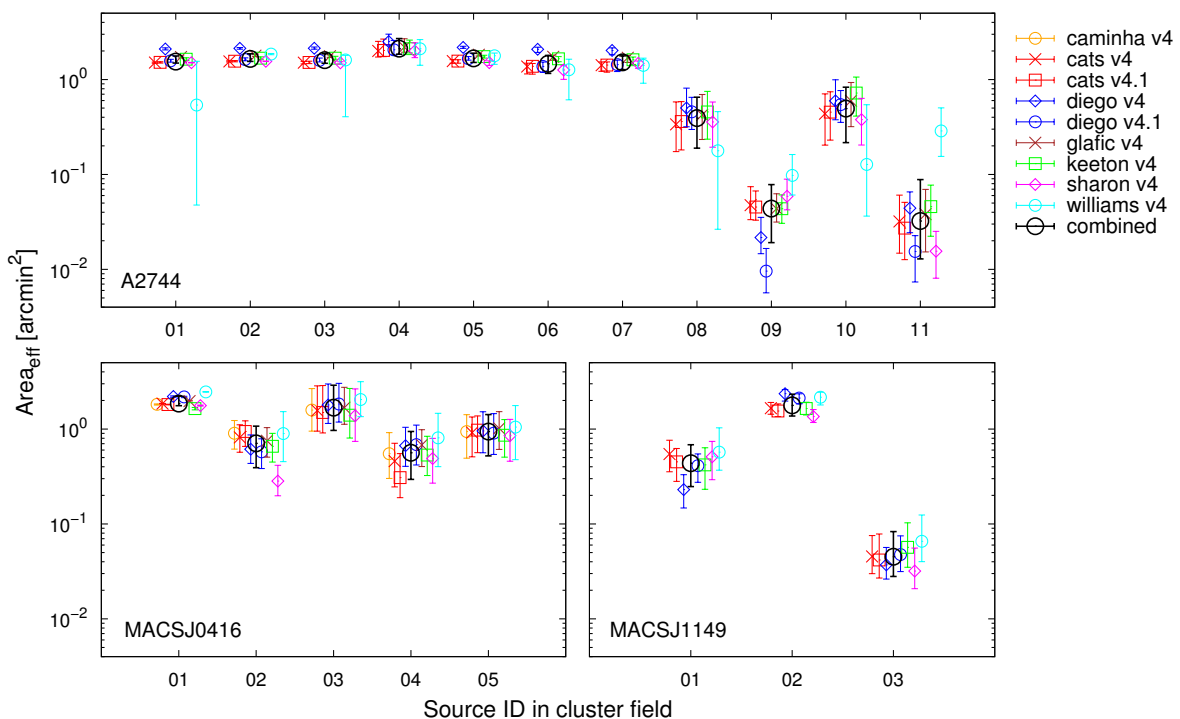
Fig. 8. Median effective area per source for the lens models listed in Table 2 (colored symbols), and also combining all models for each cluster field (large black circles). Error bars indicate the 16th and 84th percentiles. Values for each model have been offset around the source ID for clarity.

of $\approx 5.3 \, \text{arcmin}^2$. This source-plane area is around 2.6 times smaller than the total image-plane coverage (see §2.1.1). In the low peak intensity regime, lower source redshifts give smaller effective areas, while at $S_{\text{demag,peak}} \gtrsim 0.2 \, \text{mJy beam}^{-1}$ the opposite occurs. At $0.06 - 0.1 \, \text{mJy beam}^{-1}$, the steepness of the A_{eff} vs $S_{\text{demag,peak}}$ curves in a log-log scale are such that uncertainties of for instance 0.2 dex in source peak intensity lead to uncertainties around 0.5 dex in source effective area. However, the curves become shallower below $0.06 \, \text{mJy beam}^{-1}$, giving a scatter in effective area of around the same order of magnitude (or below) than that in peak intensity. We find a similar qualitative behavior in the rest of the lens models used in this work, changing the numbers in the aforementioned effective areas and peak intensities.

Figure 8 shows the median effective area for each source, computed from both the distributions obtained for each model and joining all of them for each cluster field. Median (combined) effective areas range from ~ 0.03 to $2.1 \, \text{arcmin}^2$. Within the uncertainties, combined effective areas cover around 2.5 orders of magnitude. In Fig. 9, we compare the uncertainties in the (combined) median S_{demag} and A_{eff} values for our sources. In the bright end ($\gtrsim 0.3 \, \text{mJy}$) we find that sources lie at the A_{eff} plateau, thus uncertainties in effective areas are less affected by uncertainties in S_{demag} and more by the scatter across lens models. At $\approx 0.06 - 0.3 \, \text{mJy}$, sources with a S_{demag} error of such as 0.3 dex have an A_{eff} error close to 0.5 dex. Below $0.06 \, \text{mJy}$, uncertainties in both of those quantities remain comparable in terms of order of magnitude, reaching even 1 dex.

3.6. Monte Carlo simulation for source counts

We combine the techniques explained in previous sections to estimate demagnified source counts that take into account the uncertainties in observed flux densities (see Table 1), adopted
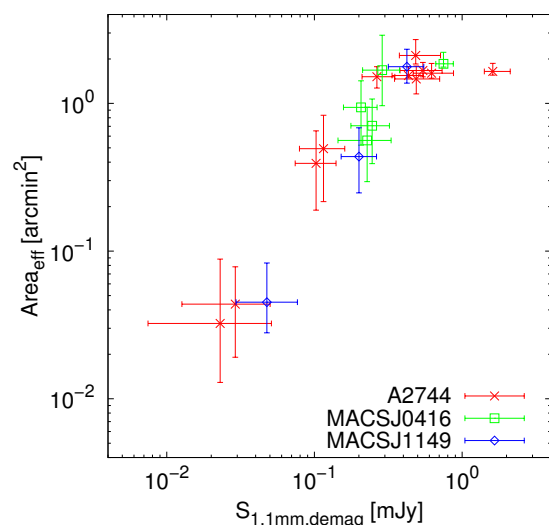

Fig. 9. Median effective area as a function of demagnified integrated flux density for A2744 (red crosses), MACSJ0416 (green squares), and MACSJ1149 (blue diamonds). Median values are obtained combining all models for each cluster field. Error bars correspond to the 16th and 84th percentiles. For comparing uncertainty values, both axes cover the same interval in order of magnitude. Within the errors, both demagnified flux densities and effective areas span around 2.5 orders of magnitude.

redshifts and modeled magnifications. We achieve this using a Monte Carlo approach. A diagram for the way in which this Monte Carlo simulation runs is shown in Fig. 10. For a given galaxy cluster field and lens model, we run a total of 1000 realizations. In each of them, we compute the number counts as follows.

We generate a simulated source catalog comprised of 19 sources, keeping the same coordinates as the true detections. For each source i, we draw a random observed integrated flux density $S_{\text{obs},i}$ from a Gaussian distribution centered at S_{obs} with standard deviation δS_{obs}; we proceed similarly for obtaining a random observed peak intensity $S_{\text{obs,peak},i}$. We also draw a random redshift from the source redshift probability distribution (see §2.2), and use its approximated value z_i as described in §3.3. We then draw a random magnification μ_i as in §3.5, that is, using the identifiers of the "range" maps at the selected z_i (and keeping a record of the selected map identifier). We also obtain the source signal-to-noise ratio as $(\text{S/N})_i = S_{\text{obs,peak},i}/\delta S_{\text{obs,peak}}$, and in the following consider only sources having $(\text{S/N})_i \geq 4.5$. We then use Eq. 4 to obtain the demagnified integrated flux density and peak intensity ($S_{\text{demag},i}$ and $S_{\text{demag,peak},i}$) from S_{obs}, $S_{\text{obs,peak},i}$ and μ_i. We also obtain the completeness correction C_i and fraction of spurious sources $p_{\text{false},i}$ at the source $(\text{S/N})_i$, interpolating the curves computed in §3.1 and §3.2 (see Figs. 1 and 3). Recalling the selected map identifier at z_i, we obtain the effective area $A_{\text{eff},i}$ at the source $S_{\text{demag,peak},i}$ interpolating the curves precomputed in §3.5.

