An ultradeep multiband VLA survey of the faint radio sky (COSMOS-XS) source catalog and number counts

Vlugt, D. van der; Algera, H.S.B.; Hodge, J.A.; Novak, M.; Radcliffe, J.F.; Riechers, D.A.; ... ; Walter, F.

AN ULTRA-DEEP MULTI-BAND VLA SURVEY OF THE FAINT RADIO SKY (COSMOS-XS): SOURCE CATALOG AND NUMBER COUNTS

D. van der Vlugt[1], H. S. B. Algera[1], J. A. Hodge[1], M. Novak[2], J. F. Radcliffe[3,4], D. A. Riechers[5,2], H. Röttgering[1], V. Smolčić[6], and F. Walter[2]

Draft version September 30, 2020

ABSTRACT

We present ultra-deep, matched-resolution Karl G. Jansky Very Large Array (VLA) observations at 10 and 3 GHz in the COSMOS field: the COSMOS-XS survey. The final 10 and 3 GHz images cover $\sim$ 16 arcmin^2 and $\sim$ 180 arcmin^2 and reach median *rms* values of 0.41 μJy beam^{-1} and 0.53 μJy beam^{-1}, respectively. Both images have an angular resolution of $\sim$ 2.0$''$. To fully account for the spectral shape and resolution variations across the broad bands, we image all data with a multi-scale, multi-frequency synthesis algorithm. We present source catalogs for the 10 and 3 GHz image with 91 and 1498 sources, respectively, above a peak brightness threshold of 5σ. We present source counts with completeness corrections included that are computed via Monte Carlo simulations. Our corrected radio counts at 3 GHz with direct detections down to $\sim$ 2.8 μJy are consistent within the uncertainties with other results at 3 and 1.4 GHz, but extend to fainter flux densities than previous direct detections. The ultra-faint 3 GHz number counts are found to exceed the counts predicted by the semi-empirical radio sky simulations developed in the framework of the SKA Simulated Skies project, consistent with previous P(D) analyses. Our measured source counts suggest a steeper luminosity function evolution for these faint star-forming sources. The semi-empirical Tiered Radio Extragalactic Continuum Simulation (T-RECS) predicts this steeper evolution and is in better agreement with our results. The 10 GHz radio number counts also agree with the counts predicted by the T-RECS simulation within the expected variations from cosmic variance. In summary, the multi-band, matched-resolution COSMOS-XS survey in the well-studied COSMOS field provides a high-resolution view of the ultra-faint radio sky that can help guide next generation radio facilities.

Keywords: catalogs-cosmology: observations-radio continuum: galaxies

1. INTRODUCTION

Over the last decade, a number of multi-wavelength (UV to radio) studies have revealed that the star-formation history of the universe (SFHU, i.e., the total star-formation rate (SFR) per unit of co-moving volume) went through several phases. The SFR apparently rose after the first galaxies formed (e.g., Bouwens et al. 2014, 2015) and reached its peak during the "epoch of galaxy assembly" at $1 \lesssim z \lesssim 3$. Subsequently, the SFR density declined rapidly to the present (e.g., Madau & Dickinson 2014 and references therein). It is well established that the majority of the star-formation activity happens in galaxies that lie on the main sequence (MS), exhibiting an intrinsic scatter of $\sim$ 0.3 dex, between SFR and stellar mass (e.g., Brinchmann et al. 2004; Noeske et al. 2007; Elbaz et al. 2007). Not only does the relation exist in the local universe (e.g., Brinchmann et al. 2004; Salim et al. 2007) but is found to hold to $z \sim 4$ or even higher, albeit with a strong redshift evolution (e.g., Daddi

et al. 2007; Pannella et al. 2009; Karim et al. 2011; Speagle et al. 2014; Schreiber et al. 2015; Salmon et al. 2015; Tomczak et al. 2016; Kurczynski et al. 2016). An accurate measurement of the SFR and its relation to stellar mass at all epochs is key for a better understanding of galaxy evolution.

Several tracers across the electromagnetic spectrum can be used to measure this SFR, each with their own unique strengths and weaknesses. For example, ultraviolet (UV) light originates mainly from massive stars and thus directly traces young stellar populations. UV based observations, however, require uncertain and significant model-dependent dust corrections to account for dust obscuration (e.g., Carilli et al. 2008; Siana et al. 2008, 2009; Magdis et al. 2010; Chary & Pope 2010; Bouwens et al. 2012). Infrared (IR) observations trace the absorbed UV emission that is thermally reprocessed by dust surrounding newly formed stars. However, the resolution of IR observations is often insufficient to provide reliable multi-wavelength identifications. In addition, the presence of polycyclic aromatic hydrocarbon (PAH) emission features redshifted ($z > 0.8$) into the 24μm band, commonly used as an estimator for the total-IR emission, makes the required k-correction uncertain.

Long-wavelength radio emission is another potential tracer of recent star-formation (Condon 1992). Radio emission in galaxies below rest frame frequencies $\lesssim$ 30 GHz is dominated by synchrotron radiation arising from cosmic ray electrons gyrating in the galaxy's magnetic field (e.g., Sadler et al. 1989; Condon 1992; Clemens et al. 2008; Tabatabaei et al. 2017). These

Electronic address: dvdvlugt@strw.leidenuniv.nl
[1] Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, the Netherlands
[2] Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
[3] Jodrell Bank Centre for Astrophysics, The University of Manchester, SK11 9DL. United Kingdom
[4] Department of Physics, University of Pretoria, Lynnwood Road, Hatfield, Pretoria 0083, South Africa
[5] Department of Astronomy, Cornell University, Space Sciences Building, Ithaca, NY 14853, USA
[6] Department of Physics, University of Zagreb, Bijenička cesta 32, 10002 Zagreb, Croatia

charged cosmic-ray particles are accelerated in shocks launched by supernovae of stars with M $> 8\,M_\odot$ in SFGs. These massive stars have lifetimes of $\lesssim 3 \times 10^8$ years, so their supernova rates are proportional to the recent SFR. This is supported by the tight correlation observed in SFGs between IR emission, originating from dust that has been heated by young and massive stars, and radio emission (the IR-radio correlation; e.g. de Jong et al. 1985; Helou et al. 1985; Yun et al. 2001; Bell 2003; Dumas et al. 2011). Deep radio observations in the synchrotron regime can thus be used to constrain SFRs. Radio observations can also provide high spatial resolution to allow for reliable counterpart matching. Distant galaxies in the GHz radio regime often have spectral energy distributions (SEDs) that can be parameterised by a featureless powerlaw, leading to a simple and robust k-correction. Radio observations in the synchrotron regime thus offers a unique opportunity to study the star-formation history of the Universe (e.g., Condon et al. 2002; Seymour et al. 2008; Smolčić et al. 2009b; Jarvis et al. 2015; Calistro Rivera et al. 2017; Novak et al. 2017) at a wavelength that is free from selection biases due to dust obscuration.

However, there are two challenges in using radio emission in the synchrotron regime as a tracer of star-formation. The first is the "contamination" by active galactic nuclei (AGN). It is hard to disentangle AGN and SFGs in the radio regime as an accreting supermassive black hole (SMBH) in an AGN can also accelerate the electrons that produce synchrotron emission. To attempt to correct for this, several methods have been developed for identifying different types of AGN and seperating them from SFGs (e.g., Hickox et al. 2009; Mendez et al. 2013; Smolčić et al. 2017b; Delvecchio et al. 2017; Algera et al. 2020) using mid-IR data, far-IR data, X-ray information and multi-band optical/IR photometry. Synchrotron emission can thus not only be used to study star-formation but also to study the blackhole accretion activity in the Universe (e.g., Jarvis & Rawlings 2000; Smolčić et al. 2009a; Rigby et al. 2011; McAlpine et al. 2013; Best et al. 2014; Delvecchio et al. 2014, 2017; Sabater et al. 2019).

The second challenge is the depth achievable in radio surveys. With the improving capabilities of modern interferometers, along with sophisticated calibration techniques, surveys of the faint μJy radio emitting objects are able to constrain the faint populations (e.g., Rujopakarn et al. 2016; Murphy et al. 2017; Smolčić et al. 2017a; Owen 2018; Bondi et al. 2018; Mauch et al. 2020). At high flux densities, the source counts are well-constrained and found to be dominated by AGNs that follow a smooth power-law distribution down to $S_{1.4\mathrm{GHz}} \sim 1\,\mathrm{mJy}$ (e.g., Condon & Mitchell 1984; Windhorst et al. 1990). Below 1 mJy, the Euclidean normalised source counts flatten (e.g., Richards 2000; Huynh et al. 2005; Biggs & Ivison 2006; Owen & Morrison 2008; Bondi et al. 2008; Padovani et al. 2015). It is now widely accepted that this flattening observed is due to the emergence of SFGs and radio-quiet AGN which begin to contribute significantly (e.g., Rowan-Robinson et al. 1993; Seymour et al. 2004; Padovani et al. 2009). New deep radio observations and $P(D)$ analyses on confusion limited surveys show evidence of a further steepening of the number counts below $S_{1.4\mathrm{GHz}} \sim 50\,\mu\mathrm{Jy}$ (e.g., Condon et al. 2012;

Vernstrom et al. 2014, 2016; Smolčić et al. 2017a; Prandoni et al. 2018; Mauch et al. 2020). The composition of this ultra faint radio population is still uncertain but it is expected from simulations and observations that the fraction of SFGs will become significant ($> 60\,\%$) below $S_{1.4\mathrm{GHz}} \sim 100\,\mu\mathrm{Jy}$ (Smolčić et al. 2017b). Constraints on the ultra-faint radio populations at high resolution are useful for predictions for future radio surveys with new and upcoming facilities such the ASKAP, MeerKAT, ngVLA, and SKA, where source confusion noise may be an issue.

Survey depth is also a challenge for radio surveys at observing frequencies $\geq 10\,\mathrm{GHz}$. Such surveys measure flux densities closer to the rest-frame frequencies $\nu \geq 30\,\mathrm{GHz}$ where the total radio emission is dominated by free-free radiation (e.g Condon 1992; Murphy et al. 2011; Klein et al. 2018) which constitutes the faintest part of the radio SED. Although more difficult to detect, free-free emission provides independent information on the star-formation process. Free-free emission directly originates from the HII-regions where massive stars form and thus provides a first-hand view on star-formation. In addition, unlike UV emission, dust obscuration plays only a minor role, and therefore free-free emission is potentially the most accurate tracer of star-formation (e.g. Mezger & Henderson 1967; Turner & Ho 1983, 1985; Klein & Graeve 1986; Kobulnicky & Johnson 1999; Murphy et al. 2012, 2015; Nikolic & Bolton 2012).

An additional limitation of high frequency observations is the smaller primary beam area, as this area decreases with frequency ($\Omega_{\mathrm{pb}} \propto \nu^{-2}$). This limits the area that is covered at a certain depth. Low frequencies have therefore been favoured and radio continuum surveys probing free-free emission are still sparse in the literature. The majority of extragalactic radio surveys are instead conducted at 1.4 and 3 GHz, where the radio emission of galaxies is intrinsically brighter than at higher frequencies, and large areas can be imaged with a single pointing. In the last decade, several studies have been conducted to trace the synchrotron emission from galaxies (e.g., Schinnerer et al. 2007, 2010; Morrison et al. 2010; Smolčić et al. 2017a; Prandoni et al. 2018). The COSMOS field has been observed with the Very Large Array (VLA) at 1.4 GHz ($\sigma \sim 10 - 15\,\mu\mathrm{Jy}\,\mathrm{beam}^{-1}$, Schinnerer et al. 2010), and more recently at 3 GHz with substantially better sensitivity ($\sigma \sim 2.3\,\mu\mathrm{Jy}\,\mathrm{beam}^{-1}$), yielding about four times more radio sources compared to the 1.4 GHz data (Smolčić et al. 2017a), see Fig. 1. Although these observations provide valuable data over the entire 2 deg^2 COSMOS field, enabling some of the most comprehensive studies to-date of the SFG and radio AGN populations, they are equivalent to ~ 2 hours per pointing. The resulting sensitivity only allows for the detection of typical SFGs out to $z \sim 1.5$ (Novak et al. 2017), the epoch where the various SF tracers begin to diverge (e.g., Ilbert et al. 2013). In order to observe MS galaxies over the full epoch of galaxy assembly ($z \sim 1 - 3$) a sub-μJy survey is essential.

The significantly increased bandwidths of the upgraded NSF's Karl G. Jansky Very Large Array (VLA) now allows for observations of the radio sky down to several hundred nJy beam^{-1} sensitivities. We have taken advantage of these recent upgrades to the VLA to do an ultra-deep-matched-resolution survey in both the X- and
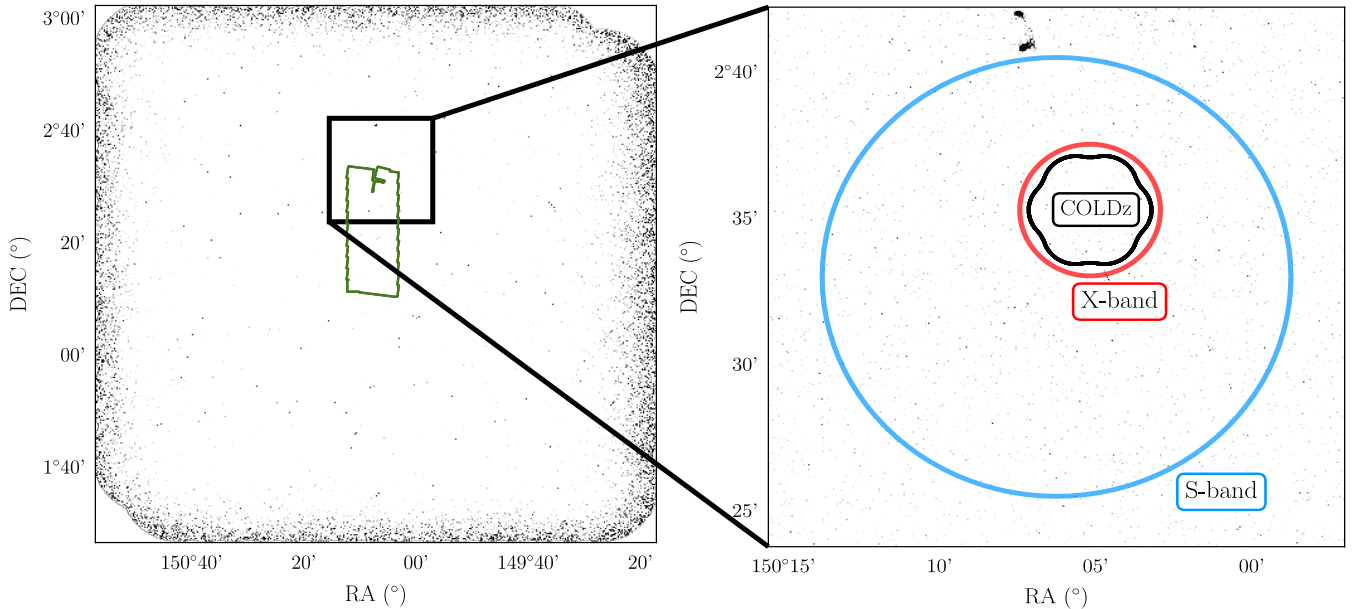
Figure 1. The full 3 GHz mosaic as imaged by Smolčić et al. (2017a) of the 2 deg² COSMOS field. The green polygon indicates the footprint of the CANDELS WFC3 imaging (Nayyeri et al. 2017). The inset of 340 arcmin² shows the position of the X- and S-band primary beam. Additionally, the seven-pointing mosaic at 34 GHz, part of the COLDz project (Pavesi et al. 2018), is shown with the black contour.

