A Multiwavelength Analysis of the Faint Radio Sky (COSMOS-XS): the Nature of the Ultra-faint Radio Population

H. S. B. Algera[1] , D. van der Vlugt[1], J. A. Hodge[1] , I. R. Smail[2] , M. Novak[3] , J. F. Radcliffe[4,5], D. A. Riechers[6,3] , H. Röttgering[1], V. Smolčić[7], and F. Walter[3]

[1] Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands; algera@strw.leidenuniv.nl
[2] Centre for Extragalactic Astronomy, Durham University, Department of Physics, South Road, Durham DH1 3LE, UK
[3] Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
[4] Department of Physics, University of Pretoria, Lynnwood Road, Hatfield, Pretoria 0083, South Africa
[5] Jodrell Bank Centre for Astrophysics, The University of Manchester, Manchester SK11 9DL, UK
[6] Department of Astronomy, Cornell University, Space Sciences Building, Ithaca, NY 14853, USA
[7] Department of Physics, University of Zagreb, Bijeniĉka cesta 32, 10002 Zagreb, Croatia

Abstract

Ultra-deep radio surveys are an invaluable probe of dust-obscured star formation, but require a clear understanding of the relative contribution from radio active galactic nuclei (AGNs) to be used to their fullest potential. We study the composition of the μJy radio population detected in the Karl G. Jansky Very Large Array COSMOS-XS survey based on a sample of 1540 sources detected at 3 GHz over an area of $\sim$350 arcmin^2. This ultra-deep survey consists of a single pointing in the well-studied COSMOS field at both 3 and 10 GHz and reaches rms sensitivities of 0.53 and 0.41 μJy beam^{-1}, respectively. We find multiwavelength counterparts for 97% of radio sources, based on a combination of near-UV/optical to sub-millimeter data, and through a stacking analysis at optical/near-IR wavelengths we further show that the sources lacking such counterparts are likely to be high-redshift in nature (typical $z \sim 4-5$). Utilizing the multiwavelength data over COSMOS, we identify AGNs through a variety of diagnostics and find these to make up $23.2 \pm 1.3\%$ of our sample, with the remainder constituting uncontaminated star-forming galaxies. However, more than half of the AGNs exhibit radio emission consistent with originating from star formation, with only $8.8 \pm 0.8\%$ of radio sources showing a clear excess in radio luminosity. At flux densities of $\sim$30 μJy at 3 GHz, the fraction of star formation-powered sources reaches $\sim$90%, and this fraction is consistent with unity at even lower flux densities. Overall, our findings imply that ultra-deep radio surveys such as COSMOS-XS constitute a highly effective means of obtaining clean samples of star formation-powered radio sources.

Supporting material: machine-readable table

1. Introduction

One of the key quests in extragalactic astronomy is to understand how the build-up and subsequent evolution of galaxies proceeds across cosmic time. Deep radio surveys offer an invaluable window onto this evolution, as they are a probe of both recent dust-unbiased star formation activity, as well as emission from active galactic nuclei (AGNs). Past radio surveys were mostly limited to probing the latter, as AGNs make up the bulk of the bright radio population (e.g., Condon 1989). Current surveys, in large part owing to the increased sensitivity of the upgraded NSF's Karl G. Jansky Very Large Array (VLA), are now changing this, and allow for joint studies of star-forming galaxies (SFGs) and faint AGNs. This, however, requires these populations be distinguished from each other, necessitating detailed studies of the multiwavelength properties of the radio-detected population (e.g., Bonzini et al. 2013; Padovani et al. 2017; Smolčić et al. 2017a; Delvecchio et al. 2017).

Radio emission from SFGs is dominated at frequencies below $\lesssim$30 GHz by non-thermal synchrotron radiation (Condon 1992), which is thought to originate mainly from the shocks produced by the supernova explosions that end the lives of massive, short-lived stars. This conclusion is supported by the existence of the far-IR$-$radio correlation (FIRC; e.g., de Jong et al. 1985; Helou et al. 1985; Yun et al. 2001; Bell 2003), which constitutes a tight correlation between the (predominantly non-thermal) radio and FIR emission of SFGs. As the FIR emission is dominated by thermal radiation from dust that has been heated by young, massive stars, this allows for the usage of radio continuum emission as a star formation tracer through the FIR$-$radio correlation. Radio observations of the star-forming population are therefore, by definition, dust-unbiased, and hence provide an invaluable probe of cosmic star formation. In particular, radio surveys directly complement rest-frame UV studies of the cosmic star formation rate (SFR) density (SFRD) which, while extending out to very high redshift ($z \sim 10$; e.g., Bouwens et al. 2015; Oesch et al. 2018), are highly sensitive to attenuation by dust. The extent of such dust attenuation remains highly uncertain beyond "cosmic noon" ($z \gtrsim 3$; e.g., Casey et al. 2018), and necessitates the additional use of dust-unbiased tracers of star formation. Recently, Novak et al. (2017) performed the first study of the radio SFRD out to $z \sim 5$, finding evidence that UV-based studies may underestimate the SFRD by 15%–20% beyond $z \gtrsim 4$. They attribute this to substantial star formation occurring in dust-obscured galaxies, which further highlights the value in carrying out deep radio surveys. Their findings are consistent with results from large sub-millimeter surveys, which are predominantly sensitive to this dusty star-forming population (see Casey et al. 2014 for a review).

In the last few decades, it has also become increasingly clear that the evolution of individual galaxies is substantially affected by the presence of supermassive black holes (SMBHs) in their center (e.g., Kormendy & Ho 2013). Among such evidence found locally is the Magorrian relation (Magorrian et al. 1998) describing the tight correlation between the mass of the central SMBH of a galaxy and of its bulge. In addition, the cosmic history of black hole growth is comparable to the growth in stellar mass (Shankar et al. 2009), suggesting strong co-evolution. This is often explained by the accretion processes onto the black hole regulating star formation through mechanical feedback, either impeding (e.g., Best et al. 2006; Farrah et al. 2012) or instead triggering (Wang et al. 2010; Reines et al. 2011) epochs of star formation. Such AGN feedback is vital in particular for numerical simulations (e.g., Springel et al. 2005; Schaye et al. 2015) in order to recover, for example, galaxy luminosity functions and local scaling relations. Furthermore, direct evidence for mechanical feedback has been observed in local systems (e.g., McNamara & Nulsen 2012; Morganti et al. 2013). These results exemplify the importance of studying both galaxies and AGNs jointly, instead of as separate entities.

Typically, two populations of AGNs can be distinguished in radio surveys: sources that can be identified through an excess in radio emission compared to what is expected from the FIRC (henceforth referred to as radio-excess AGNs) and sources that have radio emission compatible with originating from star formation, but can be identified as AGNs through any of several multiwavelength diagnostics (Padovani et al. 2011; Bonzini et al. 2013; Heckman & Best 2014; Delvecchio et al. 2017; Smolčić et al. 2017a; Calistro Rivera et al. 2017).[8] The latter class generally exhibit AGN-related emission throughout the bulk of their non-radio spectral energy distribution (SED), in the form of, e.g., strong X-ray emission or mid-IR (MIR) dust emission from a warm torus surrounding the AGN (e.g., Evans et al. 2006; Hardcastle et al. 2013). This distinction illustrates the importance of using multiwavelength diagnostics for identifying AGN activity, which forms the focus of this work.

The distribution of the star-forming and AGN populations has been well-established to be a strong function of radio flux density. At high flux densities, the radio-detected population is dominated by radio-excess AGNs (Kellermann & Wall 1987; Condon 1989), followed by a flattening of the number counts around ~ 1 mJy. This flattening was initially interpreted as the advent of purely SFGs, but subsequent studies (e.g., Gruppioni et al. 1999; Bonzini et al. 2013; Padovani et al. 2015) probing down to a few hundreds to tens of μJy at 1.4 GHz revealed that a substantial part of the sub-mJy population remains dominated by non-radio excess AGNs, with the current consensus being that SFGs only start dominating the radio source counts below $\sim 100\,\mu$Jy (Padovani et al. 2011; Bonzini et al. 2013; Padovani et al. 2015; Maini et al. 2016; Smolčić et al. 2017a), which is also in fairly good agreement with predictions from semi-empirical models of the radio source counts (Wilman et al. 2008; Bonaldi et al. 2019). This faint regime is of great interest for studies of star formation, but has not yet been widely accessible to present-day radio telescopes.

With the upgraded VLA in particular, the radio population at the μJy level can now reliably be probed, which will help constrain the relative contributions of the various radio populations to unprecedented flux densities. Historically, a "wedding-cake approach" has been the tried-and-tested design for radio surveys, incorporating both large-area observations and deeper exposures of smaller regions of the sky. By combining such observations, a clear consensus on both the bright and faint end of the radio population can be reached, which is crucial for understanding the different classes of radio-detected galaxies, as well as for the accurate determination of the radio source counts, luminosity functions, and the subsequent radio-derived cosmic star formation history.

The COSMOS-XS survey (van der Vlugt et al. 2020) was designed to explore the faint radio regime and, by construction, constitutes the top of the wedding cake, making it the natural complement to large-area surveys such as the 3 GHz VLA-COSMOS project (Smolčić et al. 2017a, 2017b). By going a factor of ~ 5 deeper than this survey, we directly probe two of the most interesting populations, namely the poorly understood radio-quiet AGNs and the faint SFGs. While radio surveys were historically limited by their inability to probe the typical star-forming population, being sensitive mostly to starburst galaxies (typical star formation rates in excess of $\gtrsim 100\,M_\odot\,\mathrm{yr}^{-1}$), the COSMOS-XS survey reaches sub-μJy depths and allows for the detection of typical star-forming sources out to redshifts of $z \lesssim 3$. For this reason, the survey is well-suited to bridge the gap between the current deepest radio surveys and those of the next-generation radio telescopes.

COSMOS-XS targets a region of the well-studied COSMOS field (Scoville et al. 2007), such that a wealth of multiwavelength data are available for the optimal classification of the radio population. Such ancillary data are crucial to place the survey into a wider astronomical context, and allows for the connection of the radio properties of the observed galaxy population to observations in the rest of the electromagnetic spectrum. With the combined COSMOS-XS survey and multiwavelength data, we ultimately aim to constrain the faint end of the high-redshift radio luminosity functions of both SFGs and AGNs, and use these to derived the corresponding dust-unbiased cosmic star formation history, as well as the AGN accretion history. Additionally, the unprecedented depth at multiple radio frequencies allows for the study of the high-redshift radio SED in great detail, and allows for the systematic isolation of the radio free–free component in SFGs to be used as a robust SFR tracer (e.g Murphy et al. 2012, 2017). All of these science goals require a good understanding of the origin of the observed radio emission and hence depend on whether it emanates from star formation or instead from an AGN, and will be tackled in forthcoming papers.

In this paper, the second in the series describing the COSMOS-XS radio survey, we address this question by studying the composition of the ultra-faint radio population detected at 3 GHz in an ultra-deep single pointing over the COSMOS field (reaching an rms of 0.53 μJy beam^{-1}). We additionally present a catalog containing the multiwavelength counterparts of the radio sources selected at 3 GHz, utilizing the wealth of X-ray to radio data available over COSMOS. The paper is structured as follows. In Section 2 we summarize the COSMOS-XS observations, the radio-selected catalog as well as the multiwavelength ancillary data. In Section 3 we describe the association of counterparts to the radio sample. The

---

[8] These sources are also often referred to as radio-loud and radio-quiet AGNs respectively (e.g., Heckman & Best 2014), though we follow here the terminology from Delvecchio et al. (2017) and Smolčić et al. (2017a).

decomposition of the radio population into SFGs and AGNs is laid out in Section 4, and we present the multiwavelength counterpart catalog in Appendix D. Finally, we present the ultra-faint radio soure counts, separated into SFGs and AGNs, in Section 5 and we summarize our findings in Section 6. Throughout this work, we assume a flat Λ-CDM cosmology with $\Omega_m = 0.30$, $\Omega_\Lambda = 0.70$ and $H_0 = 70 \ \mathrm{km \, s^{-1} \, Mpc^{-1}}$. Magnitudes are expressed in the AB system (Oke & Gunn 1983), and a Chabrier (2003) initial mass function is assumed. The radio spectral index, α, is defined through $S_\nu \propto \nu^\alpha$ where S_ν represents the flux density at frequency ν.

2. DATA

2.1. Radio Data

The COSMOS-XS survey consists of a single ultra-deep VLA pointing in the well-studied COSMOS field at both 3 and 10 GHz of ∼100 and ∼90 hr of observation time, respectively. The full survey is described in detail in Paper I, but we summarize the key procedures and parameters here. The 3 GHz observations (also known as the S-band) were taken in B-array configuration, and span a total bandwidth of 2 GHz. The effective area of the pointing—measured up to 20% of the peak primary beam sensitivity at the central frequency—is approximately 350 arcmin². The 10 GHz pointing (X-band) was taken in C-array configuration, and spans a frequency range of 4 GHz around the central frequency. The total survey area is approximately 30 arcmin² at 10 GHz. At both frequencies, roughly 20% of the bandwidth was lost due to excessive radio frequency interference.

Imaging of both data sets was performed using the standalone imager WSCLEAN (Offringa et al. 2014), incorporating w-stacking to account for the non-coplanarity of our baselines. Both images were created via Briggs weighting, with a robust parameter set to 0.5. This resulted in a synthesized beam of $2\rlap{.}''14 \times 1\rlap{.}''81$ at 3 GHz and $2\rlap{.}''33 \times 2\rlap{.}''01$ at 10 GHz. The near-equal resolution of ∼2″0 was chosen to be large enough to avoid resolving typical faint radio sources, which are generally subarcsecond in size at the μJy level (Bondi et al. 2018; Cotton et al. 2018). In turn, this resolution allows for the cleanest measurement of their radio flux densities, while ensuring that confusion noise remains negligible (Paper I).

The final rms noise levels of the S- and X-band images are 0.53 μJy beam⁻¹ and 0.41 μJy beam⁻¹ at their respective pointing centers. For the S-band, the rms noise corresponds to a brightness temperature of $T_B \simeq 20$ mK. This implies we are sensitive even to face-on star-forming spiral galaxies, which have a typical $T_B = 0.75 \pm 0.25$ K (Hummel 1981).

The locations of both the S- and X-band pointings within the COSMOS field are shown in Figure 1, and were explicitly chosen to match the pointing area of the COLDz survey (Pavesi et al. 2018; Riechers et al. 2019). A zoomed-in view of the radio maps themselves are presented in Paper I.

Source extraction on both images was performed with PYBDSF (Mohan & Rafferty 2015), down to a 5σ peak flux threshold. PYBDSF operates through identifying islands of contiguous emission around this peak value, and fitting such islands with elliptical Gaussians to obtain peak and integrated flux measurements. In total, we obtain 1498 distinct radio sources within 20% of the peak primary beam sensitivity, of which 70 (∼5%) consist of multiple Gaussian components. While our survey is far from being confusion limited, a fraction of islands deemed robustly
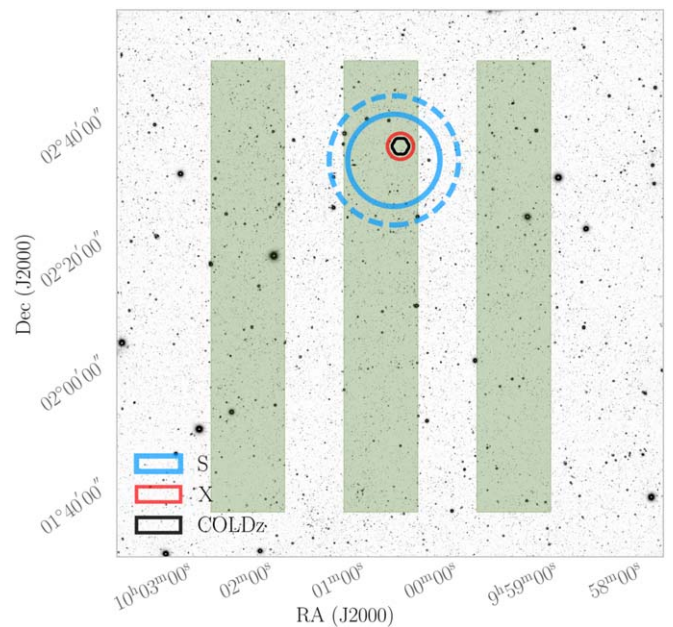

Figure 1. Layout of the S- (blue) and X-band (red) pointings, out to the half-power point of the primary beam (solid circles). For the S-band, we further highlight the field of view out to 20% of the peak primary beam sensitivity (dashed circle), which defines our total survey area, over the 2 deg² COSMOS field. Also shown is the 34 GHz COLDz mosaic (Pavesi et al. 2018), which overlaps in its entirety with the X-band field of view. The background image consists of Subaru observations in the i+ filter, with the UltraVISTA ultra-deep stripes overlaid in green. For a zoomed-in view of the radio maps themselves, see van der Vlugt et al. (2020).

resolved in the source detection are in fact artificially extended as a result of source blending. In order to disentangle true extended sources from blended ones, we examined whether the Gaussian components making up an island can be individually cross-matched to separate sources in the recent Super-deblended catalog over COSMOS (Jin et al. 2018, see Section 2.2), which contains MIR to radio photometry based on positional priors from a combination of K_s, Spitzer/MIPS 24 μm, and VLA 1.4 and 3 GHz observations. As these radio data are both shallower and higher resolution than our observations, they are suitable for assessing any source blending. Furthermore, radio sources such as FR-II AGNs (Fanaroff & Riley 1974) are not expected to have MIR counterparts for their individual lobes as these sources, by definition, have their radio emission spatially offset from the host galaxy. Hence, when all components can be individually associated to different multiwavelength counterparts, we deem these associations robust, and define the Gaussian components to be separate radio sources. Altogether, we find that 40 of the initial 70 multi-Gaussian sources separate into 82 single "deblended" components. In total, our radio survey therefore consists of 1540 individual sources detected at 3 GHz, within 20% of the peak primary beam sensitivity. Altogether, this results in a ∼4% increase in the number of radio sources that is cross-matched to a multiwavelength counterpart (Section 3). This is the result of two effects: first, we find an overall larger number of radio sources now that blended ones have been separated and, second we find a larger fractional number of cross-matches, as blended sources are assigned a flux-weighted source center that could be substantially offset from the true centers of the individual Gaussian components, preventing reliable cross-matching. We note that this procedure slightly deviates from that in Paper I, where we focused instead on the radio properties of this sample and

therefore refrained from invoking multiwavelength cross-matching.

In addition to the 1540 S-band detections, a total of 90 sources are detected at 10 GHz within 20% of the maximum primary beam sensitivity at $\geqslant 5\sigma$ significance (948 and 60 sources lie within the half-power point of the primary beam).[9] The S-band sample comprises the main radio sample used in the subsequent analysis described in this paper. We assign the radio sources detected at 3 GHz either their peak or integrated flux density, following the method from Bondi et al. (2008) described in detail in Paper I. We then ensure that we take the same flux density measurement for the X-band sources, as all sources detected at 10 GHz can be cross-matched to an S-band counterpart (Section 3) and the radio data have a similar resolution at both frequencies. In Paper I, we additionally investigated the completeness and reliability of the radio sample, for which we repeat the main conclusions here. We found that the catalogs are highly reliable with a low number of possible spurious detections ($\lesssim 2\%$). We further investigate possible spurious sources by visually inspecting all detections within $30''$ of a bright (signal-to-noise ratio (S/N) $\geqslant 200$) radio source. From these we flag, but do not remove, eight sources that are potentially spurious or have their fluxes affected as a result of the characteristic VLA dirty beam pattern around the nearby bright object. The S-band sample was further determined to be $\geqslant 90\%$ complete above integrated flux densities of 15 μJy, with the completeness dropping to 50% at $\sim 10\,\mu$Jy due mainly to the primary beam attenuation reducing the survey area. In our derivation of the radio number counts for star-forming sources and AGNs (Section 5) all of these completeness considerations are taken into account.

Additional radio data over our pointings exist as part of the 1.4 GHz VLA-COSMOS survey (Schinnerer et al. 2007, 2010), which reaches an rms of $\sim 12\,\mu$Jy beam^{-1}. Accounting for the frequency difference through scaling with $\alpha = -0.70$, our S-band observations are a factor ~ 13 deeper than these lower-frequency data, and hence the 1.4 GHz observations are mostly useful for the brightest sources detected at 3 and 10 GHz. Additionally, a seven-pointing mosaic at 34 GHz (rms $\sim 1.4\,\mu$Jy beam^{-1}, area ~ 10 arcmin^2) exists as part of the COLDz project (Pavesi et al. 2018; Riechers et al. 2019; Algera et al. 2020). The COSMOS-XS 10 GHz data is directly centered on this mosaic, allowing for a detailed analysis of the long-wavelength spectrum of faint radio sources with up to four frequencies. We defer this analysis to a future paper.

2.2. NUV to FIR Data

The COSMOS field has been the target of a considerable number of studies spanning the full electromagnetic spectrum. We complement our radio observations with NUV to FIR data that have been compiled into various multiwavelength catalogs: (i) the Super-deblended MIR to FIR catalog (Jin et al. 2018) containing photometry ranging from IRAC 3.6 μm to 20 cm (1.4 GHz) radio observations, (ii) the $z^{++}YJHK_s$-selected catalog compiled by Laigle et al. (2016) (hereafter COS-MOS2015), and (iii) the i-band selected catalog by Capak et al. (2007).