Having all the needed properties, we compute the differential and cumulative number counts using Eqs. 5 to 7. We adopt $\Delta \log(S) = 0.5$ and use the same flux limits for all realizations, in order to combine them later. We follow this procedure for all lens models and cluster fields. Using a given lens model, the set of realizations samples the probability distribution for the number counts in each flux bin, such that we can compute median number counts per flux bin. However, for estimating the associated uncertainties, in this case, we need to take into account low number statistics. We achieve this by computing, besides the 16th and 84th percentiles in the counts per flux bin, the Poisson confidence levels for 1σ lower and upper limits. These levels are provided by Gehrels (1986) as a function of the number of events, which in our case is the median number of sources per flux bin.

We compute combined differential counts taking the median value per flux bin over the lens models listed in Table 2. For computing combined cumulative counts, we take the median value per flux density limit. In both cases, we combine the counts in each cluster field separately (i.e., considering only models available for that particular field) and also combine the counts across all cluster fields.

4. Results and discussion

4.1. Number counts

Table 3 lists the median differential and cumulative counts, combining models for each cluster field both separately and altogether. Uncertainties in the counts are obtained from the 16th and 84th percentiles, listed together with the scaled errors from 1σ lower and upper limits. When the median counts in a given flux bin are zero while having non-zero values at the 84th percentile, we only list 3σ upper limits. These counts are also presented in Figs. 11 and 12, where median counts for individual models in each cluster field are also displayed. Error bars shown in these Figs. combine the uncertainties from the 16th and 84th percentiles in quadrature with those from scaled Poisson confidence levels for 1σ lower and upper limits, respectively.

Because of the small number statistics, we expect our detections to give large error bars in the derived number counts. Uncertainties coming from our Monte Carlo simulation (i.e., using the whole probability distributions for observed flux densities, source redshifts and magnifications together) differ by a factor of

$\sim 0.05 - 7$ from that predicted from Poisson statistics. In A2744 and MACSJ1149, they dominate the upper error bars in the cumulative counts at flux densities below ~ 0.1 mJy (see Table 3). This arises from the broad magnification distributions found in some of the faintest observed sources in these cluster fields. High magnifications make them the intrinsically faintest sources, with demagnified flux densities below ~ 0.1 mJy and effective areas below 0.1 arcmin^2 (see also Fig. 9). For the faintest source in these cluster fields, the effective area distributions easily reach 0.03 arcmin^2 and below, which in turn elevates the counts at their flux levels. This combination of broad distributions both in demagnified flux and effective area makes the number counts below ~ 0.1 mJy highly uncertain.

We present counts down to the flux density where at least one cluster field has non-zero combined differential counts at the 84th percentile, that is, centered on 0.007 mJy. Combining all cluster fields, our differential counts eventually span ~ 2.5 orders of magnitude in demagnified flux density, going from the mJy level down to tens of μJy. This is ≈ 4 times deeper than the observed rms level reached in our deepest ALMA FF mosaic, A2744.

We find variations across lens models in the median counts per flux bin up to ≈ 1 dex. Despite this, in all cluster fields the median counts given by each model per flux bin are consistent within the error bars. A rough agreement among lens models was also found by Coe et al. (2015) when using models (at that time based on pre-FF data only) for predicting the $z > 6$ NIR number counts in all the FFs. They found consistency among all models on the number of faint (i.e., at the nJy level) NIR-detected galaxies expected in HST FF observations.

We explored the effect of adopting different source redshifts in the predicted counts. Within the uncertainties, our differential counts combining all cluster fields and using redshift probability distributions according to available data (as above) are consistent with those obtained assuming a Gaussian redshift distribution centered at $z = 2 \pm 0.5$ for all detections. We also obtain consistent results adopting exactly $z = 2$ for all detections, as well as when assuming a uniform redshift distribution between the cluster redshift and $z = 4$. In these three cases, variations in the median counts combining all cluster fields are only up to ≈ 0.04 dex below 1.3 mJy. Our combined counts are also in agreement within the errors with those obtained centring the Gaussian at $z = 3 \pm 0.5$ for all detections (although upper error bars assuming this higher redshift center are greater by ≈ 0.25 dex at $\lesssim 0.1$ mJy, due to the larger high-magnification regions for this redshift).

4.2. Comparison to previous ALMA studies and galaxy formation model predictions

Figure 13 shows our 1.1 mm number counts compared to results from recent ALMA observations that probe down to the sub-mJy level. These include counts derived from sources detected by serendipitous (Ono et al. 2014; Carniani et al. 2015; Fujimoto et al. 2016; Oteo et al. 2016) as well as dedicated surveys in blank fields (Hatsukade et al. 2016; Aravena et al. 2016; Dunlop et al. 2017) and around a $z = 3.09$ protocluster (Umehata et al. 2017). It should be noted that these previous works use their own source detection criterion, as well as their own choice and methodology for computing corrections to the counts (e.g., completeness, flux deboosting, fraction of spurious sources, effective areas, magnifications). Recalculating their counts matching our criteria, which would ensure a fair comparison, is beyond the scope of this work. Instead, we only apply a scaling for pre-
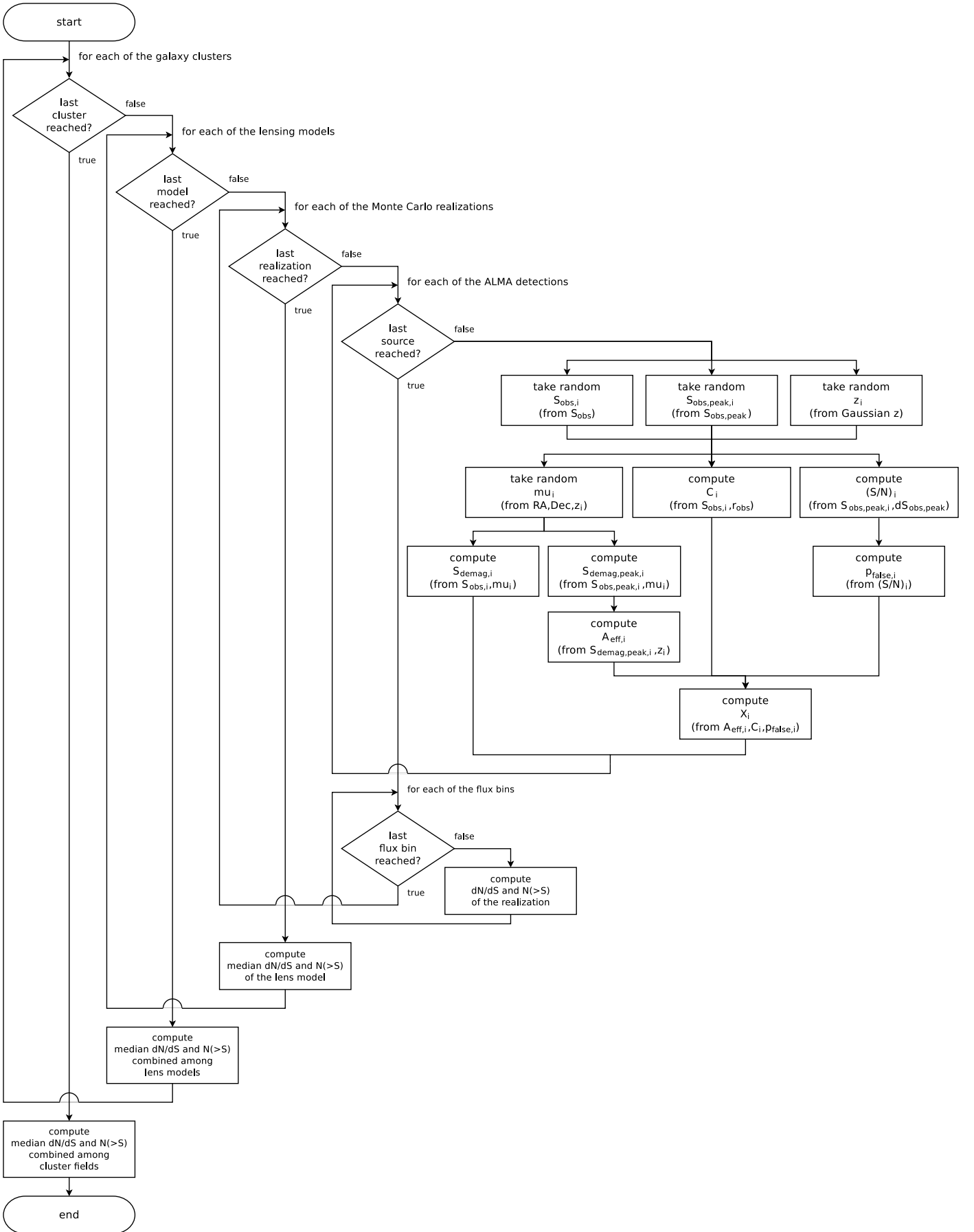
Fig. 10. Diagram of the Monte Carlo simulation developed for estimating demagnified number counts (see details in §3.6).

