A deep radio view of the evolution of the cosmic star formation rate density from a stellar-mass-selected sample in VLA-COSMOS

Eliab D. Malefahlo [1]* Matt J. Jarvis [1,2] Mario G. Santos [1,3] Sarah V. White [4] Nathan J. Adams [2] and Rebecca A. A. Bowler [2]

[1]*Department of Physics and Astronomy, University of the Western Cape, Private Bag X17, Bellville, Cape Town 7535, South Africa*
[2]*Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, UK*
[3]*South African Radio Astronomy Observatory (SARAO), 2 Fir Street, Observatory, Cape Town 7925, South Africa*
[4]*Department of Physics and Electronics, Rhodes University, PO Box 94, Makhanda 6140, South Africa*

ABSTRACT

We present the 1.4 GHz radio luminosity functions (RLFs) of galaxies in the Cosmic Evolution Survey (COSMOS) field, measured above and below the 5σ detection threshold, using a Bayesian model-fitting technique. The radio flux densities from Very Large Array (VLA)-COSMOS 3-GHz data are extracted at the position of stellar-mass-selected galaxies. We fit a local RLF model, which is a combination of active galactic nuclei and star-forming galaxies (SFGs), in 10 redshift bins with a pure luminosity evolution model. Our RLF exceeds previous determinations at low radio luminosities at $z < 1.6$ with the same radio data, due to our ability to directly constrain the knee and faint-end slope of the RLF. Beyond $z \sim 2$, we find that the SFG part of the RLF exhibits a negative evolution (L^* moves to lower luminosities) due to the decrease in low stellar-mass galaxies in our sample at high redshifts. From the RLF for SFGs, we determine the evolution in the cosmic star formation rate density (SFRD), which we find to be consistent with the established behaviour up to $z \sim 1$ using far-infrared data, but exceeds that from the previous radio-based work for the reasons highlighted above. Beyond $z \sim 1.5$ the cosmic SFRD declines. We note that the relation between radio luminosity and star formation rate is crucial in measuring the cosmic SFRD from radio data at $z > 1.5$. We investigate the effects of stellar mass on the total RLF by splitting our sample into low ($10^{8.5} \leq M/M_\odot \leq 10^{10}$) and high ($M > 10^{10} M_\odot$) stellar-mass subsets. We find that the SFRD is dominated by sources in the high stellar masses bin, at all redshifts.

Key words: methods: data analysis – galaxies: evolution – galaxies: luminosity function, mass function – radio continuum: galaxies – galaxies: star formationion.

1 INTRODUCTION

Understanding the evolution of star formation (SF) in galaxies over the history of the Universe is a key aspect of galaxy formation studies. It has the potential to tell us how, when and where, SF happened from the onset of the first galaxies within the epoch of reionization, through to the present day. Measuring the star formation rate (SFR) in galaxies can be done at a variety of wavelengths (see e.g. Kennicutt 1998). The most sensitive tracer of young massive stars within star-forming regions of galaxies comes from rest-frame ultraviolet (UV) observations, where the depth that can be reached with current telescopes means that very low SFRs can potentially be reached to the highest redshifts (e.g. McLure et al. 2013; Bouwens et al. 2015; Adams et al. 2020; Bowler et al. 2020).

However, the rest-frame UV is readily absorbed by dust, both within and along the line of sight to distant galaxies, resulting in the SFR measurements made at these wavelengths being lower limits. The dust that absorbs this UV radiation is heated and reradiates the energy at far-infrared (FIR) wavelengths with a spectrum close to a blackbody. The combination of UV through to FIR emission can therefore provide measurements of the total SFR in galaxies, both

unobscured and obscured (e.g. Burgarella, Buat & Iglesias-Páramo 2005; da Cunha, Charlot & Elbaz 2008; Berta et al. 2013; Smith & Hayward 2018).

Unfortunately, FIR observations are generally limited in their spatial resolution. For example, the *Herschel Space Observatory* has a resolution of 18 arcsec at 250 μm, leading to imaging surveys that are generally limited by source confusion (e.g. Oliver et al. 2012). Atacama Large Millimeter/submillimeter Array (ALMA) can detect this dust emission at much higher angular resolution, but these surveys are limited in area (e.g. Dunlop et al. 2017; Dudzevičiūtė et al. 2020; Franco et al. 2020; Gruppioni et al. 2020; Yamaguchi et al. 2020), or rely on pointed observations of pre-selected samples (e.g. Boogaard et al. 2019; Zavala et al. 2019; Simpson et al. 2020). Thus, it is unsurprising that over the past few years, alternative tracers of SFRs of galaxies have been considered at other wavelengths (e.g. Ouchi et al. 2010; Drake et al. 2013; Schober, Schleicher & Klessen 2015; Aird, Coil & Georgakakis 2017). Possibly the most promising one is using deep radio continuum observations.