The Super-deblended catalog contains the latest MIR–radio photometry for nearly 200,000 sources in the COSMOS field, with 12,335 located within the COSMOS-XS field of view. Due to the relatively poor resolution of FIR telescopes such as Herschel, blending of sources introduces complications for accurate photometry. Through the use of priors on source positions from higher-resolution images (Spitzer/MIPS 24 μm and VLA 1.4 and 3 GHz observations) in combination with point-spread function fitting, the contributions to the flux from various blended galaxies in a low-resolution image can be partly disentangled. This "Super-deblending" procedure is described in detail in Liu et al. (2018). The Super-deblended catalog contains photometry from Spitzer/IRAC (Sanders et al. 2007) and Spitzer/MIPS 24 μm (Le Floc'h et al. 2009) as part of the S-COSMOS survey, Herschel/PACS 100 μm and 160 μm data from the PEP (Lutz et al. 2011) and CANDELS-Herschel (PI: M. Dickinson) programs, Herschel/SPIRE images at 250, 350 and 500 μm as part of the HerMES survey (Oliver et al. 2012), and further FIR data at 850 μm from SCUBA2 as part of the Cosmology Legacy Survey (Geach et al. 2017), AzTEC 1.1mm observations (Aretxaga et al. 2011) and MAMBO 1.2 mm images (Bertoldi et al. 2007), in addition to 1.4 GHz and 3 GHz radio observations from Schinnerer et al. (2007, 2010) and Smolčić et al. (2017b), respectively. However, we use the photometry from catalogs provided directly by Schinnerer et al. (2007, 2010), as they provide both peak and integrated fluxes, whereas the Super-deblended catalog solely provides peak values. In addition, we use our ~ 5 times deeper COSMOS-XS 3 GHz observations described in Section 2.1 in favor of those from Smolčić et al. (2017b).

The Super-deblended catalog further includes photometric and spectroscopic redshifts, based on the Laigle et al. (2016) catalog where available, in addition to IR-derived photometric redshifts and star formation rates.[10]

Since a shallower radio catalog was used for the deblending procedure, this raises the concern that one of the VLA-COSMOS priors for the Super-deblended catalog is, in fact, located near a fainter radio source detected only in COSMOS-XS that contributes partially to the FIR flux at that location. In such a scenario, all the FIR emission would be wrongfully assigned to the brighter radio source, which would have an artificially boosted flux, and the fainter source may be wrongfully assigned no FIR counterpart. We verified, however, that this is not likely to be an issue, as we observe no drop in the fraction of cross-matches between COSMOS-XS and the Super-deblended catalog for radio sources with a nearby neighbor in COSMOS-XS. Hence there is no indication of any boosting in the FIR-fluxes of Super-deblended entries due to a nearby, faint radio source.

The COSMOS2015 catalog contains photometry for upwards of half a million entries over the 2 deg^2 COSMOS field, including 37,841 within the COSMOS-XS field of view. Sources are drawn from a combined $z^{++}YJHK_s$ detection image, where the deep $YJHK_s$ observations are taken from the second UltraVISTA data release (McCracken et al. 2012). The COSMOS-XS S-band pointing is largely located within one of the UltraVISTA "ultra-deep" stripes (Figure 1), which reaches a 3σ depth in magnitude of 25.8, 25.4, 25.0, and 25.2 in Y, J, H, and K_s respectively, as measured in $3''$ apertures. The

---

[9] With the adopted primary beam cut-off of 20%, the COSMOS-XS survey is deeper than the ($\sim 5\times$ shallower) 3 GHz VLA-COSMOS survey (Smolčić et al. 2017a, 2017b) across the entire field of view.

[10] We do not use FIR-derived photometric redshifts in this work, as these values are considerably more uncertain than those derived from NUV to NIR photometry (e.g., Simpson et al. 2014), but we comment on the small sample of sources with only FIR photometric redshifts in Section 5.3.

COSMOS2015 catalog further provides cross-matches with NUV, and MIR/FIR data. The former consists of GALEX observations at 1500 Å (FUV) and 2500 Å (NUV) (Zamojski et al. 2007), and the latter ranges from Spitzer/MIPS 24 μm to Herschel/SPIRE 500 μm photometry, drawn from the same programs as introduced for the Super-deblended catalog. In addition, photometric redshifts, star formation rates, and stellar masses are provided by COSMOS2015, derived through SED fitting by LEPHARE (Ilbert et al. 2006), using photometry spanning the NUV-K_s bands.

Finally, the i-band-selected catalog compiles photometry from 15 photometric bands between the u-band (0.3 μm) and the 2.5 μm K_s-band. The 5σ depth in the i-band equals 26.2 as determined within a $3''$ aperture. Photometric redshifts were derived from SED fitting with LEPHARE and were later added to the catalog by Ilbert et al. (2009).

2.3. Spectroscopic Redshifts

A substantial fraction of galaxies in the COSMOS field have been targeted spectroscopically, and therefore have a robustly determined redshift. We make use of the "master spectroscopic redshift catalog" available internally to the COSMOS team (version 2017 September 1; M. Salvato 2020, in preparation). It contains $\sim$100,000 entries, compiled from a large number of spectroscopic surveys over the COSMOS field, including zCOSMOS (Lilly et al. 2007, 2009), the VIMOS Ultra Deep Survey (Le Fèvre et al. 2015) and MOSDEF (Kriek et al. 2015). In total, there are 5074 sources with reliable spectroscopic redshifts in the COSMOS-XS field of view (see also Section 3.3).

2.4. X-Ray Data

Strong X-ray emission is a vital diagnostic for AGN activity. The most recent X-ray data over the COSMOS field is the 4.6 Ms Chandra COSMOS Legacy survey (Civano et al. 2016), covering the full 2.2 deg². Marchesi et al. (2016a) present the catalog of the optical and IR counterparts of the X-ray sources identified by the survey. This catalog includes 200 sources within our S-band field of view, and contains for each the absorption-corrected luminosity in the soft [0.5–2] keV, hard [2–10] keV, and full [0.5–10] keV bands, or the corresponding upper limit in the case of a non-detection in a given energy band.[11] Where available, X-ray sources in this catalog were assigned the spectroscopic redshift of their counterparts, taken from the COSMOS master spectroscopic catalog. For the remaining sources, photometric redshifts exist, based on template fitting making use of AGN-specific templates as described in Salvato et al. (2011). X-ray luminosities were calculated by using the best available redshift and an X-ray spectral index of $\Gamma = 1.4$ for the required K-corrections, which is a typical value for a mix of obscured and unobscured sources (Marchesi et al. 2016b). Absorption corrections to the X-ray luminosity in each energy band were calculated based on the measured hardness ratio, as described in Civano et al. (2016).

3. Multiwavelength Cross-matching

In this section, we elaborate on our multiwavelength matching process. As a brief summary of the procedure, we
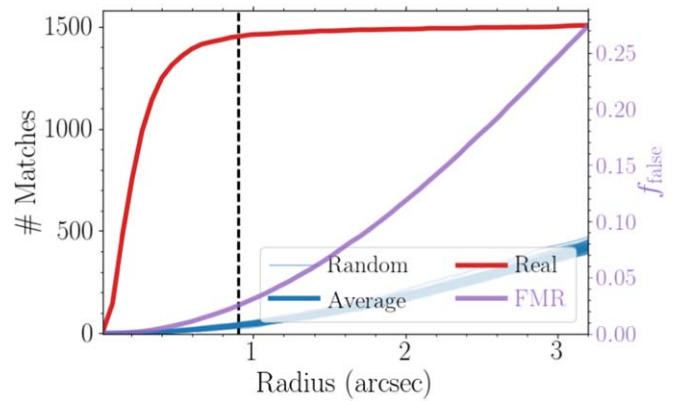

Figure 2. Illustration of the cross-matching process, for the radio and Super-deblended catalogs. The red line represents the number of matches obtained within a given matching radius, whereas the blue lines show the number of matches obtained when cross-matching with a catalog with the same source density but randomized sky positions. The purple curve, corresponding to the right ordinate axis, shows the false match fraction obtained at a given matching radius. The vertical black line indicates the typical matching radius of $0\rlap{.}''9$ adopted in this work.

cross-match catalogs based on a symmetric nearest-neighbor algorithm, whereby we search for counterparts within a given matching radius. A suitable value of this radius is determined through cross-matching with a mock version of the appropriate catalog, which contains the same sources with randomized sky coordinates. As our radio images are not uniformly sensitive across their full field of view as a result of the primary beam, we ensure mock sources can only be placed in the region where they can theoretically be detected at $\geqslant 5\sigma$, to mimic the true distribution of sources. Through cross-matching with such mock catalogs, we obtain an estimate of the number of false matches at any given matching radius, and we hence define the false matching rate (FMR) as the average number of cross-matches obtained with the randomized catalogs divided by the total number of catalog entries. The matching radius we adopt is taken to be the radius where the number of matches for the real and mock catalogs is (approximately) equal, which is generally around $0\rlap{.}''9$, and coincides with a typical FMR $\lesssim 3\%$ for all our multiwavelength cross-matching. The matching process is illustrated in Figure 2.

3.1. Radio Cross-matching

There are a total of 1540 and 90 radio sources within 20% of the peak primary beam sensitivity for the S- and X-band, respectively. We cross-match these two frequencies using a matching radius of $0\rlap{.}''9$, which yields 89 matches (FMR $\lesssim 0.7\%$). The single X-band source that could not be cross-matched to a counterpart at 3 GHz appears to be a lobed radio source where the relative brightness of the two lobes is different in the two images, causing the center of the source to be appreciably offset between the two images ($\sim 1\rlap{.}''25$). Despite this offset, visual inspection verifies that the sources are related, such that all X-band sources are assigned a counterpart at 3 GHz. Due to the relatively low density of radio sources in the VLA COSMOS 1.4 GHz catalog, we utilize a matching radius of $1\rlap{.}''2$ when cross-matching with the S-band data (FMR $\lesssim 0.1\%$). This generates 185 matches, with 12 sources being detected at all three frequencies (1.4, 3 and 10 GHz).
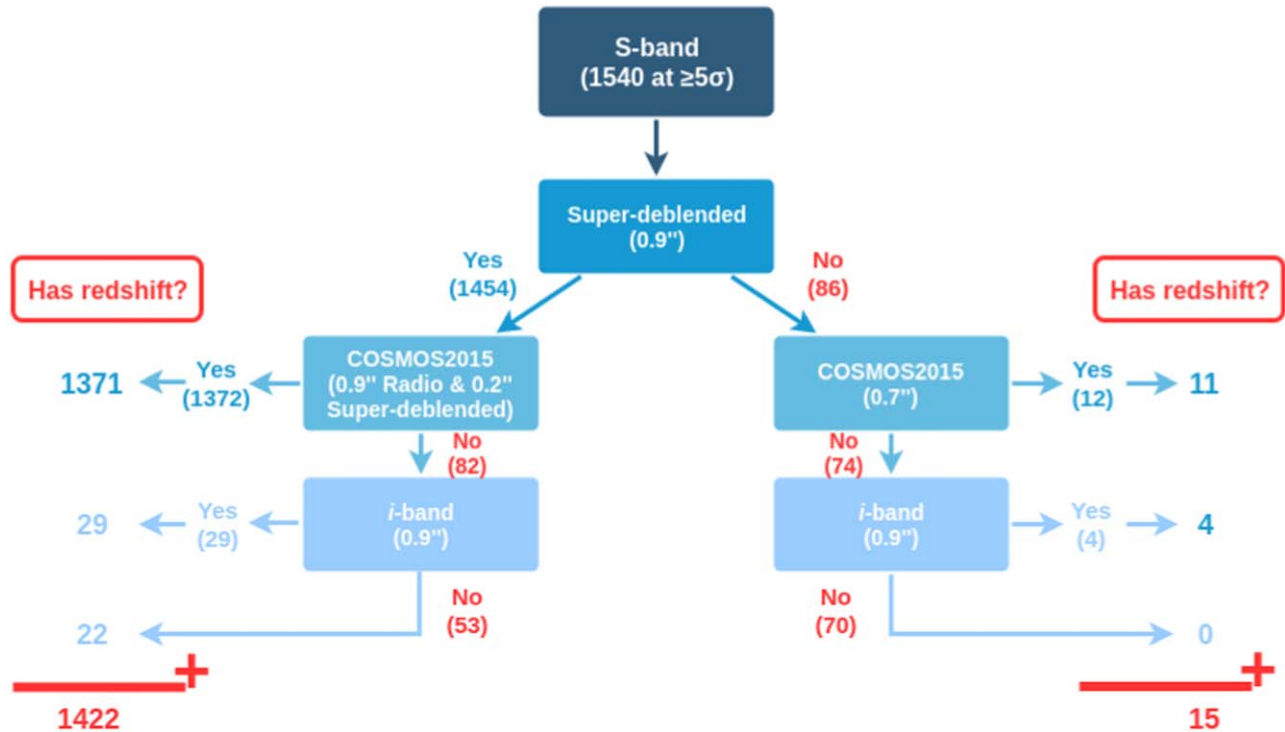
---

[11] The flux limit of the Chandra COSMOS Legacy Survey over the COSMOS-XS S-band pointing is $\sim 2 \times 10^{-15}$ erg cm^{-2} s^{-1} in the full 2–10 keV range (Civano et al. 2016).

Figure 3. Flowchart of the matching process, indicating the priority of catalogs used and number of cross-matches obtained with each. We match a total of 1470/1540 sources (95.5%) to a counterpart in at least one multiwavelength catalog. From this sample, we keep 1437 sources (93.3%) with reliable redshift information. This represents the main radio sample studied in this work.

3.2. Additional Cross-Matching

In order to construct UV/optical–FIR SEDs for the S-band-detected sources, we cross-match with three catalogs in order of decreasing priority: Super-deblended (Jin et al. 2018), COSMOS2015 (Laigle et al. 2016), and the i-band-selected catalog by Capak et al. (2007). The Super-deblended catalog contains the most up-to-date collection of FIR photometry available over the COSMOS field, but does not include optical and NIR photometry shortward of IRAC 3.6 μm. We therefore attempt to cross-match S-band sources that have Super-deblended counterparts with either COSMOS2015 or the i-band-selected catalog, which do contain photometry at these shorter wavelengths.

An overview of the matching process is presented in the form of a flowchart in Figure 3. In summary, the procedure is as follows: we first cross-match the S-band-selected catalog of 1540 sources with the Super-deblended photometric catalog, obtaining 1454 matches within $0\rlap{.}''9$ (FMR $\simeq$ 2.5%, see Figure 2). To obtain optical and NIR photometry, we subsequently cross-match with the COSMOS2015 catalog. As a result of the larger source density compared to the Super-deblended catalog, we use a matching radius of $0\rlap{.}''7$ instead of $0\rlap{.}''9$ around the radio coordinates, motivated by the number of false matches obtained at larger radii. The theoretical FMR at $0\rlap{.}''7$ equals $\sim$4.8%, which is substantial. To account for both this and the difference in matching radii between the different catalogs, which may lead to inconsistencies in the overall cross-matching process where we artificially cannot assign our sources COSMOS2015 counterparts if their offset from the radio coordinates falls into the range $0\rlap{.}''7 < r < 0\rlap{.}''9$, we therefore also utilize the Super-deblended source coordinates. As these are based in part on IR detections, they are typically more similar to the NIR-derived COSMOS2015 coordinates. We therefore perform additional cross-matching within $0\rlap{.}''2$ around these Super-deblended source positions—since such closely

associated sources are likely to be real (FMR $\lesssim$ 0.4%)—under the constraint that the offset between any of the catalogs is less than $0\rlap{.}''9$. We then acquire 1372 sources with both Super-deblended and COSMOS2015 photometry, which spans the full NUV to millimeter range. Only 25 of these cross-matches (1.8%) have COSMOS2015–Super-deblended separations $\geq 0\rlap{.}''2$, but as the median separation between the radio and Super-deblended coordinates remains small ($\sim 0\rlap{.}''3$), we expect a negligible increase in the overall FMR after cross-matching with the COSMOS2015 catalog. We additionally have 82 sources with Super-deblended cross-matches for which we did not obtain COSMOS2015 counterparts. As these sources will greatly benefit from shorter-wavelength data in our subsequent analysis, we cross-match them with the i-band-selected catalog within $0\rlap{.}''9$ and recover an additional 29 matches.[12] We thus obtain full NUV to radio photometry for 96.4% of sources cross-matched with the Super-deblended catalog.

For the 86 (5.6%) of radio-detected sources for which we could not acquire robust Super-deblended counterparts, we first cross-match with the COSMOS2015 catalog within $0\rlap{.}''7$, gaining 12 matches (four expected false matches). For the remaining sources, we obtain four matches with the i-band catalog within $0\rlap{.}''9$, containing photometry up to the K_s-band (with one expected false match). Overall, we did not obtain any non-radio counterparts for 70/1540 sources (4.5%). Two of these are detected at multiple radio frequencies and are therefore certainly real. Upon cross-matching with the S-COSMOS IRAC catalog from Sanders et al. (2007), we recover 18 additional matches within $0\rlap{.}''9$, which further indicates that a substantial number of this sample consists of

---

[12] The formal cross-matching radius adopted is $0\rlap{.}''9$ for consistency with other catalogs, but all matches lie within $0\rlap{.}''5$ of both the Super-deblended and radio sky positions, and as such we estimate the FMR to be $\lesssim$1%.

real radio sources that evade detection at shorter wavelengths. However, as these cross-matches therefore solely have IRAC photometry, they lack redshift information, which is crucial for subsequent AGN identification (Section 4). Therefore, we do not further include these sources in the characterization of our radio sample, though we return to these "optically dark" detections in Section 5.3. When accounting for these additional 18 matches with S-COSMOS, 96.6% of our radio sample can be cross-matched to a non-radio counterpart.

As reliable redshift information is crucial for investigating the physical properties of our radio-detected sources, we further remove 33 galaxies from our sample for which no redshift information was available in any of the catalogs. The majority of these sources are only present in the Super-deblended catalog, and were detected through priors from the 3 GHz VLA COSMOS project (Smolčić et al. 2017b). As for these sources no photometry exists shortward of MIPS 24 μm, no reliable photometric redshift can be obtained. Altogether, we are left with a remainder of 1437 sources (93.3% of the initial 1540 detections at 3 GHz). This constitutes the S-band-detected sample that will be used in the subsequent analysis.

We additionally recover 108 cross-matches with X-ray sources taken from the Chandra COSMOS Legacy survey (Civano et al. 2016; Marchesi et al. 2016a) based on an adopted cross-matching radius of $1\rlap{.}{''}4$ (theoretical FMR $\simeq 0.1\%$). [13] The 108 X-ray counterparts for the radio sample correspond to $\sim$51% of the X-ray sources within the COSMOS-XS field of view. This is larger than the 32% found for the 3 GHz VLA-COSMOS survey (Delvecchio et al. 2017), which is consistent with the fact that nearly one third of the radio counterparts we associate to X-ray sources lie below the theoretical $5\sigma_{\rm rms} \simeq 11.5 \mu$Jy detection limit of that survey.

Overall, we expect an FMR of $\lesssim$3% ($\sim$40 sources, which includes four sources flagged as "potentially spurious" in Section 2.1), based on the combined FMRs from the cross-matching of the individual catalogs, as well as a spurious fraction of $\lesssim$2%. We hence deem our multiwavelength catalog to be reliable.

3.3. Redshifts of the Radio Sample

Accurate redshift information for our radio sample is vital not only for the classification of AGNs, but also for subsequent studies of the star formation history of the universe. We therefore attempted to assign to each source its most reliable redshift through comparing the various redshifts (photometric or spectroscopic) we obtained from the different catalogs. First of all, we discarded all spectroscopic redshifts from the COSMOS master catalog (M. Salvato 2020, in preparation) that have a quality factor $Q_f < 3$, indicating an uncertain or poor spectroscopic redshift. All remaining sources for which a robust spectroscopic redshift is available are then assigned this value. The majority of sources additionally have photometrically determined redshifts available, from up to three different studies (e.g., Capak et al. 2007; Laigle et al. 2016; Jin et al. 2018). We prioritize the photometric redshift from the Super-deblended catalog if available, as it is determined using prior photometric redshift information from other catalogs (e.g., COSMOS2015), but is re-computed with the inclusion of longer-wavelength data. However, any differences between these

photometric redshifts are small by construction, as Jin et al. (2018) force the Super-deblended redshift to be within 10% of the prior value. If a Super-deblended redshift is unavailable, we instead use the photometric redshift from COSMOS2015 or the i-band-selected catalog, in that order. We make an exception when the source is X-ray detected, in which case we assign it the photometric redshift from the Chandra X-ray catalog (Marchesi et al. 2016a). These redshifts have been determined through SED fitting with the inclusion of AGN templates, and are therefore more appropriate for AGNs, which form the bulk of our X-ray-detected sample (Section 4.1).

We compute the reliability of our redshifts, defined as $\sigma(z) = |z_{\rm spec} - z_{\rm phot}|/(1 + z_{\rm spec})$ for the 584 sources that have both photometric and spectroscopic redshift information available. The normalized median absolute deviation (Hoaglin et al. 1983), defined as 1.48 times the median of $\sigma(z)$, is found to be 0.012, indicating a very good overall consistency between the two redshifts. The fraction of sources with $\sigma(z) > 0.15$, the common threshold for defining "catastrophic failures", equals 4.8%, with the main region of such failures being the optically fainter sources, as expected. We verified that the distribution of such failures in terms of i-band magnitude is similar to that in Figure 11 of Laigle et al. (2016), though with a slightly larger failure fraction at fainter magnitudes $i_{\rm AB} \gtrsim 23$, which can be fully explained by our small sample size at these magnitudes. [14]

In summary, our sample contains 584 sources ($\sim$41%) with spectroscopic and 853 sources with photometric redshifts (left panel of Figure 4). About two-thirds of our $z \lesssim 1$ sample is detected spectroscopically, but the fraction of spectropscopic redshifts drops dramatically toward higher redshift. Additionally, a total of 103 sources have no redshift information. The majority of these (70 sources) are not cross-matched to any multi-wavelength counterpart. Only two sources without redshift information have optical/NIR photometry, but nevertheless no robust photometric redshift could be obtained for these catalog entries. The median radio flux density (and bootstrapped uncertainty) of the sources without redshift information is $S_{\rm 3GHz} = 10.5^{+2.0}_{-1.2} \mu$Jy, similar to the median of the full sample, which equals $S_{\rm 3\,GHz} = 11.3^{+0.4}_{-0.5} \mu$Jy. The median flux density of the radio sources detected only in COSMOS-XS, i.e., without cross-matches to the Super-deblended or VLA-COSMOS cata-logs, equals $7.2^{+1.4}_{-0.7} \mu$Jy—below the formal detection limit of the latter survey, as expected.