Table 1
10 and 3 GHz pointing centres and total observing time.

Band	Central frequency [GHz]	Configuration	Centre (J2000) RA	Dec	Total integration time [hours]	Primary beam FWHP [arcmin]	central r.m.s. [μJy]
S	3	B	$10^h00^m25^s$	$+02°33'00''$	100	15	0.53
X	10	C	$10^h00^m20^s.7$	$+02°35'17''$	90	4.5	0.41

S-band (10 GHz and 3 GHz). This COSMOS-XS survey is one of the deepest radio surveys to-date, reaching sub-μJy sensitivities, and is ∼ 5 times deeper than the previous 3 GHz observations conducted in the COSMOS field (Smolčić et al. 2017a). When combined with the rich COSMOS multi-wavelength data, this survey thus yields a unique data set to test the composition of the faintest radio source populations that can currently be probed.

The combined S- and X-band observations will enable us to study the properties and importance of AGNs and SFGs in a dust-unbiased way and to make predictions for the populations to be detected by future surveys. This paper (hereafter Paper I), describes the 10 and 3 GHz observations and examines their implications for the ultra-faint source counts. In a companion paper (Algera et al. 2020; hereafter Paper II), we match the obtained catalogs with the multi-wavelength data available in the COSMOS field to distinguish between AGNs and SFGs. The obtained populations are then used to constrain the composition of the ultra-faint source counts.

This paper is organised as follows: in Section 2 we describe the VLA 10 and 3 GHz observations, calibration, imaging and catalog extraction. In Section 3, we present the final images and describe the source-detection method and the compilation of the source catalog. This section also includes an analysis of the quality of the catalog. In Section 4 we discuss the derivation of the completeness-corrected radio source counts. Finally, Section 5 summarizes and concludes this work.

Throughout this paper, the spectral index, α, is defined as $S_\nu \propto \nu^\alpha$, where S is the source flux density, and ν is the observing frequency. We assume a spectral index of -0.7 unless otherwise stated.

2. OBSERVATIONS AND DATA REDUCTION

2.1. *Observations*

The COSMOS-XS survey (see Table 1) consists of the combination of a single deep S-band pointing centered on RA=10:00:25, Dec=+02°33'00'' (see Fig. 1) in the B-configuration and an additional single X-band pointing observed in the C-array centered on the COSMOS/AzTEC-3 protocluster at $z \sim 5.3$, coordinates RA=10:00:20.7 Dec=02°35'17'' (see Fig. 1). The chosen configurations provide a resolution of $\sim 2''$ in X- and S-band which is large enough to avoid resolving out faint sources. The X-band surveys overlaps with the COLDz survey (Pavesi et al. 2018; Riechers et al. 2019), one of the deepest 34 GHz continuum surveys to date, covering $\sim 10 \, \text{arcmin}^2$. The S-band pointing is chosen to be slightly offset from the X-band pointing to overlap more with the CANDELS-COSMOS field (Nayyeri et al. 2017). The X- and S-bands were observed for 90h and 100h, respectively. These data were taken between 4 Dec 2014 and 27 Feb 2016 with the individual observations constituting either two- or five-hour observing blocks. For the X-band observations, J1024-0052 was used as the phase calibrator, while for the S-band observations, J0925+0019 was used as the phase calibrator. For both bands, 3C286 served as both the flux and

bandpass calibrator. The X-band covers a bandwidth of 4096 MHz centered at 10 GHz, and is separated into 32 spectral windows 128 MHz wide. All four polarization products were recorded and a 2s signal-averaging time was used. The S-band covers a bandwidth of 2048 MHz centered at 3 GHz, and is separated into 16 spectral windows 128 MHz wide. Again all four polarization products were recorded and a 5s signal-averaging time was used.

2.2. Calibration

The calibration of the X-band and S-band visibilities was performed using CASA[7] version 5.0.0. Extensive use was made of the NRAO VLA reduction pipeline[8]. Radio frequency interference (RFI) constitutes the major uncertainty in our data calibration. We experimented with different methods and tools (`rflag`, `AOflagger`; Offringa 2010) to remove the RFI. However, the unmodified pipeline resulted in the best image.

Before running the pipeline, Hanning smoothing was applied to lessen Gibbs ringing from strong spectral features such as strong, narrow RFI. The pipeline runs several flagging rounds to flag bad or unnecessary data such as the initial few integration points of a scan where not all antennas may be on source. To flag RFI the pipeline performed several rounds of flagging using the `rflag` algorithm. The pipeline uses the `gencal` and `gaincal` tasks to derive the necessary calibration solutions. `setjy` was used to to set the flux scale of the observations (Perley & Butler 2013) and the flux density calibrator was 3C286. During the calibration, no time or bandwidth averaging was performed so as to minimize time and bandwidth smearing effects (see Section 2.3).

In the X-band two spectral windows (SPWs), SPW 32 and SPW 33, were heavily flagged ($> 85\%$) because of RFI. For the S-band only SPW 4 was found to be heavily corrupted by RFI and was flagged almost entirely by the pipeline. After the pipeline was run on the 10 and 3 GHz data, the target field was split off using the task `split` and, respectively, 24.9% and 25.8% of the split off data was flagged.

2.3. Bandwidth and time smearing

An antenna receiver has a finite bandwidth which causes bandwidth smearing. This effect radially smears peak brightness while the integrated flux densities are conserved. Bandwidth smearing is a function of distance from the pointing center. The theoretical prediction from Condon et al. (1998) for the reduction of peak response is given by $I/I_0 = 1/\sqrt{1 + 0.46\beta^2}$, where $\beta = (\Delta\nu/\nu_0) \times (\theta_0/\theta_{HPBW})$. If we calculate the reduction at 20% of the peak primary beam sensitivity using the VLA channel width $\Delta\nu = 2$ MHz, central frequency $\nu_0 = 3$ GHz and a beam size of $\theta_{HPBW} = 2''$, we find an offset of $\sim 1\%$. When we compare the peak brightness over the total flux density for point-like sources ($0.9 \leq S_{\mathrm{peak}}/S_{\mathrm{int}}$) with S/N > 6 as a function of distance from the pointing center, we find that an offset of $\sim 4\%$ is present. This is, however, not distance dependent, and thus unlikely to be related to bandwidth smearing.

[7] http://casa.nrao.edu/
[8] https://science.nrao.edu/facilities/vla/data-processing/pipeline

Table 2
Overview of the wide-field imaging parameters for the 10 and 3 GHz image.

Band	robust	pixel size [arcsec]	w-planes	restoring beam [arcsec]
S	0.5	0.4	400	2.14 × 1.81
X	0.5	0.42	128	2.33 × 2.01

For the 10 GHz observations, we calculate a reduction of $\sim 0.01\%$ at 20% of the peak primary beam sensitivity using the VLA channel width $\Delta\nu = 2$ MHz, central frequency $\nu_0 = 10$ GHz and a beam size of $\theta_{HPBW} = 2''$. The $S_{\mathrm{peak}}/S_{\mathrm{int}}$ distribution of point-like sources ($0.9 \leq S_{\mathrm{peak}}/S_{\mathrm{int}}$) with S/N > 6 with distance also shows no distance dependent offset. Therefore, we do not apply any corrections at 10 and 3 GHz for the bandwidth smearing effect.

The individual observations were concatenated after the calibration steps using the CASA task `concat`. These data were subsequently averaged in time using `split` with an averaging time of 21s for the X-band and 6s for the S-band. Averaging visibility data in the time domain causes a similar distortion to bandwidth smearing, but in the opposite direction (i.e. tangentially). The averaging time used should lead to an amplitude loss of at most one percent for a point source located at the first null of the primary beam due to loss of coherence, in the given VLA configurations.[9]

2.4. Imaging

Imaging of the concatenated datasets was performed both with CASA task `clean` and standalone imaging algorithm `WSclean` (Offringa et al. 2014). For our purposes, the main difference between the two algorithms is the method of taking into account the interferometric w-term during deconvolution. We opted for `WSclean` to create the final images, based on its faster processing compared to CASA's `clean`. However, differences between the CASA `clean` and `WSclean` images are minimal (see also Offringa et al. 2014), hence the choice of algorithm has no effect on the end product.

`WSclean` produces images by jointly gridding and deconvolving the measurement set, which is called joined-channel deconvolution (Offringa & Smirnov 2017). Spectral behavior of sources can be captured during deconvolution by setting the parameters `-channels-out` and `-join-channels`. The data is then imaged in separate channels across the band. During deconvolution, `WSclean` finds peaks in the full-band image and deconvolves these in each channel independently.

For the 10 GHz image, we weight our image via Briggs weighting with a robust parameter of 0.5 (Briggs 1995) and apply w-stacking using the minimum recommended number of 128 layers (`-nwlayers`). We use the joined-channel deconvolution technique, specified by setting `-channels-out` to 32 (i.e. one per SPW) and specifying parameter `-join-channels`. A power law is fit to account for in-band spectral variations (similar to the multi-term multi-frequency synthesis algorithm used in CASA's `tclean` with nterms=2; Rau & Cornwell 2011). In addition, we utilize auto-masking of sources after first

[9] https://science.nrao.edu/facilities/vla/docs/manuals/oss2013B/performance/fov/t-av-loss

cleaning down to 3σ, whereupon masked sources are further cleaned down to 0.5σ (parameters `-auto-mask` and `-auto-threshold`, respectively). The combination of these settings for `-auto-mask` and `-auto-threshold` is a good general setting for `WSclean`, and leaves almost no residuals. The resulting image reaches an r.m.s.-sensitivity of $\sigma = 0.41\,\mu$Jy beam^{-1} (Table 2).

For the 3 GHz image, we take a similar approach as described above, but with a few changes in the imaging parameters. The number of layers for w-stacking is increased to 400. We find that w-term artifacts persist for imaging with 128 layers and increasing the number to 400 solves these issues. We also do not fit the spectral variations with a powerlaw because of the large noise fluctuations at the beam edge caused by a bright source (see Fig. 4). The `-channels-out` parameter is set to 16 (i.e. one per SPW). The resulting image reaches an r.m.s.-sensitivity of $\sigma = 0.53\,\mu$Jy beam^{-1} (Table 2).

Because primary beam correction via projection-based gridding is not yet operational, we take the primary beam model from the CASA `widebandpbcor` task[10] which takes the frequency dependence of the beam into account.

2.5. *Confusion*

Finally, we note that the source confusion is negligible in the deep 3 GHz VLA observation (only the S-band image is considered as it has the highest source density). The beam size is $2\rlap.''14 \times 1\rlap.''81$, which results in 3.16×10^6 beams per deg^2. At the faintest flux density bin of our number count measurements for the S-band (see Sec. 4.1.6), we find $\sim 1 \times 10^4$ sources per deg^2. This translates to one source per ~ 316 beams, and implies that confusion is not an issue. Following Condon et al. (2012), source confusion becomes important at one source per 25 beams. This confusion limit depends on the slope of the scale-free power law approximation for the number counts

$$n(S) = kS^{-\gamma}, \tag{1}$$

where K is the count normalization and $1 < \gamma < 3$ is the differential count slope (Condon et al. 2012). For the calculation of the confusion limit we assumed a slope of $\gamma = 2.0$. We expect source confusion to contribute approximately $0.01\,\mu$Jy beam^{-1} to the noise following Equation (27) from Condon et al. (2012).

3. FINAL IMAGE AND CATALOGING

The final 10 and 3 GHz images are shown in Fig. 3 and Fig. 5, respectively. The central r.m.s. noise level is relatively smooth for both images (see also Fig. 2 and 4). The 3 GHz image shows a small number of artifacts (see e.g., the Northern part of the image shown in Fig. 5). The artifacts are localized around bright sources and have little impact on the majority of the map, as can be seen from Fig. 4.

3.1. *Source detection and characterization*

We compiled a source catalog using PyBDSF[11] (Mohan & Rafferty 2015) to detect and characterize sources. We

---

[10] This task was run by means of CASA version 5.3, which has an updated beam shape model.
[11] http://www.astron.nl/citt/pybdsf/

ran PyBDSF on the final image, using the pre-primary-beam-corrected image as the detection image.

PyBDSF identifies peaks of emission above a given threshold (`thresh_pix`) that are surrounded by contiguous pixels with emission greater than a minimum (`thresh_isl`). PyBDSF fits the identified island with one or more Gaussians, which are subsequently grouped into sources. This happens if all the pixels on the line joining their centers have a value greater than `thresh_isl` and if the length of this line is less than half the sum of their FWHMs. The total flux density of the sources is estimated by adding those from the individual Gaussians, while the central position and source size are determined via moment analysis.

The spatial variation of the image noise was estimated by sliding a box across the image in overlapping steps, calculating the root mean square (r.m.s.) of the pixels within the box and interpolating the values measured from each step. The resulting r.m.s. map provides PyBDSF with an estimate of the spatial variation of the image noise for detection thresholding purposes. The r.m.s. maps for the 10 (see Fig. 2) and 3 GHz image (see Fig. 4) are determined with a sliding box of `rms_box` = (120, 60) pixels and `rms_box` = (400, 100) pixels (i.e. a box size of 400 pixels every 100 pixels), respectively.

For source extraction, we used `thresh_pix` = 5.0σ and `thresh_isl` = 3.0σ (i.e. the limit at which flux density is included in the source for fitting). Fig. 2 and 4 illustrate the variation in r.m.s. noise determined across the image and shows the increase in local r.m.s. as a result of calibration artifacts near bright sources. We used the `group_tol` parameter with a value of 1.0 as a larger value (allowing for larger sources to be fitted) was not necessary. Sources are classified as 'S' for single sources and 'M' for multiple Gaussian sources. The parameters of the Gaussian fitted to the source are reported by PyBDSF.

For the 10 GHz (3 GHz) images, the total number of sources detected by PyBDSF within 20% of the peak primary beam sensitivity is 93 (1498), of which 90 (1392) are single-component sources; sources fitted by a single Gaussian.

3.2. *Resolved sources*

In order to determine whether our identified source components are resolved, we make use of the ratio between integrated flux density ($S_{\rm int}$) and peak brightness ($S_{\rm peak}$) which is a direct measure of the extension of a radio source. For an unresolved source, the peak brightness equals the total flux density. Noise influences the measurements of $S_{\rm int}$ and $S_{\rm peak}$ and, therefore, the $S_{\rm int}/S_{\rm peak}$ distribution gets broadened towards the low signal-to-noise end. The effect of noise on the $S_{\rm int}/S_{\rm peak}$ ratio can be determined by performing Monte Carlo simulations in which simulated sources are added to the pre-primary-beam-corrected image and then retrieved the same way as the observed data. We simulate 100 mock sources and insert these in the real image. This is repeated 200 times, simulating 20000 sources in total. The method described below is used to derive the upper envelope of the $S_{\rm int}/S_{\rm peak}$ distribution for both the 10 and 3 GHz image.