Table 3. Demagnified 1.1 mm number counts.

Cluster field	$S_{1.1\,mm}$ [mJy]	$dN/d\log(S)$ [deg^{-2}]	# sources	$S_{1.1\,mm}$ [mJy]	$N(>S)$ [deg^{-2}]	# sources
A2744	0.024	$< 1.522 \times 10^6$	< 3.0	0.013	$(1.371^{+1.769}_{-1.091}\,^{+0.625}_{-0.446}) \times 10^5$	$9.0^{+1.0}_{-0.0}\,^{+4.1}_{-3.9}$
	0.075	$(2.932^{+14.30}_{-2.932}\,^{+6.710}_{-2.421}) \times 10^4$	$1.0^{+1.0}_{-1.0}\,^{+2.3}_{-0.8}$	0.042	$(3.444^{+6.750}_{-1.729}\,^{+1.696}_{-0.887}) \times 10^4$	$8.0^{+1.0}_{-1.0}\,^{+3.9}_{-2.8}$
	0.237	$(1.853^{+1.189}_{-1.241}\,^{+1.793}_{-1.005}) \times 10^4$	$3.0^{+1.0}_{-1.0}\,^{+2.9}_{-1.6}$	0.133	$(1.733^{+0.672}_{-0.279}\,^{+0.929}_{-0.636}) \times 10^4$	$7.0^{+1.0}_{-1.0}\,^{+3.8}_{-2.6}$
	0.750	$(1.634^{+0.458}_{-0.731}\,^{+1.286}_{-0.779}) \times 10^4$	$4.0^{+1.0}_{-2.0}\,^{+3.1}_{-1.9}$	0.422	$(9.931^{+2.485}_{-3.483}\,^{+6.685}_{-4.273}) \times 10^3$	$5.0^{+1.0}_{-2.0}\,^{+3.4}_{-2.2}$
	2.371	$(4.090^{+0.465}_{-0.671}\,^{+9.361}_{-3.377}) \times 10^3$	$1.0^{+0.0}_{-0.0}\,^{+2.3}_{-0.8}$	1.334	$(2.039^{+2.237}_{-0.341}\,^{+4.666}_{-1.684}) \times 10^3$	$1.0^{+0.0}_{-0.0}\,^{+2.3}_{-0.8}$
MACSJ0416	0.024	$0.000^{+0.000}_{-0.000}$	$0.0^{+0.0}_{-0.0}\,^{+1.8}_{-0.0}$	0.013	$(2.097^{+1.364}_{-0.773}\,^{+1.650}_{-1.000}) \times 10^4$	$4.0^{+1.0}_{-1.0}\,^{+3.1}_{-1.9}$
	0.075	$0.000^{+0.000}_{-0.000}$	$0.0^{+0.0}_{-0.0}\,^{+1.8}_{-0.0}$	0.042	$(2.097^{+1.364}_{-0.773}\,^{+1.650}_{-1.000}) \times 10^4$	$4.0^{+1.0}_{-1.0}\,^{+3.1}_{-1.9}$
	0.237	$(3.145^{+1.743}_{-1.344}\,^{+3.044}_{-1.706}) \times 10^4$	$3.0^{+1.0}_{-1.0}\,^{+2.9}_{-1.6}$	0.133	$(1.892^{+0.916}_{-0.773}\,^{+1.489}_{-1.000}) \times 10^4$	$4.0^{+1.0}_{-1.0}\,^{+3.1}_{-1.9}$
	0.750	$(3.915^{+0.557}_{-0.636}\,^{+8.959}_{-3.232}) \times 10^3$	$1.0^{+0.0}_{-0.0}\,^{+2.3}_{-1.9}$	0.422	$(1.952^{+0.230}_{-0.318}\,^{+4.466}_{-1.611}) \times 10^3$	$1.0^{+0.0}_{-0.0}\,^{+2.3}_{-0.8}$
	2.371	$0.000^{+0.000}_{-0.000}$	$0.0^{+0.0}_{-0.0}\,^{+1.8}_{-0.0}$	1.334	$0.000^{+0.000}_{-0.000}$	$0.0^{+0.0}_{-0.0}\,^{+0.8}_{-0.0}$
MACSJ1149	0.024	$0.000^{+0.000}_{-0.000}$	$0.0^{+0.0}_{-0.0}\,^{+1.8}_{-0.0}$	0.013	$(5.032^{+13.67}_{-4.121}\,^{+6.605}_{-3.244}) \times 10^4$	$2.0^{+1.0}_{-1.0}\,^{+2.6}_{-2.6}$
	0.075	$< 7.449 \times 10^5$	< 3.0	0.042	$(1.862^{+1.193}_{-1.046}\,^{+3.244}_{-1.199}) \times 10^4$	$2.0^{+1.0}_{-1.0}\,^{+2.6}_{-2.3}$
	0.237	$(1.835^{+1.306}_{-1.074}\,^{+4.199}_{-1.515}) \times 10^4$	$1.0^{+0.0}_{-0.0}\,^{+2.3}_{-0.8}$	0.133	$(1.048^{+0.699}_{-0.590}\,^{+2.398}_{-0.865}) \times 10^4$	$1.0^{+1.0}_{-0.0}\,^{+2.3}_{-0.8}$
	0.750	$< 1.281 \times 10^4$	< 3.0	0.422	$< 5.787 \times 10^3$	< 3.0
	2.371	$0.000^{+0.000}_{-0.000}$	$0.0^{+0.0}_{-0.0}\,^{+1.8}_{-0.0}$	1.334	$0.000^{+0.000}_{-0.000}$	$0.0^{+0.0}_{-0.0}\,^{+0.0}_{-0.0}$
Combined	0.024	$< 7.530 \times 10^5$	< 3.0	0.013	$(3.157^{+17.40}_{-1.795}\,^{+2.484}_{-1.505}) \times 10^4$	$4.0^{+5.0}_{-2.0}\,^{+3.1}_{-1.9}$
	0.075	$< 4.782 \times 10^5$	< 3.0	0.042	$(2.501^{+7.015}_{-0.828}\,^{+1.968}_{-1.294}) \times 10^4$	$4.0^{+4.0}_{-2.0}\,^{+3.1}_{-1.9}$
	0.237	$(2.207^{+1.781}_{-1.133}\,^{+2.896}_{-1.421}) \times 10^4$	$2.0^{+2.0}_{-1.0}\,^{+2.6}_{-1.3}$	0.133	$(1.645^{+0.828}_{-0.675}\,^{+1.294}_{-0.784}) \times 10^4$	$4.0^{+3.0}_{-2.0}\,^{+3.1}_{-1.9}$
	0.750	$(4.025^{+1.261}_{-1.377}\,^{+9.212}_{-3.323}) \times 10^3$	$1.0^{+3.0}_{-0.0}\,^{+2.3}_{-0.8}$	0.422	$(1.994^{+8.675}_{-1.994}\,^{+4.563}_{-1.646}) \times 10^3$	$1.0^{+2.0}_{-1.0}\,^{+2.3}_{-0.8}$
	2.371	$< 1.237 \times 10^4$	< 3.0	1.334	$< 6.162 \times 10^3$	< 3.0