We show the detection limit of the COSMOS-XS survey as function of redshift in the right panel of Figure 4. For ease of comparison with previous surveys, which have predominantly been performed at lower frequencies, we show the distribution of rest-frame 1.4 GHz luminosities, computed using a standard spectral index of $\alpha = -0.7$ even where multiple radio fluxes were available (see Section 4.2). These luminosities are converted into star formation rates adopting the conversion from Bell (2003) under the assumption that the radio emission is fully powered by star formation. In the following section, we will instead adopt a redshift-dependent conversion factor, as it has recently been shown to evolve with cosmic time (Magnelli et al. 2015; Calistro Rivera et al. 2017; Delhaize et al. 2017).

Finally, in Figure 5, we show the overall counterpart completeness of the 1540 S-band sources in 12 flux density bins.

---

[13] We adopted a cross-matching radius of $1\rlap{.}{''}4$ as Chandra astrometry is accurate to 99% within this radius, see http://cxc.harvard.edu/cal/ASPECT/celmon/.

[14] While most of our photometric redshifts are from the Super-deblended catalog, and not from COSMOS2015, the former are by construction similar to the redshift adopted for the deblending prior, and as such comparing our photometric redshift accuracy with Laigle et al. (2016) is justified.
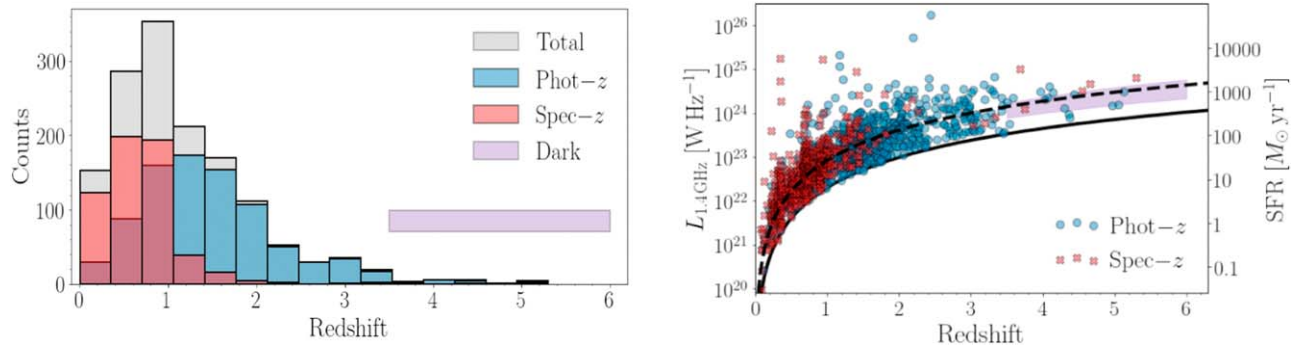
Figure 4. Left: distribution of the radio-selected sample in redshift for all 1437 sources with redshift information. Red and blue bars indicate spectroscopic and photometric redshifts respectively, and the gray histogram represents the full distribution. We additionally show the expected redshift range populated by the 29 "optically dark" sources—those without optical counterparts and redshift information, analyzed in Section 5.3—via the purple bar. Out to $z \sim 1$, nearly two-thirds of our redshifts are spectroscopic. Right: rest-frame 1.4 GHz luminosity vs. redshift for the COSMOS-XS 3 GHz-detected sample with reliable redshift information. The flux limit of the survey is indicated through the solid black line. Blue circles and red crosses indicate photometric and spectroscopic redshifts, respectively. The dashed, black line represents the flux limit of the 3 GHz VLA-COSMOS survey (Smolčić et al. 2017b) and is included for comparison. The COSMOS-XS survey constitutes a factor ∼5 increase in sensitivity compared to these wider radio data. We additionally show the redshift and luminosity range likely populated by the radio-detected sources without optical/NIR photometry through the shaded purple area (Section 5.3). Star formation rates (right ordinate axis) are computed assuming the radio emission is fully powered by star formation, and adopting a conversion that is independent of redshift, based on the local galaxy sample studied by Bell (2003); see also Section 4.2. Radio luminosities are computed at a rest-frame frequency of 1.4 GHz using a fixed spectral index $\alpha = -0.70$. The theoretical detection limit (solid black line) is computed using five times the rms in the S-band image center and scaled to rest-frame 1.4 GHz using an identical spectral index.
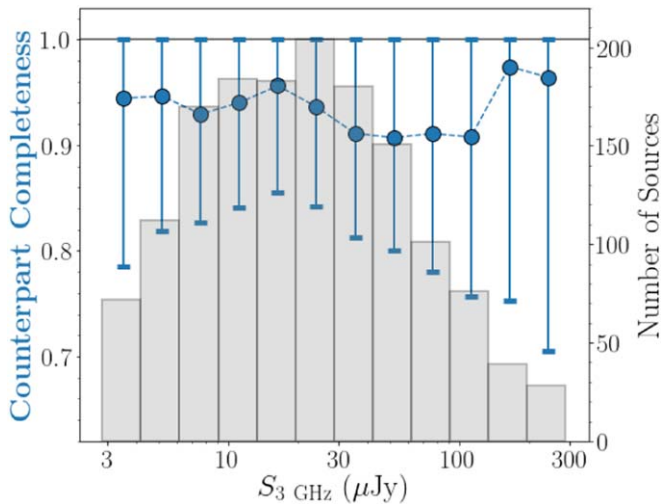

Figure 5. Counterpart completeness as a function of flux density for the S-band sample (left ordinate axis). Our S-band-selected catalog contains counterparts with reliable redshifts for 93.3% of all S-band-detected sources, and all flux density bins are ≳90% complete. No clear trend with flux density is seen, making it unlikely that a substantial number of low-S/N S-band sources are spurious, as for these no counterpart would be expected. The background histogram and corresponding right ordinate axis indicate the total number of sources in a given flux density bin.

Uncertainties on the counting statistics were calculated following Gehrels (1986) for bins with fewer than 10 sources without optical counterparts, and were assumed to be Poissonian otherwise. We adopt these confidence limits for sparsely populated bins throughout this work. The completeness in all bins is upward of 90%, and no trend with radio flux density can be seen, indicating that the association of counterparts to our radio sources is not limited by the depth of the multiwavelength photometry. Additionally, this indicates that it is unlikely that there are a substantial number of spurious radio sources within our 3 GHz image, as these would likely populate the low flux density bins and would typically not be associated to a multiwavelength counterpart. Despite the lack of multiwavelength counterparts for ∼4.5% of radio sources, a substantial fraction of these are simply faint at

optical/NIR wavelengths, and as such are not spurious (see the discussion in Section 5.3).

4. AGN Identification

In this section, we outline the criteria used for identifying AGNs among our radio-detected sources, making use of the wealth of data available over the COSMOS field. In the literature, numerous such multiwavelength identifiers exist, based on both broadband photometry and spectroscopy. In order to avoid potential biases due to incomplete coverage in various bands, we only utilize AGN diagnostics that are available for the vast majority of our radio-selected sample, which are outlined below. This excludes such diagnostics as inverted radio spectral indices (e.g., Nagar et al. 2000), very long baseline interferometry (e.g., Herrera Ruiz et al. 2017), and optical spectra (e.g., Baldwin et al. 1981).[15]

In our panchromatic approach to AGN identification, we make use of the SED-fitting code MAGPHYS (da Cunha et al. 2008, 2015) to model the physical parameters of our galaxy sample. MAGPHYS imposes an energy balance between the stellar emission and absorption by dust, and is therefore well-suited to model the dusty star-forming populations to which radio observations are sensitive. We stress that MAGPHYS does not include AGN templates, and is therefore only suitable for modeling sources whose SEDs are dominated by star formation-related emission. However, we can use this to our advantage when identifying AGNs based on their radio emission in Section 4.2. Additionally, we use AGNFITTER (Calistro Rivera et al. 2016, 2017), a different SED-fitting routine appropriate for AGNs, in Section 4.1.3 to further mitigate this issue. We fit our full radio sample with MAGPHYS, including all FUV to millimeter data. The radio data are not fitted, as an excess in radio emission is indicative of AGN activity and could therefore bias our results (Section 4.2).

---

[15] Radio spectral indices are available for 255 sources (∼18% of the sample with redshift information). Eight of these have inverted spectral indices (seven from 1.4–3 GHz data, and a single source detected at 3 and 10 GHz), though seven were previously identified as AGNs through other multiwavelength diagnostics. The inclusion of an inverted spectral index as AGN diagnostic will therefore have a negligible effect on the overall AGN identification.

We will follow the terminology introduced in Delvecchio et al. 2017 and Smolčić et al. 2017a, and divide the radio-detected AGNs into two classes: the moderate-to-high luminosity AGNs (HLAGNs) and low-to-moderate luminosity AGNs (MLAGNs). These definitions refer to the radiative luminosity of AGNs, resulting from accretion onto the SMBH, which traces the overall accretion rate.[16] For efficiently accreting AGNs ($\dot{m} \gtrsim 0.01\,\dot{m}_{\rm Edd}$, where $\dot{m}$ is the mass accretion rate and $\dot{m}_{\rm Edd}$ the Eddington rate), the bulk of the radiative luminosity is emitted by the accretion disk (UV) as well as in X-rays (Lusso et al. 2011; Fanali et al. 2013). Depending on both the orientation and the optical depth of the obscuring torus, this radiation may be (partially) attenuated and re-emitted in the MIR (e.g., Ogle et al. 2006). X-ray- and MIR-based tracers of AGN activity therefore preferentially select high radiative luminosity AGNs, and hence imply higher overall accretion rates. Conversely, low accretion rates ($\dot{m} \ll 0.01\,\dot{m}_{\rm Edd}$) are associated with radiatively inefficient accretion, whereby the accretion disk is generally truncated and advection-dominated accretion takes over in the vicinity of the black hole (e.g., Heckman & Best 2014). As the timescale of such inflows is much shorter than the cooling time of the material, such ineffecient accretion produces little UV and X-ray emission. A recent study by Delvecchio et al. (2018), employing X-ray stacking on the 3 GHz-selected radio-excess AGN sample from VLA-COSMOS indeed finds that below $z \lesssim 2$ the accretion rates of such AGNs are $\dot{m} \lesssim 0.01\,\dot{m}_{\rm Edd}$, with only 16% of this sample being individually X-ray detected, implying overall inefficient accretion for the typical radio-excess AGNs. They additionally do not find any correlation between AGN X-ray and radio luminosity at a fixed redshift. Instead, the identification of AGNs with lower accretion rates and radiative luminosities, referred to here as MLAGNs, thus relies predominantly on radio-based diagnostics. These are effectively based on the fact that, for such AGNs, the multiwavelength star formation rate indicators are discrepant, as will be clarified in Section 4.2. It must be noted, however, that a hard division between high- and moderate-luminosity AGNs does not exist, and we therefore follow Delvecchio et al. (2017) by applying the tags "moderate-to-high" and "low-to-moderate" to indicate the overlap between the classes. This further serves to illustrate that there is no one-to-one relation between the class an AGN belongs to and its accretion rate. We will study the various sets of AGNs in more detail in a future paper, and instead focus on the classification in this work.

4.1. HLAGNs

In the context of this paper, we identify a source as an HLAGN if it satisfies any of the following criteria.

1. The source shows a $\geqslant 2\sigma$ excess in X-ray luminosity compared to its FIR-derived star formation rate, based on the relations from Symeonidis et al. (2014).
2. The source exhibits MIR IRAC colors that place it within the Donley et al. (2012) wedge, provided it lies at $z \leqslant 2.7$.
3. The source shows a significant AGN component in the form of a dusty torus or accretion disk, based on SED fitting.

We expand on each of these criteria in the following subsections.

4.1.1. X-Ray AGNs

A subset of HLAGNs is characterized by a high X-ray luminosity, which is thought to originate from the accretion disk around the central SMBH. In the conventional picture (Heckman & Best 2014), this disk is surrounded by a hot corona, which boosts the energy of the seed photons from the accretion process through inverse Compton scattering into the X-ray regime. The accretion disk is further thought to be obscured by a dusty torus, which—if sufficiently dense—may absorb even the hard X-rays produced by the AGN. Nevertheless, in the scenario of low-obscuration, AGN-powered X-ray emission can be orders of magnitude brighter than X-rays expected from star formation-related processes, which arise primarily from high-mass X-ray binaries (Fabbiano 2006). In order to classify our X-ray-detected sources as AGNs, we make use of the X-ray properties of SFGs derived by Symeonidis et al. (2014). They found that typical SFGs have a relation between their FIR luminosity and soft band ([0.5–2] keV) X-ray luminosity given by $\log L_{\rm FIR} = \log L_{[0.5-2]\,{\rm keV}} + 4.55$, with a 2σ scatter around this relation of 0.74 dex. We classify sources with an X-ray excess above this 2σ scatter as HLAGNs.

For the AGN classification we extract [0.5–10] keV obscuration-corrected X-ray luminosities from the Chandra COSMOS Legacy catalog presented in Marchesi et al. (2016a).[17] We scale X-ray luminosities between different bands with a power-law index of $\Gamma = 1.4$, defined such that the X-ray luminosity follows $L_{\rm X} \propto \nu^{1-\Gamma}$. In the following, we will quote X-ray luminosities in the [0.5–8] keV range. Out of the 108 cross-matches we obtained within $1''.4$ with this catalog, we identify 106 sources as X-ray AGNs (Figure 6).[18] Had we adopted a fixed X-ray luminosity threshold of $L_{\rm X} = 10^{42}$ erg s^{-1}, as is common in the literature (e.g., Wang et al. 2013; Delvecchio et al. 2017; Smolčić et al. 2017a), we would have missed an additional seven X-ray sources at low redshift that are only modestly X-ray luminous, but are nonetheless substantially offset from the relations from Symeonidis et al. (2014). Our main motivation for adopting a threshold dependent on $L_{\rm FIR}$ is, however, to avoid the misclassification of highly starbursting sources, as SFRs upwards of $\sim 300\,M_\odot$ yr^{-1} are expected to generate X-ray luminosities in excess of $L_{\rm X} \sim 10^{42}$ erg s^{-1}. We therefore regard a selection based on the comparison of FIR- and X-ray luminosities to be more robust in general. Based on the discussion in Appendix C.1, where we employ X-ray stacking, we further conclude that we are minimally affected by incompleteness issues, which may arise from the relatively shallow X-ray data, resulting in rather unconstraining upper limits on the X-ray luminosities of the typical high-redshift ($z \gtrsim 2$) source.

4.1.2. MIR AGNs

Sources that fall within the class of high-luminosity AGNs are believed to be surrounded by a warm, dusty torus, which will absorb and re-radiate emission emanating from the region

---

[16] Delvecchio et al. (2017) further argue that the HLAGN/MLAGN definitions closely resemble the widely used HERG/LERG nomenclature.

[17] In the case where an X-ray source was assigned a different redshift than given in the Chandra catalog, we re-computed its X-ray luminosity using the updated value.

[18] Upon adopting a steeper X-ray photon index of $\Gamma = 1.7$, as may be more appropriate for SFGs (e.g., Lehmer et al. 2010), we instead identify one additional X-ray AGN. However, including this additional source as an X-ray AGN has no impact on our overall conclusions.
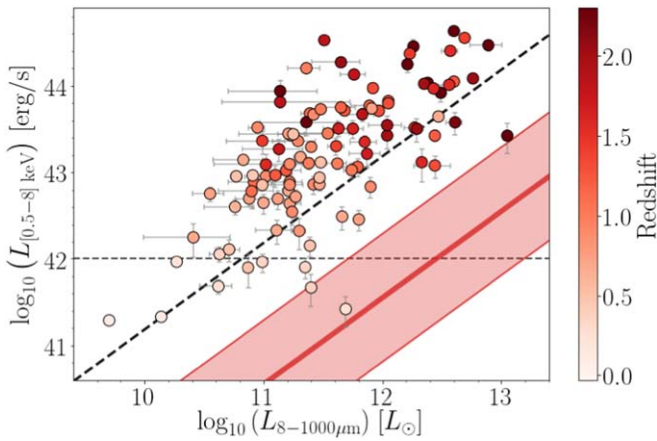
Figure 6. X-ray luminosity in the [0.5–8] keV energy band vs. total IR (8–1000 μm) luminosity for our X-ray-detected sample, color-coded by redshift. The solid diagonal line represents the X-ray vs. total IR luminosity relations for star-forming sources from Symeonidis et al. (2014), scaled to a full [0.5–8] keV using a power-law exponent $\Gamma = 1.4$. The red shaded region indicates the 2σ scatter around this relation (equivalent to 0.74 dex). All soures that fall outside of this scatter are identified as X-ray active galactic nuclei (AGNs) in this work. The diagonal black line is shown for comparison and represents the AGN luminosity threshold from Alexander et al. (2005), which is more conservative in identifying X-ray AGNs than the threshold we adopt here. The dashed horizontal line indicates a constant X-ray luminosity threshold of 10^{42} erg s^{-1}, which is a typically used threshold for identifying AGNs in the literature at low redshift. All but two of our X-ray detected sources are substantially offset from the Symeonidis et al. (2014) relations, including seven low-redshift sources with $L_X \lesssim 10^{42}$ erg s^{-1}.

around the central SMBH. This gives rise to a specific MIR-continuum signature associated to predominantly dusty and obscured AGNs. Early work in the identification of AGNs based on MIR colors was done by Lacy et al. (2004), based mostly on the Spitzer/IRAC colors of local Seyfert galaxies. Due to intrinsic reddening of high-redshift sources, these criteria are not optimized for galaxies at moderate redshift ($z \gtrsim 0.5$), and we therefore use the adapted criteria from Donley et al. (2012) to identify obscured HLAGNs. We locate sources within the Donley et al. (2012) wedge, defined through their Equations (1) and (2), and identify such sources as MIR AGNs. As the MIR colors of dusty SFGs at high redshift ($z \gtrsim 3$) closely resemble those of obscured AGNs (e.g., Stach et al. 2019), we restrict our analysis to $z \leqslant 2.7$, as Donley et al. (2012) are increasingly biased above this redshift. We note these MIR criteria are somewhat conservative in order to minimize the occurrence of false positives, and the MIR identification becomes less complete for X-ray-faint AGNs. Overall, we recover 28 AGNs based on their Spitzer/IRAC colors. While only ~60% of our sample has reliable IRAC photometry in all four channels, and hence we cannot robustly place the remaining sources within the Donley et al. (2012) wedge, this incompleteness has a negligible effect on our overall AGN identification (Appendix C.1).

4.1.3. AGN SED Fitting

HLAGNs are expected to show a composite multiwave-length SED, exhibiting signs of both star formation and AGN-related processes. A spectral decomposition will therefore detail the relative contribution of these two components, allowing AGNs to be identified as such if their emission dominates over the contribution from star formation at certain wavelengths. We use AGNFITTER (Calistro Rivera et al. 2016)

to fit the FUV to FIR SEDs of our radio-selected sample. AGNFITTER is a publicly available python-based SED-fitting algorithm implementing a Bayesian Markov chain Monte Carlo method to fit templates of star-forming and AGN components to observed multiwavelength galaxy photometry. Two such AGN components are fitted: an accretion disk, which predominantly emits at UV and optical wavelengths, and a warm, dusty torus that contributes mostly to the MIR continuum. The SED fitting further includes UV/optical emission from direct starlight, as well as dust-attenuated stellar emission in the IR.

As AGNFITTER utilizes a Monte Carlo method in its SED-fitting procedure, its output includes realistic uncertainties on any of its computed parameters, such as the integrated luminosities in the various stellar and AGN components. These uncertainties are particularly informative for galaxies with no or little FIR photometry, as in this case the long-wavelength SED is largely unconstrained. This is in contrast to SED-fitting codes that impose energy balance between the stellar and dust components, such as MAGPHYS (da Cunha et al. 2008, 2015) and SED3FIT (Berta et al. 2013), which is built upon the former and extended to include AGN templates. We opt for a Bayesian algorithm without energy balance, as it has been shown that dust and stellar emission can be spatially offset in high-redshift dusty galaxies such that imposing energy balance may be inaccurate (e.g., Hodge et al. 2016). In addition, it allows for the comparison of realistic probability distributions for the integrated luminosities of the various galaxy and AGN components, enabling us to separate the populations based on physical properties, rather than based on the goodness of fit. We compare our results with those obtained with SED3FIT by Smolčić et al. (2017a) in Appendix B.2.

Prior to the fitting, we account for uncertain photometric zero-point offsets and further potential systematic uncertainties by adding a relative error of 5% in quadrature to the original error on all photometric bands between U and MIPS 24 μm, similar to, e.g., Battisti et al. (2019). This further serves to guide the fitting process into better constraining the spectrum at FIR wavelengths, where photometric uncertainties are gener-ally large. Without such an adjustment, the fitting would be dominated by the small uncertainties on the short-wavelength photometry, and occasionally fail to accurately model the FIR component.