The radio SFR estimate relies on the FIR SFR through the far-infrared–radio correlation (FIRC). This is a tight correlation between the radio luminosity and the total infrared (IR) luminosity of galaxies (e.g. van der Kruit 1971; de Jong et al. 1985; Condon, Anderson & Helou 1991; Jarvis et al. 2010; Delhaize et al. 2017). The correlation spans over five orders of magnitude and its existence

has been attributed to young massive stars. After their short life span, of a few Myr, the massive stars reach a catastrophic end in a supernova explosion, which accelerates electrons that then emit synchrotron radiation observed in the radio. During their short lifetimes these same massive stars emit optical and UV radiation that is then absorbed and reradiated into the IR by surrounding dust. Thus resulting in a correlation between the radio synchrotron emission and the dust continuum emission. In recent years, it has become apparent that the form of the FIRC may also depend on other properties of the galaxy (e.g. Molnár et al. 2018; Read et al. 2018; Delvecchio et al. 2021; Smith et al. 2021). Departures from the FIRC could be due to excess radio emission due to active galactic nuclei (AGN) activity, or that the FIR emission is not fully accounting for the total SFR in some galaxies. Indeed, using a total SFR from full spectral energy distribution (SED) modelling or by combining UV and FIR emission may alleviate some of these concerns, or possibly complicate them further for certain types of galaxy (e.g. Gürkan et al. 2018).

The most reliable estimates for the SFR based on radio emission are those where the contribution from a central accreting black hole is thought to be negligible. One way to do this is to select galaxies based on their optical properties, rather than use a radio selected sample. This mitigates against the inevitably bias for 'normal' galaxies with low-level AGN-related radio emission to be boosted above the flux-density limit of the radio survey. Whereas those galaxies with the same SFR but no AGN-related emission fall below the same flux limit.

Therefore, in this work we measure the radio luminosity function (RLF) of near-infrared (NIR)-selected galaxies below the nominal flux limit by applying the technique developed by Zwart, Santos & Jarvis (2015), extended upon in Malefahlo et al. (2020), and used in a similar way to measure the H I mass function (Pan et al. 2020).

We use a set of [star-forming galaxy (SFG) and AGN] models to fit the RLF directly to the radio data using a full Bayesian approach. Based on these RLFs we derive the evolution of the cosmic star formation rate density (SFRD) for the complete mass-limited sample and also for samples splits into a high and low stellar-mass bin, allowing us to understand the contribution from galaxies of different mass to the SFRD as a function of redshift. In Section 2, we describe the radio, optical, and NIR survey data we use, along with the photometric redshifts and the derived stellar masses. In Section 3, we provide a description of the Bayesian methodology we use to model the RLF below the nominal detection threshold as a function of redshift. In Section 4, we present the results of the various RLF model forms, and in Section 5, we use the most appropriate RLF models to calculate the evolution of the cosmic SFRD, and compare this with other studies in the literature. Section 6 summarizes our conclusions.

Throughout the paper we use the following Λ cold dark matter (ΛCDM) cosmology, with $H_0 = 70$ km^{-1} Mpc^{-1}, $\Omega_\Lambda = 0.7$, and $\Omega_M = 0.3$. All quoted optical and NIR magnitudes are in the AB system (Oke & Gunn 1983). We assume a spectral index, defined as $\alpha \equiv \log(S/S_0)/\log(\nu/\nu_0)$, with $\alpha = -0.7$, when converting flux density to luminosity and one reference frequency to another.

2 DATA

2.1 Near-infrared data

In order to select the galaxies for this study, we use the near-infrared (NIR) imaging in Y, J, H, and K_s bands taken with the VISTA InfraRed CAMera (VIRCAM; Emerson & Sutherland 2010) as part of the ultradeep survey on the Visible and Infrared Survey Telescope for Astronomy (VISTA), UltraVISTA (McCracken et al.

2012), and the deep optical data from Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) and the HyperSuprimeCam (HSC) Strategic Survey Programme (Aihara et al. 2018a,b) over the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007) field. Additionally, we use mid-infrared data from *Spitzer*/Infrared Array Camera (IRAC; Sanders et al. 2007; Steinhardt et al. 2014; Ashby et al. 2018). In the fourth data release (DR4) the survey covers a total area of ~1.9 deg^2 that is reduced to an effective area of ~1.8 deg^2 when masked regions (saturated by stars, regions of high noise) are excluded.[1] The overlapping effective area between DR4, IRAC, and CFHTLS or HSC is ~1.45 deg^2. The flux densities were extracted from a 2-arcsec-diameter aperture in each band using the K_s band as the detection image (a rough proxy for stellar mass over the redshift range we are interested), and extracting the flux at these positions across the other NIR and optical data (following Adams et al. 2020; Bowler et al. 2020). The catalogue has a minimum 5σ detection threshold of $K_s = 24.5$.