Sources are injected as Gaussians with their FWHM equal to the beam size and they are thus unresolved by construction. We only insert point sources to quantify
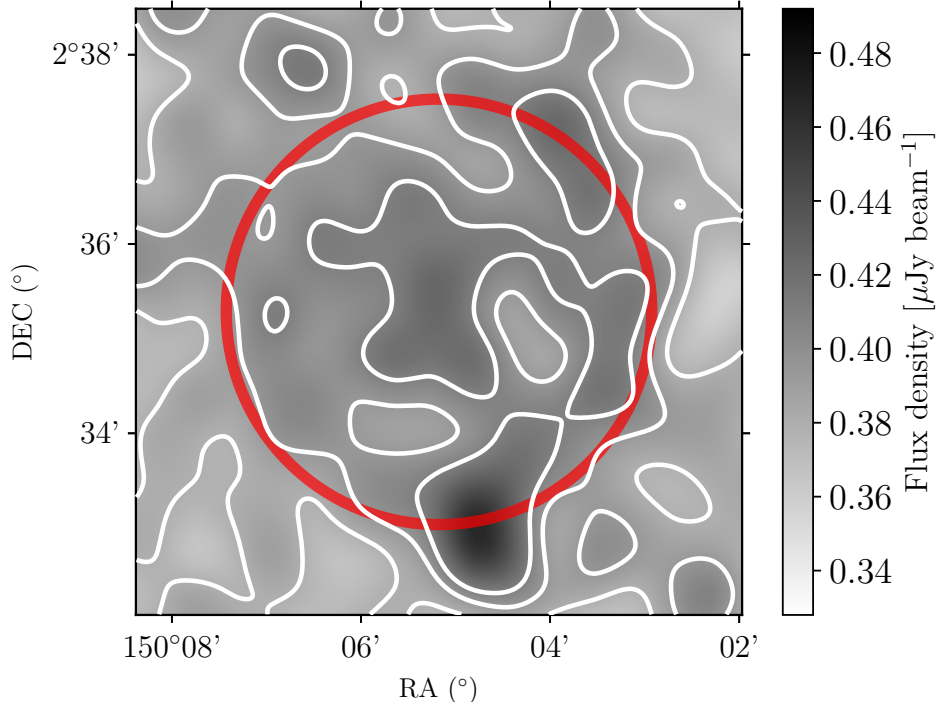
Figure 2. The r.m.s. map of the 10 GHz observation before primary beam correction. The image size is 41 arcmin². The r.m.s. map is created with PyBDSF. The red circle indicates the HPBW of the primary beam at 10 GHz, which corresponds to 4.5′. The grey scale shows the r.m.s. noise from 0.8σ to 1.2σ, where $\sigma = 0.41\,\mu$Jy beam⁻¹ The contours are plotted at [0.38, 0.39, 0.41] μJy beam⁻¹.
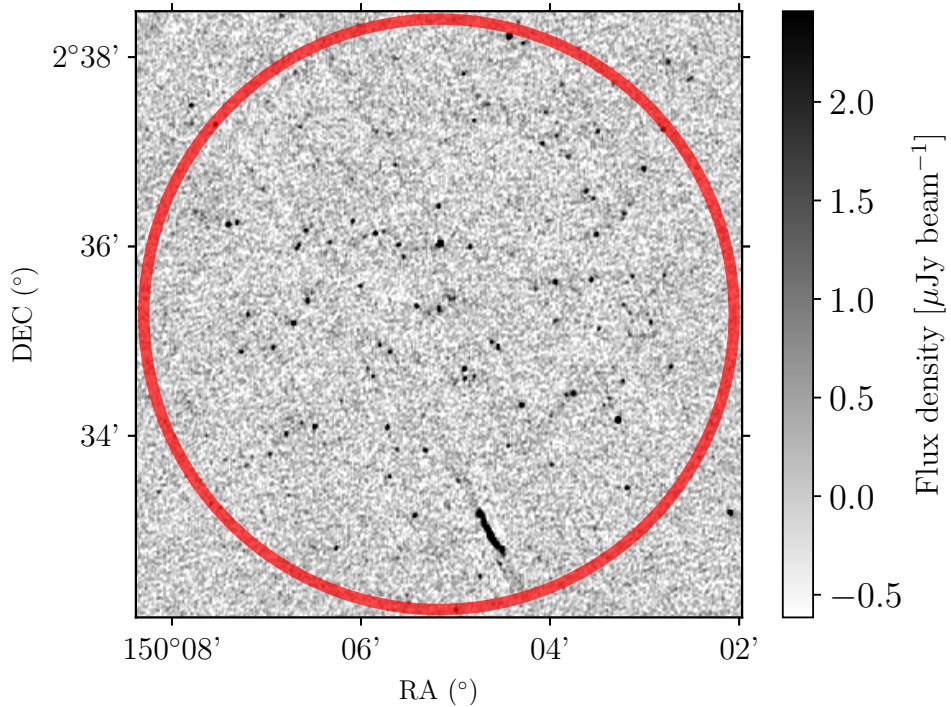
Figure 3. Final calibrated 10 GHz image before primary beam correction. The size is the same as in Fig. 2. The red circle indicates the point where the primary beam sensitivity is 20% of its peak. The grey-scale shows the flux density from -1.5σ to 6σ, where $\sigma = 0.41\,\mu$Jy beam⁻¹ is the median r.m.s. within the primary beam FWHP. The corresponding brightness temperature r.m.s. value is $\sigma = 1.25$ mK. The image is matched in resolution and depth for a spectral index of -0.7 with the 3 GHz image.
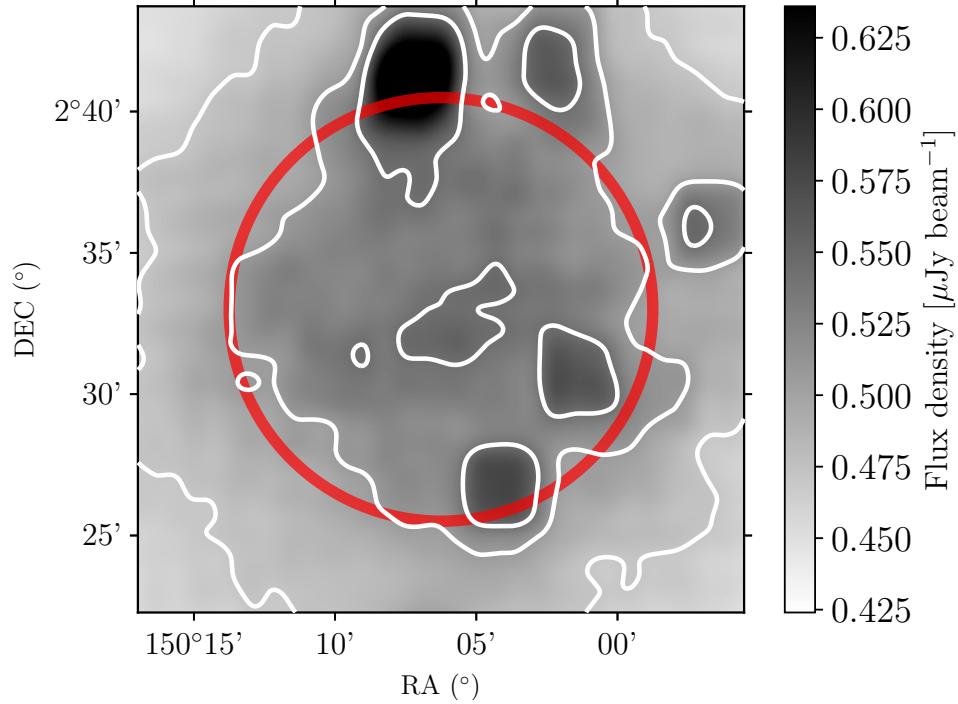
Figure 4. The r.m.s. map of the 3 GHz observation before primary beam correction. The image size is $460\,\mathrm{arcmin}^2$. The r.m.s. map is created with PyBDSF. The red circle indicates the HPBW of the primary beam at 3 GHz, which corresponds to $15'$. The grey scale shows the r.m.s. noise from 0.8σ to 1.2σ, where $\sigma = 0.53\,\mu\mathrm{Jy\,beam}^{-1}$. The contours are plotted at $[0.47, 0.51, 0.54]\,\mu\mathrm{Jy\,beam}^{-1}$.
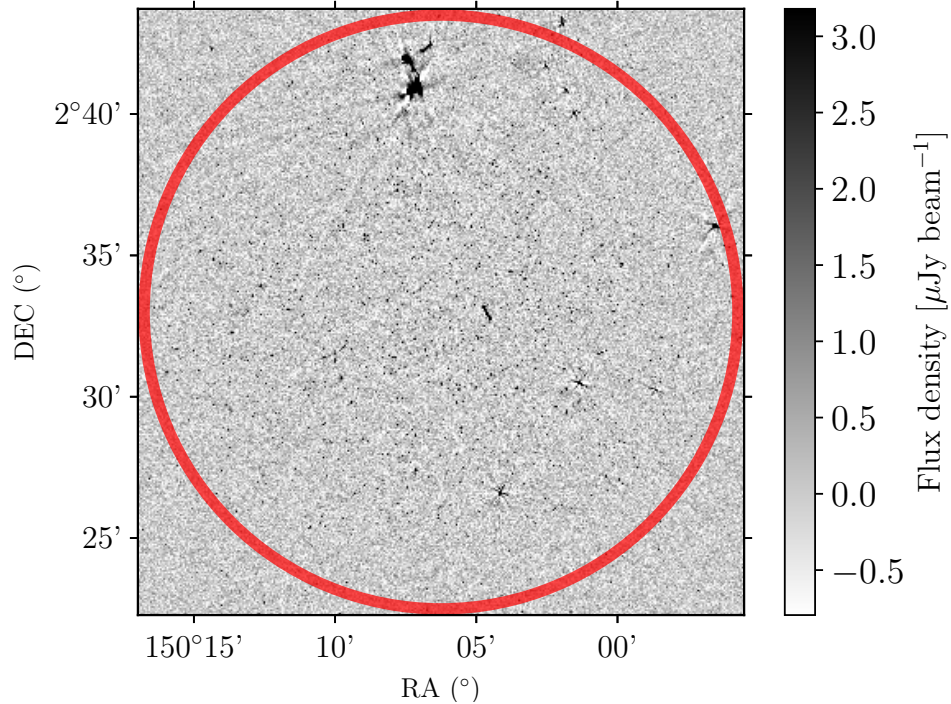
Figure 5. Final calibrated 3 GHz image before primary beam correction. The size is the same as in Fig. 4. The red circle indicates the point where the primary beam sensitivity is 20% of its peak. The grey-scale shows the flux density from -1.5σ to 6σ, where $\sigma = 0.53\,\mu\mathrm{Jy\,beam}^{-1}$ is the median r.m.s. within the primary beam FWHP. The corresponding brightness temperature r.m.s. value is $\sigma = 18.0\,\mathrm{mK}$.

how the noise 'resolves' unresolved sources. The peak brightness of the sources are drawn from the real source distribution to generate a realistic mock catalog. We fitted a power law of the form $n = a \times S^{-\gamma} + b$ (see Equation 1) to the binned measured peak brightness distribution, and draw fluxes between 3σ and 60σ randomly from this distribution. The mock sources are assigned a position that is at least 20 pixels ($\sim 8''$) away from both real sources and other mock sources. The position of the mock source also has to lie within the 20 per cent power point of the primary beam. The position was randomly chosen until both restrictions are satisfied.

After all mock sources are inserted, we run PyBDSF with the exact parameters as performed for the real sources. Since the extraction is carried out on a map containing both real and mock emission, the real sources are always recovered and had to be filtered out in order to keep only simulated sources in the extracted catalog.

The recovered $S_{\text{int}}/S_{\text{peak}}$ distribution as a function of S/N is shown in Fig. 6. To determine the 95 per cent envelope in Fig. 6, a curve (red line) is fitted to the 95th percentile of logarithmic bins across the S/N, where N $= \sigma_{\text{local}}$ as measured by PyBDSF. The shape of the envelope was chosen following Bondi et al. (2008). The fit for the 10 GHz image simulations is given by $S_{\text{int}}/S_{\text{peak}} = 1.14 + 11.6 \times (\text{S/N})^{-1.64}$. The fit for the 3 GHz image simulations is given by $S_{\text{int}}/S_{\text{peak}} = 1.07 + 14.0 \times (\text{S/N})^{-1.69}$.

We consider sources from our catalog that lie above this envelope to be resolved. For the 10 GHz image there are 12 (13%) resolved sources and for the 3 GHz image there are 475 (32%) resolved sources. For sources that lie under the envelope, the integrated flux density was set equal to the peak brightness. The resolved sources are flagged as resolved in the final catalog presented in Section 3.8. Note that not all the PyBDSF sources with multiple Gaussian components are resolved by this criterion as each component is considered separately. Conversely, not all single-component sources are unresolved.

3.3. Flux boosting

The noise fluctuations in the image may influence the flux measurements of the extracted sources. Since the counts of faint sources increase with decreasing flux density, there should be a sea of faint sources below the noise level that may influence the source extraction. There is therefore a probability that intrinsically faint sources are detected at higher flux because of noise fluctuations. This effect, called flux boosting, is extremely important at low S/N where flux measurement can be overestimated.

The degree of flux boosting (the probability that faint sources are detected at higher flux density because of noise fluctuations) can be estimated by examining the output-to-input flux density of the simulations described in the above section. Fig. 7 shows the distribution of the recovered flux density minus the inserted flux density normalized by the inserted flux density as a function of recovered flux density, where we calculated the mean and standard deviation in logarithmic bins.

The effect for flux boosting is, as expected, greatest at the flux limit of our survey. For the 10 GHz (3 GHz) image, sources with S/N $\simeq 5$ are boosted by 11% (15%) on average. The boosting effect quickly decreases with

S/N, we find that sources with S/N $\simeq 10$ are boosted by less then 5% on average. Therefore, we do not correct for flux boosting.

3.4. Flux density uncertainties at 3 GHz

In order to determine any systematic offsets, we have compared our flux densities to those of the VLA-COSMOS 3 GHz Large Project. For the comparison, we selected only sources that could be detected at high signal to noise in the VLA-COSMOS 3 GHz Large Project catalog which has a flux density limit of $S/\sigma > 5 = 11.5 \mu\text{Jy}$. We also consider only unresolved sources in the VLA-COSMOS 3 GHz Large Project catalog to rule out resolution effects. This yielded a sample of 250 objects. For this subsample of sources, we determined the ratio of peak brightness between the COSMOS-XS survey and the VLA-COSMOS 3 GHz Large Project $f_S = S_{\text{COSMOS-XS}}/S_{\text{Smolcic+2017}}$. The result of this comparison is shown in Fig. 8. We measured a median ratio of 0.98 with a standard deviation of 0.18. The plot shows that the flux scale is in good agreement with the VLA-COSMOS 3 GHz Large Project one over the entire flux range probed. We see also no systematic offsets in this subset with distance from the phase centre.

3.5. Astrometry

To assess our astrometric accuracy, we have compared the positions of 30 sources at 3 GHz with S/N > 20 with the positions detected in the Very Long Baseline Array (VLBA)- COSMOS 1.4 GHz survey (Herrera Ruiz et al. 2017). The matching radius used to find the 3 GHz sources in the Herrera Ruiz et al. (2017) catalog was $0.4''$. The results, shown in Fig. 9, yield an excellent agreement with a mean offset of $-0.0001''$ in ΔRA and $0.02''$ in ΔDec. We find a standard deviation of $0.04''$ for ΔRA and $0.05''$ for ΔDec. We note that we did not correct the catalog entries for the offsets found.