We then identify AGNs via a comparison of the integrated luminosities in both the torus and accretion disk components with, respectively, the stellar-heated dust continuum and the direct optical and NUV stellar light, taking into account the probability distributions of these integrated luminosities. This comparison then directly takes into account the reliability of the photometry, as large photometric uncertainties will naturally lead to a broad probability distribution in the integrated luminosities of the various components. Therefore, this procedure only includes AGNs that can reliably be identified as such, similar to, e.g., Delvecchio et al. (2014, 2017), who compare the best-fitting SEDs with and without AGN templates, and require the former to be a better fit at the 99% confidence level in order to identify it as an AGN. We slightly modify this procedure for sources without any FIR photometry, and expand on the exact criteria we employ in Appendix A. Overall, we identify 149 sources as HLAGNs based on SED fitting, with 51 (78) being identified solely through a MIR torus
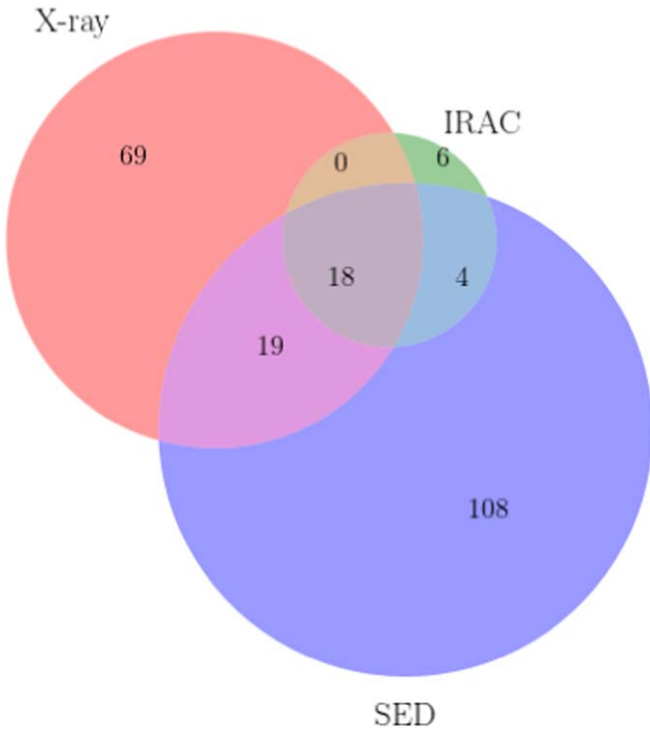
Figure 7. Venn diagram showing the overlap between the various AGNs diagnostics used for identifying moderate-to-high luminosity AGNs. The areas of circles and overlapping regions are roughly proportional to the number of sources within this category.

(accretion disk) component. A further 20 sources are classified as AGNs based on both of these features.

We show the overlap between the three different methods utilized for identifying HLAGNs in Figure 7. As expected, the subset of AGNs identified through MIR colors largely overlaps with those found through SED fitting, such that the X-ray- and SED-fitted AGNs make up the bulk of the total set of HLAGNs.

4.2. MLAGNs

Whereas AGNs selected through X-ray emission, MIR colors, and SED-fitting diagnostics preferentially identify HLAGNs, which form the subset of AGNs powered by efficient accretion, a second, low-luminosity AGN population is most readily detected through its radio properties. We assign sources to the class of MLAGNs if they are not identified as HLAGNs and satisfy one of the following criteria.

1. The source exhibits radio emission that exceeds (at the 2.5σ level) what is expected from star formation, based on the radio–FIR correlation.
2. The source exhibits red rest-frame $[NUV - r^+]$ colors, corrected for dust attenuation, typically indicating a lack of star formation.

4.2.1. Radio-excess AGNs

The FIR–radio correlation describes a tight interconnection between the dust luminosity of an SFG and its low-frequency radio luminosity. This connection arises because the same population of massive stars that heats up dust, causing it to re-radiate its energy in the FIR, produces supernovae that generate

relativistic particles emitting synchrotron radiation at radio frequencies. However, galaxies that host an AGN may have their radio emission dominated instead by the active nucleus, and will therefore be offset from the FIRC. To quantify this, we define the correlation parameter q_{TIR} as the logarithmic ratio of a galaxy's total-IR luminosity L_{TIR}, measured between (rest-frame) 8–1000 μm, and its monochromatic radio luminosity at rest-frame 1.4 GHz, $L_{1.4GHz}$ (following e.g., Bell 2003; Thomson et al. 2014; Magnelli et al. 2015; Delhaize et al. 2017; Calistro Rivera et al. 2017; Algera et al. 2020:

$$q_{TIR} = \log_{10}\left(\frac{L_{TIR}}{3.75 \times 10^{12} \, W}\right) - \log_{10}\left(\frac{L_{1.4GHz}}{W/Hz}\right). \quad (1)$$

The factor 3.75×10^{12} is the central frequency of the total-IR continuum (8–1000 μm) in hertz and serves as the normalization. There is now a growing consensus that q_{TIR} is a function of redshift (Magnelli et al. 2015; Delhaize et al. 2017; Calistro Rivera et al. 2017), for reasons that are still rather poorly understood. Nevertheless, we utilize a redshift-dependent threshold in terms of q_{TIR} to identify galaxies with radio excess based on what is expected from the FIRC. We show the distribution of q_{TIR} as a function of redshift for our sample of radio-detected sources in Figure 8, with the FIR luminosities obtained from MAGPHYS.[19] Rest-frame 1.4 GHz luminosities are determined using the measured spectral index for the required K-corrections if available. When only a single radio flux is available, a spectral index of $\alpha = -0.7$ is assumed instead. The luminosities are then calculated through

$$L_{1.4 \, GHz} = \frac{4\pi D_L^2}{(1+z)^{1+\alpha}}\left(\frac{1.4 \, GHz}{3 \, GHz}\right)^\alpha S_{3 \, GHz}. \quad (2)$$

Here D_L is the luminosity distance at redshift z and S_{3GHz} is the observed flux density at 3 GHz. The uncertainty on the luminosity is computed by propagating the error on the flux density and—if the source is detected at $\geqslant 2$ radio frequencies —the spectral index, i.e., the source redshift is taken to be fixed. The error on q_{TIR} further includes the propagated uncertainty on the FIR luminosity returned by MAGPHYS.

To quantify radio excess, we adopt the redshift-dependent q_{TIR} determined for SFGs by Delhaize et al. (2017). They determine a best-fitting trend of $q_{TIR,D17} = 2.86 \times (1+z)^{-0.19}$, with an intrinsic scatter around the correlation of $\sigma_q = 0.31$. Their best fit takes into account the sample selection at both radio and FIR wavelengths through a two-sided survival analysis. As such, Delhaize et al. (2017) take into account that radio-faint SFGs— those with a value of q_{TIR} above the typical correlation—are preferentially missed in radio-selected samples, in particular at high redshift due to the negative radio K-correction. We therefore adopt their median redshift-dependent value for q_{TIR} appropriate for SFGs, minus $2.5 \times \sigma_q$, as our threshold for identifying radio-excess AGNs. That is, our threshold identifies a radio source as an AGN if it lies below the median FIR–radio correlation for SFGs at more than 2.5σ significance. Adopting the Delhaize et al. (2017) results appropriate for star-forming sources, and taking into account the intrinsic scatter about the correlation, minimizes the effect of selection biases and incompleteness in our

---

[19] Using energy balance is appropriate here, as it allows us to associate FIR luminosities even to sources without good photometric coverage at these wavelengths. Adopting a code without energy balance would instead result in artificially low FIR luminosities and as such bias sources toward being radio-excess AGNs. We study the effect of incompleteness in our FIR photometry and the assumption of energy balance in Appendix C.3.
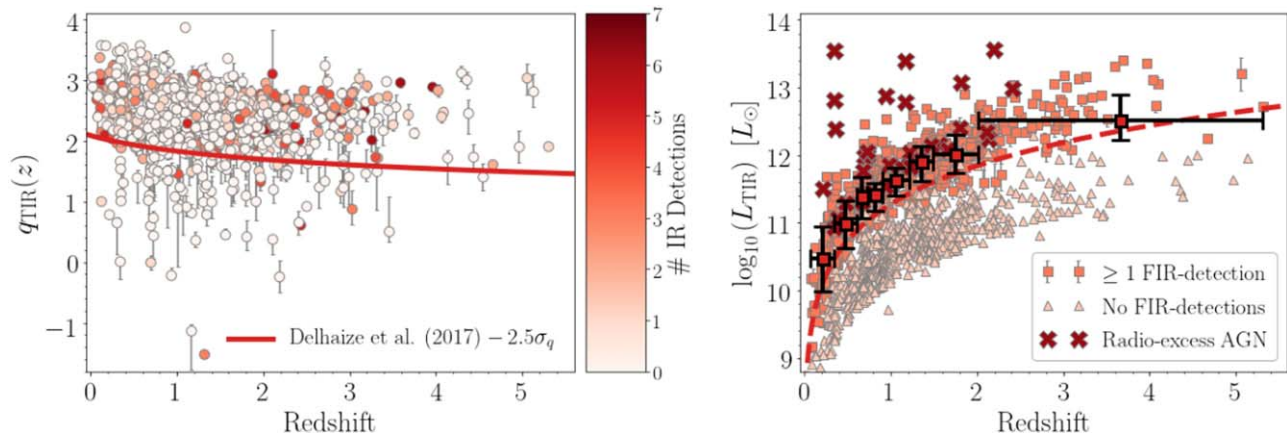
Figure 8. Left: FIRC parameter q_{TIR} as a function of redshift for the S-band-detected sample of 1437 sources. Sources are colored by the number of detections in the FIR (including Herschel/PACS & SPIRE, as well as SCUBA2, JCMT/AzTEC and IRAM/MAMBO). The solid red line constitutes the redshift-dependent trend of $q_{TIR}(z)$ as determined for star-forming galaxies by Delhaize et al. (2017), minus $2.5\times$ the intrinsic scatter of 0.31 dex, which constitutes our redshift-dependent threshold for radio excess. The 110 sources below this line are then identified as radio-excess AGNs. Right: empirically determined sensitivity curve of our Herschel observations, showing the redshift-dependent FIR luminosity of our sample, either computed through MAGPHYS (orange squares) or via the FIRC, assuming a conservative $q_{TIR}(z)$ (triangles and crosses). The sources with good FIR photometry are binned in redshift (large red squares). The empirical detection limit is then determined via a fit through the lower errorbars, and is shown via the dashed red line. Radio sources without FIR photometry that fall above this detection threshold are identified as AGNs. This diagnostic allows us to quantify radio excess for sources lacking FIR photometry, and identifies a total of 62 radio-excess AGNs.

multiwavelength photometry on the AGN classification. Overall, we identify a total of 110 radio-excess AGNs via this method.

However, this number of AGNs may be affected by the fact that we do not have radio spectral indices for ~80% of our sample (see also Gim et al. 2019). To test the effect of assuming a fixed spectral index for these sources, we re-compute the number of radio-excess AGNs by assigning every source detected solely at 3 GHz a spectral index drawn from a normal distribution, centered around $\alpha = -0.70$, with a standard deviation of $\sigma = 0.30$ (similar to the intrinsic scatter in the radio spectral indices found by Smolčić et al. 2017b in the 3 GHz VLA-COSMOS survey). We then run this procedure 200 times and find the mean number of radio-excess AGNs to be $N_{AGN} = 120$, with a standard deviation of five sources. It is unsurprising that the typical number of radio-excess AGNs increases slightly when a distribution of spectral indices is assumed, as the number of star-forming sources is ~12× greater than the number of AGNs. As such, it is more likely for a galaxy initially classified as star-forming, when $\alpha = -0.70$ is assumed, to scatter below the AGN threshold than for an AGN to scatter into the star-forming regime. However, the minor increase of ~10 radio-excess AGNs we find when adopting such a distribution of spectral indices does not change our conclusions that radio-excess AGNs make up only a small fraction of the μJy radio population.

While the radio-excess criterion described above constitutes a clear way to identify AGNs for sources with well-constrained FIR luminosities, only 50% of our sample is detected in the FIR at $\geqslant 3\sigma$. To improve the completeness of our sample of radio-excess AGNs, we utilize the distribution of FIR luminosities for the sources with at least one 3σ detection at any of the FIR wavelengths, and compare this with expected FIR luminosities of the Herschel-undetected sources. For each of the latter, we compute the FIR luminosity through the FIR radio correlation, assuming the previously mentioned relation for star-forming sources by Delhaize et al. (2017), again adopting their normalization minus 2.5 times their intrinsic scatter about the correlation. Effectively, we thus compute a conservative FIR luminosity of our Herschel-undetected

sample at the 2.5σ level, assuming all radio emission is powered by star formation. We compare the FIR luminosities derived in this manner with the distribution of luminosities for our sources with well-constrained FIR photometry in the lower panel of Figure 8. We fit a power law through the 16th percentile of the distribution of $\log L_{TIR}$ in each redshift bin for sources with FIR detections, and thus empirically determine the detection threshold of sources with a given dust luminosity. Sources that fall above the median FIR luminosity determined for the sample with Herschel detections, yet are themselves undetected in the FIR, are also identified as "inverse" radio-excess AGNs. This constitutes a total of 62 sources, shown via the red crosses in the lower panel of Figure 8. A substantial number of these, 46 in total, were previously identified through the normal radio-excess criterion. This substantiates that the energy balance MAGPHYS applies to determine FIR luminosities is typically a good assumption for these sources (see also Dudzevičiūtė et al. 2019).

4.2.2. Red, Quiescent AGNs

The class of MLAGNs can be further extended by including red galaxies, as—once obscuration by dust has been corrected for—such colors indicate a cessation of star formation. We quantify this through the rest-frame [NUV $- r^+$] colors of our sources, which we model through integrating the best-fitting, unattenuated MAGPHYS SED over the GALEX NUV 2300 Å and Subaru r^+-band filters. We follow Ilbert et al. (2010) and define sources with [NUV $- r^+$] > 3.5 as quiescent galaxies, but limit this analysis to sources at $z \leqslant 2$, as we do not accurately measure the rest-frame UV emission for sources at high redshift. As radio emission traces star formation, and quiescent sources by definition lack significant star formation activity, we identify radio sources with red rest-frame colors as MLAGNs. We find 50 such sources and note that 56% of these were already previously identified through the (inverse) radio-excess criterion, similar to what is found by Smolčić et al. (2017a). We verified that there is no trend between the NUV/optical colors and redshift, as such a trend may be indicative of an inaccurate extrapolation of a galaxy's SED by MAGPHYS to
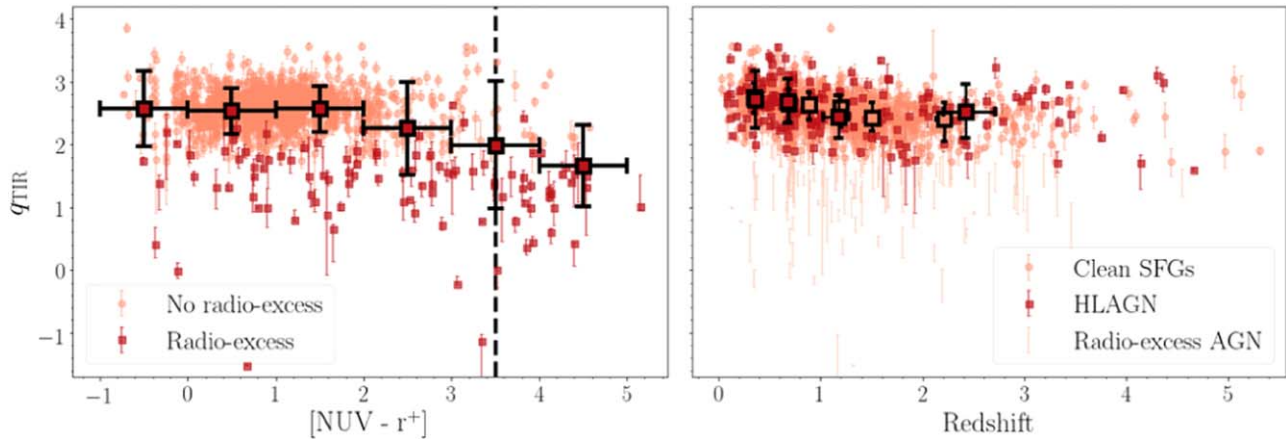
Figure 9. Left: radio–FIR correlation parameter q_{TIR} (Equation (1)) as a function of rest-frame $[NUV - r^+]$ colors for all sources at $z \leqslant 2$. The sample is divided into radio-excess AGNs (red squares) and sources primarily powered by star formation (orange circles). The dashed black line represents the threshold for identifying sources as red, quiescent low-to-moderate luminosity AGNs (MLAGNs). The large squares indicate the mean values per bin, and the bin width is shown through the horizontal errorbars. The vertical errorbars represent the 1σ standard deviation per bin. On average, redder sources show stronger radio emission for a given FIR luminosity, which is consistent with a substantial fraction of these sources being radio-excess AGNs. Right: q_{TIR} as a function of redshift for the clean star-forming sample (orange circles) and moderate-to-high luminosity AGNs (HLAGNs) without radio excess (red squares). The large orange and red squares represent the binned values for the clean SFGs and HLAGNs, respectively. No significant difference is present between the distribution of q_{TIR} for the two samples, indicating that AGNs without radio excess have radio properties consistent with arising solely from star formation.

rest-frame NUV wavelengths, resulting in the misclassification of red, quiescent MLAGNs.

We show the relation between the FIRC parameter q_{TIR} and rest-frame $[NUV - r^+]$ colors for sources at $z \leqslant 2$ in Figure 9 (left panel). On average, redder sources exhibit lower values of q_{TIR}, and hence constitute a higher fraction of radio-excess AGNs. A total of 22 sources—nearly half of the sources with $[NUV - r^+] > 3.5$—nevertheless show radio emission consistent with originating solely from star formation, which may indicate that these sources are in fact low-level (dusty) SFGs without substantial AGN activity in the radio (though five are identified as HLAGNs instead). The determination of whether these objects have an AGN contribution is, however, complicated by the fact that only four of them have detections in the FIR, such that the modeled dust continuum emission of these objects is determined solely through the energy balance that MAGPHYS imposes. This adds an additional layer of uncertainty to their distribution of q_{TIR}. For this reason, as well as the general observation that red, early-type galaxies are typically linked with radio-bright AGN hosts (e.g., Rovilos & Georgantopoulos 2007; Smolčić 2009; Cardamone et al. 2010; Delvecchio et al. 2017), we identify these sources as quiescent (ML)AGNs nevertheless. Ultimately, the sample of red rest-frame optical/NUV sources not identified as AGNs through other means only concerns a small number of sources (only 1.5% of our radio sample, or 6.6% of all AGNs), and their inclusion has a negligible effect on the number counts we derive in Section 5, with all results being consistent within the uncertainties if we include these sources in the clean SFG sample instead. In fact, omitting these sources from the sample of MLAGNs further strengthens our conclusions that AGNs make up only a small fraction of the μJy radio population.

4.3. AGN Identification Summary

The results of the AGN identification process are listed in Table 1, and are additionally summarized as a Venn diagram in Figure 10. We find a total of 334 AGNs in our sample ($23.2 \pm 1.3\%$, where the error represents the 1σ Poissonian uncertainty), using the five different diagnostics detailed in the previous sections. Combined, our AGN sample contains 224

Table 1
Summary of the AGN Identification

Method	HLAGNs	Fraction	MLAGNs[a]	Fraction
X-ray	106	7.4%	⋯	⋯
IRAC	28	1.9%	⋯	⋯
SED fitting	149	10.4%	⋯	⋯
•Torus	71	4.9%	⋯	⋯
•Disk	98	6.8%	⋯	⋯
Radio excess	25	1.7%	85	5.9%
Inverse radio excess	19	1.3%	43	3.0%
$[NUV - r^+]$	5	0.3%	45	3.1%
Total[b]	224	15.6%	110	7.7%

Notes.
[a] MLAGNs are, by definition, not identified through the X-ray, IRAC or SED-fitting criteria.
[b] The total does not equal the sum of all rows, as a single AGN may be identified through multiple diagnostics.

HLAGNs and 110 MLAGNs. Overall, $\sim$64% of the sample was identified using just a single AGN tracer, whereas the remaining AGNs were identified as such with up to four diagnostics. This exemplifies the importance of using various tracers of AGN activity, as the different diagnostics trace intrinsically different populations.

Less than half of the AGNs we identify show an excess in radio emission with respect to the radio–FIR correlation, as is shown in the Venn diagram in Figure 10 by the red circle. We overplot all AGNs without radio excess on the radio–FIR correlation in Figure 9 (right panel), which shows that these sources have radio emission that is fully consistent with originating from star formation. Therefore, only $8.8 \pm 0.8\%$ of our radio sample has radio emission that is clearly not from a star-forming origin.

5. Composition of the Ultra-faint Radio Population

5.1. The Ultra-faint Radio Population

In Figure 11, we show both the fractional and cumulative contribution of the different radio populations as a function of
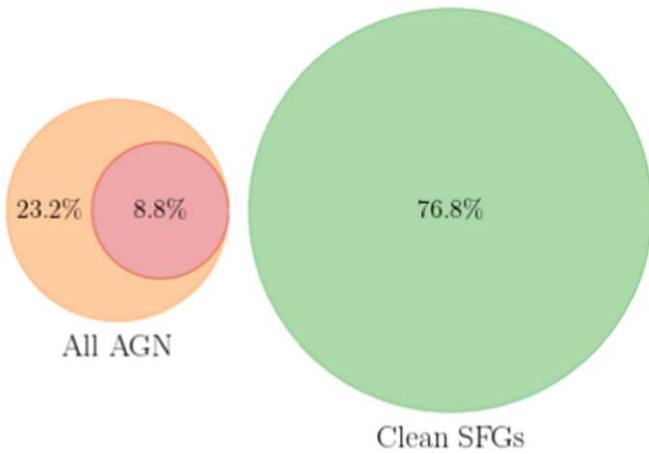
Figure 10. Visual representation of the different radio populations examined in this paper. The subset of all AGNs is shown in orange, and combines both MLAGNs and HLAGNs. AGNs exhibiting a radio excess are shown in red, and form less than half of the total AGN population. Clean star-forming sources are shown in green, and form the remainder of the radio sample.