2.1.1 Photometric redshifts and stellar masses

The photometric redshifts are the same as those used by Adams et al. (2021) and are measured by fitting the multiband data available in the COSMOS field to a synthetic library of galaxy templates using LePHARE (Arnouts et al. 2002; Ilbert et al. 2009). In summary they follow Ilbert et al. (2013) using several synthetic galaxy multiband templates from Bruzual & Charlot (2003) and Polletta et al. (2007) [generated using the stellar population synthesis model of Bruzual & Charlot (2003) assuming a Chabrier (2003) initial mass function (IMF)] to compare with the observed photometry.

Comparing the photometric redshifts to the spectroscopic redshifts available in the literature (Lilly et al. 2009; Coil et al. 2011; Cool et al. 2013; Le Fèvre et al. 2013; Alam et al. 2015; Hasinger et al. 2018), Adams et al. (2021) report an outlier rate of 828/19 752 (4.2 per cent) and a normalized mean absolute deviation (NMAD) of 0.0312.

The stellar masses are determined again by using LePHARE to fit the multiband data with templates but with the redshift fixed at the best-fitting photometric redshift. χ^2 minimization is used to find the best-fitting template from the Bruzual & Charlot (2003) models on the total flux measurements in each filter.
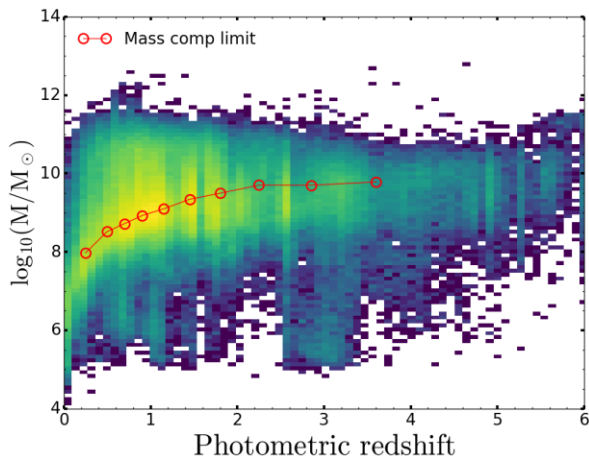
2.1.2 Sample

Our goal in this paper is to measure the evolution of the RLF and thus the cosmic SFRD using SFGs, which means removing contamination from stars and emission from AGN. Sources are classified as a star if (1) the best-fitting star template has higher probability than the best-fitting galaxy template, and (2) the source does not meet the *BzK* colour–colour selection criteria from Daddi et al. (2004), which combines the B and z optical bands with NIR K band to identify stars. Passive galaxies (predominantly AGN) are traditionally identified using colour–colour plots with $(U - V)$ versus $(V - J)$, usually referred to as *UVJ* (e.g. Wuyts et al. 2007; Williams et al. 2009). However, several studies have shown that ~10–20 per cent of passive sources have significant SF in the host galaxy (e.g. Belli et al. 2017; Merlin et al. 2018; Leja, Tacchella & Conroy 2019).

Therefore, we choose not to separate the galaxies in our sample into quiescent and SF, and instead aim to detect radio emission from SF for all galaxies that lie above our flux/mass limit.

---

[1]http://ultravista.org/release4/dr4_release.pdf

Table 1. The redshift bins along with the median redshift of the data in each bin. $\log_{10}[L_{det}/(\text{W Hz}^{-1})]$ is the minimum 1.4 GHz luminosity in each bin corresponding to the detection threshold (5σ). N_{Tot} is the total number of galaxies in each bin. We show the stellar-mass completeness limit that contains 90 per cent of the galaxies stellar mass completeness. We also present the number of galaxies with stellar mass above the stellar-mass completeness limit (N, our sample) and the number of sources in our sample that have VLA-COSMOS 3-GHz counterparts (N_{VLA}).

Redshift bin	z_M	$\log_{10}[L_{det}/(\text{W Hz}^{-1})]$	N_{Tot}	$\log_{10}(M_{lim}/M_\odot)$	N	N_{VLA}
$0.1 < z < 0.4$	0.32	22.12	27 890	8.0	17 759	485
$0.4 < z < 0.6$	0.53	22.62	21 738	8.5	15 943	584
$0.6 < z < 0.8$	0.7	22.90	32 649	8.7	22 481	675
$0.8 < z < 1.0$	0.9	23.16	36 510	8.9	26 461	764
$1.0 < z < 1.3$	1.12	23.38	43 971	9.1	26 593	872
$1.3 < z < 1.6$	1.45	23.64	29 912	9.3	15 848	622
$1.6 < z < 2.0$	1.74	23.82	33 868	9.5	14 057	778
$2.0 < z < 2.5$	2.24	24.06	15 167	9.7	5344	349
$2.5 < z < 3.2$	2.82	24.28	20 353	9.7	6398	257
$3.2 < z < 4.0$	3.42	24.47	9108	9.8	2878	45

Figure 1. The stellar mass of the galaxies in the UltraVISTA DR4 sample as a function of photometric redshifts. The red circles connected by lines represent the stellar mass completeness limit.