Notes. Column 1: cluster field where the counts are computed. In the bottom row group, counts combining all cluster fields are listed. Column 2: flux density bin for differential counts. Column 3: median differential counts per flux bin. For non-zero median counts, uncertainties are given separately using the 16th (84th) percentiles and scaled Poisson confidence levels for 1σ lower (upper) limits. For flux density bins having zero median counts and non-zero values at the 84th percentile, only 3σ upper limits are provided. Column 4: median number of sources per flux bin. Uncertainties are given separately using the 16th (84th) percentiles and Poisson confidence levels for 1σ lower (upper) limits. Column 5: flux density limit for cumulative counts. Column 6: median cumulative counts per flux limit. Uncertainties and upper limits are as in Column 3. Column 7: median number of sources per flux limit. Uncertainties are as in Column 4.
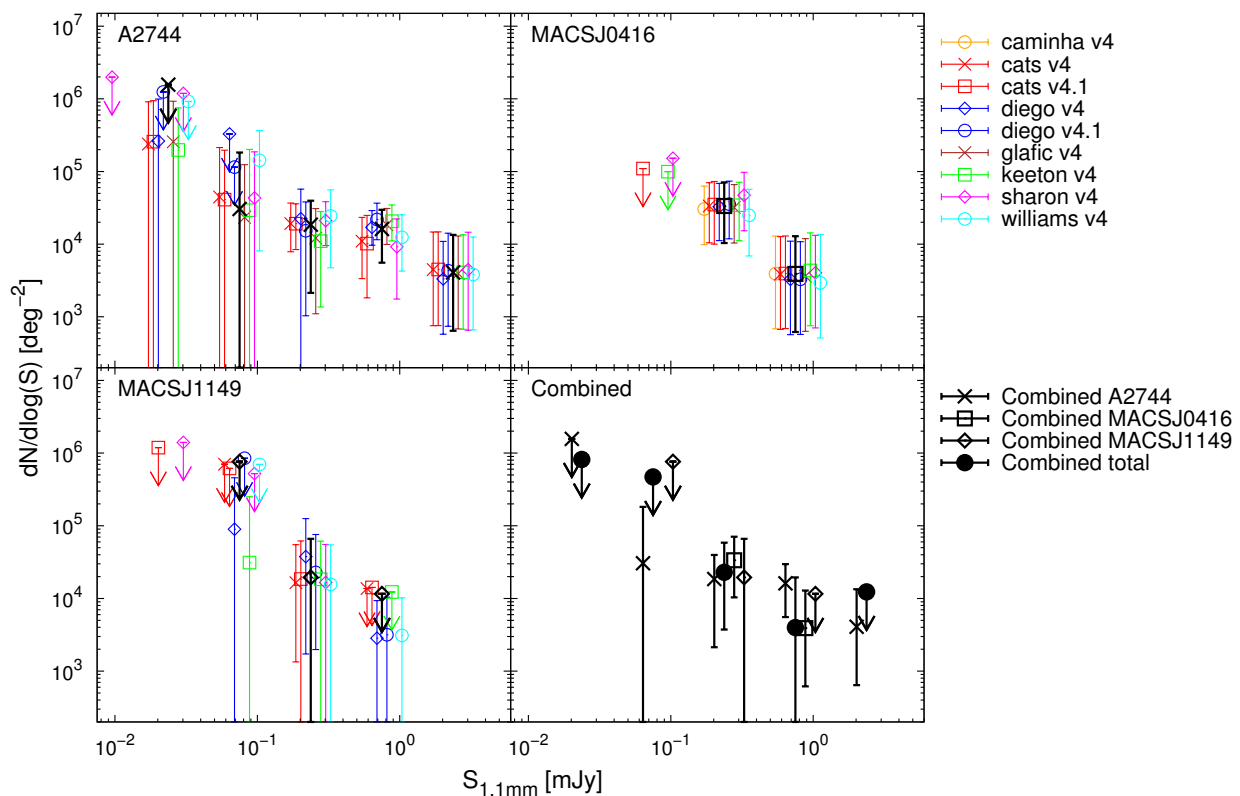
Fig. 11. Demagnified differential counts at 1.1 mm, for each cluster (see legends at top-left) and combining all cluster fields (bottom-right panel). Values correspond to median counts for the lens models listed in Table 2 (colored symbols), combining all models for each cluster field (large black crosses, squares and diamonds) and combining all models for all cluster fields (large black filled circles). Error bars indicate the 16th and 84th percentiles, adding the scaled Poisson confidence levels for 1σ lower and upper limits respectively in quadrature. Arrows indicate 3σ upper limits for flux density bins having zero median counts and non-zero values at the 84th percentile. In the first three panels, counts for each model have been offset in flux around the combined counts for clarity. In the bottom-right panel, this is done for each galaxy cluster field around the counts that combine all models for all cluster fields.

vious counts derived at wavelengths other than 1.1 mm. In those cases, we scale their estimates as $S_{1.1\,\mathrm{mm}} = 1.29 \times S_{1.2\,\mathrm{mm}}$ and $S_{1.1\,\mathrm{mm}} = 1.48 \times S_{1.3\,\mathrm{mm}}$. These conversion factors are derived by assuming a characteristic modified blackbody spectrum (following Hatsukade et al. 2016), and are adopted for consistency with previous works (which assume distinct SED templates).