3 GHz flux density. We restrict our analysis to sources with flux densities $<100\,\mu$Jy, which constitute $\sim$98% of our sample, because of poor statistics at our bright end. At relatively high flux densities between 50 and $100\,\mu$Jy, the radio population remains fairly equally split among the combination of MLAGNs and HLAGNs and clean SFGs, though our modest sample size at these fluxes results in significant uncertainties. Nevertheless, the class of MLAGNs dominates the population of AGNs at $S_\nu \gtrsim 50$–$100\,\mu$Jy, which is unsurprising as the bulk of this population is made up of sources that show radio excess, and are therefore radio-bright by definition. At flux densities $\lesssim 30\,\mu$Jy, which constitutes 86% percent of our sample, we observe a clear increase in star-forming sources, reaching a fractional contribution of $\gtrsim$80% in the lowest flux density bins. Cumulatively, our sample reaches 50% star-forming sources at flux densities $\sim 10\,\mu$Jy, and overall is made up of $\sim$75% by sources with no hints of AGN activity. At our detection limit of $\sim 2.7\,\mu$Jy, approximately 85% of the sample is made up of clean SFGs.

Instead of adopting the MLAGN and HLAGN terminology, which includes sources with signs of AGN activity across their full SED, we consider in Figure 12 the fractional contribution of sources with and without radio excess. The latter class includes galaxies that exhibit AGN-like activity in their X-ray to MIR SEDs, but show no sign of AGN activity at radio wavelengths. While at fluxes above $\gtrsim 100\,\mu$Jy sources with radio excess dominate the population, their fractional contribution declines steeply toward lower flux densities, and below $\lesssim 20\,\mu$Jy the contribution of galaxies without any radio excess is $\gtrsim$95%. If we adopt the definition that—despite any other AGN signatures—galaxies without any radio excess are star-forming, this implies that the fraction of SFGs is nearly unity below $\lesssim 20\,\mu$Jy. Overall, the fraction of sources without AGN signatures in the radio in the COSMOS-XS survey equals 91.2 ± 0.8%. We verify in Appendix C.4 that there is no dependence of any of the AGN diagnostics as a function of flux density, indicating that the increased fractional contribution of star-forming sources with decreasing flux density is robust.

In Figure 13, we show the distribution of the sample with redshift, and the fractional contribution of each population per redshift bin. The median redshift for all populations is

approximately $z \sim 1$, illustrating the well-known result that the redshift distribution of radio sources is nearly independent of flux density (Condon 1989). In addition, the overall fraction of the various source populations remains fairly constant with redshift. This likely indicates that there are no obvious biases in the AGN selection as a function of redshift, which we investigate further for each of the AGN diagnostics individually in Appendix C.4.

5.2. Euclidean-normalized Number Counts

In this section, we translate our observed sample, which may be parameterized as N_i radio-detected sources within the ith flux density bin $S_{\nu,i}$, into the completeness-corrected Euclidean-normalized number counts. The completeness corrections are required to reconstruct the intrinsic number density of radio sources from our observed sample. Our modus operandi has been to cross-match the 3 GHz-detected radio population with various multiwavelength catalogs, and to use this information to classify this radio sample into AGNs and star-forming sources. The main source of incompleteness is the primary beam attenuation, decreasing our sensitivity to faint radio sources toward the edge of the pointing. We additionally correct for spurious detections in the original 3 GHz map, as well as our incompleteness in assigning multiwavelength counterparts to real radio sources. The magnitude of the former two completeness corrections are detailed in Paper I, and the incompleteness in counterpart association was determined in Section 3.3 (see also Figure 5).

In the following, we will assume that the completeness corrections we have derived apply uniformly to the various radio populations—that is, these corrections are a function of observed flux density only, and not of any additional source properties. The full completeness correction C_i applied to the ith flux density bin is then given by

$$C_i(S_{\nu_i}) = \frac{1 - f_{\mathrm{spurious}}(S_{\nu_i})}{f_{\mathrm{flux}}(S_{\nu_i}) \times f_{\mathrm{ctrpt}}(S_{\nu_i})}. \qquad (3)$$

Here f_{spurious} is the fraction of spurious sources expected in the given flux density bin, f_{flux} is the fractional flux density completeness of our sample, taking into account our declining sensitivity to sources away from the primary beam center, and f_{ctrpt} is the fraction of sources in the given flux density bin for which we have obtained reliable non-radio counterparts. The Euclidean source counts in the ith bin are then computed via

$$S_{\nu,i}^{5/2} n(S_{\nu,i}) = \frac{C_i N_i}{\Delta S_{\nu,i} \Omega} S_{\nu,i}^{5/2}, \qquad (4)$$

where $\Omega \sim 350\,\mathrm{arcmin}^2$ is the field of view of the S-band survey area, out to 20% of the maximum primary beam sensitivity, and $\Delta S_{\nu,i}$ is the width of the ith flux density bin. The normalization with $S_\nu^{5/2}$, applied to the center of the bin, has historically been used, and translates into a flat slope of the number counts with flux density for a fully Euclidean universe. The completeness-corrected Euclidean source counts for the different radio populations are shown in Figure 14 and tabulated in Table 2. The total uncertainties on our measurements combine uncertainties on the counting statistics with the propagated errors on the various completeness corrections. The effects of cosmic variance are not included in the uncertainties,
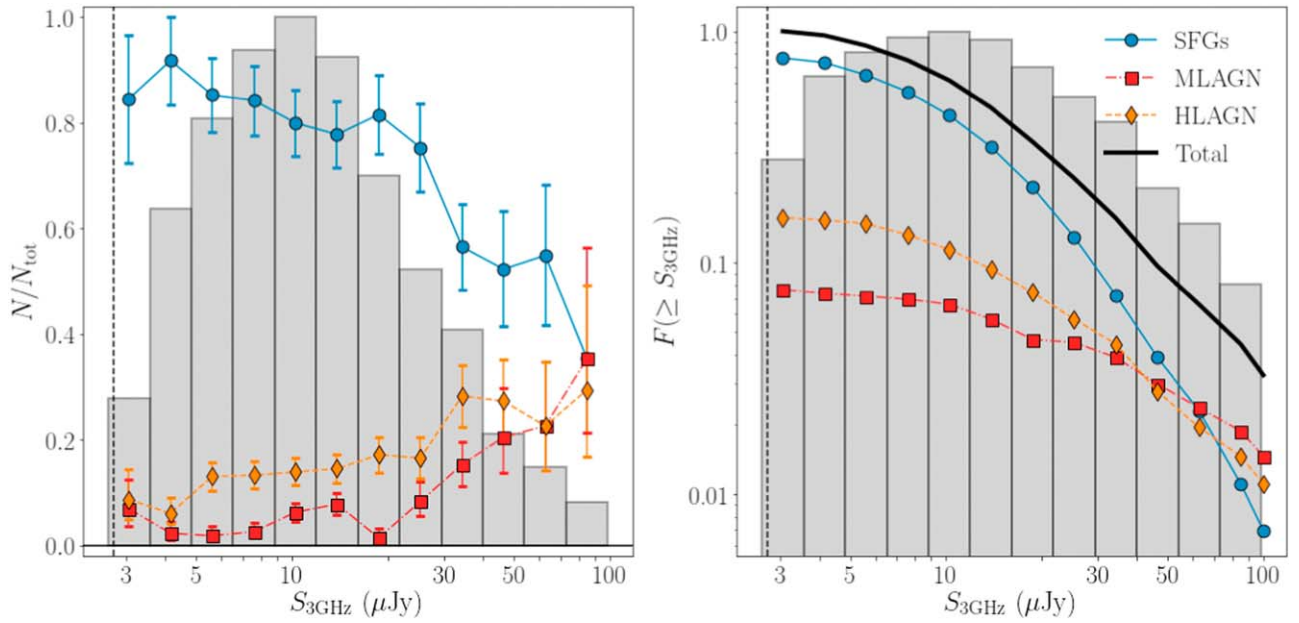
Figure 11. Left: fraction of sources of a given type (SFGs, MLAGNs, and HLAGNs) as a function of flux density. Errorbars represent the Poissonian uncertainties for >10 sources per bin, or the confidence limits from Gehrels (1986) otherwise. In both panels, the vertical dashed gray line indicates the 5σ detection limit of the COSMOS-XS survey, and the gray histogram shows the normalized number of sources in a given logarithmic 3 GHz flux density bin. Right: cumulative fraction of sources as a function of flux density, defined to increase toward lower flux densities. Overall, ~75% of sources make up the sample of clean, SFGs, while below ~30 μJy, the fraction of such SFGs increases to $\gtrsim$80%.
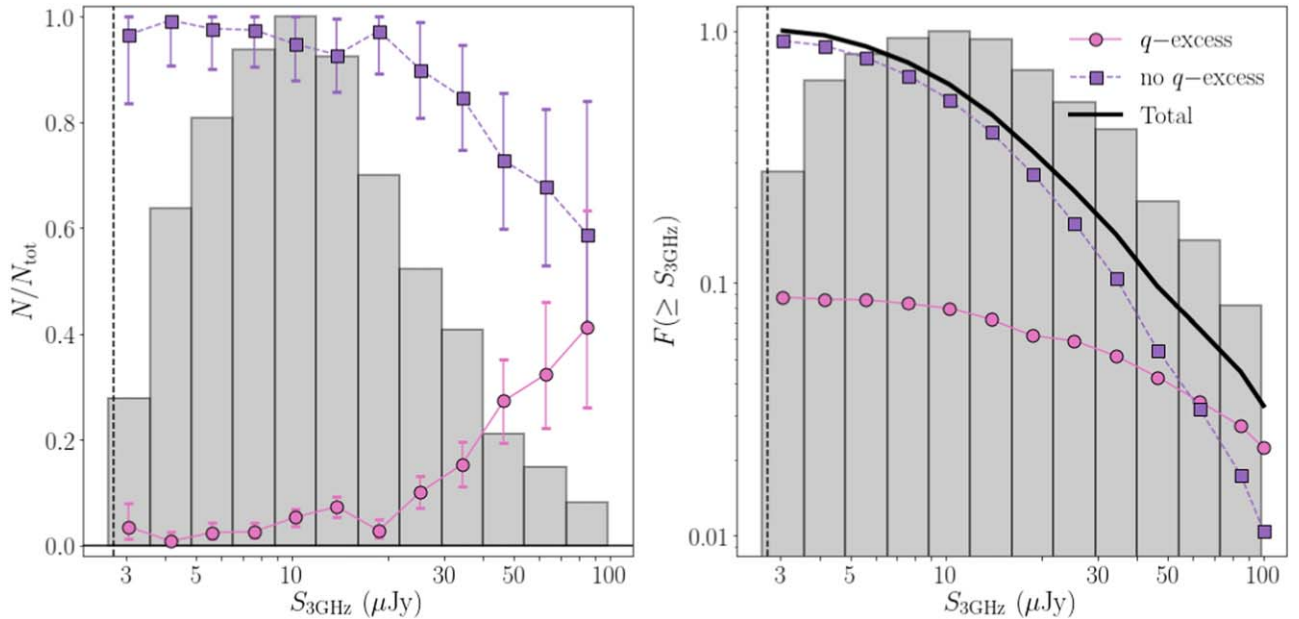

Figure 12. Left: fraction of sources with (without) radio excess in pink (purple) as a function of 3 GHz flux density. The data points are shown with Poissonian uncertainties when there are >10 sources per bin. The confidence limits from Gehrels (1986) are adopted otherwise. In both panels, the gray histogram shows the normalized number of sources in a given logarithmic flux density bin and the vertical dashed gray line indicates the 5σ detection limit of the COSMOS-XS survey. Right: cumulative fraction of sources with and without radio excess vs. flux density, defined to increase toward lower flux densities. The fraction of sources powered by star formation reaches near unity below ~20 μJy. Cumulatively, ~90% of our sample shows radio emission that is star formation powered.

but we quantify its contribution in Section 5.2.1 and overplot the results as shaded regions in the figure.

We compare our results with the number counts from Smolčić et al. (2017a), who cover a larger area to shallower depths. Focusing first on the sample of MLAGNs, HLAGNs, and clean SFGs (upper panel of Figure 14), we observe a good match between the number counts of the two types of AGNs

between our data and the Smolčić et al. (2017a) sample at the flux densities we have in common ($25 \lesssim S_\nu \lesssim 100\,\mu$Jy), whereas we find a slight increase in clean SFGs at the fainter flux densities. This may be explained by cosmic variance (Section 5.2.1), or by uncertainties in the completeness corrections at the faint end of the VLA-COSMOS survey ($\lesssim$50 μJy).
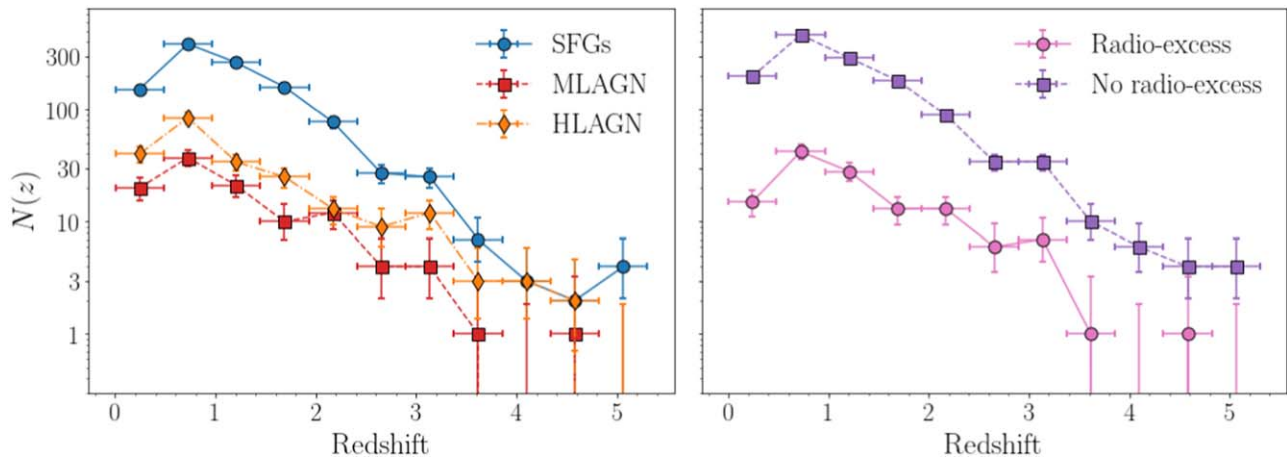
15

Figure 13. Left: distribution of SFGs, MLAGNs, and HLAGNs with redshift. Right: distribution of radio-excess AGNs and sources without radio excess with redshift. Errorbars represent the uncertainties on the counting statistics in each bin. The median redshift of the sample is $z \sim 1.0$, and the fractional contribution of the SFGs and various types of AGN remains roughly constant as a function of redshift.

In the lower panel of Figure 14, we compare our number counts for sources with and without radio excess with the VLA-COSMOS sample. We find an overall agreement, although our fraction of sources without radio excess at flux densities $\lesssim 50\,\mu$Jy is slightly larger than what is found by Smolčić et al. (2017a), similar to our increase in the counts for clean SFGs. The number counts of radio-excess AGNs are further in good agreement at the flux densities the two surveys have in common. Combined, this is fully consistent with our results from Paper I, where we find a slight increase in the overall radio number counts compared to the 3 GHz VLA-COSMOS sample. Similar to the above, these differences may be explained by cosmic variance or uncertainties in the completeness corrections. Overall, our 3 GHz source counts are broadly consistent with the VLA-COSMOS data, and as ultimately the population of sources with and without radio excess are used to determine cosmic SFR densities, this agreement is encouraging.

We additionally compare our results to recent simulations from Bonaldi et al. (2019), who simulate the radio continuum emission from radio-excess AGNs and SFGs. They model the star-forming population by converting galaxy luminosity functions from UV, Lyα, and IR observations from Cai et al. (2013, 2014)—corrected for dust extinction where necessary— into the star formation history of the universe (SFHU). As star formation is directly coupled to radio emission, this SFHU is sampled and converted into a radio luminosity function using the local calibration from Murphy et al. (2012), slightly adapted to compensate for the over-prediction of the faint end of the local radio luminosity functions (see the discussion in Mancuso et al. 2015). For the radio-excess AGNs, Bonaldi et al. (2019) start with radio luminosity functions for three different types of AGNs (steep-spectrum sources, flat-spectrum radio quasars, and BL Lac objects), which they evolve through cosmic time using evolutionary parameters motivated by the literature. Additionally, different spectral indices are assumed for the three sets of AGNs, including a Gaussian distribution around the mean spectral index per class of AGN. Their resulting definitions of the AGNs and star-forming populations are largely consistent with our definition of sources with and without radio excess, respectively, and hence we show the source counts determined from their simulated catalog in the lower panel of Figure 14. For this, we use the 25 deg^2

simulations at 3 GHz that probe down to flux densities of 10 nJy. Our number counts for star formation powered sources are in excellent agreement with these simulations across the full range of flux densities we cover, including at the faint end, where we are probing a fully new parameter space. Our number counts for radio-excess AGNs are additionally consistent with the predictions from Bonaldi et al. (2019), although the last two bins for radio-excess AGNs ($S_{3\,\mathrm{GHz}} \lesssim 10\,\mu$Jy) lie slightly below the expected value from the simulations. However, our number counts are still consistent with the predictions within the 1σ scatter due to cosmic variance, which we quantify in the following section.

5.2.1. Cosmic Variance

To quantify how cosmic variance may influence our observed number counts, we make use of the $5 \times 5\,\mathrm{deg}^2$ simulations by Bonaldi et al. (2019), which model the populations of SFGs and radio-excess galaxies. We draw radio sources from non-overlapping circular regions from their full simulated cosmic volume, whereby we take into account that faint sources can only be recovered in the central regions of the primary beam. For example, a source with a flux density of $10 \times \sigma_{\mathrm{rms}}$ can be detected at 5σ significance out to the half-power point of the primary beam. Such sources are therefore drawn from a circular region with a diameter equal to the FWHM of the S-band primary beam. The brightest sources can be detected in the full field of view of our observations, which equals $\Omega \simeq 350\,\mathrm{arcmin}^2$ out to 20% primary beam sensitivity. As such, our cosmic variance calculation takes into account that faint sources, while more numerous than brighter ones, can only be detected within a smaller region in our pointing.

For each of the resulting 225 regions, we compute the Euclidean number counts and the corresponding 1σ and 2σ confidence intervals, which are shown as the shaded regions in Figure 14. These confidence intervals hence reflect two effects: at low flux densities the effective field of view of our observations is small, such that the uncertainties resulting from cosmic variance are large. At high flux densities, the effect of cosmic variance is similarly large as bright sources are relatively rare, and this outweighs the increased effective field of view.
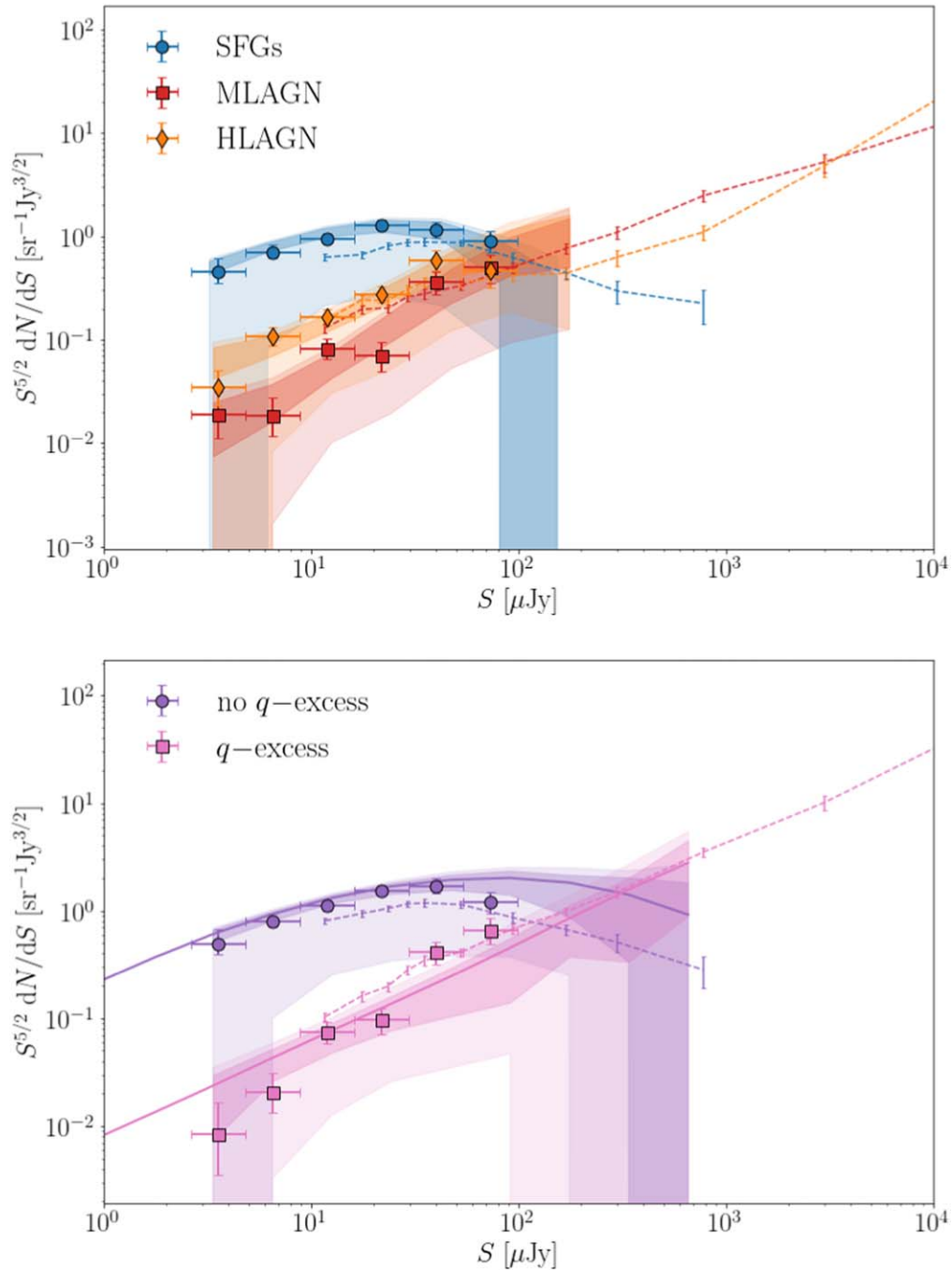
Figure 14. Euclidean-normalized source counts for the faint radio population. Upper: source counts for the HLAGNs (orange), MLAGNs (red) and clean SFGs (blue). Our data are represented by the colored markers, and we compare with the recent 3 GHz VLA-COSMOS observations (Smolčić et al. 2017b, 2017a), represented by the dashed lines. The expected scatter as a result of cosmic variance is shown by the shaded regions (dark: 1σ, light: 2σ). Lower: source counts for the radio sources with and without radio excess (pink and purple, respectively). The source counts from the Bonaldi et al. (2019) simulations are shown by the solid lines. We omit these from the upper plot, as their simulations do not model the MLAGN and HLAGN populations separately.