3.6. Completeness

To quantify the completeness of the catalog, we performed another set of Monte Carlo simulations where we added simulated sources to the pre-primary-beam-corrected image. Injecting sources into the image allows us to account for the varying noise across the field. We use the same approach as in Section 3.2 but when the position of the source was determined, the input flux density of the source was reduced, dependent on position within the primary beam. By doing this we account for the decreasing sensitivity with increasing distance from the pointing centre due to the primary beam attenuation. Because the beam sensitivity is not uniform and decreasing, the effective area over which we are sensitive to a given flux density S decreases rapidly with the flux density itself. By inserting primary beam corrected flux densities, the derived completeness correction automatically includes the effect of variation of sensitivity as a function of distance from the map centre.

To allow for a better estimate of the completeness in terms of integrated flux densities, the mock sources also include a percentage of extended sources - Gaussians with FWHM larger than the beam size. The mock sources are therefore assigned, in addition to a peak flux density, a major axis, minor axis and an integrated flux
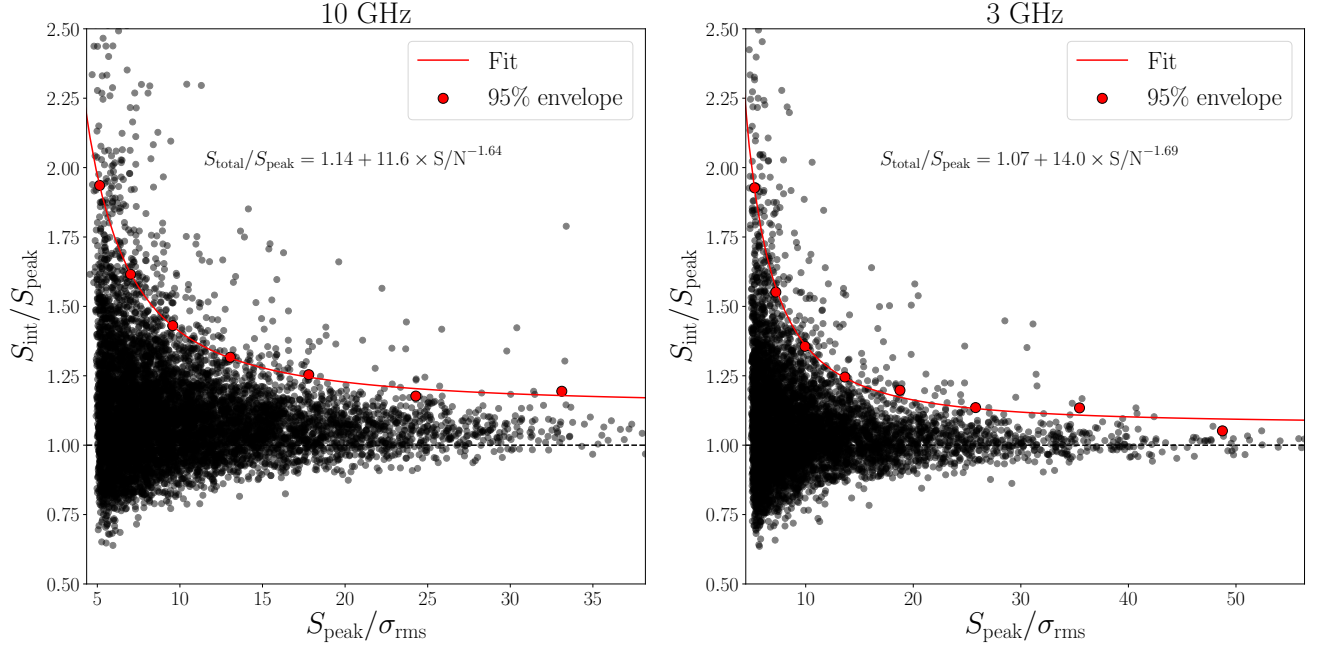
Figure 6. The simulated ratio of integrated flux density to peak brightness as a function of signal-to-noise ratio for unresolved sources from the 200 Monte Carlo simulations for the 10 GHz image (left) and the 3 GHz image (right). For logarithmic bins in signal-to-noise ratio, the red points show the threshold below which 95 per cent of the sources lie in that bin. The red line shows the fit to this upper envelope. Applying these thresholds to the real data, we find that, respectively, 13% and 32% of the sources are resolved.
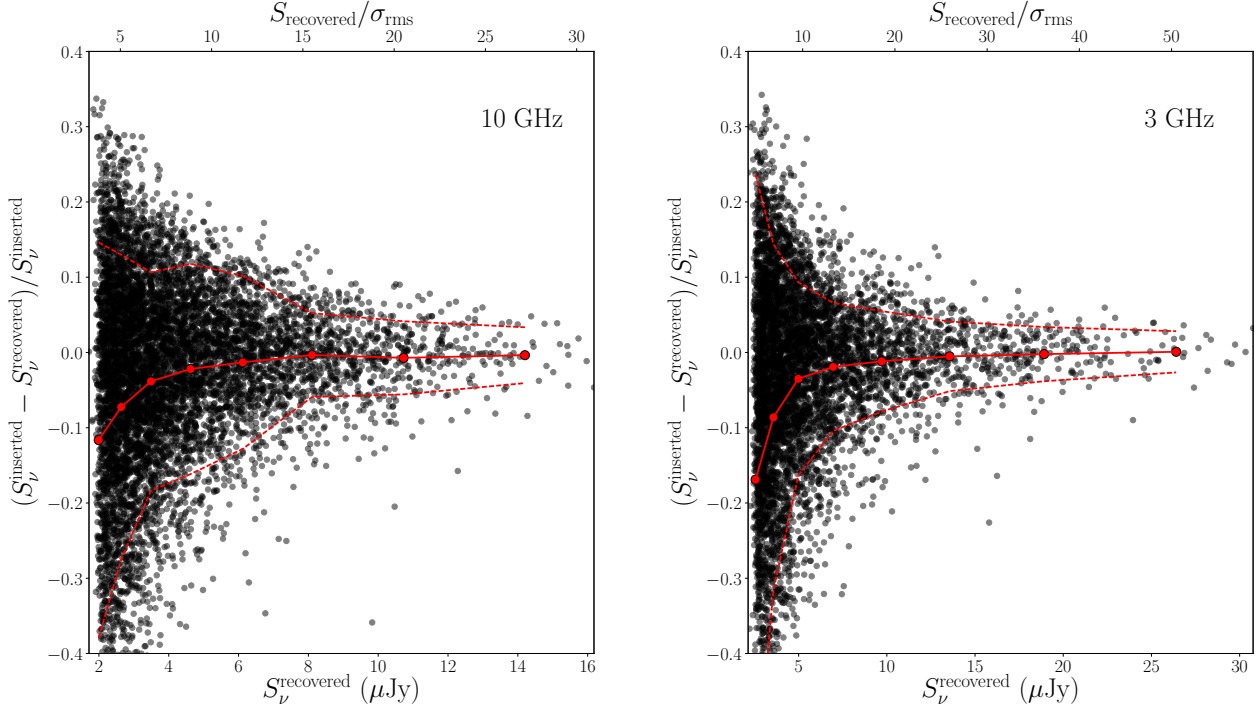
Figure 7. The simulated ratio of (inserted flux density - recovered flux density) to inserted flux density as a function of the recovered flux density. The left panel shows the distribution for the 10 GHz image and the right panel shows the distribution for the 3 GHz image. The solid red line denotes the median of 8 logarithmic bins (indicated by the red points) across the flux density range and the dashed lines mark the 1σ upper and lower bounds in those bins. The effect of flux boosting at the faint end is illustrated by the rapid downturn below about $3\,\mu$Jy for 10 GHz and 3 GHz.
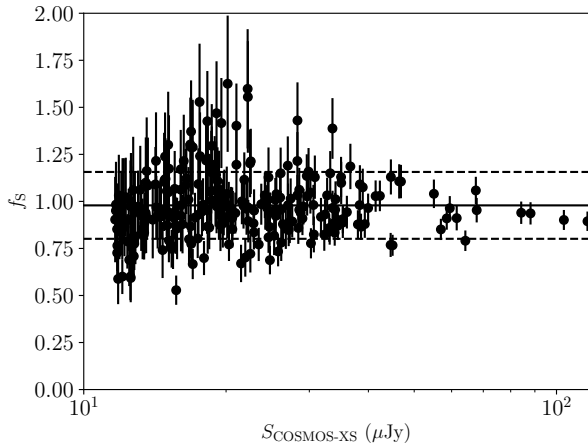
Figure 8. Flux comparison between our sample and the VLA-COSMOS 3 GHz Large Project (Smolčić et al. 2017a). The median ratio of 0.98 with a standard deviation of 0.18 shows that the flux densities are on the same flux scale.
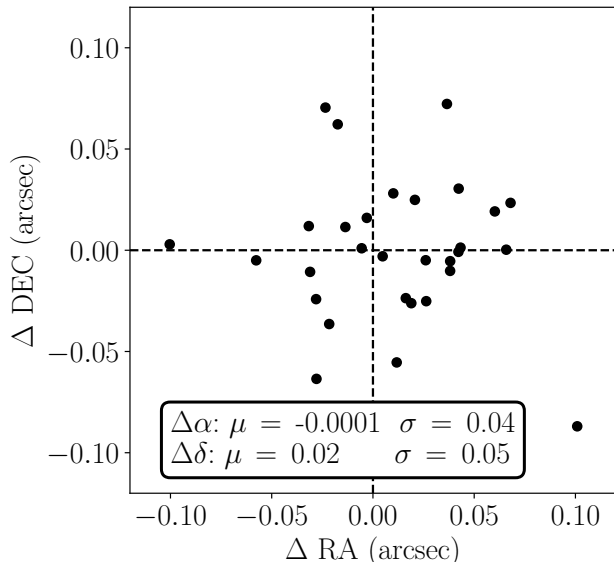

Figure 9. Astrometry comparison between 3 GHz and 1.4 GHz VLBA data for 30 VLBA sources (Herrera Ruiz et al. 2017). Median and standard deviation for ΔRA and ΔDec are reported in the panel. The positions of sources detected at 3 GHz are in excellent agreement with the posistions as reported in the Very Long Baseline Array (VLBA)- COSMOS 1.4 GHz survey (Herrera Ruiz et al. 2017).

density. To generate an angular size distribution for the mock sources, we draw randomly from two skewed Gaussians, one for the major axis and one for the minor axis. These Gaussians are determined by fitting to the normalized distribution of the fitted Gaussian parameters as measured for the real sources by PyBDSF.

The total flux density was chosen to be either their integrated flux density if resolved, or their peak brightness if unresolved. To determine whether a simulated source was resolved or unresolved, we use the same $S_{\mathrm{int}}/S_{\mathrm{peak}}$ envelope as described in Section 3.2.

Recovered mock sources are found by matching the
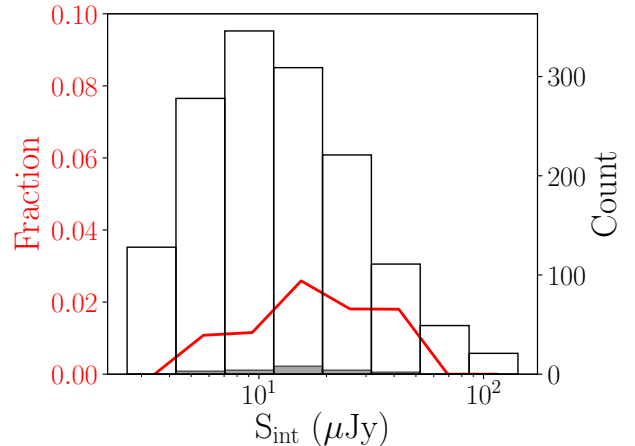
Figure 10. Fraction of false detections (red line) as a function of flux density. The open (filled) histogram shows the number of components cataloged in the observed 3 GHz map (detected in the inverted map). These data are also listed in Table 4. The false detection rate is always smaller than 3%, and is 1% overall.

retrieved catalog to the mock catalog. A retrieved source was considered to be matched with the inserted mock source if it was found within $0.5''$ of the inserted mock source position. Sources with a counterpart were flagged as recovered sources.

The completeness of a catalog represents the probability that all sources above a given flux density are detected. We have estimated this by giving the fraction of mock sources that are recovered using the same detection parameters. In Fig. 11 we plot the fraction of detected sources in our simulation as a function of integrated flux density, accounting for the primary beam. This detection fraction is largely driven by the variations in r.m.s. across the image and the primary beam. The error on the correction is the standard deviation of the calculated correction over all 200 realizations of the mock catalog. We thus estimate that the catalog is 85 per cent complete above a flux density of $10\,\mu$Jy for the 10 GHz image and above a flux density of $15\,\mu$Jy for the 3 GHz image.

3.7. *Reliability*

The reliability of a source catalog indicates the probability that all sources above a given flux density are real sources and not accidental detections of background features or noise. To assess the false detection rate of our source extraction, we ran PyBDSF on the inverted (i.e., multiplied by -1) continuum map, within 20% of the peak primary beam sensitivity, with the same settings used for the main catalog. Since there is no negative emission on the sky, every source detected in the inverted map is by definition a noise peak (i.e., a false detection). The false detection rate is determined from the number ratio of negative sources over positive sources per flux density bin. Errors are calculated based on Poissonian errors on the number of sources per flux density bin.

For the inverted 10 GHz image, no sources were detected above our catalog threshold of 5σ, and thus the false detection rate was determined to be zero. The false detection rate found for the 3 GHz image is shown in Fig. 10 with the red line. The number counts for both the real and falsely detected sources are also shown. In total there
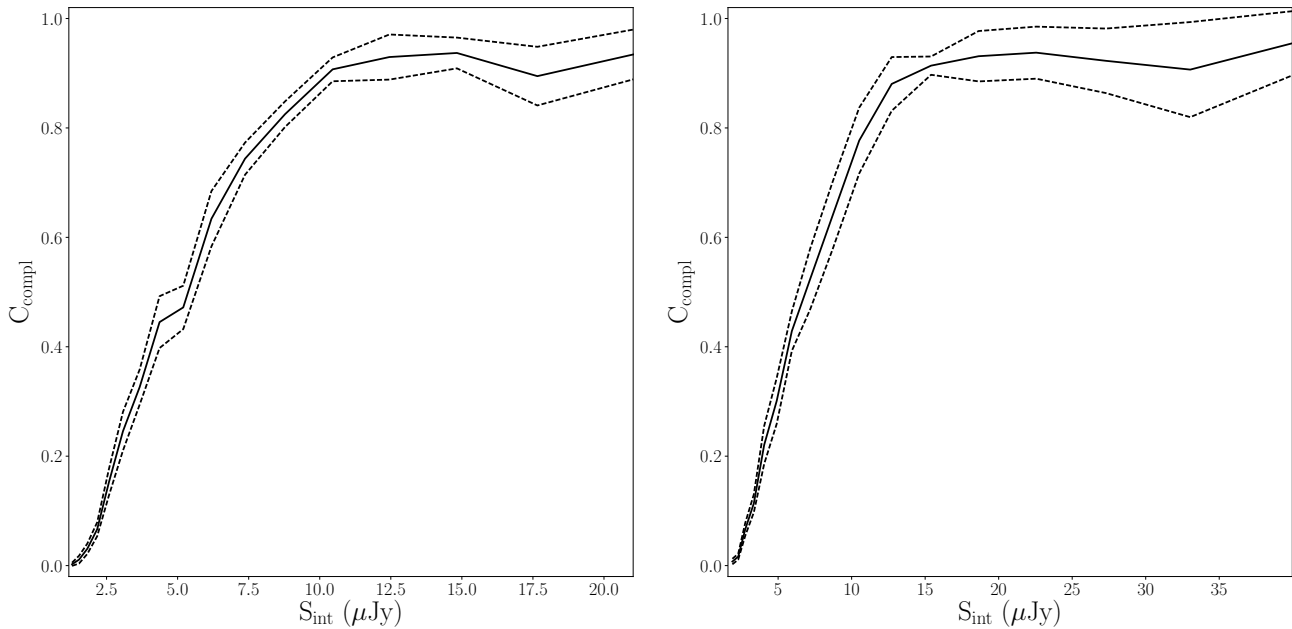
Figure 11. Completeness of the 10 GHz (left) and the 3 GHz (right) source catalog as a function of flux density. The solid black line shows the mean completeness of all Monte Carlo runs and the dotted black lines show the standard deviation. The completeness is above $\sim 85\%$ for $10\,\mu$Jy and for $15\,\mu$Jy for the 10 GHz and the 3 GHz source catalog, respectively. Note that the completeness does not reach 100% because effects of the primary beam are also included.

are 22 negative detections for the 3 GHz image, which results in a total false detection rate of $\sim 1\%$. These sources are located around bright sources and are thus likely caused by the artifacts surrounding these sources.