Within the uncertainties, our estimates for both differential and cumulative number counts are in good agreement with all the aforementioned works for the two or three brightest bins, that is, down to 0.422 mJy. At $0.133 - 0.422$ mJy, our number counts are consistent within 1σ with all but Fujimoto et al. (2016) and Hatsukade et al. (2016) data. At flux densities fainter than 0.133 mJy, however, the derived 3σ upper limits to our differential counts are consistent with both the Fujimoto et al. (2016) data and their Schechter (1976) best-fitting function. Also below this flux density, our cumulative counts are lower by ≈ 1 dex than Aravena et al. (2016) data, being inconsistent with their results at a 1σ level. Our findings suggest a flattening of the number counts.

The counts derived from serendipitously detected sources are based on detections in fields that targeted a previously defined set of sources. These counts are expected to be biased, as the detections might be clustered around the original targets (Hatsukade et al. 2016). Restricting only to flux densities above 0.133 mJy, we are not able to quantify this bias, given the large uncertainties in our derived counts. Neither can we make a strong distinction between our counts, which are based solely on observations

lensed by galaxy clusters, and those derived from blank-field observations. Intriguingly, our counts in the brighter flux density bins are consistent with those found by Umehata et al. (2017) toward the SSA22 protocluster, both including and not including their detections having spectroscopic redshifts coincident with the protocluster (in Fig. 13 we show only the first case).

In addition to recent ALMA data, Fig. 13 shows the counts predicted by galaxy formation models down to $\lesssim 0.1$ mJy. In particular, Cowley et al. (2015) use the semi-analytic model GALFORM (Lacey et al. 2016) to predict the submm counts. We compare our results to their cumulative number counts at 1.1 mm for their simulated lightcones down to 0.1 mJy. On the other hand, Béthermin et al. (2017) and Schreiber et al. (2017) present simulations of the extragalactic sky (SIDES and EGG, respectively). For Béthermin et al. (2017), we compare our results to their cumulative number counts at 1.2 mm for their "intrinsic" simulation down to 0.01 mJy, while for Schreiber et al. (2017) we compare to their differential and cumulative number counts at 1.2 mm down to 10^{-8} mJy. In these last two cases we rescale their counts to 1.1 mm as was done for 1.2 mm observations above.

Cowley et al. (2015) obtain the dust SED per galaxy in a self-consistent way (see also Lacey et al. 2016), using a simplified model based in the spectrophotometric code GRASIL (Silva et al. 1998) that agrees with GRASIL predictions at rest-frame wavelengths $> 70\,\mu$m. We note that the constraints for their model parameters include the observed cumulative counts
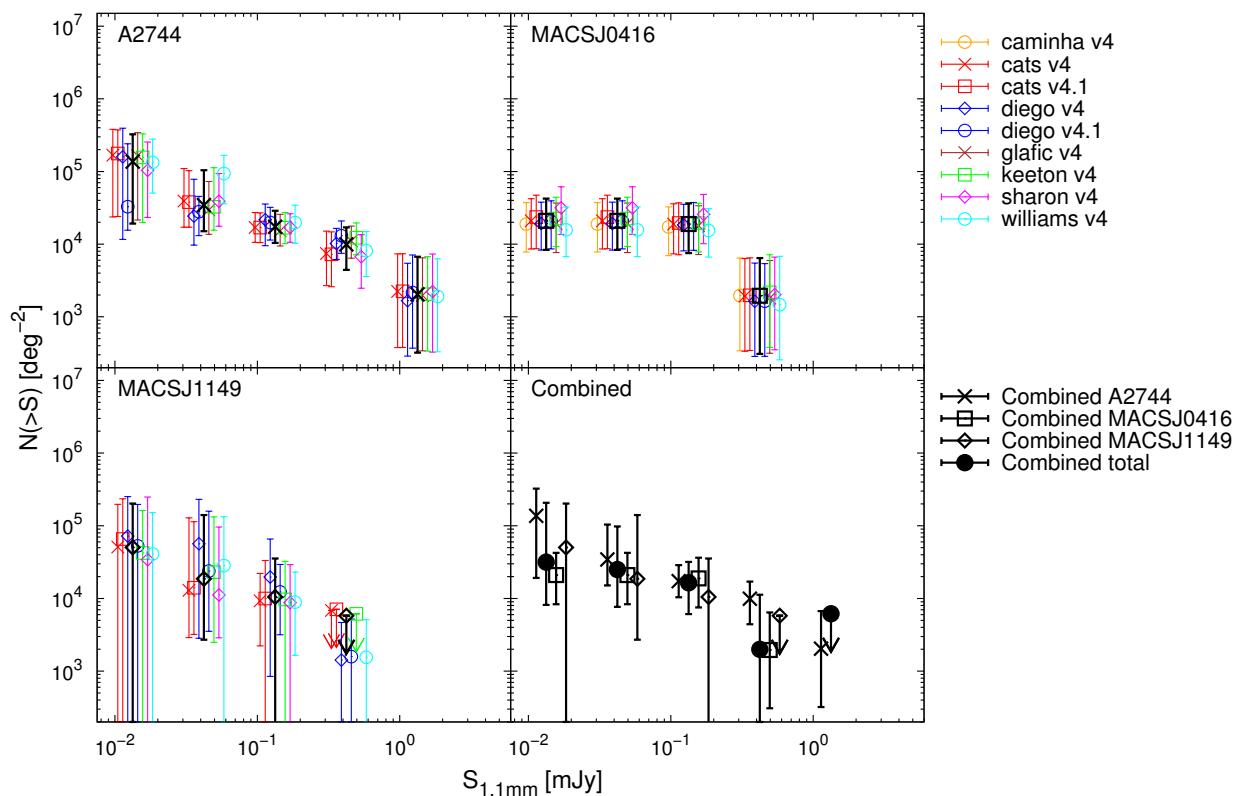
Fig. 12. As in Fig. 11, but for the demagnified cumulative number counts at 1.1 mm.

at 850 μm and the redshift distribution of sources with flux density > 5 mJy at 850 μm (see Lacey et al. 2016). Béthermin et al. (2017) and Schreiber et al. (2017) use the phenomenological model 2SFM (two star formation modes; Sargent et al. 2012), which is based on the observed evolution of the main sequence with redshift. Both groups add their own assumptions for constructing the mock catalogs and estimating further source properties from empirical prescriptions. They also choose particular SED libraries (which cover the FIR-to-submm wavelengths) for assigning spectra to model sources based on these properties (see Béthermin et al. 2017; Schreiber et al. 2017). Both groups calibrate their models using particular sets of observational constraints for the SED evolution, based on stacking analyses. For Béthermin et al. (2017), these constraints include LABOCA 870 μm and AzTEC 1.1 mm data (see Béthermin et al. 2015a). Schreiber et al. (2017) note that at 1.2 mm they do not calibrate their FIR SEDs nor prescriptions, although their constraints include ALMA 870 μm data in the Extended Chandra Deep Field South (Extended CDFS). From Fig. 13, we note that these three models agree well in the displayed flux range. Although none of the models predict number counts as shallow as our estimates, at flux densities < 0.1 mJy they predict counts around 1σ lower than Fujimoto et al. (2016) and Aravena et al. (2016) values, and agree with our predictions within 1σ uncertainties.