As the simulations by Bonaldi et al. (2019) do not explicitly model the population of AGNs without radio excess, we combine our observed number counts with their results to obtain an estimate of the effect of cosmic variance for this sub-population of AGNs. We extrapolate our measured number counts for clean SFGs, HLAGNs, and MLAGNs to fainter flux densities by fitting a quadratic function in log-space to the measured number counts of SFGs, and a linear function to the MLAGNs and HLAGNs. We then draw sources from the Bonaldi et al. (2019) simulations based on the expected ratios

of these three radio populations, and repeat the cosmic variance calculation as for the star-forming and radio-excess samples. We caution that our extrapolations are not based on any physical model, but we judge the cosmic variance results to be robust in the range of flux densities we probe ($\sim$2–100 μJy) based on the good correspondence between the fits, our data, and the shape of the number counts as obtained from the Bonaldi et al. (2019) simulations. The typical 1σ uncertainties as a result of cosmic variance are $\sim$0.1 dex for SFGs, and $\sim$0.3 dex for AGNs—substantial compared to the formal

Table 2
Euclidean-normalized Source Counts for the Various Radio Populations

Flux (μJy)			Completeness				SFGs	MLAGNs	HLAGNs	No $q-$excess	$q-$ excess
S_{low}	S_{high}	S_{center}	f_{flux}[a]	f_{ctrpt}[b]	$f_{spurious}$[c]	C_{full}[d]	Counts[e]	Counts	Counts	Counts	Counts
2.64	4.82	3.56	0.162	0.909	0.010	$7.11^{+2.37}_{-1.45}$	$0.45^{+0.16}_{-0.10}$	$0.02^{+0.01}_{-0.01}$	$0.03^{+0.02}_{-0.01}$	$0.50^{+0.17}_{-0.10}$	$0.01^{+0.01}_{-0.00}$
4.82	8.82	6.52	0.444	0.935	0.011	$2.45^{+0.41}_{-0.29}$	$0.70^{+0.13}_{-0.09}$	$0.02^{+0.01}_{-0.01}$	$0.11^{+0.02}_{-0.02}$	$0.80^{+0.15}_{-0.10}$	$0.02^{+0.01}_{-0.01}$
8.82	16.12	11.92	0.828	0.938	0.017	$1.28^{+0.17}_{-0.10}$	$0.94^{+0.13}_{-0.09}$	$0.08^{+0.02}_{-0.02}$	$0.17^{+0.03}_{-0.03}$	$1.11^{+0.15}_{-0.10}$	$0.07^{+0.02}_{-0.02}$
16.12	29.49	21.81	0.934	0.947	0.015	$1.13^{+0.14}_{-0.07}$	$1.29^{+0.18}_{-0.13}$	$0.07^{+0.02}_{-0.02}$	$0.28^{+0.05}_{-0.05}$	$1.54^{+0.21}_{-0.15}$	$0.10^{+0.03}_{-0.03}$
29.49	53.94	39.88	0.940	0.900	0.020	$1.17^{+0.19}_{-0.13}$	$1.16^{+0.23}_{-0.19}$	$0.36^{+0.10}_{-0.08}$	$0.59^{+0.14}_{-0.12}$	$1.70^{+0.32}_{-0.26}$	$0.41^{+0.11}_{-0.09}$
53.94	98.65	72.95	1.000	0.920	0.000	$1.09^{+0.20}_{-0.09}$	$0.89^{+0.24}_{-0.20}$	$0.50^{+0.16}_{-0.14}$	$0.46^{+0.16}_{-0.14}$	$1.20^{+0.29}_{-0.24}$	$0.66^{+0.20}_{-0.17}$

Notes.
[a] Flux completeness of the given flux density bin, including the incompleteness resulting from reduced primary beam sensitivity.
[b] Fraction of radio sources in the given flux density bin assigned a multiwavelength counterpart with a robustly measured redshift.
[c] Expected fraction of spurious sources in the given flux density bin.
[d] Overall completeness correction applied to the bin, as defined in Equation (3), with the propagated uncertainty.
[e] All the number counts are given in units of $Jy^{3/2}$ sr^{-1}.

uncertainties on the derived Euclidean number counts. The additional uncertainty due to cosmic variance must therefore be taken into account when comparing our results to different radio surveys.

5.3. Optically Dark Sources

In total, our radio sample comprises 70 sources that were not matched to a counterpart in any of the three catalogs used for cross-matching. As we expect only ~20 spurious detections, most of the unmatched sources are likely to be real. About half of these radio-only-detected sources have Super-deblended counterparts within $2\rlap{.}''0$, which is beyond our adopted cross-matching radius of $0\rlap{.}''9$. However, visual inspection of these sources at shorter wavelengths shows that the nearby Super-deblended entries are potentially associated to the same source, but exhibit spatial offsets between the infrared (K_s and IRAC 3.6 μm) and radio emission, similar to what has been observed at millimeter wavelengths (e.g., Hodge et al. 2012). Nevertheless, a substantial number of detections have S/N $\gtrsim 6$ yet have no counterpart within $\gtrsim 5''$, making it likely that these are indeed optically dark radio sources.

For an additional 53 S-band-detected sources we obtained a counterpart in the Super-deblended catalog, but no counterpart at shorter wavelengths. A subset of 23 of these sources was previously cataloged in the K_s-band-selected catalog by Muzzin et al. (2013), but was excluded from the COSMOS2015 catalog as they are located close to saturated optical stars. As this might affect their optical and NIR photometry and consequently any derived physical quantities such as their photometric redshifts, we have not further analyzed this sample. The remainder of the sources we solely identify with a counterpart from the Super-deblended catalog are, however, based on a single detection at 3 GHz in the VLA-COSMOS survey, which we verify in this work through our deeper radio observations. These 30 sources have no existing photometric detections below MIPS 24 μm, which is a potential sign of obscured AGNs at $z \gtrsim 6$ (e.g., Rujopakarn et al. 2018) or highly dusty SFGs (e.g., Simpson et al. 2014; Dudzevičiūtė et al. 2019), both of which evade detection at short wavelengths. Whereas no robust photometric redshifts for these sources exist due to the lack of optical/NIR photometry, Jin et al. (2018) have employed photometric redshift fitting based on the FIR data available for these sources. While these redshifts are rather uncertain as a result of

a degeneracy between redshift and dust temperature (Jin et al. 2018 quote a relative uncertainty of up to ~$0.1 \times (1 + z_{IR})$ on the derived values), 13 sources are placed beyond $z \geqslant 3$, including eight sources at $z \geqslant 4$. Additional data are needed to robustly ascertain whether these sources are indeed at such high redshift, but we verified that the sample of $z \geqslant 4$ sources are all roughly consistent with the radio–FIR correlation as obtained in Section 4.2, extrapolated to the FIR-derived redshifts. This indicates that—if these photometric redshifts are correct—the majority of this subsample is likely to have radio emission of a star-forming origin. This is in agreement with our results from Section 5.1, as we expect around three out of 30 sources (10%) to be radio-excess AGNs based on their 3 GHz flux densities.

We compile a sub-sample of optically dark detections, consisting of sources detected either only at 3 GHz, or cross-matched to a Super-deblended counterpart which in turn was based on a prior position from the VLA COSMOS survey. For the former set, we further ensure that there is no COSMOS2015 counterpart within $2\rlap{.}''0$ around any of the detections. This requirement ensures that the remaining sources are indeed "optically dark," as entries in the COSMOS2015 catalog are—by construction—detected at short (NIR) wavelengths. The separation of $2\rlap{.}''0$ is motivated by the fact that, beyond this distance, the distribution of separations between S-band sources and COSMOS2015 counterparts is consistent with being solely the result of chance associations. We further remove all sources with a peak S/N at 3 GHz below 6, to ensure the optically dark sample is robust against spurious detections. We additionally remove any objects flagged as due to being located near a bright radio source, leaving us with a sample of 46 sources that we inspect visually in the Subaru/ Suprime-cam $i+$, VISTA/Vircam K_s, and Spitzer/IRAC 3.6 μm bands. From this subset, we then discard radio sources that largely overlap with bright or saturated objects in the optical/NIR bands, as source blending at these wavelengths is likely to have affected the source detection. The remaining 29 optically dark sources are compiled in Table 3.

We investigate the multiwavelength properties of these sources by stacking on their radio positions in optical/NIR bands, as well as at 1.4 GHz. We do not attempt to stack in any of the Herschel bands, due to the large point-spread function at these wavelengths. All bands used for stacking, the corresponding photometry, and the limiting magnitudes of the stacks are listed in Table 4. We

Table 3
Optically Dark Sources Detected at 3 GHz

R.A.	Decl.	Flux[a]	Error[a]	S/N	S17[b]
		(μJy)	(μJy)		
$10^h00^m12.05^s$	$02°36'48.95''$	33.73	0.76	44.5	1
$10^h00^m00.40^s$	$02°36'12.40''$	32.66	1.72	28.3	1
$10^h00^m03.82^s$	$02°26'31.33''$	36.81	1.39	26.5	1
$10^h00^m38.06^s$	$02°28'06.06''$	20.27	0.83	24.5	1
$09^h59^m58.46^s$	$02°30'34.58''$	26.06	1.07	24.3	1
$10^h00^m03.28^s$	$02°34'48.64''$	20.06	0.84	24.0	1
$10^h00^m26.66^s$	$02°31'26.71''$	16.93	1.38	21.5	1
$10^h00^m54.48^s$	$02°34'35.86''$	28.08	1.90	17.7	1
$10^h00^m03.27^s$	$02°29'42.80''$	23.03	1.65	16.9	1
$09^h59^m58.79^s$	$02°34'57.61''$	15.41	1.02	15.2	1
$10^h00^m19.89^s$	$02°33'31.45''$	8.56	0.57	15.1	0
$09^h59^m59.81^s$	$02°34'54.66''$	17.69	1.55	12.0	1
$10^h00^m07.39^s$	$02°42'03.22''$	59.45	3.69	12.0	0
$10^h00^m20.11^s$	$02°39'39.99''$	12.01	1.02	11.8	1
$10^h00^m24.44^s$	$02°37'49.34''$	6.77	0.74	9.1	0
$10^h00^m29.93^s$	$02°29'18.17''$	5.86	0.65	9.1	0
$10^h00^m09.52^s$	$02°26'48.42''$	10.07	1.13	8.9	0
$10^h01^m03.04^s$	$02°29'11.96''$	29.77	3.52	8.4	1
$10^h00^m39.20^s$	$02°40'52.58''$	27.73	2.20	7.7	0
$10^h00^m25.08^s$	$02°26'07.27''$	7.62	1.00	7.6	1
$10^h00^m35.34^s$	$02°28'27.00''$	5.83	0.77	7.6	0
$10^h00^m15.11^s$	$02°38'17.58''$	6.50	0.88	7.4	0
$10^h00^m04.85^s$	$02°35'59.51''$	6.21	0.85	7.3	0
$10^h00^m08.64^s$	$02°32'50.79''$	4.50	0.68	6.6	0
$10^h00^m24.56^s$	$02°39'11.57''$	6.34	0.97	6.5	0
$10^h00^m41.82^s$	$02°25'47.13''$	8.80	1.35	6.5	1
$10^h00^m48.46^s$	$02°36'41.25''$	6.54	1.01	6.4	0
$10^h00^m11.73^s$	$02°34'25.69''$	4.27	0.66	6.4	0
$10^h00^m34.83^s$	$02°28'35.83''$	4.41	0.73	6.1	0

Notes.
[a] Peak or integrated flux density and its corresponding uncertainty, based on the analysis in Paper I.
[b] Boolean flag indicating whether the source is also present in Smolčić et al. (2017a).

perform aperture photometry on the optical/NIR stacks using a $1''.5$ diameter aperture for all bands up to Vista/Vircam K_s, and adopt a $3''.0$ diameter aperture for the Spitzer/IRAC channels owing to the larger point-spread function at these wavelengths. We compute uncertainties on the fluxes by placing the apertures on random locations in the 100×100 pixel^2 stack, away from the central region and computing the flux within these apertures. The median and standard deviation of these measured aperture fluxes are then taken to be the background and typical flux uncertainty in the stack, respectively.

We determine the photometric redshift of the stacked SED through EAZY (Brammer et al. 2008), using the standard set of templates provided with the code. The stack is assigned a best-fitting redshift of $z_{stack} = 4.65$, with a 16th–84th confidence interval of $z = (3.22–5.63)$. This interval was computed via a bootstrap analysis, whereby we sampled, with replacement, from the 29 optically dark sources, generating a total of 200 bootstrap samples. We then stacked all of these samples, and for each performed aperture photometry and photometric redshift fitting in the same way as for the original stack. The given confidence interval on the best-fitting photometric redshift as such represents the 16th–84th percentile of the bootstrapped redshift distribution. To further verify the robustness of SED-fitting our stacked sample, we additionally stack an identical number of sources that do have

optical/NIR photometry, and have a robustly measured photometric or spectroscopic redshift, to see if the average redshift is indeed correctly recovered. For this, we draw 29 sources from our radio sample at random, within three different redshift ranges, requiring only that the sources have no additional COSMOS2015 counterpart within $2''.0$ to minimize source blending. Three sets of radio sources are drawn within a redshift of $1.5 \leqslant z \leqslant 3.0$ (median redshift $z = 1.99$), $2.5 \leqslant z \leqslant 4.0$ (median redshift $z = 2.93$), and $z \geqslant 3$ (median redshift $z = 3.42$). Our radio sample does not contain sufficiently many sources to further probe higher redshift ranges. The SEDs of both the optically dark and the control samples are shown in Figure 15. The recovered photometric redshifts and bootstrapped uncertainties are $z = 1.76^{+0.28}_{-0.14}$, $z = 3.10^{+0.17}_{-0.21}$, and $z = 3.28^{+0.10}_{-0.28}$ for these three stacked samples, respectively, and hence recover the median redshift of the control sample relatively well. As the redshift probability distribution of the true optically dark sample is substantially broader than that of the control samples, it is likely that it consists of sources across a wider range of redshifts, with a typical value of $z \sim 4$–5. Nevertheless, if we make the assumption that all sources in the stack lie at the best-fitting photometric redshift, this would imply $\sim 90\%$ of our sample at $z \sim 5$ is undetected in optical and NIR photometry. As a sanity check, we compare with the IR-derived photometric redshifts for the 15 optically dark sources detected only in the Super-deblended catalog from Jin et al. (2018), corresponding to roughly half of the total optically dark sample. This subset has a mean redshift of $z_{IR} = 3.3$ (with a scatter of $\sigma_z = 1.9$), substantiating the typical high-redshift nature of these detections. We note, however, that these IR-derived redshifts are highly uncertain and therefore mostly indicative, and as such we do not use them in this work. This is further illustrated by the recent work from Jin et al. (2019) who, upon spectroscopic confirmation, find significant differences ($\Delta z \gtrsim 1$) between the true and IR-derived redshifts for a sample of four dusty galaxies in the COSMOS field. We further note that, at sub-millimeter wavelengths, the typical redshift of the optically dark population is indeed also higher than that of the typical population of sub-millimeter galaxies (e.g., Simpson et al. 2014; Dudzevičiūtė et al. 2019), similar to what we observe here in the radio.

In addition to stacking in optical/NIR data, the optically dark sources are also detected in the 1.4 GHz stack. The spectral index between the 1.4 and 3 GHz data is found to be $\alpha_3^{1.4} = -0.65^{+0.26}_{-0.22}$, consistent with a typical radio spectrum of $\alpha \sim -0.70$. Based on the observed 3 GHz peak flux density of $S_{3GHz} = 10.7 \pm 0.20 \, \mu$Jy, we expect $\lesssim 1$ radio-excess AGN to be present in our sample, based on the results in Section 5, which in turn implies radio surveys may miss a large fraction of the high-redshift star-forming population when optically dark sources are simply discarded. The consequences of this on the COSMOS-XS-derived cosmic star formation history will be addressed in a forthcoming paper (D. Van der Vlugt et al. in preparation). If we assume the radio emission from the optically dark sample to be fully powered by star formation, we expect their average SFRs to be ~ 500–$1500 M_\odot \, yr^{-1}$ when assuming the local value of q_{TIR} derived by Bell (2003), and placing the sources at a redshift between $z \sim 3.5$–6. Adopting the redshift-dependent conversion from Delhaize et al. (2017) instead results in more modest SFRs between 150–$350 M_\odot \, yr^{-1}$, though we stress that both the large spread in the redshift range of the optically dark sample, as well as the unknown value of q_{TIR} for typical SFGs above $z \gtrsim 3$, are large factors of uncertainty in the derived SFRs. The range of radio luminosities and SFRs spanned by this optically dark sample is
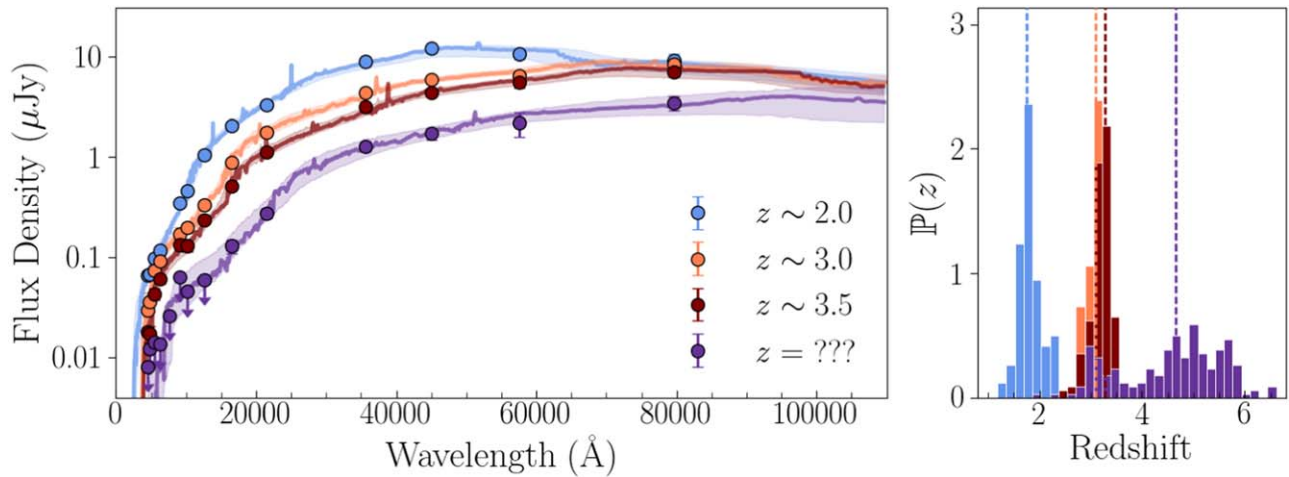
Figure 15. Left: best-fitting SED for a stack of 29 high-redshift radio detections, with a median redshift of $z = 1.99$ (blue), $z = 2.93$ (orange), and $z = 3.42$ (red), as well as the stack for our optically dark sample (purple). The shaded regions indicate the 16th–84th confidence limits around the stacked SED, as obtained from bootstrap re-sampling. Right: redshift distributions of the bootstrapped optically dark and control samples. The vertical dashed lines indicate the best-fitting redshift for the original stacks. The median redshifts of the control samples are fairly well recovered, strengthening the conclusion that the "optically dark" sample consists predominantly of high-redshift ($z \gtrsim 4$) sources.

Table 4
Data Used for Stacking of Optically Sark Sample

Band[a]	Wavelength (μm)	Depth[b]	Flux Density (μJy)[c]	Error (μJy)	S/N
Subaru/Suprime-cam B	0.4458	28.6	0.003	0.003	1.2
Subaru/Suprime-cam $g+$	0.4777	28.1	0.004	0.004	1.0
Subaru/Suprime-cam V	0.5478	28.0	0.005	0.005	1.0
Subaru/Suprime-cam $r+$	0.6288	28.0	0.013	0.005	2.8
Subaru/Suprime-cam $i+$	0.7683	27.3	0.022	0.009	2.5
Subaru/Suprime-cam $z+$	0.9037	26.4	0.063	0.020	3.2
Vista/Vircam Y	1.0214	26.7	0.046	0.015	3.0
Vista/Vircam J	1.2534	26.5	0.060	0.019	3.2
Vista/Vircam H	1.6454	26.4	0.13	0.020	6.6
Vista/Vircam K_s	2.1540	26.0	0.27	0.030	9.1
Spitzer/IRAC CH1	3.5634	24.2	1.29	0.15	8.9
Spitzer/IRAC CH2	4.5110	23.8	1.72	0.22	7.7
Spitzer/IRAC CH3	5.7593	22.7	2.20	0.60	3.7
Spitzer/IRAC CH4	7.9595	22.8	3.47	0.57	6.1
VLA/S-band	3.0 GHz (10 cm)	⋯	10.7	0.20	53.5
VLA/L-band	1.4 GHz (21 cm)	⋯	17.6	3.1	5.7

Notes.
[a] References: Subaru/Suprime-cam images from Capak et al. (2007); Vista/Vircam data are UltraVISTA DR3 (McCracken et al. 2012); Spitzer/IRAC images are from S-COSMOS (Sanders et al. 2007); the VLA L-band mosaic is from Schinnerer et al. (2007).
[b] For the optical/NIR bands, this corresponds to the 5σ limiting magnitude in $1''\!5$ apertures ($3''\!0$ for Spitzer/IRAC). For the radio data, see the error on the flux density as measure of the depth of the stack.
[c] For the two radio bands, these are the peak flux densities in μJy beam^{-1}.

shown via the shaded purple region on the COSMOS-XS sensitivity curve in the right panel of Figure 4.