3.8. *Source catalog*

The final 10 GHz catalog consists of 91 sources and the final 3 GHz catalog consists of 1481 sources. Both catalogs are available as a table in FITS format as part of the online version of this article. Resolved sources are identified as described in Section 3.2. Errors given in the catalog are the nominal fit errors reported by PyBDSM. A sample of the 3 GHz catalog is shown in Table 3.

4. RESULTS

In this section, we report two results based on the produced catalogs: the 10 and 3 GHz faint source counts. Further analysis of these data will be presented in future publications.

4.1. *Radio source counts*

We use the 10 and 3 GHz catalogs to compute the source counts down to integrated flux densities $\sim 2\,\mu$Jy and $\sim 2.5\,\mu$Jy respectively.

The source counts are computed using the integrated flux densities (which means peak brightness for unresolved and integrated flux density for resolved), but sources are detected based on their measured peak brightness over the local noise level. The completeness of the source counts will thus depend both on the variation of the noise in the image and on the relation between integrated flux densities and peak brightness. In the following, we discuss these effects and how we correct for them in deriving the source counts.

4.1.1. *Multi-component sources*

In order to derive the source counts, we need to take into account that a single source may be made up of multiple components. For example, radio sources associated with radio galaxies can be made up of a nucleus with hot spots along, or at the end of, one or two jets. When jets are detected, it is relatively easy to recognize the components belonging to the same source. When a jet is missing, the radio-lobes are detected as two separated sources. We apply the statistical technique described by Magliocchetti et al. (1998) and White et al. (2012) to find these double component sources. We consider the separation of the nearest neighbour of each component and the summed flux density of the source and its neighbour. Multi-components are combined as single sources if the ratio of their flux densities is between 0.25 and 4 and their separation is less than a critical value dependent on their integrated flux density given by:

$$\theta_{\rm crit} = 100 \left(\frac{S_{\rm sum}}{10} \right)^{0.5} , \qquad (2)$$

where $S_{\rm sum}$ is in mJy and $\theta_{\rm crit}$ is in arcsec. This maximum separation is shown in Fig. 12. We analyzed both the 10 and 3 GHz image but only found multi-component sources for the 3 GHz image. The 3 GHz sources that meet both requirements are shown in Fig. 12 as red circles. From this analysis we could identify 17 multi-component sources. See also Fig. 17 for an example of a multi-component source.

4.1.2. *Completeness and reliability*

We consider a correction for the completeness of the catalog as derived in Section 3.6. Using the derived completeness corrections we can calculate the source counts

Table 3
Sample of the 3 GHz COSMOS-XS catalog.

Source ID	RA [deg]	$\sigma_{\rm RA}$ [arcsec]	Dec [deg]	$\sigma_{\rm Dec}$ [arcsec]	$S_{\rm int}$ [μJy]	$S_{\rm peak}$ [μJy beam^{-1}]	$\sigma_{\rm local}$ [μJy beam^{-1}]	Gaussians	Resolved
(1)	(2)	(3)	(4)	(5)	(6-7)	(8-9)	(10)	(11)	(12)
COSMOS-XS J100027.60+023634.21	150.11501	0.14907	2.6095	0.19592	4.04 ± 1.11	3.52 ± 0.67	0.65	S	U
COSMOS-XS J100014.35+023147.58	150.0598	0.03873	2.52988	0.04186	15.58 ± 1.03	13.74 ± 0.62	0.6	S	U
COSMOS-XS J100055.00+022849.80	150.22917	0.02603	2.4805	0.03019	52.54 ± 2.4	46.48 ± 1.44	1.41	S	R
COSMOS-XS J100031.26+023642.93	150.13026	0.03604	2.61192	0.04184	17.77 ± 1.16	16.04 ± 0.69	0.68	S	U
COSMOS-XS J100004.85+023559.51	150.02019	0.12507	2.59986	0.12439	7.01 ± 1.43	6.21 ± 0.85	0.84	S	U
COSMOS-XS J100056.66+022635.42	150.23607	0.02513	2.44317	0.03391	101.17 ± 3.56	76.85 ± 2.22	2.14	S	R
COSMOS-XS J100029.99+022903.64	150.12494	0.06062	2.48434	0.1729	5.85 ± 1.63	3.19 ± 0.66	0.66	M	U
COSMOS-XS J100046.61+023443.89	150.1942	0.2192	2.57886	0.10006	7.25 ± 1.37	5.54 ± 0.84	0.83	S	U
COSMOS-XS J100013.14+022550.66	150.05477	0.06633	2.43074	0.09338	22.27 ± 2.2	17.69 ± 1.34	1.32	S	R
COSMOS-XS J100004.59+023301.50	150.01914	0.05262	2.55042	0.06861	11.75 ± 1.31	11.05 ± 0.77	0.76	S	U

Notes. The format is the following: Column (1): Source name. Column (2) and (3): flux density-weighted right ascension (RA) and uncertainty. Column (4) and (5): flux density-weighted declination (Dec.) and uncertainty. Column (6) and (7): integrated source flux density and uncertainty, in μJy. Column (8) and (9): peak brightness and uncertainty, in μJy beam^{-1}. Column (10): the local r.m.s. noise. Column (11): number of Gaussian components. S refers to a single-Gaussian source. M refers to a a multi-Gaussian source. Column (12): a flag indicating the resolved parametrization of the source. 'U' refers to unresolved sources and 'R' to resolved sources.
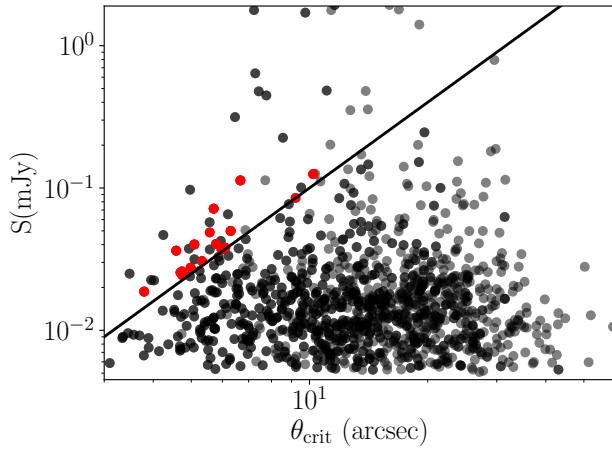
Figure 12. Sum of the flux densities of nearest neighbor pairs as a function of their separation. Source pairs that have separation less then the critical value given by the solid line and flux densities that differ by less than a factor of 4 are considered as double source candidates. 17 sources are identified as multi-component sources and are indicated with red points.

$(n(S))$ in each flux density bin using

$$n(S) = \frac{1}{A} \sum_{i=1}^{N} \frac{1}{C_{\rm compl}(S_i)}, \quad (3)$$

where $A \sim 350\,{\rm arcmin}^2$ is the solid angle considered for the source counts, N is the number of sources in the flux density bin and $C_{\rm compl}(S_i)$ is the derived completeness for the given flux density.

Additionally, we make a correction for the reliability for the source counts at 3 GHz by applying the false detection rate derived in Section 3.7. This correction acts in the opposite direction to the completeness correction and can be added to Equation 3:

$$n(S) = \frac{1}{A} \sum_{i=1}^{N} (1 - F_{\rm false-det}(S_i)) \frac{1}{C_{\rm compl}(S_i)}, \quad (4)$$

where $F_{\rm false-det}(S_i)$ is the correction for the false detection rate for the given flux density bin.

4.1.3. Resolution bias

Sources in our image are found by identifying peaks of emission above a given threshold. This means that a resolved source of a given integrated flux density will be missed more easily than a point source of the same integrated flux density. This incompleteness is called the resolution bias and causes the number of sources to be underestimated, particularly near the detection limit of the survey. We correct for the resolution bias by adopting the analytic method used in Prandoni et al. (2001) and Williams et al. (2016).

The following relation can be used to calculate the maximum angular size that a source can have and still be detected for a given integrated flux density:

$$S_{\rm int}/5\sigma = \frac{\theta_{\rm min}\theta_{\rm maj}}{b_{\rm min}b_{\rm maj}}, \quad (5)$$

where $\theta_{\rm min}$ and $\theta_{\rm maj}$ are the fitted source axes as measured by PyBDSF, $b_{\rm min}$ and $b_{\rm min}$ are the synthesized beam axes and 5σ is the peak brightness detection limit (where σ is the local r.m.s. in the image). We use this relation to calculate the maximum size that a source can have and still be detected. This relation is shown in Fig. 13, where the range reflects the range of r.m.s. noise in the image. Fig. 13 also shows the distribution of θ, the
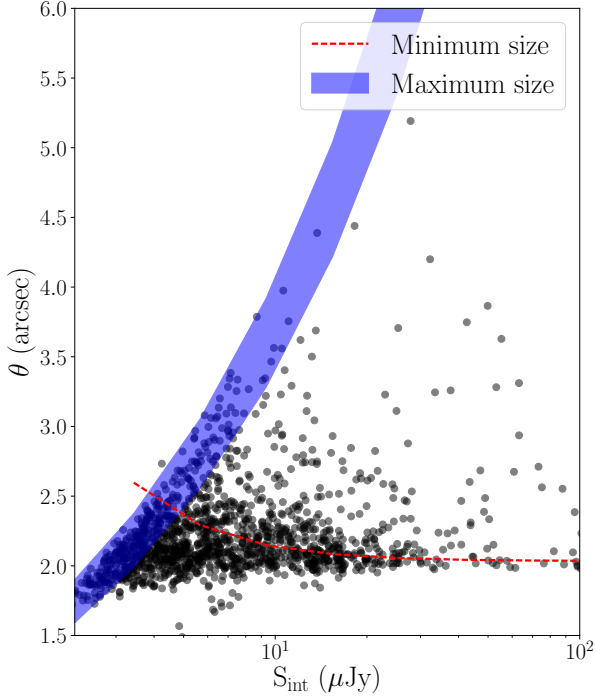
Figure 13. The fitted angular size, θ (geometric mean), as a function of integrated flux density at 3 GHz. Unresolved sources fall below the red dotted line giving the minimum size of a source. The minimum size was derived using the envelope from Section 3.2. The blue shaded region shows the maximum size a source of a given integrated flux density can have before dropping below the peak brightness detection threshold where the range reflects the range of r.m.s. noise in the 3 GHz image.

geometric mean of the source major and minor axes, as a function of integrated flux density for all sources in the 3 GHz catalog. Also shown is the minimum size a source can have before it is deemed to be unresolved. We use the maximum size to calculate the fraction of sources expected to be larger than this value, following Windhorst, Mathis & Neuschaefer (1990), using

$$h(> \theta_{\max}) = \exp\left[-\ln(2)\left(\frac{\theta_{\max}}{\theta_{\mathrm{med}}}\right)^{0.62}\right], \qquad (6)$$

where $\theta_{\max}$ is the maximum angular size and θ_{med} the median angular size. $\theta_{\max}$ can be calculated by rewriting Equation 5:

$$\theta_{\max} = \sqrt{b_{\min}b_{\mathrm{maj}} \times \frac{S_{\mathrm{int}}}{5\sigma}} \qquad (7)$$

We use two different versions of θ_{med} for comparison; the first, given by Windhorst et al. 1990,

$$\theta_{\mathrm{med}} = 2(S_{1.4\mathrm{GHz}})^{0.3\,\prime\prime} \qquad (8)$$

with $S_{1.4\mathrm{GHz}}$ in mJy (flux densities are scaled from 10 GHz and 3 GHz to 1.4 GHz using a spectral index of -0.7), and the second adopting a constant size of $0.35''$ below 1 mJy (based on recent results from Cotton et al. 2018 and Bondi et al. 2018) as follows:
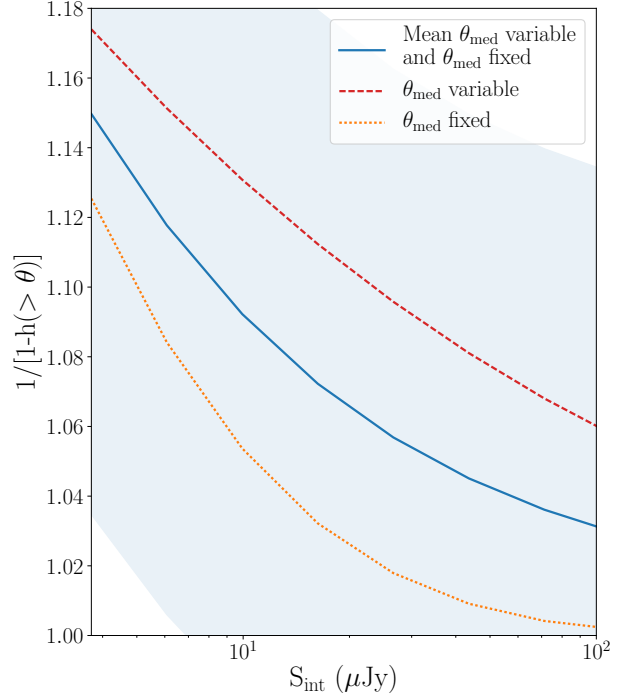
Figure 14. Resolution bias correction $1/[1 - h(> \theta_{\max})]$ for the fraction of sources with angular size larger than $\theta_{\max}$ at a given integrated flux density at 3 GHz. For the faintest sources, three curves are shown: the red dashed curve shows θ_{med} as a function of S_{int}, the orange dotted curve shows the result assuming $\theta_{\mathrm{med}} = 0.35''$ at these flux densities, and the blue solid curve shows the mean of the two curves. The range reflects the assumed 10 per cent uncertainty following Windhorst et al. (1990). The resolution bias correction is found to be 1.15 for the lowest flux bin.