4.3. Effect of source sizes

For exploring the effect of varying the image-plane source sizes on the demagnified number counts, we test the following extreme cases, which should bracket our expectations. Firstly, assuming that $4.5 \leq S/N < 5$ sources are point sources. This assumption is supported by recent results regarding DSFG sizes, both from several publicly available ALMA maps at 1 mm (Fu-

jimoto et al. 2017) and an ALMA follow-up of SCUBA2 sources in the CDFS at 850 μm (González-López et al. 2018, in prep.). Fujimoto et al. (2017) find a positive correlation between source size and bolometric luminosity in the FIR; an extrapolation of this trend to lower luminosities may suggest that our sources at lower S/N are more compact than high-significance detections. Similarly, González-López et al. (2018, in prep.) find that robustly selected DSFGs at a few mJy (equivalent to $\gtrsim$ 0.5 mJy at 1.1 mm) in the CDFS have compact sizes on average, with a median effective radius $\approx$ 0$\overset{''}{.}$08. And secondly, assuming that $4.5 \leq S/N < 5$ sources have more extreme observed effective radii of $r_{\rm eff,obs} = 0.5''$. This value is motivated by the 1σ uncertainty found for the largest image-plane scale radius among the high-significance detections. In this case, we obtain the integrated flux densities of the low-significance detections scaling the peak intensities by the typical ratios 0.55, 0.84, and 0.84 in A2744, MACSJ0416, and MACSJ1149, respectively.

Estimated number counts for these cases are shown in Fig. 14, together with our fiducial case. We find that assuming $r_{\rm eff,obs} = 0.5''$ for low-significance sources leads to an agreement with Aravena et al. (2016) at 1σ. Assuming that our low-significance detections are point sources disagrees with their estimates at 3σ, although remains consistent with Fujimoto et al. (2016) counts assuming our 3σ upper limit.

Below 0.133 mJy, our fiducial number counts are consistent with available data from both serendipitous and blank-field surveys only at a 3σ level. The discrepancy with our median counts could be attributed to cosmic variance or to the aforementioned observational biases. However, it may also reveal the need for further corrections in our number counts. More specifically, we may require a proper treatment for the stretching that source shapes experience in the image plane. There may be sources missed because their high magnifications led them to have lensed
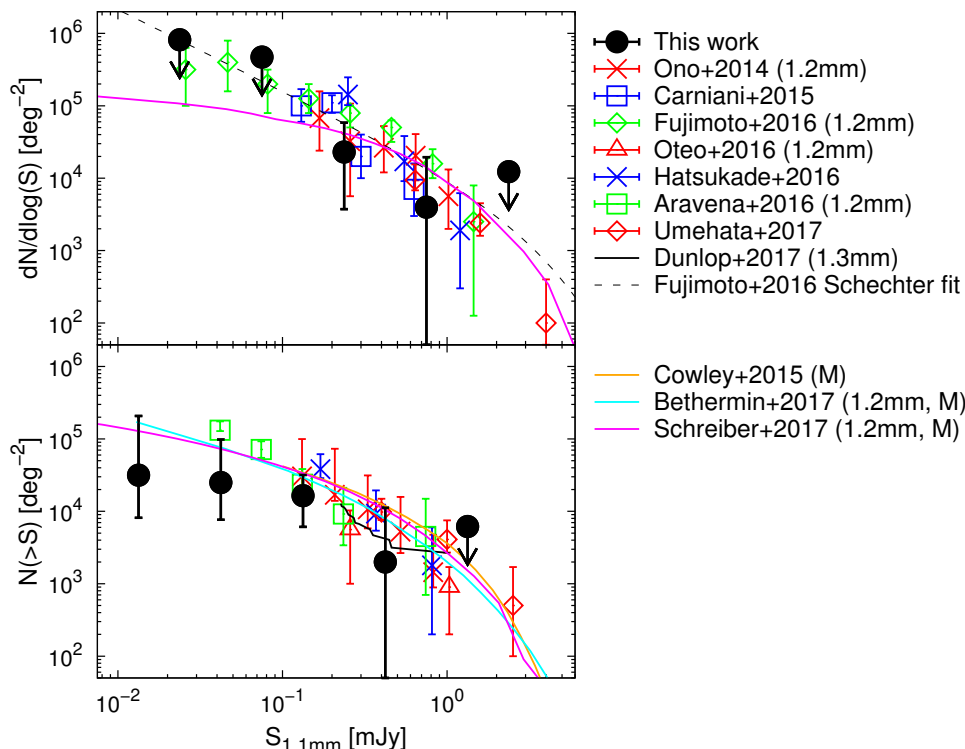
Fig. 13. Differential (top) and cumulative (bottom) counts at 1.1 mm compared to ALMA results and galaxy formation model predictions from the literature. Our counts (large black filled circles) correspond to median values combining all models for all cluster fields. Error bars indicate the 16th and 84th percentiles, adding the scaled Poisson confidence levels for 1σ lower and upper limits respectively in quadrature. Arrows indicate 3σ upper limits for flux densities having zero median counts and non-zero values at the 84th percentile. We show previous results reported by Ono et al. (2014) as red crosses, Carniani et al. (2015) as blue squares, Fujimoto et al. (2016) as green diamonds (with their Schechter fit shown as a black dashed line), Oteo et al. (2016) as red triangles, Hatsukade et al. (2016) as blue crosses, Aravena et al. (2016) as green squares, Umehata et al. (2017) as red diamonds and Dunlop et al. (2017) as a black solid curve. We show number counts predicted by the galaxy formation models from Cowley et al. (2015) (orange line), Béthermin et al. (2017) (cyan line) and Schreiber et al. (2017) (magenta line). We scale the counts derived at other wavelengths as $S_{1.1\,mm} = 1.29 \times S_{1.2\,mm}$ and $S_{1.1\,mm} = 1.48 \times S_{1.3\,mm}$ (following Hatsukade et al. 2016).

angular sizes greater than a synthesized beam. In order to take these effects into account, we would need to assume a distribution of source sizes at several redshifts and a set of different intrinsic source geometries, as well as passing them through the uv and lens modeling. Accounting for this could elevate the derived number counts in the faint end, if the dust emission from low-significance detections is more extended than suggested by extrapolations to current observational data (Fujimoto et al. 2017 and González-López et al. 2018 in prep., see §2.1.2). We leave a detailed analysis about the impact of demagnified source geometries on the number counts for future work, so at low flux densities our reported counts are strictly computed for the beam size quoted for each ALMA mosaic.

4.4. Contribution to the extragalactic background light

We use the Monte Carlo realizations of the differential number counts to compute the contribution to the EBL provided by each of them, adding up the contribution contained in each flux bin. From this procedure, we estimate a median contribution of $8.222^{+8.837}_{-4.188}$ ($3.861^{+7.847}_{-2.059}$) Jy deg^{-2} resolved in our demagnified sources at 1.1 mm down to 0.013 (0.133) mJy, with uncertainties computed from the 16th and 84th percentiles.

We compare our estimate with the total EBL measurement at that wavelength estimated by the Planck collaboration using their best-fit extended halo model. Following Aravena et al. (2016), we interpolate the Planck estimate (see

Planck Collaboration et al. 2014, Table 10) finding an EBL of $19.143^{+0.751}_{-0.723}$ Jy deg^{-2} at 263.14 GHz, which is the set Local Oscillator frequency for our observations (see Paper I). The contribution provided by our demagnified sources represents $43^{+46}_{-22}\%$ ($20^{+41}_{-11}\%$) of this EBL at 1.1 mm down to 0.013 (0.133) mJy. As expected from Fig. 13, this contribution is lower than (although consistent to $\approx 1.5\sigma$ with) results by Carniani et al. (2015) and Hatsukade et al. (2016), both at 1.1 mm. Carniani et al. (2015) found an estimate of 17^{+10}_{-5} Jy deg^{-2} down to 0.1 mJy, while a value around 12 (14) Jy deg^{-2} is obtained when we extrapolate the Schechter (double power law) best-fitting function by Hatsukade et al. (2016) down to 0.1 mJy.