We lastly investigate the average X-ray properties of the optically dark sample through an X-ray stacking analysis using C-STACK,[20] and find that the radio sample is not detected in the X-ray stack, with the count rate in both the soft ([0.5–2] keV) and hard ([2–10] keV) bands being consistent with zero at the 1σ level. This corresponds to an upper limit of $L_{[0.5–8]\ \mathrm{keV}} \lesssim 2 \times 10^{43}$ erg s^{-1}, assuming a redshift of $z = 5.0$ for the stack.

Based on this fairly shallow upper limit, we cannot place any constraints on whether the typical source in the stack exhibits X-ray emission suggestive of AGN activity. Nevertheless, the upper limit is consistent with the typical radio-derived SFR for the stack, assuming the FIRC from Delhaize et al. (2017) and the X-ray−SFR relations from Symeonidis et al. (2014), further substantiating the notion that the optically dark sample is dominated by high-redshift SFGs.

5.4. Implications for Next-generation Radio Surveys

The two largest upcoming radio telescopes are the next-generation VLA (ngVLA) and the Square Kilometer Array

---

[20] C-STACK is an online X-ray-stacking tool developed by Takamitsu Mijayi, and can be accessed via http://cstack.ucsd.edu/ or http://lambic.astrosen.unam.mx/cstack/.

(SKA). Both are expected to revolutionize the radio view on galaxy evolution owing to their greatly improved sensitivity and resolution compared to present-day radio interferometers. For example, the ngVLA will be able to reach the COSMOS-XS 3 GHz rms sensitivity of $\sim 0.53\,\mu$Jy beam^{-1} at a similar frequency within an hour of observing time (Selina et al. 2018).

Additionally, assuming a spectral index of $\alpha = -0.7$ and an observing frequency of ~ 1 GHz, the SKA will reach a similar sensitivity as our 3 GHz observations as part of the SKA Wide Survey, which is expected to cover an immense area of $\sim 10^3\,\mathrm{deg}^2$ (Prandoni & Seymour 2015). This is a factor $\sim 10^4$ larger than our surveyed area, such that a zeroth-order scaling of our number counts predicts that the SKA Wide Survey will detect upward of $\gtrsim 10^7$ faint radio sources.

Both our results for the faint radio source counts and predictions from simulations (Wilman et al. 2008; Bonaldi et al. 2019) are consistent with a continuous increase in the fraction of sources without radio excess toward lower flux densities. This is in further agreement with the recent work from Novak et al. (2018), who extrapolate the VLA-COSMOS luminosity functions to predict the relative contributions of SFGs and AGNs. Based on a simple power-law extrapolation of our radio number counts below $30\,\mu$Jy, we find that at $\sim 1\,\mu$Jy, the fractional contribution of radio-excess sources is $\lesssim 1\%$, such that the next-generation radio surveys will be highly sensitive to the faint, star-forming population. In fact, simple flux cuts on the radio-detected sample are likely to be sufficient for obtaining a highly pure sample of star-forming sources.

6. Summary and Future

We have presented a multiwavelength analysis of the faint radio population identified in the COSMOS-XS survey, which is the deepest multi-frequency radio survey to date, reaching a 5σ flux limit of $\sim 2.7\,\mu$Jy beam^{-1} within the center of the 3 GHz image (rms of $0.53\,\mu$Jy beam^{-1}). This image, which covers a total area of ~ 350 arcmin^2, is a factor ~ 5 deeper than the previously deepest radio data over COSMOS at this frequency, and enables the direct detection of the typical star-forming population (SFR $\lesssim 100\,M_\odot$ yr^{-1}) out to $z \lesssim 3$ (Figures 1 and 4). To characterize the observed faint radio population, we associated multiwavelength counterparts to our radio sample by drawing from the several photometric catalogs available over the well-studied COSMOS field. In total, we associated such counterparts to 96.6% of our radio sources. We restricted ourselves to the cross-matches with robustly determined photometric (59%) or spectroscopic (41%) redshifts available, accounting for a total of 1437 sources (93.3% of the total radio sample). The median redshift of the population is $\langle z \rangle = 0.97 \pm 0.03$, and the sample further includes 51 high-redshift sources with $3.0 \leqslant z \leqslant 5.3$.

We separate this faint radio population into star-forming sources and AGNs, where the latter are again divided into different subclasses, either based on their radiative luminosities or on an excess in radio emission based on what is predicted from star formation-related emission alone. Sources with moderate-to-high radiative luminosities (HLAGNs) exhibit AGN-related emission throughout their multiwavelength SEDs, and are identified through strong X-ray emission, AGN-specific MIR colors or based on clear AGN-like components from SED-fitting. On the other hand, low-to-moderate radiative luminosity sources (MLAGNs) are identified through an

observed lack of star formation, based on red rest-frame UV/optical colors, or an excess in radio emission compared to their full 8–1000 μm FIR luminosity.

The full catalog of sources detected at 3 GHz, including their multiwavelength source identifications, and the results from our AGN identification, is available with this paper in standard FITS format. We show a sample of the catalog in Table D1, and elaborate on its contents in Appendix D.

Overall, these multiwavelength diagnostics identify $23.2 \pm 1.3\%$ of the COSMOS-XS sample as AGNs, with only $8.8 \pm 0.8\%$ of the radio-detected sources exhibiting AGN-related emission at the observed radio frequencies. Radio emission from the remainder of the sample therefore appears consistent with originating from star formation (Figure 8). The incidence of the various types of AGN is a strong function of flux density (Figures 11 and 12); the relative contributions of clean SFGs, MLAGNs, and HLAGNs are similar for $S_{3\,\mathrm{GHz}} \gtrsim 50\,\mu$Jy, but toward lower flux densities the fraction of clean star-forming sources rises rapidly, reaching $\gtrsim 80\%$ at $S_{3\,\mathrm{GHz}} \lesssim 20\,\mu$Jy. Moreover, sources without a radio excess form the overwhelming majority of our radio sample for $S_{3\,\mathrm{GHz}} \lesssim 50\,\mu$Jy and reach a fraction of $\gtrsim 90\%$–95% below $S_{3\,\mathrm{GHz}} \lesssim 30\,\mu$Jy. Our radio sample is therefore strongly dominated by the star-forming population. Next-generation radio surveys, capable of reaching sensitivities similar to COSMOS-XS in an hour, are therefore ideally suited to probe this faint population and constrain the dust-insensitive star-formation history of the universe accordingly.

We additionally observe an interesting subset of radio-detected sources with multiwavelength counterparts solely at FIR wavelengths, or no such counterparts at all. A stacking analysis at optical, NIR and radio wavelengths indicates these sources are likely to be high redshift in nature ($z \sim 5$, Figure 15). If powered entirely through star formation, this implies radio-based SFRs of hundreds of solar masses per year. As such, these "optically dark" sources may contribute appreciably to cosmic star formation at high redshift.

We further present the completeness-corrected Euclidean-normalized 3 GHz radio number counts within the faint regime ($2\,\mu$Jy $\lesssim S_{3\,\mathrm{GHz}} \lesssim 100\,\mu$Jy) for the various radio population in Figure 14, and show that the COSMOS-XS sample forms a natural extension of the previous literature toward fainter flux densities. Additionally, our number counts are in good agreement with recent simulations of the radio sky within the expected scatter of cosmic variance.

Overall, COSMOS-XS provides the deepest multi-frequency radio survey to date, and probes a new, faint parameter space in the radio. In this work—the second paper in the COSMOS-XS series—we have presented the decomposition of the faint radio population into SFGs and AGNs. It will be followed by additional works studying the radio-derived star formation history and AGN accretion history and, additionally, we will leverage the multi-frequency data over COSMOS to further study the radio properties of the COSMOS-XS sample, including the evolution of the FIR–radio correlation, and the intrinsic radio spectra of the faint, star-forming population.

Appendix A
AGN Diagnostics from SED Fitting

We utilize the Bayesian Markov chain Monte Carlo (MCMC) SED-fitting code AGNFITTER (Calistro Rivera et al. 2016, 2017) to identify AGNs based on a comparison between the integrated luminosities in two sets of components: the direct stellar light and the accretion disk in the wavelength range 0.1–1 μm, and the MIR-continuum emission from a warm, dusty torus and the stellar-heated dust in the wavelength range 1–30 μm. As AGNFITTER returns realistic uncertainties on these luminosities through an MCMC approach, we compare the resulting probability distributions instead of simply comparing the luminosities from the best fit. We write $L_{C,p}$ to be the pth percentile of the integrated luminosity in component C. A source is then identified as an AGN if

$$(L_{\mathrm{disk},97.5} \geqslant L_{\mathrm{stellar},2.5}) \vee (L_{\mathrm{torus},97.5} \geqslant L_{\mathrm{cold\ dust},2.5}).$$

Here $\vee$ denotes the logical or operator, and the 2.5th and 97.5th percentiles are equivalent to Gaussian confidence intervals 2σ below and above the mean, respectively. We slightly modify this procedure for sources without any FIR photometry, as for this subset of our radio sample the MIR to FIR SED is nearly fully unconstrained as AGNFITTER does not impose energy balance. As a result, the MIR continuum of such sources is potentially fully fitted by a torus component, whereas the FIR luminosity is artificially low, as illustrated in the left panel of Figure A1. To avoid identifying such sources as AGNs, we define the parameter $f_{\mathrm{bolo}} = L_{\mathrm{torus},2.5}/(L_{\mathrm{torus},2.5} + L_{8-1000\ \mu\mathrm{m}})$, which measures the fractional contribution of the torus to the total FIR luminosity.

Here the luminosity $L_{8-1000\ \mu\mathrm{m}}$ is—by definition—not well-constrained, and therefore we utilize the upper limit on this value as determined empirically in Section 4.2, Figure 8 (lower panel), where we fit the sensitivity curve of our Herschel observations, and hence find a redshift-dependent upper limit on the FIR luminosity that a source without any FIR detections can have. Based on visual inspection of the SEDs, as well as the distribution of f_{bolo} for sources that do have accurately constrained FIR luminosities, we then require that sources without FIR photometry have $f_{\mathrm{bolo}} \geqslant 0.20$ in order to be considered an AGN. This, in turn, identifies sources such as the example shown in the right panel of Figure A1 as an AGN based on a strong torus component and stringent upper limits on the FIR photometry. Through this procedure, we can then accurately assess whether sources without robust FIR detections are likely to be AGNs based on a substantial torus component, without having to invoke the uncertainties associated with energy balance. Additionally, the separation between SFGs and AGNs is done based on physically motivated properties, instead of the comparison of goodness-of-fit.

Appendix B
Comparison to 3 GHz VLA-COSMOS

The 3 GHz VLA-COSMOS project (Smolčić et al. 2017a, 2017b; Delvecchio et al. 2017) covers the full 2 deg² COSMOS field at an identical frequency, allowing us to compare the multiwavelength AGN identification for the bright end of our radio sample. Overall, we have 471 radio sources in common (based on cross-matching within 1″0), subject to two independent analyses. While our overall AGN diagnostics are similar, there are a few differences in the overall analysis. Primarily, we are using improved FIR photometry taken from the recent Super-deblended catalog (Jin et al. 2018), which includes photometry up to 1.2 mm, whereas the COSMOS2015 catalog (Laigle et al. 2016) used in the 3 GHz VLA-COSMOS project contains photometry only up to 500 μm (Herschel/ SPIRE). In addition, the deblending procedure employed by Jin et al. (2018) is more detailed than previous techniques, and is optimized for highly confused images. Nevertheless, based on the sample both surveys have in common, the overall number of AGNs identified is highly similar: we identify 176 out of these 471 sources as AGNs, compared to 193 AGNs in VLA-COSMOS. The overlap between these two AGN samples is
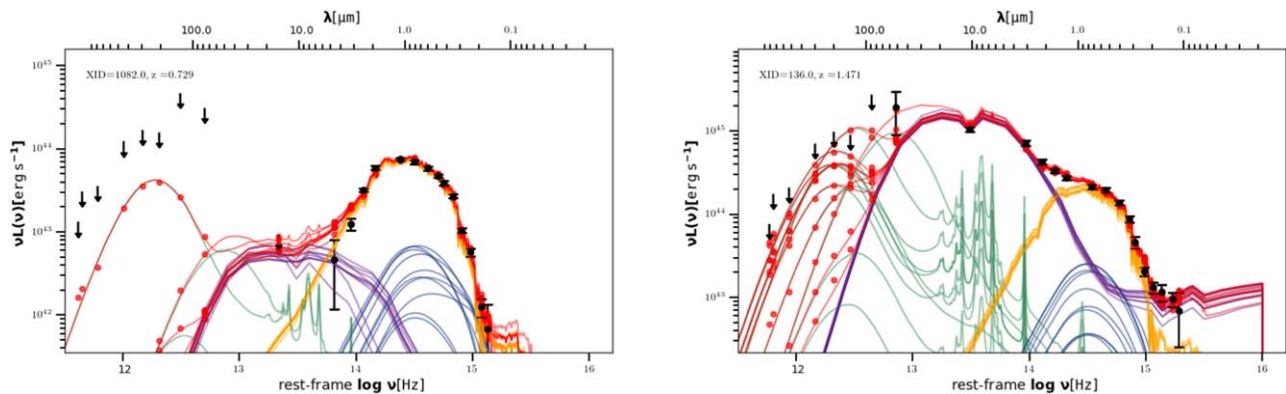

Figure A1. Example SEDs from AGNFITTER. The photometric data are shown through the black points, and upper limits are indicated by downward arrows. The fitted direct stellar contribution is shown in yellow, the accretion disk in blue, the dusty torus in purple, and the stellar-heated dust in green. The overall fit is shown in red. A total of 10 random Markov chain Monte Carlo realizations are plotted. Left: SED of a source without FIR photometry, which shows a substantial torus component in the 1–30 μm range, and has negligible cold dust component despite having very unconstraining upper limits. Sources such as these are not classified as AGNs. Right: SED of a source without reliable ($\geqslant 3\sigma$) FIR photometry which shows a large torus component and additionally has stringent upper limits on the total FIR component. Sources such as these are classified as AGNs.

substantial, with 147 of these AGNs (∼84%) being identified in both surveys. The small remaining differences can be explained by the slightly different criteria and multiwavelength data sets employed in the AGN identification process. We expand upon all of these in the next sections, and summarize the results in Table B1.

B.1. Radio-excess AGNs

In the 3 GHz VLA-COSMOS project, radio-excess AGNs are identified through a redshift-dependent threshold in $\log_{10}(L_{1.4\,\mathrm{GHz}}/\mathrm{SFR}_\mathrm{IR})$. This quantity is equal to minus q_TIR, up to a constant. The threshold for radio-excess AGNs was then determined by considering the scatter toward values below the median in the distribution of $\log_{10}(L_{1.4\,\mathrm{GHz}}/\mathrm{SFR}_\mathrm{IR})$, as this regime should predominantly be populated by SFGs, and will hence trace the intrinsic scatter in the FIR radio correlation. Radio AGNs should instead only scatter toward larger values. However, since only the observed radio sources are used to determine the spread in the distribution, any upper limits or non-detections will affect the observed scatter. Radio-faint SFGs in particular are potentially missed, such that the observed spread about the correlation is lower than the intrinsic scatter. As a result, the adopted threshold for classifying AGNs is likely to be somewhat conservative, whereby some star formation powered sources with low q_TIR will be classified as AGNs. In a subsequent paper as part of the 3 GHz VLA COSMOS survey, Delhaize et al. (2017) re-calculated the distribution of q_TIR for SFGs, now including detection limits in both the FIR and radio. We adopted their best-fitting trend for star-forming sources, and classified radio-excess AGNs as the sources that are $>2.5\sigma$ offset from the correlation. In addition, we take into account sources without FIR photometry by comparing the expected IR luminosity, assuming the radio emission is powered by star formation only, with the Herschel sensitivity curves (see the lower panel of Figure 8). Nevertheless, the difference in the total number of radio-excess AGNs identified in both surveys is small: we find 94 such AGNs in COSMOS-XS, whereas there are 102 identified in VLA-COSMOS, among the 471 radio sources in common. This small difference likely reflects the slightly different thresholds adopted.

B.2. SED-fitted AGNs

In our analysis, we adopted the AGNFITTER SED-fitting code to decompose the full SED of our radio-detected sample into star-forming and AGN components. This code is fundamentally different from SED3FIT used in the 3 GHz VLA-COSMOS project, as it does not adopt energy balance and is based on an MCMC method. As a result, AGNFITTER is designed to distinguish between AGNs and star-forming sources not based on a goodness-of-fit test, but instead on physical properties, taking into account the robust uncertainties obtained from MCMC fitting (see the discussion in Appendix A).

Overall, we recover a similar number of AGNs through SED-fitting, namely 59 in COSMOS-XS versus 65 in VLA-COSMOS, though only 26 of these are in common between the two radio samples. This difference may most likely be explained by the different fitting codes adopted, in particular whether or not the assumption of energy balance is made, and by the different criteria used for establishing whether a source

Table B1
Comparison of the COSMOS-XS AGN Sample with the 3 GHz VLA-COSMOS Survey, Based on 471 Sources in Common

Diagnostic	COSMOS-XS	VLA-COSMOS	Overlap
X-ray	61	61	56
IRAC	20	26	20
SED-fitting	59	65	26
Radio-excess	116	102	71
$[\mathrm{NUV} - r^+]$	33	52	28
Overall[a]	190	193	154

Note.
[a] Represents the total number of AGNs identified across all five diagnostics in the two surveys.

is an AGN. In addition, the slight differences between the two surveys further highlights that there is no clear distinction between AGNs and SFGs, and instead indicates that a large fraction of sources is composite in nature.

B.3. Red, Quiescent AGNs

In this paper we follow the method of Ilbert et al. (2010), where galaxies are classified as star-forming or quiescent based on their $[\mathrm{NUV} - r^+]$ colors. In particular, both in this work and in Smolčić et al. (2017a), sources are identified as quiescent AGNs when $[\mathrm{NUV} - r^+] > 3.5$, where the colors are corrected for dust attenuation. Despite these identical criteria, there is some variation in the AGNs identified through this method as in the VLA-COSMOS survey the number of AGNs found through red rest-frame UV/optical colors is ∼58% larger than found in this work (33 in COSMOS-XS versus 52 in VLA-COSMOS, with 28 sources in common). This cannot be explained by our requirement that sources identified through red colors lie below $z \leqslant 2$, as this does not discard any additional red objects from the sample. Nevertheless, in both our and the VLA-COSMOS survey the red, quiescent AGNs are also identified through radio excess in the majority of the cases (56% in COSMOS-XS versus 64% in VLA-COSMOS), which indicates that sources identified solely through red rest-frame optical/NUV-colors are only a small fraction of the overall AGN population.

Appendix C
Incompleteness in the Multiwavelength Photometry

As we utilize multiwavelength photometry spanning X-ray to radio wavelengths in order to identify AGNs, this identification may be affected by incompleteness in the various data sets we employ. We here study the effect this has on our sample of AGNs.

C.1. X-ray AGNs

We identify HLAGNs through a combination of X-ray, MIR, and SED-fitting techniques. Our X-ray data in particular are relatively shallow—we are already incomplete to X-ray sources with a luminosity of $L_{[0.5-8]\,\mathrm{keV}} \sim 10^{42}$ erg s^{-1} beyond $z \sim 0.5$ (see also Figure 7 in Marchesi et al. 2016a). While we hence cannot place any direct constraints on the X-ray luminosities of sources beyond this redshift, we employ X-ray stacking to investigate the typical X-ray properties of our clean, SFG

sample, making use of the X-ray stacking tool C-STACK. By stacking in four redshift bins between $0.5 \lesssim z \lesssim 2.5$, we find that the average X-ray luminosities of this clean SFG sample are fully consistent with the star formation−X-ray luminosity relations from Symeonidis et al. (2014). We will further address this, as well as the stacked X-ray properties for the different classes of AGNs, in a forthcoming paper. Overall, this implies that our sample of clean SFGs is likely minimally contaminated by X-ray AGNs.

C.2. MIR AGNs

In our analysis, we identify 28 sources as MIR AGNs based on their Spitzer/IRAC colors, following the criteria from Donley et al. (2012) for sources with $z \leqslant 2.7$. However, we only have IRAC photometry in all four channels for ∼60% of our radio sample, which will cause us to miss some MIR AGNs simply based on lack of photometry. In Figure C1 we show the fraction of radio sources with with IRAC photometry as a function of 3 GHz flux density, which decreases to ∼30% for our lowest flux density bin ($\sim 4\ \mu$Jy).[21] The total number of MIR AGNs expected in a given radio flux density bin is then the observed number divided by the fraction of sources with full IRAC photometry. The corrected fraction of MIR AGNs is shown with the red data points in Figure C1. After applying this completeness correction, the fraction of MIR AGNs is still decreasing toward lower radio flux densities, which is a trend we observe for HLAGNs in general. Overall, we expect to find 53 ± 14 MIR AGNs in our sample of radio sources, after correcting for the IRAC incompleteness, which implies we are missing 28 ± 8 additional MIR AGNs. However, as the majority (∼80%) of IRAC AGNs are also identified through either our X-ray or SED-fitting criteria, we expect to miss only 6 ± 2 AGNs based on our incomplete IRAC photometry, accounting for only ∼3% of the total number of HLAGNs. We are therefore not limited by incompleteness in the MIR photometry in our identification of (HL)AGNs.