$$\theta_{\mathrm{med}} = \begin{cases} 0.35\,'', & \text{for } S_{1.4\mathrm{GHz}} < 1 \text{ mJy,} \\ 2(S_{1.4\mathrm{GHz}})^{0.3}\,'', & \text{otherwise.} \end{cases} \qquad (9)$$

We can calculate the resolution-bias correction factor to be applied to the source counts using:

$$c = 1/[1 - h(> \theta_{\max})]. \qquad (10)$$

The correction factors calculated using the two different size distributions are plotted as a function of the integrated flux density at 3 GHz S_{int} in Fig. 14. We use the mean of the two correction factors to correct the source counts in this paper, and use an identical procedure to obtain the correction at 10 GHz. We use the uncertainty in the forms of θ_{med} and in $\theta_{\max}$ to estimate the uncertainty in the resolution bias correction. We further include an overall 10 per cent uncertainty following Windhorst et al. (1990) which dominates the error budget. We then determine the source counts, including a correction factor $c(S_i)$ for the resolution bias in each bin, via

$$n(S) = \frac{1}{A}\sum_{i=1}^{N}(1 - F_{\mathrm{false-det}}(S_{\mathrm{i}}))c(S_{\mathrm{i}})\frac{1}{C_{\mathrm{compl}}(S_{\mathrm{i}})} . \qquad (11)$$

ΔS_ν [μJy]	S_ν [μJy]	Counts [Jy$^{1.5}$ sr^{-1}]	Error [Jy$^{1.5}$ sr^{-1}]	N	$F_{\text{false}-\text{det}}$	C_{compl}	c
2.82 − 4.61	3.72	0.56	0.12	161	0.0	0.17	1.14
4.61 − 7.55	6.08	0.84	0.12	287	0.02	0.4	1.11
7.55 − 12.35	9.95	1.13	0.14	349	0.01	0.74	1.09
12.35 − 20.2	16.27	1.55	0.19	292	0.02	0.91	1.07
20.2 − 33.04	26.62	1.97	0.25	184	0.02	0.93	1.05
33.04 − 54.04	43.54	2.23	0.36	101	0.02	0.94	1.04
54.04 − 88.41	71.22	2.25	0.44	48	0.0	0.94	1.03
88.41 − 144.62	116.51	2.33	0.59	24	0.0	0.94	1.03

Notes. The format is the following: Column (1): flux interval. Column (2): bin centre. Column (3): differential counts normalized to a non evolving Euclidean mode. Column(4): error on differential counts. Column (5): number of sources detected. Column (6): false detection rate. Column (7): completeness correction. Column (8): resolution bias correction. The listed differential counts were corrected for completeness and resolution bias (C_{compl} and c), as well as false detection fractions ($F_{\text{false}-\text{det}}$), by multiplying the raw counts by the correction factor which is equal to $(c(S_i)*(1-F_{\text{false}-\text{det}}(S_i)))/C_{\text{compl}}(S_i)$. The source count errors take into account the Poissonian errors and completeness and bias correction uncertainties (see text for details).

4.1.4. *The sub-μJy source counts at* 3 GHz *compared to observations*

The Euclidean normalized counts are shown in Fig. 15(a) and tabulated in Table 4. Uncertainties on the final normalized source counts are propagated from the errors on the correction factors and the Poisson errors on the raw counts per bin. For the uncorrected data points, these errors are only the Poisson errors. Fig. 15(a) illustrates our results compared to other observational results, including the 3 GHz surveys by Condon et al. (2012), Vernstrom et al. (2014), Smolčić et al. (2017a).

Condon et al. (2012) used the $P(D)$ (probability distribution of peak flux densities) analysis technique to statistically estimate the radio number counts down to $\sim 1\,\mu$Jy from VLA 3 GHz observations. $P(D)$ analyses tend to be less prone to resolution effects, as these study the statistical properties of sources well below the confusion limit of the survey. This approach results in statistical estimates of the source counts that are much fainter than the faintest sources that can be counted individually. Vernstrom et al. (2014) also used the $P(D)$ analysis technique to re-analyse the 3 GHz observations from Condon et al. (2012) and estimate the radio number counts down to $\sim 50\,$nJy. They used a different approach from Condon et al. (2012) that allowed for more flexibility in accurately modelling the true source counts. Our completeness corrected points are slightly higher than the Condon et al. (2012) results and are more consistent with the Vernstrom et al. (2014) results.

Smolčić et al. (2017a) derived source counts from 3 GHz observations of the two square degree COSMOS field from the VLA-COSMOS 3 GHz Large Project. For comparison we show both the uncorrected number counts and those corrected for completeness, resolution bias and false detections. Our completeness corrected points consistently lie a factor ~ 1.4 above the source counts from Smolčić et al. (2017a).

One possible cause for the offset between our source

counts and those from Smolčić et al. (2017a) is an underestimated correction for the resolution bias. In particular, our survey has a lower resolution than the Smolčić et al. (2017a) observations (2″ versus 0.75″) and is therefore less likely to miss sources due to the resolution bias (see also the small correction derived in Section 4.1.3). Smolčić et al. (2017a) performed extensive Monte Carlo simulations to account for the resolution bias. They modeled the intrinsic angular sizes of mock sources using a simple power-law parametrization of the angular size distribution as a function of their total flux density, as derived by Bondi et al. (2008). They further applied a minimum angular size for faint mock sources to ensure the modelled angular size distribution matched the observed size distribution. We investigated whether the difference in resolution bias correction methods between Smolčić et al. (2017a) and the one used in this work (described in Section 4.1.3) is significant, finding that when applied to the same data, our method results in a higher normalization by a factor of ~ 1.1 (i.e., $\sim 25\%$ of the observed difference). Therefore, the difference in resolution bias correction method may partly explain the offset between our number counts and those reported in Smolčić et al. (2017a).

To attempt to gain some additional insight, we also compare our results with those from recent 1.4 GHz (or similar) observations in Fig. 15(a). We scale those flux densities to the 3 GHz observed frame assuming a spectral index of -0.84.[12] Bondi et al. (2008) derive their counts in the inner 1 deg^2 region of the COSMOS field from the VLA-COSMOS 1.4 GHz Large Project (Schinnerer et al. 2007) catalog, which required a somewhat uncertain correction for the effect of bandwidth smearing, while Prandoni et al. (2018) derived their counts from 1.4 GHz mosaic observations obtained with the Westerbork Synthesis Radio Telescope (WSRT). Because the latter cover an area of 6.6 square degrees, the source catalogue contains ~ 6000 sources (note that the errorbars in Fig. 15(a) are smaller than the symbols). Finally, Mauch et al. (2020) used the $P(D)$ analysis technique to derive source counts from $0.25\,\mu$Jy to $10\,\mu$Jy using confusion-limited 1.28 GHz GHz MeerKAT observations. They further derive the source counts between $10\,\mu$Jy and $2.5\,$mJy from individual detected sources. The direct source counts (scaled to 3 GHz) are shown in Fig. 15(a).

With the assumed spectral index of $\alpha = -0.84$, we find that the source counts of Bondi et al. (2008), Prandoni et al. (2018), and Mauch et al. (2020) generally fall between the Smolčić et al. (2017a) counts and those derived here, and are on average lower than our source counts (particularly the Mauch et al. (2020) counts, which also extend to the lowest flux densities). As already noted by Smolčić et al. (2017a), the comparison between the 1.4 and 3 GHz source counts is complicated by the potentially overly simplistic scaling of the 3 GHz counts to 1.4 GHz using a single spectral index value (in addition to the varying resolution bias and bandwidth smearing

---

[12] Smolčić et al. (2017a) found this spectral index by performing a Gaussian fit to the spectral index distribution of 3 GHz sources which were also detected at 1.4 GHz. This spectral index is different from the value derived using a survival analysis which takes non-detections into account. The difference can be explained by selection effects as discussed in Condon (1984).
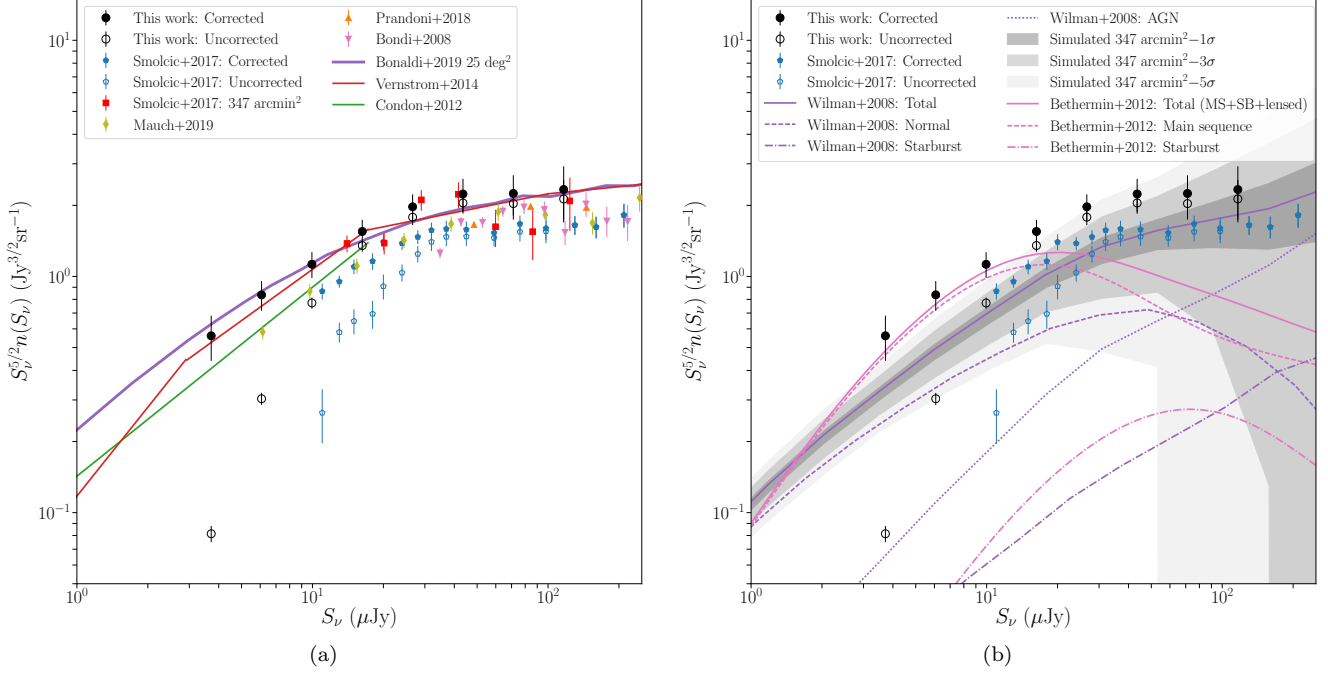
Figure 15. Euclidean-normalized radio source counts at 3 GHz (filled black circles). Open black circles show the source counts without corrections. Error bars correspond to the Poisson error for the uncorrected points. For the corrected points the errors are propagated from the errors on the correction factors and the Poisson errors on the raw counts per bin. Corrected (closed) and uncorrected (open) VLA-COSMOS 3 GHz measurements from Smolčić et al. (2017a) are shown with blue points. Panel (a) also shows the number counts from $P(D)$ analysis by Condon et al. (2012) (green line) and Vernstrom et al. (2014) (red line). The purple line shows the results from the simulation of Bonaldi et al. (2019). The red filled squares show the source counts for sources in the Smolčić et al. (2017a) catalog within the area of our pointing. Panel (a) also shows the results from 1.4 GHz observations from Bondi et al. (2008) (filled pink downward triangles) and Prandoni et al. (2018) (filled orange upward triangles) and 1.28 GHz observations from Mauch et al. (2020) (filled olive diamonds). The flux densities are shifted to the 3 GHz observed frame using a spectral index of -0.84. Panel (b) shows the comparison to the number counts from the simulation of Wilman et al. (2008) (purple lines) and Béthermin et al. (2012) (pink lines). Different line styles correspond to different source types as defined in the legend. The solid lines show the total source counts. The shaded region demonstrates the effect of cosmic variance and correspond to 1, 3 and 5 standard deviations in the source counts measurements of the simulations by Wilman et al. (2008).

effects present). We therefore investigate one final effect that may influence the measured source counts in the following section.

4.1.5. *Cosmic variance*

A final source of uncertainty that may affect the measured source counts is cosmic variance. In particular, if we observe overdensities in our single pointing, the resulting number counts will be higher than the number counts averaged over a larger area (as in Smolčić et al. 2017a).

Heywood et al. (2013) developed a method to assess the influence of source clustering on radio source counts, by extracting a series of independent samples from the models of Wilman et al. (2008). We used a similar approach with both the Wilman et al. (2008) simulation and the Bonaldi et al. (2019) simulation, which are further discussed in the next section, to estimate the uncertainty induced by sample variance on a survey with properties similar to ours. Following Heywood et al. (2013), we extract multiple non-overlapping sky patches with areas of $0.31 \times 0.31 \deg^2$ from the Wilman et al. (2008) simulation (comparable to the effective area of our observations). We used a simulated area of $4 \times 4 \deg^2$. This process resulted in 169 source catalogs. For each of these simulated source subsets, we compute the Euclidean normalized differential source counts. Fig. 15(b) shows the

mean value of the simulated counts from the independent distributions in each bin with the solid purple line. The shaded regions surrounding this correspond to one, three and five times the standard deviation of the count measurements. Fig. 15(b) shows that count fluctuations found in our observed survey area are significant enough to dominate the observed scatter at flux densities above $100 \, \mu$Jy, and contribute significantly below this.

A similar approach was used to produce the shaded regions in Fig. 16(a). Here, we extract multiple non-overlapping sky patches with areas of $0.31 \times 0.31 \deg^2$ from the simulation of Bonaldi et al. (2019). We used a simulated area of 25 deg², resulting in 289 source catalogs. For each of these simulated source subsets, we compute the Euclidean normalized differential source counts. The shaded regions in Fig. 15(b) and Fig. 16(a) show that the offset between our number counts and the Smolčić et al. (2017a) counts could be at least partly explained by cosmic variance.

We are able to further test the magnitude of cosmic variance by using the fact that the Smolčić et al. (2017a) counts are derived from an area that also covers our pointing. By counting the sources in the Smolčić et al. (2017a) catalog within the area of our pointing and assuming the completeness corrections described in Smolčić et al. (2017a), we can compare their source density directly with our number counts. This comparison is shown

in Fig. 15(a). The number counts in the Smolčić et al. (2017a) catalog over our pointing are noisy due to the small number statistics in this (shallower) sample, but indicate a slightly higher source density (on average) within the area of our survey. We therefore conclude that the systematic offset observed between our 3 GHz number counts and the 3 GHz direct detections from Smolčić et al. (2017a) can most likely be explained by a combination of cosmic variance in our pointing and an underestimated resolution bias correction in Smolčić et al. (2017a).

In summary, we find that our 3 GHz source counts agree with those at both 3 and 1.4 GHz within the various uncertainties, while extending down to typically lower flux densities than previous direct detections. In the following section, we will therefore explore the implications of our derived source counts on the modeling of the ultra-faint radio population.