2.1.3 Completeness

In a magnitude-limited survey, the stellar mass completeness is a function of the mass-to-light ratio (which depends on a galaxy template) and redshift. In light of this we divide the data into 10 redshift bins (from $z = 0.1$ to $z = 4$, ensuring that the redshift bins are large enough to not be compromised by photometric redshift uncertainties, i.e. the photometric redshift uncertainty $\ll$ redshift bin width) and estimate a conservative stellar-mass completeness limit (M_{lim}) at each redshift. To estimate this stellar-mass completeness limit we follow Ilbert et al. (2013). We start by computing the stellar-mass limit (M_{min}) for each galaxy, and the limit is the stellar mass that a galaxy at a certain redshift with stellar mass (M) would have if it was observed at the 5σ flux limit ($K_s = 24.5$),

$$\log(M_{min}) = \log(M) + 0.4(K_s - 24.5). \tag{1}$$

The completeness limit is then given by the stellar mass that is above 90 per cent of the stellar-mass limits in the redshift bin (see Table 1 and Fig. 1). This stellar-mass completeness limit takes into account the different galaxy templates (and their corresponding mass-to-light ratio) that then ensures that not more than 10 per cent of the low-mass galaxies are missing in our sample. Applying the stellar-mass

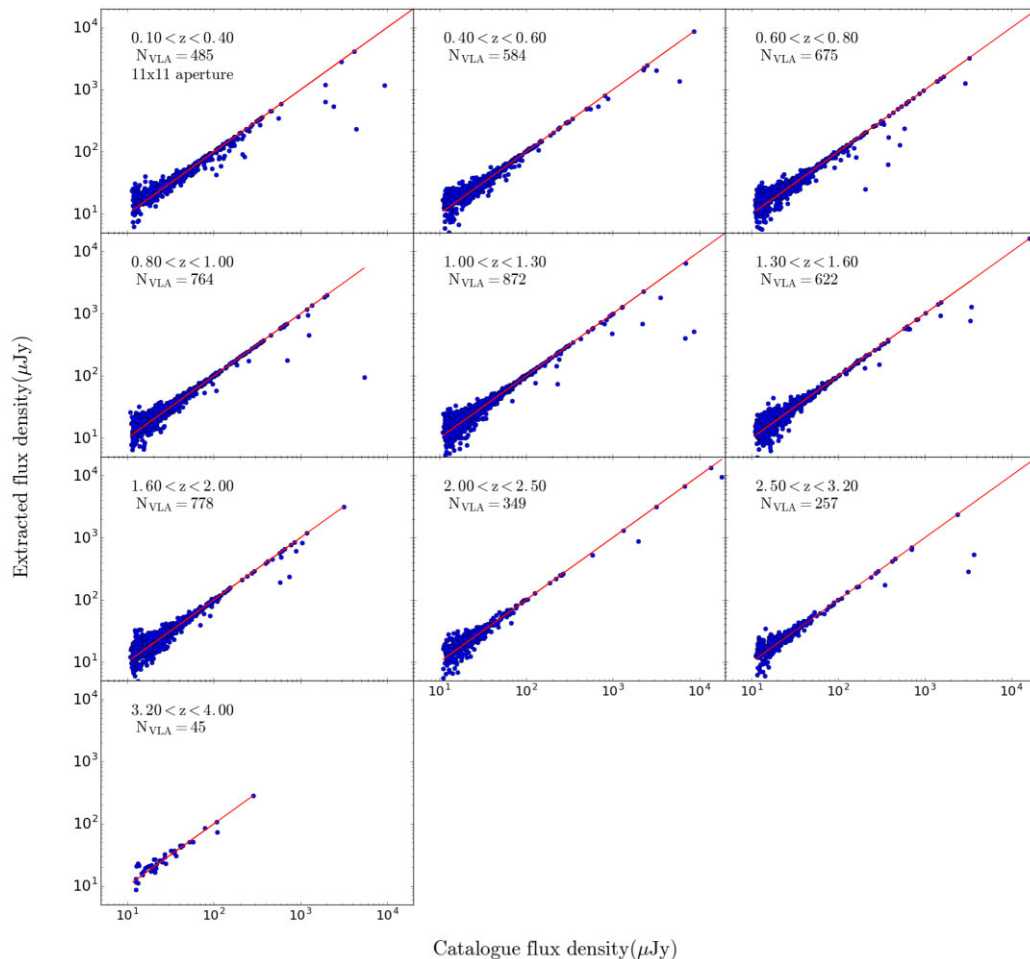cut results in a total galaxy sample size of 171 621, over the redshift range $0.1 < z < 4$.