5. Summary

We have derived lensing-corrected number counts at 1.1 mm exploiting: 1) the high resolution and depth reached in a dedicated ALMA survey of three galaxy clusters (i.e., A2744, MACSJ0416, and MACSJ1149) as part of the Frontier Fields program, and 2) the public availability of several models for the mass reconstruction of these clusters. This is the first time that the surface density of DSFGs is estimated around three well-studied galaxy clusters using ALMA data. Our source catalog includes S/N ≥ 5 detections already introduced with the ALMA Frontier Fields Survey (Paper I), plus $4.5 \leq$ S/N < 5 detections reported in the present work. We correct the counts for completeness and fraction of spurious sources. Moreover, we develop a
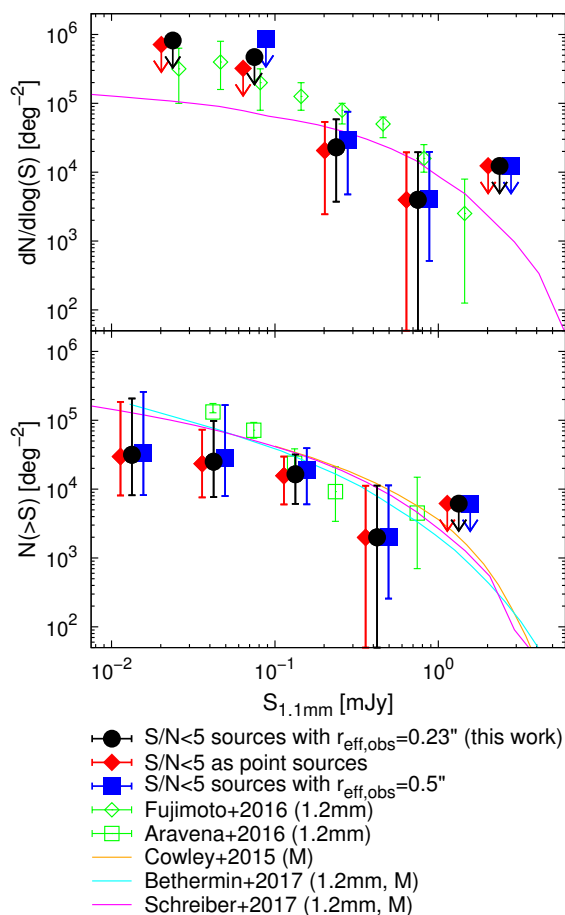
Fig. 14. Differential (top) and cumulative (bottom) counts at 1.1 mm for different assumptions regarding the image-plane source scale radii for low-significance sources: adopting $r_{eff,obs} = 0''23$ (black filled circles, fiducial); assuming they are point sources (red filled diamonds); and adopting $r_{eff,obs} = 0''5$ (blue filled squares). Our counts correspond to median values combining all models for all cluster fields. Error bars indicate the 16th and 84th percentiles, adding the scaled Poisson confidence levels for 1σ lower and upper limits respectively in quadrature. Arrows indicate 3σ upper limits for flux densities having zero median counts and non-zero values at the 84th percentile. We show previous results reported by Fujimoto et al. (2016) as green diamonds and Aravena et al. (2016) as green squares. We show number counts predicted by the galaxy formation models from Cowley et al. (2015) (orange line), Béthermin et al. (2017) (cyan line) and Schreiber et al. (2017) (magenta line). We scale the counts derived at other wavelength as $S_{1.1\,mm} = 1.29 \times S_{1.2\,mm}$ (following Hatsukade et al. 2016).

careful treatment to fold the magnification uncertainties in the derived counts using a Monte Carlo simulation.

Our ALMA mosaics of the three FF galaxy clusters cover a total observed area of $\sim 14\,arcmin^2$, which results in a smaller effective area in the source plane once a lens model is applied (e.g., the total area is reduced by ~ 2.6 times in the CATS v4 model for a source-plane $z = 2$). Combining all cluster fields, our differential number counts span ~ 2.5 orders of magnitude in demagnified flux density, going from the mJy level down to tens of μJy. We find an overall agreement between the counts derived for different lens models in a given cluster field. Within the error bars in our number counts (coming from both Poisson errors and lensing model uncertainties) our results are consistent at 3σ with recent estimates from deep ALMA observations (Ono et al. 2014; Carniani et al. 2015; Fujimoto et al. 2016; Oteo et al.

2016; Hatsukade et al. 2016; Aravena et al. 2016; Umehata et al. 2017; Dunlop et al. 2017). However, below ≈ 0.1 mJy our cumulative number counts are ≈ 1 dex lower than previous estimates. Our work suggests a flattening of the number counts, and implies that we may finally be seeing a turn over.

Using publicly available lens models and a statistical approach, we are able to derive 1.1 mm number counts around three galaxy clusters, down to demagnified flux densities ≈ 4 times fainter than the rms level reached in our deepest ALMA mosaic. This highlights the potential of finding even fainter sources in these FFs with deeper ALMA data, suggesting that future 1.1 mm observations reaching an rms of such as $10\,\mu$Jy could yield number counts down to $\approx 2.5\,\mu$Jy in these fields. Additionally, further spectroscopic redshift determinations for our detections could serve as new constraints for lensing models, helping to increase the accuracy in the magnification estimates (Johnson & Sharon 2016) and hence in the number counts derived from future deep surveys.

References

Acebron, A., Jullo, E., Limousin, M., et al. 2017, MNRAS, 470, 1809
Ammons, S. M., Wong, K. C., Zabludoff, A. I., & Keeton, C. R. 2014, ApJ, 781, 2
Aravena, M., Decarli, R., Walter, F., et al. 2016, ApJ, 833, 68
Balestra, I., Mercurio, A., Sartoris, B., et al. 2016, ApJS, 224, 33
Béthermin, M., Daddi, E., Magdis, G., et al. 2015a, A&A, 573, A113
Béthermin, M., De Breuck, C., Sargent, M., & Daddi, E. 2015b, A&A, 576, L9
Béthermin, M., Wu, H.-Y., Lagache, G., et al. 2017, A&A, 607, A89
Blain, A. W., Kneib, J.-P., Ivison, R. J., & Smail, I. 1999, ApJ, 512, L87
Blain, A. W., Smail, I., Ivison, R. J., Kneib, J.-P., & Frayer, D. T. 2002, Phys. Rep., 369, 111
Bouwens, R. J., Oesch, P. A., Illingworth, G. D., Ellis, R. S., & Stefanon, M. 2017, ApJ, 843, 129
Bussmann, S., Leung, T. K. ., & Conley, A. 2016, uvmcmcfit: Parametric models to interferometric data fitter, Astrophysics Source Code Library
Caminha, G. B., Grillo, C., Rosati, P., et al. 2017, A&A, 600, A90
Carniani, S., Maiolino, R., De Zotti, G., et al. 2015, A&A, 584, A78
Casey, C. M., Narayanan, D., & Cooray, A. 2014, Phys. Rep., 541, 45
Castellano, M., Amorín, R., Merlin, E., et al. 2016, A&A, 590, A31
Chapman, S. C., Blain, A. W., Smail, I., & Ivison, R. J. 2005, ApJ, 622, 772
Coe, D., Bradley, L., & Zitrin, A. 2015, ApJ, 800, 84
Coe, D., Fuselier, E., Benítez, N., et al. 2008, ApJ, 681, 814
Condon, J. J. 1997, PASP, 109, 166
Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693