4.1.6. *The sub-µJy source counts at* 3 GHz *compared to simulations*

In Fig. 15(b) we compare our number counts with semi-analytical models from Wilman et al. (2008) and Béthermin et al. (2012). Wilman et al. (2008) developed a semi-empirical simulation that used observed and extrapolated luminosity functions to populate an evolving dark matter skeleton with various galaxy types (normal, starburst and AGN). The adopted distribution of sources on the underlying dark matter density was chosen to reflect their measured large-scale clustering. The Béthermin et al. (2012) model uses two main ingredients to predict the number counts: the evolution of main-sequence galaxies and starburst galaxies based on the 2-Star-Formation-Mode framework of Sargent et al. (2012) and main-sequence and starburst SED's to predict the shape of the IR luminosity function at $z \leq 2$. Béthermin et al. (2012) also included dust attenuation and strong lensing in their model. To get to 1.4 GHz radio source counts, they assumed a non-evolving IR-radio correlation and a synchrotron spectral slope of $\alpha = -0.8$.

Our derived source counts deviate from those predicted by the Wilman et al. (2008) model. In particular, they tend to be higher than the predicted model from $\sim 40\,\mu$Jy downward (and we note that the Smolčić et al. (2017a) counts are also systematically above this model at low flux densities despite a possibly underestimated resolution bias correction – see Section 4.1.4). On the other hand, our 3 GHz source counts are in good agreement with the Béthermin et al. (2012) model below $\sim 10\,\mu$Jy. The AGN population is not included in the Béthermin et al. (2012) model and this explains the decrease in the source counts above $20\,\mu$Jy as AGN start contributing significantly at these flux densities. Below $20\,\mu$Jy the star-forming population dominates the number counts, as can be seen from Fig. 16(a) and Fig. 15(b). Béthermin et al. (2012) model this population using a framework that uses more recent results to predict the LF of SFGs at $z \leq 2$ compared to the Wilman et al. (2008) model. Béthermin et al. (2012) use a combination of mid-IR and UV data (Daddi et al. 2007; Elbaz et al. 2007; Noeske et al. 2007), far-IR data (Elbaz et al. 2011; Pannella et al. 2015), and radio continuum imaging (Karim et al. 2011). Wilman et al. (2008) assume pure luminosity evolution out to $z = 1.5$ of the local LF

derived from the IRAS 2-Jy sample by Yun et al. (2001). Our results suggest that the Wilman et al. (2008) model could be improved by using the most recent observations to derive a model for the LF for SFGs.

Bonaldi et al. (2019) developed an improved simulation in the spirit of that of Wilman et al. (2008): the Tiered Radio Extragalactic Continuum Simulation (T-RECS). Bonaldi et al. (2019) model the radio sky in terms of two main populations, AGNs and SFGs, and use these to populate an evolving dark matter skeleton. To describe the cosmological evolution of the LF of AGNs, they adopted an updated model of Massardi et al. (2010), with the revision of Bonato et al. (2017). The LF of SFGs is derived from the evolving SFR function as the radio continuum emission is correlated with the SFR. The SFR function gives the number density of galaxies per logarithmic bin of SFR at a given redshift z. The evolution of the synchrotron-SFR relation accounts for the evolving radio-FIR correlation. Bonaldi et al. (2019) use the SFR function with redshift derived by Cai et al. (2013) and Cai et al. (2014) with an extension derived by Mancuso et al. (2015). The resulting SFR function extends to $z \sim 10$ and includes effects from strong gravitational lensing.

We compare our 3 GHz source counts to the simulation from Bonaldi et al. (2019), using their medium tier catalog of 25 deg^2 containing 3 GHz flux densities. Our number counts, shown in Fig. 16(a), match the total number counts from the simulations. Fig. 16(a) also shows the comparison between our source counts and those from the model by Mancuso et al. (2017). This model includes three populations: radio loud (RL) AGN, radio quiet (RQ) AGN and SFG. Using the same prescriptions as Bonaldi et al. (2019), they convert SFRs into 10 and 3 GHz luminosities. The SFR function used is, however, slightly different from the function used in Bonaldi et al. (2019) (see Section 4.1.7). The RL AGNs are modeled the same way as in Bonaldi et al. (2019). To model RQ AGNs, the SFR function is mapped into the AGN luminosity function. The obtained AGN bolometric luminosity is subsequently converted to X-ray luminosity. The AGN radio power is then derived using the observed relation between rest-frame X-ray and 1.4 GHz radio luminosity for RQ AGNs.

The Mancuso et al. (2017) model uses an intrinsic SFR function and therefore claims to be less sensitive to dust extinction effects. The derived intrinsic SFR function implies a heavily dust-obscured galaxy population at high redshifts ($z > 4$) and at large SFRs (10^2 M$_\odot$ yr^{-1}) (Mancuso et al. 2016). Our number counts, shown in Fig. 16(a), fall well above the total number counts predicted by Mancuso et al. (2017) at the faint end. In addition, there is a clear difference between the Bonaldi et al. (2019) simulations and the Mancuso et al. (2017) model, which is surprising as they use similar assumptions. The difference between the two simulations is further discussed in Section 4.1.7.

To summarize, Wilman et al. (2008) model SFGs assuming pure luminosity evolution out to $z = 1.5$ for the local LF, while thereafter no further evolution is assumed. This model reproduces the sources in the local universe as it matches the observed number density, with these sources dominating the flux densities at > 2 mJy. However, for flux densities $< 40\,\mu$Jy we find that the

ΔS_ν [μJy]	S_ν [μJy]	Counts [Jy$^{1.5}$ sr^{-1}]	Error [Jy$^{1.5}$ sr^{-1}]	N	$F_{\rm false-det}$	$C_{\rm compl}$	c
$2.71 - 8.36$	5.53	0.53	0.1	52	0.0	0.53	1.1
$8.36 - 25.8$	17.08	0.97	0.22	31	0.0	0.9	1.06
$25.8 - 79.66$	52.73	1.2	0.65	7	0.0	0.88	1.04

Notes. The format is the following: Column (1): flux interval. Column (2): bin centre. Column (3): differential counts normalized to a non evolving Euclidean mode. Column(4): error on differential counts. Column (5): number of sources detected. Column (6): false detection rate. Column (7): completeness correction. Column (8): resolution bias correction. The listed differential counts were corrected for completeness and resolution bias ($C_{\rm compl}$ and c), as well as false detection fractions ($F_{\rm false-det}$), by multiplying the raw counts by the correction factor which is equal to $(c(S_{\rm i}) * (1 - F_{\rm false-det}(S_{\rm i})))/C_{\rm compl}(S_{\rm i})$. The source count errors take into account the Poissonian errors and completeness and bias correction uncertainties (see text for details).

Béthermin et al. (2012) model and the Bonaldi et al. (2019) simulation are in better agreement with our observations. These models are able to produce the observed excess with respect to the source counts predicted by the Wilman et al. (2008) simulation. Our number counts support the steeper LF evolution for SFGs that is used in these models.

4.1.7. *The sub-μJy source counts at 10 GHz*

The Euclidean normalized counts at 10 GHz are shown in Fig. 16(b). Uncertainties on the final normalized source counts are propagated from the errors on the correction factors and the Poisson errors on the raw counts per bin. The source counts are tabulated in Table 5.

With an observing frequency of 10 GHz, we measure flux densities closer to the rest-frame frequencies $\nu \geq 30$ GHz where the total radio emission is dominated by free-free radiation (e.g Condon 1992; Murphy et al. 2011; Klein et al. 2018). The calibration of the free-free radiation and SFR in the Bonaldi et al. (2019) simulations follows Mancuso et al. (2015) and Murphy et al. (2011). The Mancuso et al. (2017) model used the same calibration for SFGs.

We can compare the 10 GHz source counts to the simulation from Bonaldi et al. (2019) by interpolating between the flux densities of their simulated sources given at 9.2 GHz and 12.5 GHz. Our number counts fall slightly above the total number counts from the simulations. The discrepancy is however within the σ derived for the cosmic variance, indicated through the gray shaded regions in Fig. 16(b). The shaded regions are derived in the same way as described in Section 4.1.5 by extracting sky patches of $0.09 \times 0.09\,{\rm deg}^2$ from the Bonaldi et al. (2019) simulations. This process results in 3025 source catalogs.

Our number counts are systematically higher than those predicted by Mancuso et al. (2017) at 10 GHz, especially at the faint end. This would indicate that Mancuso et al. (2017) underestimates the flux density produced at 10 GHz, especially by SFGs, as these contribute most at low flux densities as can be seen in Fig. 16(b).

As seen in Fig. 16(a) and Fig. 16(b), the Bonaldi et al. (2019) and Mancuso et al. (2017) simulations predict

roughly the same general shape of the number counts, but they predict a different normalization. The Mancuso et al. (2017) model includes the specific modeling of RQ AGN. In the Bonaldi et al. (2019) simulations, RQ AGN are not specifically modeled, and Bonaldi et al. (2019) mention that RQ AGNs would contribute part of the flux density of those sources that, in T-RECS, are modelled as SFGs. This can therefore not explain the difference in normalization between the total source counts.

Instead, the difference in normalization may be partially explained by how the two models treat the evolution of the radio-FIR correlation. Bonato et al. (2017) already showed that the radio source counts support an evolving radio-FIR correlation. In their study, the source counts found without an evolving radio-FIR correlation are found to be substantially below the observational determinations. The Mancuso et al. (2017) model does not include the increase of the ratio between sychrotron luminosity and far-IR luminosity as was recently reported by Delhaize et al. (2017) and Magnelli et al. (2015) and falls thus below the observed source counts. Bonaldi et al. (2019) evolve the synchrotron-SFR relation to account for the evolving radio-FIR correlation, yielding a very good fit to the observational estimates of the radio luminosity function of SFGs (Novak et al. 2017).

In addition to their different treatment of the radio-FIR correlation, both the Mancuso et al. (2017) and Bonaldi et al. (2019) studies also use a slightly different SFR function to make number counts predictions. In particular, both studies use a smooth, analytic representation of the SFR function derived from IR and dust-extinction-corrected UV, Lyα and Hα data. However, Mancuso et al. (2017) use a standard Schechter shape characterized by three evolving parameters: the normalization, the characteristic SFR and the faint-end slope. Meanwhile, Bonaldi et al. (2019) use the modified Schechter function introduced by Aversa et al. (2015), with two evolving characteristic slopes. Our observed number counts are in better agreement with the SFR function used by Bonaldi et al. (2019), which was derived by Mancuso et al. (2015), although the function used in Mancuso et al. (2017) was derived using the most recent IR and UV data.

In summary, as seen in Sections 4.1.6 and 4.1.7, models typically assume either an evolving LF or an evolving SFR function to make number counts predictions at the μJy level. We find that our source counts are in agreement with the Bonaldi et al. (2019) model that uses an evolving radio-FIR relation and the modified Schechter function to describe the SFR function. We also find that the source counts predictions from the evolving LF for SFGs used by Béthermin et al. (2012) result in a good match with the observed number counts below 10 μJy.

The relation between the LF and the SFR functions will be investigated in a future publication. The luminosity-SFR relation is of crucial importance for measuring the cosmic star-formation history, which will also be studied in a future publication. In addition, we can utilize the multi-wavelength information available in the COSMOS region to determine the composition of the faint radio population responsible for the measured number counts. This will be fully discussed and presented in a companion paper (Algera et al. 2020).
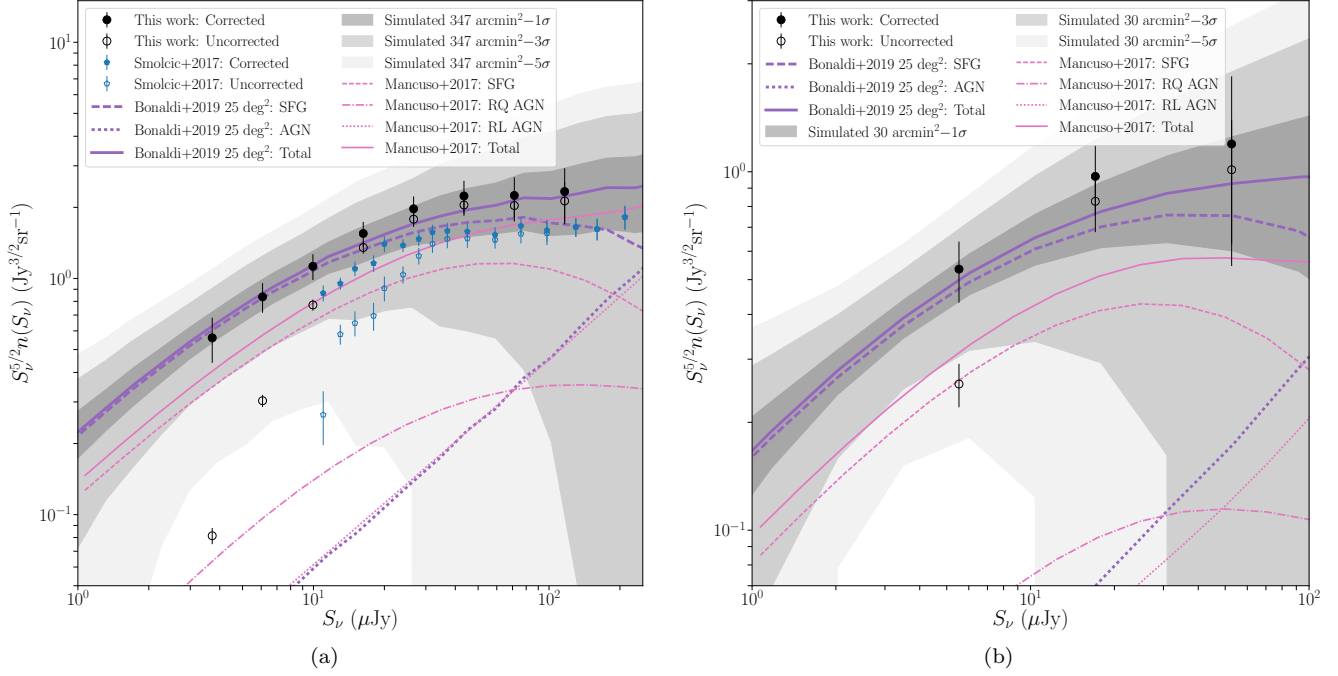
Figure 16. Euclidean-normalized radio source counts at $3\,\mathrm{GHz}$ (Panel (a)) and $10\,\mathrm{GHz}$ (Panel (b)). The filled black circles show the source counts from COSMOS-XS. Open black circles show the source counts without corrections. Error bars correspond to the Poisson error for the uncorrected points. For the corrected points the errors are propagated from the errors on the correction factors and the Poisson errors on the raw counts per bin. Corrected (closed) and uncorrected (open) VLA-COSMOS $3\,\mathrm{GHz}$ measurements from Smolčić et al. (2017a) are shown with blue points. The number counts from the simulation of Bonaldi et al. (2019) and predictions from Mancuso et al. (2017) are shown in purple and pink lines, respectively. Different line styles corresponding to different source types as defined in the legend. The solid lines show the total source counts from the simulation. The shaded regions demonstrate the effect of cosmic variance and correspond to 1, 3 and 5 standard deviations in the source counts measurements of the simulations by Bonaldi et al. (2019).

5. SUMMARY & CONCLUSIONS

In this paper, we presented the details of the COSMOS-XS survey: an ultra–deep pointing in the COSMOS field at 10 and 3 GHz with the Karl G. Jansky VLA. The final 10 and 3 GHz images have a resolution of $\sim 2''$ and reach a median r.m.s. of $0.41\,\mu\mathrm{Jy\,beam^{-1}}$ and $0.53\,\mu\mathrm{Jy\,beam^{-1}}$, respectively. The two radio catalogs contain sources detected within 20% of the peak primary beam sensitivity with peak brightness exceeding 5σ. The 10 GHz catalog contains 91 sources and the 3 GHz catalog contains 1498 sources.