2.2 Radio data

We use radio data from the Very Large Array (VLA)-COSMOS 3-GHz survey (Smolčić et al. 2017a). The survey covers 2.6 deg^2 with a resolution of 0.75 arcsec and an rms sensitivity with a median value of 2.3 μJy. 10 830 detected sources were extracted in the central 2 deg^2 using BLOBCAT (Hales et al. 2012) with 67 found to be multicomponent. The multicomponent sources are visually confirmed and most of them are galaxies with resolved structures such as jet/lobe/core. A small portion of the multicomponent sources are SFGs with disc-like structures (Smolčić et al. 2017a).

2.3 Flux-density extraction

For detected sources, flux densities are usually extracted by running a source finder that identifies a source lying significantly above the noise. The second step is then to quantify the integrated flux density of the sources, either using the fitting of multiple Gaussians (e.g. PYBDSF; Mohan & Rafferty 2015) or by flood filling to a certain level above the background noise [e.g. BLOBCAT (Hales et al. 2012) and PROFOUND (Robotham et al. 2018; Hale et al. 2019)].

The challenge here is that most of the stellar-mass selected sources do not have a radio counterpart above the detection threshold. The simplest approach would be to use an aperture centred at the NIR position and measure total flux density by summing the individual flux densities per pixel, accounting for the beam area. The size of the aperture plays an important role because if it is too big compared to the projected size of the galaxy, then there will be increased contribution from noise, and there is a greater probability that the measured flux will also include a contribution from nearby objects (that can potentially introduce a bias to our results, as our technique, described in Section 3.2, assumes that only one galaxy is contained in the aperture). If the size is too small, then the flux density of the galaxy might be underestimated. The extraction aperture should therefore be as close as possible to the expected size of the galaxies. In this paper, we use a square with size of 9×9 pixels (1.8×1.8 arcsec^2), which is large enough to fully contain galaxies over the redshift range of interest ($z > 0.5$), based on several studies of radio continuum sizes of μJy radio sources (e.g. Guidetti et al. 2017; Murphy et al. 2017;

Figure 2. Comparison between the VLA-COSMOS 3-GHz flux densities extracted around the position of the stellar-mass-selected galaxy sample and those from Smolčić et al. (2017b), for sources that are formally detected. The solid lines denote the one-to-one relation.

Bondi et al. 2018; Cotton et al. 2018; Jiménez-Andrade et al. 2019), and small enough to avoid contamination from background sources.

In Fig. 2, we compare our extracted flux densities at the positions of our stellar-mass-selected galaxy sample with the detected VLA-COSMOS 3-GHz (Smolčić et al. 2017a) sources. Our measured flux densities scatter uniformly (in log-scale) around the one-to-one line at faint flux densities (i.e. $S_{3\,GHz} < 30$ μJy). Above $S_{3\,GHz} \sim 40$ μJy our extracted flux densities are in good agreement with the VLA-COSMOS 3-GHz flux densities with a few outliers. The outliers are resolved sources that are larger than our aperture (predominantly large low-redshift SFGs and more distant extended AGN). We obtain a better estimate of the flux density at $S_{3\,GHz} > 40$ μJy when we use a larger extraction box, however, we then overestimate the flux density for sources below $S_{3\,GHz} \sim 20$ μJy, which are the main focus of this paper.
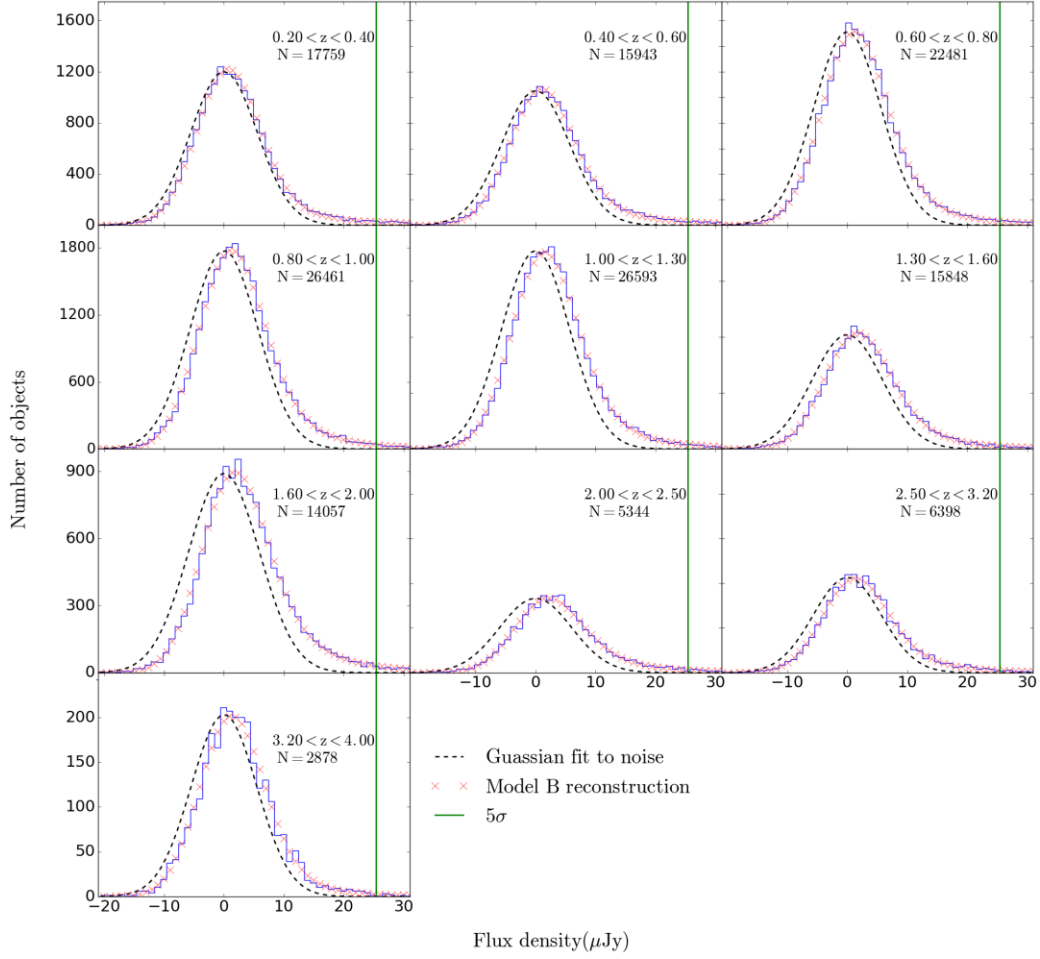
Since our focus in this paper is on sources below the nominal detection threshold we use the 9×9 pixels box, and for sources above 0.5 mJy we use the flux densities measured by Smolčić et al. (2017a). Fig. 3 shows the extracted flux densities for all of the stellar-mass-selected sources in each redshift bin. We also show the reconstructed flux-density distribution obtained using Model B (the best-fitting model in Section 4.2 and shown in Fig. 4). We note that the 3-GHz flux densities follow a Gaussian distribution with an offset from zero and a tail to brighter flux densities, as one would