C.3. Radio-excess AGNs

Radio-excess AGNs are identified through an excess in radio emission compared to what is expected based on their total FIR luminosity, with this additional radio emission being ascribed to AGN activity. While our full sample is—by definition—detected at 3 GHz, we only have FIR photometry for ∼50% of our radio sources, which complicates the determination of accurate total (8–1000 μm) IR luminosities. In order to identify radio-excess AGNs, we therefore assume energy balance between the UV to NIR data and the longer-wavelength emission, as the photometry at short wavelengths is typically better constrained. We test the reliability of the FIR luminosities derived in this way by artificially removing the FIR photometry of sources that do have detections at such wavelengths and subsequently re-fitting their SEDs with MAGPHYS. For this, we limit ourselves to the 1371 radio sources that we cross-matched with both COSMOS2015 and the Super-deblended catalog. Among these sources, we identified 118 radio-excess AGNs (103 and 61 via the normal and "inverse" radio excess criteria, respectively, with substantial overlap; see Figure 12).
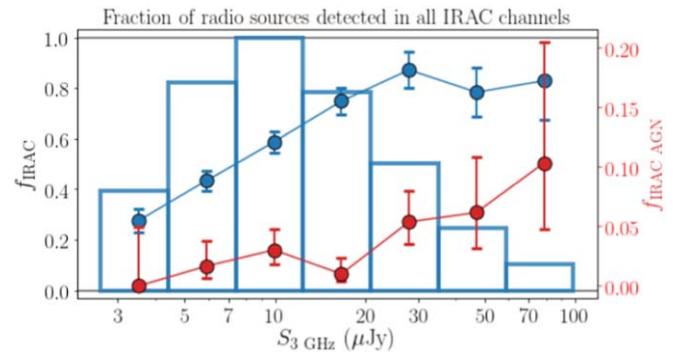

Figure C1. Fraction of sources with measured IRAC flux densities in all four channels as a function of 3 GHz flux density (blue line and points). The background histogram shows the overall distribution of radio flux densities, and the red line—corresponding to the right ordinate axis—denotes the fraction of MIR AGNs we expect when correcting for the incompleteness in IRAC photometry. Similar to our overall results for HLAGNs, the fraction of MIR AGNs decreases toward fainter flux densities, even after correcting for the incompleteness.

After removing all FIR detections, we recover a total of 143 radio-excess AGNs instead (118 and 71 sources through the regular and inverse criteria, respectively; note that we use the empirical detection limit in Figure 8 for the latter, as before), ∼20% larger than the number of AGNs we find when we do include available FIR photometry. This indicates that the energy balance MAGPHYS employs works well for the majority of sources, as even without FIR information accurate FIR luminosities, and hence q_{TIR}, are predicted. We further note that the "inverse" radio-excess criterion we apply requires a priori information on the FIR properties of a subset of our radio sample, as these are used to determine an appropriate sensitivity limit for the various Herschel photometric bands. When removing all FIR information, we can no longer directly apply this criterion. Using only the standard radio-excess diagnostic, we recover a total of 118 AGNs—equal to the number identified through the combined radio-excess and "inverse" radio-excess criteria when FIR data are included. Overall, we therefore conclude that the lack of FIR photometry for half of our radio sample does not substantially impede our classification of radio-excess AGNs.

C.4. Summary

In the previous subsections, we have established that incompleteness issues are not expected to have a substantial effect on the AGN classification in this work. What remains is then to show that any of the trends we see for the SFGs and various types of AGNs with either flux density (Figures 11, 12 and 14) or redshift (Figure 13) are not caused by any of the AGN diagnostics being more efficient in identifying sources within a given flux density or redshift interval. In Figure C2, we therefore plot the fractional contribution of each of these diagnostics to the total AGN counts versus redshift (left panel) and flux density (right panel). No strong trends are visible in either of the panels, which indicates that we are not biased toward finding a specific subset of AGNs in a given bin of our redshift- and flux density space.[22] In particular, the fractional contribution of the various types of AGNs is a near-constant

---

[21] We limit this analysis to the 1390 radio sources (96.7%) with $S_{3GHz} \leqslant 100\ \mu$Jy due to poor statistics at the brighter end. This includes 25 of the in total 28 MIR-AGNs.

[22] We, however, caution that one cannot robustly disentangle whether a trend with redshift is a physical effect or the result of a selection bias. Nevertheless, a lack of any given trend substantiates that no such bias is present.
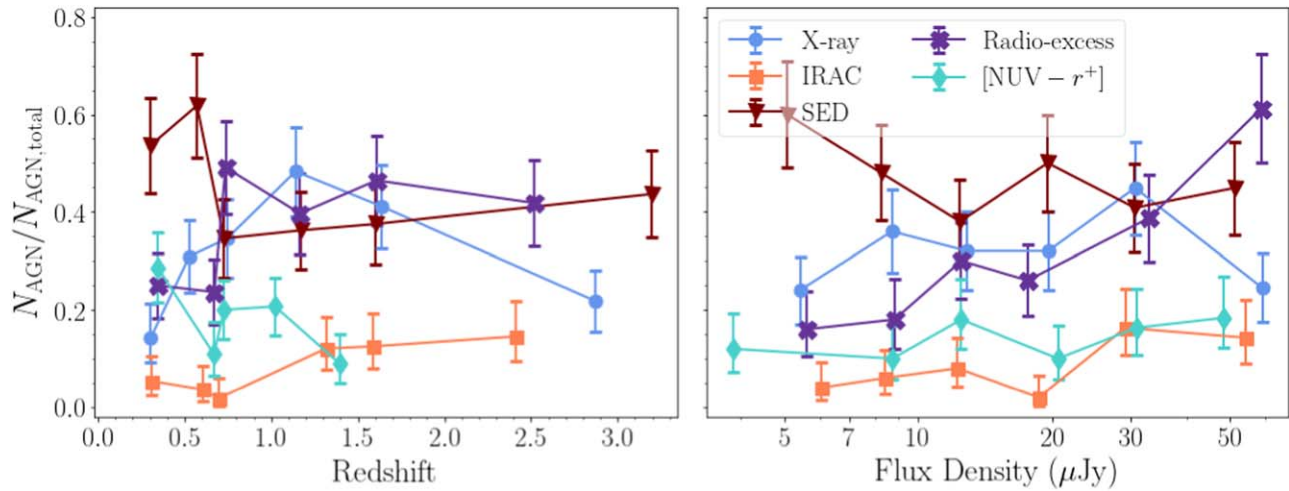
Figure C2. Left: fractional contribution of the different AGN diagnostics (X-ray emission, Spitzer/IRAC colors, SED-fitting, radio excess, and red rest-frame NUV/optical colors) to the overall sample of AGNa as a function of redshift. The data were binned into six fixed redshift ranges, containing an approximately equal number of AGNa in total. The redshift value of each data point marks the median value per bin, and the errorbars represent Poissonian uncertainties. Note that the fractional contribution sums to a number greater than unity, as a single AGN may be identified through multiple diagnostics. The lack of strong trends with redshift indicate that we do not suffer from substantial biases as function of redshift (e.g., luminosity incompleteness) in our identification of AGNa. Right: fractional contribution of the different AGN diagnostics to the overall sample of AGN as a function of flux density. The data follow a similar binning as used in the left panel. With exception of the radio-excess AGNa, which we established show a real and significant trend with flux density in the main text, no trends of the other AGN diagnostics are seen to correlate with radio flux density. This indicates that our various methods of AGN identification are reliable across the full range of fluxes probed in the COSMOS-XS survey.

function of flux density, which substantiates our results in Section 5 that AGNs make up a comparatively smaller fraction of the faint radio population.

Appendix D
Final Counterpart Catalog

The catalog of sources detected at 3 GHz, including their multiwavelength source identifications, is available with this paper in standard FITS format. We show a sample of the catalog in Table D1, and elaborate here on its contents.

1. *Column 1*. Source ID, equal to the ID assigned in Paper I.
2. *Columns 2, 3*. Radio coordinates of the source as determined by PyBDSF.
3. *Columns 4, 5*. Optimal redshift of the radio source, including a Boolean indicating whether it is spectroscopic or photometric.
4. *Columns 6–11*. Flux densities and errors of the source at, respectively, 1.4, 3, and 10 GHz, in μJy. In the absence of a radio counterpart at 1.4 or 10 GHz, both the flux and error are set to -99.
5. *Columns 12–14*. ID of the source in, respectively, the Super-deblended catalog, COSMOS2015, and the i-band-selected catalog. If no counterpart is found, this value is set to -99.
6. *Column 15*. Boolean indicating whether the source is flagged as "potentially spurious" based on the discussion in Section 2.1, in which case it is set to True.
7. *Columns 16–18*. The rest-frame 1.4 GHz luminosity of the source and its 16th and 84th percentiles (equivalent to 1σ confidence intervals), using the measured spectral index if available, or a fixed value of $\alpha = -0.7$ otherwise. The luminosity is given in W Hz^{-1}.
8. *Columns 19–21*. The value of the radio–FIR correlation parameter q_{TIR} as defined in Equation (1), and its 16th and 84th percentiles.

9. *Column 22*. Boolean indicating whether the source is identified as an X-ray AGN, in which case it is set to True.
10. *Column 23*. Boolean indicating whether the source is identified as an AGN through its MIR colors, by means of the Donley et al. (2012) wedge.
11. *Column 24*. Boolean indicating whether the source is identified as an AGN based on SED fitting.
12. *Column 25*. Boolean indicating whether the source is identified as an SED AGN based on a MIR torus component.
13. *Column 26*. Boolean indicating whether the source is identified as an SED AGN based on a UV/optical accretion disk component.
14. *Column 27*. Boolean indicating whether the source is identified as an AGN based on red rest-frame NUV and optical colors.
15. *Column 28*. Boolean indicating whether the source is identified as an AGN based on an excess in radio emission from what is expected from the radio–FIR correlation, with FIR luminosities calculated using MAGPHYS.
16. *Column 29*. Boolean indicating whether the source is identified as an AGN based on an excess in radio emission from what is expected from the radio–FIR correlation, in the absence of Herschel FIR photometry, through comparison with the detection limit of Herschel.
17. *Column 30*. Boolean indicating whether the source is identified as an HLAGN, which is True when either of columns 22–26 is True.
18. *Column 31*. Boolean indicating whether the source is identified as an MLAGN, which is True when all of columns 22–26 are False and either of columns 27–29 is True.
19. *Column 32*. Boolean indicating whether the source is a radio-excess AGN, which is True when either of columns 28, 29 is True.

Table D1
Sample of the COSMOS-XS Multiwavelength and AGN Catalog

ID	R.A.	Decl.	Redshift	Spec $-z$		$S_{3\,GHz}$	$\sigma_{3,GHz}$		No-qE AGN
	(deg)	(deg)				(μJy)	(μJy)		
CXS J100016.55+023309.17	150.06894	2.55255	4.36	0	...	4.95	0.59	...	1
CXS J095952.41+023148.56	149.96837	2.53016	1.60	0	...	8.26	1.39	...	0
CXS J100034.78+022849.87	150.14493	2.48052	1.56	0	...	10.56	0.71	...	0
CXS J100036.03+023937.91	150.15012	2.66053	1.22	0	...	507.22	1.16	...	0
CXS J100019.40+022936.63	150.08083	2.49351	0.21	0	...	5.22	0.62	...	0
CXS J100057.22+023644.73	150.23842	2.61243	1.77	0	...	10.11	1.56	...	0
CXS J100028.30+024008.46	150.11790	2.66902	0.35	1	...	13.27	1.38	...	0
CXS J100000.45+022739.28	150.00188	2.46091	0.73	0	...	12.79	1.37	...	1
CXS J100016.38+023620.36	150.06823	2.60565	2.26	0	...	3.90	0.69	...	0
CXS J100015.53+023742.11	150.06469	2.62836	0.69	0	...	11.87	1.31	...	0

Notes. The contents of the COSMOS-XS multiwavelength catalog, including a description of the various columns, are elaborated on in the text.

(This table is available in its entirety in machine-readable form.)

20. *Column 33*. Boolean indicating whether the source is an AGN without radio excess, which is True when either of columns 22–27 is True and both columns 28, 29 are False.

ORCID iDs

H. S. B. Algera ⓘ https://orcid.org/0000-0002-4205-9567
J. A. Hodge ⓘ https://orcid.org/0000-0001-6586-8845
I. R. Smail ⓘ https://orcid.org/0000-0003-3037-257X
M. Novak ⓘ https://orcid.org/0000-0001-8695-825X
D. A. Riechers ⓘ https://orcid.org/0000-0001-9585-1462
F. Walter ⓘ https://orcid.org/0000-0003-4793-7880

References

Alexander, D. M., Bauer, F. E., Chapman, S. C., et al. 2005, ApJ, 632, 736
Algera, H.S.B., Smail, I., Dudzevičiūtė, U., et al. 2020, arXiv:2009.06647
Aretxaga, I., Wilson, G. W., Aguilar, E., et al. 2011, MNRAS, 415, 3831
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Battisti, A. J., da Cunha, E., Grasha, K., et al. 2019, arXiv:1908.00771
Bell, E. F. 2003, ApJ, 586, 794
Berta, S., Lutz, D., Santini, P., et al. 2013, A&A, 551, A100
Bertoldi, F., Carilli, C., Aravena, M., et al. 2007, ApJS, 172, 132
Best, P. N., Kaiser, C. R., Heckman, T. M., & Kauffmann, G. 2006, MNRAS, 368, L67
Bonaldi, A., Bonato, M., Galluzzi, V., et al. 2019, MNRAS, 482, 2
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
Bonzini, M., Padovani, P., Mainieri, V., et al. 2013, MNRAS, 436, 3759
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
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., Lusso, E., Hennawi, J. F., & Hogg, D. W. 2016, ApJ, 833, 98
Calistro Rivera, G., Williams, W. L., Hardcastle, M. J., et al. 2017, MNRAS, 469, 3468
Capak, P., Aussel, H., Ajiki, M., et al. 2007, ApJS, 172, 99
Cardamone, C. N., Urry, C. M., Schawinski, K., et al. 2010, ApJL, 721, L38
Casey, C. M., Narayanan, D., & Cooray, A. 2014, PhR, 541, 45
Casey, C. M., Zavala, J. A., Spilker, J., et al. 2018, ApJ, 862, 77
Chabrier, G. 2003, ApJL, 586, L133
Civano, F., Marchesi, S., Comastri, A., et al. 2016, ApJ, 819, 62
Condon, J. J. 1989, ApJ, 338, 13
Condon, J. J. 1992, ARA&A, 30, 575
Cotton, W. D., Condon, J. J., Kellermann, K. I., et al. 2018, ApJ, 856, 67
da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595
da Cunha, E., Walter, F., Smail, I. R., et al. 2015, ApJ, 806, 110

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
Delvecchio, I., Smolčić, V., Zamorani, G., et al. 2018, MNRAS, 481, 4971
Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142
Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2019, arXiv:1910.07524
Evans, A. S., Solomon, P. M., Tacconi, L. J., Vavilkin, T., & Downes, D. 2006, AJ, 132, 2398
Fabbiano, G. 2006, ARA&A, 44, 323
Fanali, R., Caccianiga, A., Severgnini, P., et al. 2013, MNRAS, 433, 648
Fanaroff, B. L., & Riley, J. M. 1974, MNRAS, 167, 31P
Farrah, D., Urrutia, T., Lacy, M., et al. 2012, ApJ, 745, 178
Geach, J. E., Dunlop, J. S., Halpern, M., et al. 2017, MNRAS, 465, 1789
Gehrels, N. 1986, ApJ, 303, 336
Gim, H. B., Yun, M. S., Owen, F. N., et al. 2019, arXiv:1903.07632
Gruppioni, C., Mignoli, M., & Zamorani, G. 1999, MNRAS, 304, 199
Hardcastle, M. J., Ching, J. H. Y., Virdee, J. S., et al. 2013, MNRAS, 429, 2407
Heckman, T. M., & Best, P. N. 2014, ARA&A, 52, 589
Helou, G., Soifer, B. T., & Rowan-Robinson, M. 1985, ApJL, 298, L7
Herrera Ruiz, N., Middelberg, E., Deller, A., et al. 2017, A&A, 607, A132
Hoaglin, D. C., Mosteller, F., & Tukey, J. W. 1983, Understanding Robust and Exploratory Data Anlysis (New York: Wiley)
Hodge, J. A., Carilli, C. L., Walter, F., et al. 2012, ApJ, 760, 11
Hodge, J. A., Swinbank, A. M., Simpson, J. M., et al. 2016, ApJ, 833, 103
Hummel, E. 1981, A&A, 93, 93
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236
Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, ApJ, 709, 644
Jin, S., Daddi, E., Liu, D., et al. 2018, ApJ, 864, 56
Jin, S., Daddi, E., Magdis, G. E., et al. 2019, arXiv:1906.00040
Kellermann, K. I., & Wall, J. V. 1987, in IAU Symp., 124, Observational Cosmology, ed. A. Hewitt, G. Burbidge, & L. Z. Fang (Dordrecht: Reidel), 545
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, ApJS, 218, 15
Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, ApJS, 154, 166
Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24
Le Fèvre, O., Tasca, L. A. M., Cassata, P., et al. 2015, A&A, 576, A79
Le Floc'h, E., Aussel, H., Ilbert, O., et al. 2009, ApJ, 703, 222
Lehmer, B. D., Alexander, D. M., Bauer, F. E., et al. 2010, ApJ, 724, 559
Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, ApJS, 184, 218
Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70
Liu, D., Daddi, E., Dickinson, M., et al. 2018, ApJ, 853, 172
Lusso, E., Comastri, A., Vignali, C., et al. 2011, A&A, 534, A110
Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, A&A, 532, A90
Magnelli, B., Ivison, R. J., Lutz, D., et al. 2015, A&A, 573, A45
Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285
Maini, A., Prandoni, I., Norris, R. P., Giovannini, G., & Spitler, L. R. 2016, A&A, 589, L3
Mancuso, C., Lapi, A., Cai, Z.-Y., et al. 2015, ApJ, 810, 72
Marchesi, S., Civano, F., Elvis, M., et al. 2016a, ApJ, 817, 34

Marchesi, S., Lanzuisi, G., Civano, F., et al. 2016b, ApJ, 830, 100
McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156
McNamara, B. R., & Nulsen, P. E. J. 2012, NJPh, 14, 055023
Mohan, N., & Rafferty, D. 2015, PyBDSF: Python Blob Detection and Source Finder, Astrophysics Source Code Library, ascl:1502.007
Morganti, R., Fogasy, J., Paragi, Z., Oosterloo, T., & Orienti, M. 2013, Sci, 341, 1082
Murphy, E. J., Bremseth, J., Mason, B. S., et al. 2012, ApJ, 761, 97
Murphy, E. J., Momjian, E., Condon, J. J., et al. 2017, ApJ, 839, 35
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJS, 206, 8
Nagar, N. M., Falcke, H., Wilson, A. S., & Ho, L. C. 2000, ApJ, 542, 186
Novak, M., Smolčić, V., Delhaize, J., et al. 2017, A&A, 602, A5
Novak, M., Smolčić, V., Schinnerer, E., et al. 2018, A&A, 614, A47
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labbé, I., & Stefanon, M. 2018, ApJ, 855, 105
Offringa, A. R., McKinley, B., & Hurley-Walker, N. 2014, MNRAS, 444, 606
Ogle, P., Whysong, D., & Antonucci, R. 2006, ApJ, 647, 161
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Oliver, S. J., Bock, J., Altieri, B., et al. 2012, MNRAS, 424, 1614
Padovani, P., Alexander, D. M., Assef, R. J., et al. 2017, A&ARv, 25, 2
Padovani, P., Bonzini, M., Kellermann, K. I., et al. 2015, MNRAS, 452, 1263
Padovani, P., Miller, N., Kellermann, K. I., et al. 2011, ApJ, 740, 20
Pavesi, R., Sharon, C. E., Riechers, D. A., et al. 2018, ApJ, 864, 49
Prandoni, I., & Seymour, N. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 67, arXiv:1412.6512
Reines, A. E., Sivakoff, G. R., Johnson, K. E., & Brogan, C. L. 2011, Natur, 470, 66
Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, ApJ, 872, 7

Rovilos, E., & Georgantopoulos, I. 2007, A&A, 475, 115
Rujopakarn, W., Nyland, K., & Kimball, A. E. 2018, arXiv:1810.07098
Salvato, M., Ilbert, O., Hasinger, G., et al. 2011, ApJ, 742, 61
Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86
Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
Schinnerer, E., Sargent, M. T., Bondi, M., et al. 2010, ApJS, 188, 384
Schinnerer, E., Smolčić, V., Carilli, C. L., et al. 2007, ApJS, 172, 46
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Selina, R. J., Murphy, E. J., McKinnon, M., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conf. Ser., 10700, Society of Photo-Optical Instrumentation Engineers (SPIE) Conf. Ser.107001O
Shankar, F., Weinberg, D. H., & Miralda-Escudé, J. 2009, ApJ, 690, 20
Simpson, J. M., Swinbank, A. M., Smail, I., et al. 2014, ApJ, 788, 125
Smolčić, V. 2009, ApJL, 699, L43
Smolčić, V., Delvecchio, I., Zamorani, G., et al. 2017b, A&A, 602, A2
Smolčić, V., Novak, M., Bondi, M., et al. 2017a, A&A, 602, A1
Springel, V., Di Matteo, T., & Hernquist, L. 2005, MNRAS, 361, 776
Stach, S. M., Dudzeviciute, U., Smail, I., et al. 2019, MNRAS, 487, 4648
Symeonidis, M., Georgakakis, A., Page, M. J., et al. 2014, MNRAS, 443, 3728
Thomson, A. P., Ivison, R. J., Simpson, J. M., et al. 2014, MNRAS, 442, 577
van der Vlugt, D., Algera, H. S. B., Hodge, J. A., et al. 2020, arXiv:2009.13528
Wang, R., Carilli, C. L., Neri, R., et al. 2010, ApJ, 714, 699
Wang, S. X., Brandt, W. N., Luo, B., et al. 2013, ApJ, 778, 179
Wilman, R. J., Miller, L., Jarvis, M. J., et al. 2008, MNRAS, 388, 1335
Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, ApJ, 554, 803
Zamojski, M. A., Schiminovich, D., Rich, R. M., et al. 2007, ApJS, 172, 468