Comparing the positions of our 3 GHz sources with those from the high-resolution VLBA imaging at 1.4 GHz, we estimated that the astrometric uncertainties are negligible. Completeness corrections are calculated using Monte Carlo simulations whereby the effect of the primary beam is taken into account. We find that the completeness is above $\sim 85\%$ at $10\,\mu\mathrm{Jy}/15\,\mu\mathrm{Jy}$ for the 10 GHz/3 GHz source catalogs, respectively. We correct for the resolution bias by utilising an analytic method as used in Prandoni et al. (2001) and Williams et al. (2016).

The deep 10 and 3 GHz observations enable us to investigate the Euclidean normalized source counts down to the μJy level. Our corrected radio counts at 3 GHz with direct detections down to $\sim 2.8\,\mu$Jy are consistent within the uncertainties with other results at 3 and 1.4 GHz, but extend to fainter flux densities than previous direct detections.

In comparison to simulations developed in the framework of the SKA Simulated Skies project (Wilman et al. 2008), our source counts show an excess which cannot be accounted for by cosmic variance. Our measured source counts suggest a steeper LF evolution for SFGs for low flux densities ($< 2\,\mathrm{mJy}$). The T-RECS simulations (Bonaldi et al. 2019) and the Béthermin et al. (2012) model both produce this steeper evolution and are in better agreement with our measured source counts.

Our corrected radio counts at 10 GHz with direct detections down to $\sim 2.7\,\mu$Jy are systematically higher than those predicted by the T-RECS simulations (Bonaldi et al. 2019), but we demonstrate that this falls within the expected variations from cosmic variance. The more significant offset observed between our counts and Mancuso et al. (2017) support an evolving radio-FIR relation and a modified Schechter function to describe the SFR function, as done in Bonaldi et al. (2019).

The COSMOS-XS survey is one of the deepest radio continuum surveys to-date, providing valuable information for next generation facilities, such as the ngVLA and SKA, which will achieve much deeper imaging. These radio data, in conjunction with the vast panchromatic COSMOS data sets, will allow us to study the dust-unbiased star-formation history, place critical constraints on dust attenuation and conduct a study of the long-wavelength spectra of faint radio sources.
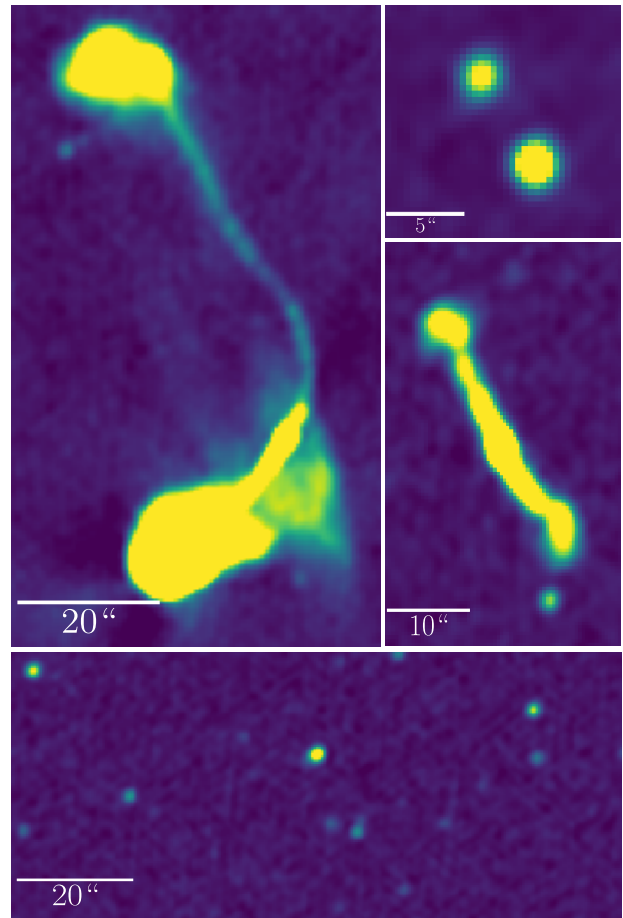
APPENDIX

A. POSTAGE STAMP IMAGES

Examples of extended, multi-component and compact sources are shown in Fig. 17.

Figure 17. Stamps from the 3 GHz continuum map showing examples of extended (left and lower right image), multi-component (upper right) and compact (lower panel) radio sources. The colorscale shows the flux density from $-3\,\sigma_{\rm local}$ to $50\sigma_{\rm local}$, where $\sigma_{\rm local}$ is the local r.m.s. noise.

REFERENCES

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, aap, 558, A33

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123

Aversa, R., Lapi, A., de Zotti, G., Shankar, F., & Danese, L. 2015, ApJ, 810, 74

Bell, E. F. 2003, ApJ, 586, 794

Best, P. N., Ker, L. M., Simpson, C., Rigby, E. E., & Sabater, J. 2014, MNRAS, 445, 955

Béthermin, M., Daddi, E., Magdis, G., et al. 2012, ApJ, 757, L23

Biggs, A. D., & Ivison, R. J. 2006, MNRAS, 371, 963

Bonaldi, A., Bonato, M., Galluzzi, V., et al. 2019, MNRAS, 482, 2

Bonato, M., Negrello, M., Mancuso, C., et al. 2017, MNRAS, 469, 1912

Bondi, M., Ciliegi, P., Schinnerer, E., et al. 2008, ApJ, 681, 1129

Bondi, M., Zamorani, G., Ciliegi, P., et al. 2018, A&A, 618, L8

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012, ApJ, 754, 83

Bouwens, R. J., Bradley, L., Zitrin, A., et al. 2014, ApJ, 795, 126

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34

Briggs, D. S. 1995, in American Astronomical Society Meeting Abstracts, Vol. 187, 112.02

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Cai, Z.-Y., Lapi, A., Bressan, A., et al. 2014, ApJ, 785, 65

Cai, Z.-Y., Lapi, A., Xia, J.-Q., et al. 2013, ApJ, 768, 21

Calistro Rivera, G., Williams, W. L., Hardcastle, M. J., et al. 2017, MNRAS, 469, 3468

Carilli, C. L., Lee, N., Capak, P., et al. 2008, ApJ, 689, 883

Chary, R.-R., & Pope, A. 2010, ArXiv e-prints, arXiv:1003.1731

Clemens, M. S., Vega, O., Bressan, A., et al. 2008, A&A, 477, 95

Condon, J. J. 1984, ApJ, 287, 461

Condon, J. J. 1992, ARA&A, 30, 575

Condon, J. J., Cotton, W. D., & Broderick, J. J. 2002, AJ, 124, 675

Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693

Condon, J. J., & Mitchell, K. J. 1984, AJ, 89, 610

Condon, J. J., Cotton, W. D., Fomalont, E. B., et al. 2012, ApJ, 758, 23

Cotton, W. D., Condon, J. J., Kellermann, K. I., et al. 2018, ApJ, 856, 67

Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156

de Jong, T., Klein, U., Wielebinski, R., & Wunderlich, E. 1985, A&A, 147, L6

Delhaize, J., Smolčić, V., Delvecchio, I., et al. 2017, A&A, 602, A4

Delvecchio, I., Gruppioni, C., Pozzi, F., et al. 2014, MNRAS, 439, 2736

Delvecchio, I., Smolčić, V., Zamorani, G., et al. 2017, A&A, 602, A3

Dumas, G., Schinnerer, E., Tabatabaei, F. S., et al. 2011, AJ, 141, 41

Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33

Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119

Helou, G., Soifer, B. T., & Rowan-Robinson, M. 1985, ApJ, 298, L7

Herrera Ruiz, N., Middelberg, E., Deller, A., et al. 2017, A&A, 607, A132

Heywood, I., Jarvis, M. J., & Condon, J. J. 2013, MNRAS, 432, 2625

Hickox, R. C., Jones, C., Forman, W. R., et al. 2009, ApJ, 696, 891

Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90

Huynh, M. T., Jackson, C. A., Norris, R. P., & Prandoni, I. 2005, AJ, 130, 1373

Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55

Jarvis, M., Seymour, N., Afonso, J., et al. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 68

Jarvis, M. J., & Rawlings, S. 2000, MNRAS, 319, 121

Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open source scientific tools for Python

Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, ApJ, 730, 61

Klein, U., & Graeve, R. 1986, A&A, 161, 155

Klein, U., Lisenfeld, U., & Verley, S. 2018, A&A, 611, A55

Kobulnicky, H. A., & Johnson, K. E. 1999, ApJ, 527, 154

Kurczynski, P., Gawiser, E., Acquaviva, V., et al. 2016, ApJ, 820, L1

Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415

Magdis, G. E., Elbaz, D., Hwang, H. S., et al. 2010, ApJ, 720, L185

Magliocchetti, M., Maddox, S. J., Lahav, O., & Wall, J. V. 1998, MNRAS, 300, 257

Magnelli, B., Ivison, R. J., Lutz, D., et al. 2015, A&A, 573, A45

Mancuso, C., Lapi, A., Shi, J., et al. 2016, ApJ, 823, 128

Mancuso, C., Lapi, A., Cai, Z.-Y., et al. 2015, ApJ, 810, 72

Mancuso, C., Lapi, A., Prandoni, I., et al. 2017, ApJ, 842, 95

Massardi, M., Bonaldi, A., Negrello, M., et al. 2010, MNRAS, 404, 532

Mauch, T., Cotton, W. D., Condon, J. J., et al. 2020, ApJ, 888, 61

McAlpine, K., Jarvis, M. J., & Bonfield, D. G. 2013, MNRAS, 436, 1084

Mendez, A. J., Coil, A. L., Aird, J., et al. 2013, ApJ, 770, 40

Mezger, P. G., & Henderson, A. P. 1967, ApJ, 147, 471

Mohan, N., & Rafferty, D. 2015, PyBDSF: Python Blob Detection and Source Finder, Astrophysics Source Code Library, ascl:1502.007

Morrison, G. E., Owen, F. N., Dickinson, M., Ivison, R. J., & Ibar, E. 2010, ApJS, 188, 178

Murphy, E. J., Momjian, E., Condon, J. J., et al. 2017, ApJ, 839, 35

Murphy, E. J., Condon, J. J., Schinnerer, E., et al. 2011, ApJ, 737, 67

Murphy, E. J., Bremseth, J., Mason, B. S., et al. 2012, ApJ, 761, 97

Murphy, E. J., Dong, D., Leroy, A. K., et al. 2015, ApJ, 813, 118

Nayyeri, H., Hemmati, S., Mobasher, B., et al. 2017, The Astrophysical Journal Supplement Series, 228, 7

Nikolic, B., & Bolton, R. C. 2012, MNRAS, 425, 1257

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ, 660, L43

Novak, M., Smolčić, V., Delhaize, J., et al. 2017, A&A, 602, A5

Offringa, A. R. 2010, AOFlagger: RFI Software, Astrophysics Source Code Library, ascl:1010.017

Offringa, A. R., & Smirnov, O. 2017, MNRAS, 471, 301

Offringa, A. R., McKinley, B., Hurley-Walker, N., et al. 2014, WSClean: Widefield interferometric imager, Astrophysics Source Code Library, ascl:1408.023

Owen, F. N. 2018, ApJS, 235, 34

Owen, F. N., & Morrison, G. E. 2008, AJ, 136, 1889

Padovani, P., Bonzini, M., Kellermann, K. I., et al. 2015, MNRAS, 452, 1263

Padovani, P., Mainieri, V., Tozzi, P., et al. 2009, ApJ, 694, 235

Pannella, M., Carilli, C. L., Daddi, E., et al. 2009, ApJ, 698, L116

Pannella, M., Elbaz, D., Daddi, E., et al. 2015, ApJ, 807, 141

Pavesi, R., Sharon, C. E., Riechers, D. A., et al. 2018, ApJ, 864, 49

Perley, R. A., & Butler, B. J. 2013, ApJS, 204, 19

Prandoni, I., Gregorini, L., Parma, P., et al. 2001, A&A, 365, 392

Prandoni, I., Guglielmino, G., Morganti, R., et al. 2018, MNRAS, 481, 4548

Rau, U., & Cornwell, T. J. 2011, A&A, 532, A71

Richards, E. A. 2000, ApJ, 533, 611

Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, ApJ, 872, 7

Rigby, E. E., Best, P. N., Brookes, M. H., et al. 2011, MNRAS, 416, 1900

Rowan-Robinson, M., Benn, C. R., Lawrence, A., McMahon, R. G., & Broadhurst, T. J. 1993, MNRAS, 263, 123

Rujopakarn, W., Dunlop, J. S., Rieke, G. H., et al. 2016, ApJ, 833, 12

Sabater, J., Best, P. N., Hardcastle, M. J., et al. 2019, A&A, 622, A17

Sadler, E. M., Jenkins, C. R., & Kotanyi, C. G. 1989, MNRAS, 240, 591

Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267

Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, ApJ, 799, 183

Sargent, M. T., Béthermin, M., Daddi, E., & Elbaz, D. 2012, ApJ, 747, L31

Schinnerer, E., Smolčić, V., Carilli, C. L., et al. 2007, ApJS, 172, 46

Schinnerer, E., Sargent, M. T., Bondi, M., et al. 2010, ApJS, 188, 384

Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74

Seymour, N., McHardy, I. M., & Gunn, K. F. 2004, MNRAS, 352, 131

Seymour, N., Dwelly, T., Moss, D., et al. 2008, MNRAS, 386, 1695

Siana, B., Teplitz, H. I., Chary, R.-R., Colbert, J., & Frayer, D. T. 2008, ApJ, 689, 59

Siana, B., Smail, I., Swinbank, A. M., et al. 2009, ApJ, 698, 1273

Smolčić, V., Zamorani, G., Schinnerer, E., et al. 2009a, ApJ, 696, 24

Smolčić, V., Schinnerer, E., Zamorani, G., et al. 2009b, ApJ, 690, 610

Smolčić, V., Novak, M., Bondi, M., et al. 2017a, A&A, 602, A1

Smolčić, V., Delvecchio, I., Zamorani, G., et al. 2017b, A&A, 602, A2

Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15

Tabatabaei, F. S., Schinnerer, E., Krause, M., et al. 2017, ApJ, 836, 185

Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29

Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2016, ApJ, 817, 118

Turner, J. L., & Ho, P. T. P. 1983, ApJ, 268, L79

Turner, J. L., & Ho, P. T. P. 1985, ApJ, 299, L77

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engineering, 13, 22

Vernstrom, T., Scott, D., Wall, J. V., et al. 2016, MNRAS, 462, 2934

Vernstrom, T., Scott, D., Wall, J. V., et al. 2014, MNRAS, 440, 2791

White, G. J., Hatsukade, B., Pearson, C., et al. 2012, MNRAS, 427, 1830

Williams, W. L., van Weeren, R. J., Röttgering, H. J. A., et al. 2016, MNRAS, 460, 2385

Wilman, R. J., Miller, L., Jarvis, M. J., et al. 2008, MNRAS, 388, 1335

Windhorst, R., Mathis, D., & Neuschaefer, L. 1990, in Astronomical Society of the Pacific Conference Series, Vol. 10, Evolution of the Universe of Galaxies, ed. R. G. Kron, 389–403

Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, ApJ, 554, 803