expect. Malefahlo et al. (2020) showed that the lower the noise is relative to the average flux density then the larger the offset is between the noisy flux densities and the intrinsic Gaussian noise. In our stellar-mass-selected sample, the offset is smaller at the lower redshifts and increases with redshift. This increase in the offset with redshift suggests that the average flux density of undetected sources is increasing.

To measure noise of the image in our chosen aperture, we extract the flux densities from 9×9 pixels boxes centred 50 arcsec from the NIR positions of the galaxies and then fit a Gaussian to the flux-density histogram, which as expected is centred at 0 μJy. We find that the width of the best-fitting Gaussian is $\sigma_{3\,GHz} = 5.08 \pm 0.07$ μJy (equivalent to noise at 1.4 GHz of $\sigma_{1.4\,GHz} = 8.66$ μJy).

3 BAYESIAN FRAMEWORK FOR MEASURING THE RLF

We use a 'stacking' analysis to probe the RLF below the 3-GHz VLA-COSMOS nominal detection threshold. In this analysis we start with a model for the RLF as a function of redshift, which is then converted into a sources-count model that is directly compared/fitted to the binned flux densities (Malefahlo et al. 2020). Below is a summary of the technique.

Figure 3. The histograms of the VLA-COSMOS 3 GHz integrated flux density extracted from boxes (9×9 pixels) centred at the NIR positions. The black dashed curve centred at zero is a Gaussian fit to the flux densities extracted from boxes centred 50 arcsec from the NIR positions in each redshift bin. The Gaussians have mean $\sigma = 5.08 \pm 0.07$ μJy (8.66 ± 0.21 μJy at 1.4 GHz) over all the redshift bins. The red crosses represent the reconstruction of the flux-density distribution using Model B. The green vertical line in each panel represents the $5\sigma = 25.4$ μJy (43.3 μJy at 1.4 GHz) limit of our extracted flux densities.

3.1 Bayesian analysis

The stacking technique is set in a full Bayesian framework (henceforth, we will call the Bayesian stacking technique BAYESTACK). Hence, we start off with Bayes' theorem, which states the relationship between parameters (θ), the model ($\mathcal{M}$), and data ($\mathcal{D}$),

$$P(\theta|\mathcal{D}, \mathcal{M}) = \frac{P(\mathcal{D}|\theta, \mathcal{M})P(\Theta|\mathcal{M})}{P(\mathcal{D}|\mathcal{M})} = \mathcal{P} = \frac{\mathcal{L}\Pi}{\mathcal{Z}}, \quad (2)$$

where $\mathcal{P}$, $\mathcal{L}$, Π, and $\mathcal{Z}$ represent the posterior distribution of the parameters, the likelihood, the prior, and the Bayesian evidence, respectively.