Coppin, K., Chapin, E. L., Mortier, A. M. J., et al. 2006, MNRAS, 372, 1621
Cowley, W. I., Lacey, C. G., Baugh, C. M., & Cole, S. 2015, MNRAS, 446, 1784
Di Criscienzo, M., Merlin, E., Castellano, M., et al. 2017, A&A, 607, A30
Diego, J. M., Broadhurst, T., Molnar, S. M., Lam, D., & Lim, J. 2015, MNRAS, 447, 3130
Diego, J. M., Protopapas, P., Sandvik, H. B., & Tegmark, M. 2005, MNRAS, 360, 477
Diego, J. M., Tegmark, M., Protopapas, P., & Sandvik, H. B. 2007, MNRAS, 375, 958
Dunlop, J. S., McLure, R. J., Biggs, A. D., et al. 2017, MNRAS, 466, 861
Ebeling, H., Ma, C.-J., & Barrett, E. 2014, ApJS, 211, 21
Eddington, A. S. 1913, MNRAS, 73, 359
Fujimoto, S., Ouchi, M., Ono, Y., et al. 2016, ApJS, 222, 1
Fujimoto, S., Ouchi, M., Shibuya, T., & Nagai, H. 2017, ApJ, 850, 83
Gehrels, N. 1986, ApJ, 303, 336
González-López, J., Bauer, F. E., Romero-Cañizales, C., et al. 2017, A&A, 597, A41
Greve, T. R., Bertoldi, F., Smail, I., et al. 2005, MNRAS, 359, 1165
Grillo, C., Karman, W., Suyu, S. H., et al. 2016, ApJ, 822, 78
Hatsukade, B., Kohno, K., Umehata, H., et al. 2016, PASJ, 68, 36
Hayward, C. C., Narayanan, D., Kereš, D., et al. 2013, MNRAS, 428, 2529
Hezaveh, Y. D. & Holder, G. P. 2011, ApJ, 734, 52
Hoag, A., Huang, K.-H., Treu, T., et al. 2016, ApJ, 831, 182
Hodge, J. A., Karim, A., Smail, I., et al. 2013, ApJ, 768, 91
Hogg, D. W. & Turner, E. L. 1998, PASP, 110, 727
Hsu, L.-Y., Cowie, L. L., Chen, C.-C., Barger, A. J., & Wang, W.-H. 2016, ApJ, 829, 25
Hughes, D. H., Serjeant, S., Dunlop, J., et al. 1998, Nature, 394, 241
Jauzac, M., Clément, B., Limousin, M., et al. 2014, MNRAS, 443, 1549
Jauzac, M., Richard, J., Jullo, E., et al. 2015, MNRAS, 452, 1437
Jauzac, M., Richard, J., Limousin, M., et al. 2016, MNRAS, 457, 2029
Johansson, D., Sigurdarson, H., & Horellou, C. 2011, A&A, 527, A117
Johnson, T. L & Sharon, K. 2016, ApJ, 832, 82
Johnson, T. L., Sharon, K., Bayliss, M. B., et al. 2014, ApJ, 797, 48
Jullo, E. & Kneib, J.-P. 2009, MNRAS, 395, 1319
Jullo, E., Kneib, J.-P., Limousin, M., et al. 2007, New Journal of Physics, 9, 447
Karim, A., Swinbank, A. M., Hodge, J. A., et al. 2013, MNRAS, 432, 2
Kawamata, R., Ishigaki, M., Shimasaku, K., et al. 2018, ApJ, 855, 4
Kawamata, R., Oguri, M., Ishigaki, M., Shimasaku, K., & Ouchi, M. 2016, ApJ, 819, 114
Keeton, C. R. 2010, General Relativity and Gravitation, 42, 2151
Kneib, J.-P. & Natarajan, P. 2011, A&A Rev., 19, 47
Knudsen, K. K., van der Werf, P. P., & Kneib, J.-P. 2008, MNRAS, 384, 1611
Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2016, MNRAS, 462, 3854
Laporte, N., Bauer, F. E., Troncoso-Iribarren, P., et al. 2017, A&A, 604, A132
Liesenborgs, J., De Rijcke, S., & Dejonghe, H. 2006, MNRAS, 367, 1209
Liesenborgs, J., de Rijcke, S., Dejonghe, H., & Bekaert, P. 2007, MNRAS, 380, 1729
Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, ApJ, 837, 97
Mahler, G., Richard, J., Clément, B., et al. 2018, MNRAS, 473, 663
McCully, C., Keeton, C. R., Wong, K. C., & Zabludoff, A. I. 2014, MNRAS, 443, 3631
Meneghetti, M., Natarajan, P., Coe, D., et al. 2017, MNRAS, 472, 3177
Molino, A., Benítez, N., Ascaso, B., et al. 2017, MNRAS, 470, 95
Muñoz Arancibia, A. M., Navarrete, F. P., Padilla, N. D., et al. 2015, MNRAS, 446, 2291
Oguri, M. 2010, PASJ, 62, 1017
Ono, Y., Ouchi, M., Kurono, Y., & Momose, R. 2014, ApJ, 795, 5
Oteo, I., Zwaan, M. A., Ivison, R. J., Smail, I., & Biggs, A. D. 2016, ApJ, 822, 36
Owers, M. S., Randall, S. W., Nulsen, P. E. J., et al. 2011, ApJ, 728, 27
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2012, ArXiv e-prints [arXiv:1201.0490]
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A30
Pope, A., Scott, D., Dickinson, M., et al. 2006, MNRAS, 370, 1185
Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25
Priewe, J., Williams, L. L. R., Liesenborgs, J., Coe, D., & Rodney, S. A. 2017, MNRAS, 465, 1030
Rawle, T. D., Altieri, B., Egami, E., et al. 2016, MNRAS, 459, 1626
Richard, J., Jauzac, M., Limousin, M., et al. 2014, MNRAS, 444, 268
Sargent, M. T., Béthermin, M., Daddi, E., & Elbaz, D. 2012, ApJ, 747, L31
Schechter, P. 1976, ApJ, 203, 297
Schreiber, C., Elbaz, D., Pannella, M., et al. 2017, A&A, 602, A96
Sebesta, K., Williams, L. L. R., Mohammed, I., Saha, P., & Liesenborgs, J. 2016, MNRAS, 461, 2126
Silva, L., Granato, G. L., Bressan, A., & Danese, L. 1998, ApJ, 509, 103
Smail, I., Ivison, R. J., & Blain, A. W. 1997, ApJ, 490, L5
Smolčić, V., Aravena, M., Navarrete, F., et al. 2012, A&A, 548, A4
Trenti, M. & Stiavelli, M. 2008, ApJ, 676, 767
Treu, T., Brammer, G., Diego, J. M., et al. 2016, ApJ, 817, 60
Treu, T., Schmidt, K. B., Brammer, G. B., et al. 2015, ApJ, 812, 114
Umehata, H., Tamura, Y., Kohno, K., et al. 2017, ApJ, 835, 98
Walter, F., Decarli, R., Aravena, M., et al. 2016, ApJ, 833, 67
Weiß, A., Kovács, A., Coppin, K., et al. 2009, ApJ, 707, 1201
Younger, J. D., Fazio, G. G., Huang, J.-S., et al. 2007, ApJ, 671, 1531
Zemcov, M., Blain, A., Halpern, M., & Levenson, L. 2010, ApJ, 721, 424
Zitrin, A., Fabris, A., Merten, J., et al. 2015, ApJ, 801, 44