3.2 The likelihood function

The first term in the numerator of equation (2) means that we need the likelihood for the data given a model, where the data we have are the binned extracted flux densities (e.g. Fig. 3). These extracted flux densities (S_E) are the sum of the true flux density (S) of the stellar-mass-selected galaxy and the associated noise (n) in the radio image. For this analysis, we assume that n follows a Gaussian distribution, centred at 0 μJy with a constant variance σ_n^2, which we found to be

$\sigma_n = 5.08$ μJy for our data set[2] (Section 2.3). Since we are working with binned data, we consider a Poisson distribution to be our likelihood function. As such, the likelihood of getting k_i galaxies in the ith measured flux-density bin $[S_{E_i}, S_{E_i} + \Delta S_E]$ follows a Poisson distribution:

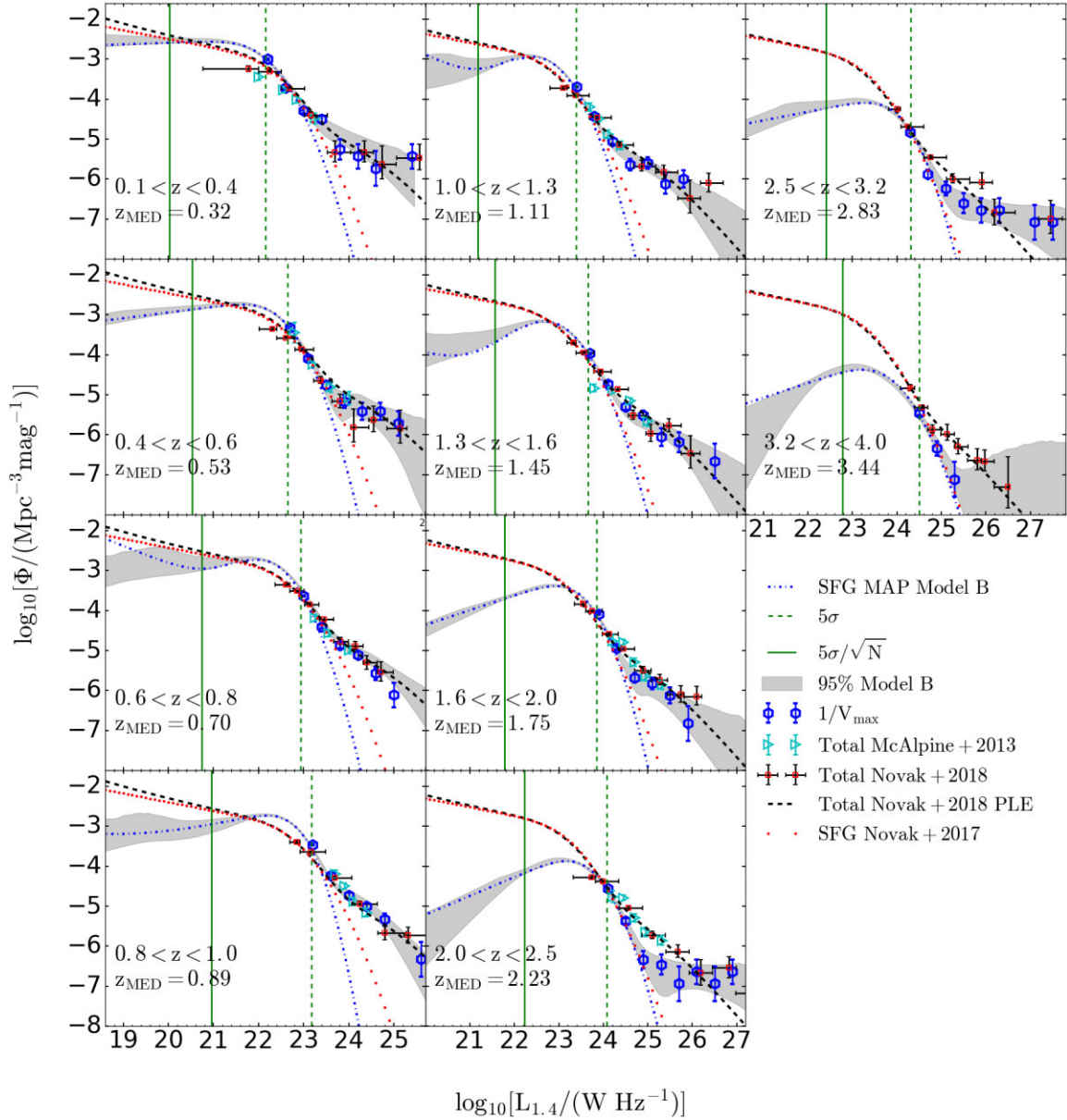$$\mathcal{L}_i(k_i|\boldsymbol{\Theta}) = \frac{I_i^{k_i} e^{-I_i}}{k_i!}, \quad (3)$$

where I_i is the expected number of sources in the ith measured flux-density bin, and is given by (Mitchell-Wynne et al. 2014)

$$I_i = \int_{S_{min}}^{S_{max}} dS \frac{dN(S)}{dS} \int_{S_{E_i}}^{S_{E_i} + \Delta S_{E_i}} dS_E \frac{1}{\sigma_n \sqrt{2\pi}} e^{-\frac{(S - S_E)^2}{2\sigma_n^2}}. \quad (4)$$

Here S and S_E are the intrinsic and extracted flux densities, respectively, S_{min} and S_{max} are the minimum and maximum flux density wherein the model is valid, dN/dS is the model for the number of sources per flux-density bin (source-count model), and σ_n^2 is variance of the noise. Assuming that the likelihood in each bin is independent,

---

[2]This assumption is only valid in the central 2 deg^2 of the COSMOS field where the noise of the VLA-COSMOS 3-GHz is relatively homogeneous (Smolčić et al. 2017a).

Figure 4. The rest-frame 1.4-GHz RLF for both AGN and SFGs in the COSMOS field. The blue dash–dotted is the SF RLF reconstructed using maximum a posteriori (MAP) parameters from the lognormal power-law (Model B) fit to each redshift bin. The grey region is the 95 per cent confidence interval of the distribution of reconstructions of models in the posterior. The blue hexagons represent $1/V_{max}$ estimates for our detected sources. The red squares are radio-selected RLF data points from Novak et al. (2018), with the curved black dashed line showing a pure luminosity evolution (PLE) fit to them. The cyan triangles represent the total RLF from McAlpine et al. (2013). The red dots show a PLE fit to SFGs from Novak et al. (2017). The vertical green dashed lines correspond to the detection threshold (5σ). The vertical green solid lines represent the $5\sigma/\sqrt{N}$, which is an indicator of the luminosity below which the 95 per cent region is expected to be larger than shown. Both 5σ and $5\sigma/\sqrt{N}$ are calculated using the median redshift for each redshift bin.

the combined likelihood for the all N bins is then

$$\mathcal{L}\left(\boldsymbol{k}|\boldsymbol{\Theta}\right) = \prod_{i=1}^{N} \mathcal{L}_i\left(k_i|\boldsymbol{\theta}\right). \tag{5}$$

3.3 Radio luminosity function models

The radio luminosity functions (RLFs) in this paper are all presented at 1.4 GHz for ease of comparison to the literature. The luminosity of a radio source, $L_{1.4\,GHz}$, is related to the observed 3 GHz flux density

through

$$L_{1.4\,GHz} = 4\pi D_L^2 (1+z)^{-\alpha-1} (1.4/3.0)^\alpha S_{3.0\,GHz}, \tag{6}$$

where α is the spectral index of the source (assumed to be $\alpha = -0.7$, which is typical of SFGs), D_L is the luminosity distance, and z is the (photometric) redshift of the galaxy. The parametric models for the RLF we use all consist of two functions, one for AGN (depicted by subscript 1) and the second for SFGs (depicted by subscript 2). The RLF of AGN is well known to follow a double power law (e.g. Willott et al. 2001; Mauch & Sadler 2007; Prescott et al. 2016)

and we therefore parametrize it as such. We consider two models to describe the RLF of SFG: a double power law and a modified Schechter function (lognormal power law).

Model A: the parametric function for this model consists of a double power law for both the AGN and SFG,

$$\Phi(L)_A = \frac{\Phi_1^*}{\left(L/L_1^*\right)^{\alpha_1} + \left(L/L_1^*\right)^{\beta_1}} + \frac{\Phi_2^*}{\left(L/L_2^*\right)^{\alpha_2} + \left(L/L_2^*\right)^{\beta_2}}. \quad (7)$$

Model B: the parametric form for this model consists of a lognormal power law for the SFG (e.g. Tammann, Yahil & Sandage 1979) and a double power law for the AGN,

$$\Phi(L)_B = \frac{\Phi_1^*}{\left(L/L_1^*\right)^{\alpha_1} + (L/L_1*)^{\beta_1}}$$
$$+ \Phi_2^* \left(\frac{L}{L_2^*}\right)^{1-\delta} \exp\left[-\frac{1}{2\sigma_{LF}^2} \log_{10}^2\left(1 + \frac{L}{L_2^*}\right)\right]. \quad (8)$$

Models A and B are both fit to the individual redshift bins, rather than assuming a fixed shape and adopting an evolution term to fit the same parametric model across all redshift bins. In order to explore a model of fixed functional form that evolves with redshift, and to facilitate comparison with previous work (e.g. McAlpine, Jarvis & Bonfield 2013; Novak et al. 2017, 2018), we also adopt an additional model, 'Model C'. Model C has a total RLF of fixed shape, defined by combining the local SFG and AGN RLFs, but allowed to evolve with redshift. We use the local AGN RLF model (Φ_0^{AGN}) and parameters from Mauch & Sadler (2007), where they constrain both the bright and faint ends of the AGN population. They fit their RLF with a double power law (first function of equation 8), with best-fitting parameters, $\phi_1^* = 10^{-5.5}$ Mpc^{-3} mag^{-1}, $L_1^* = 10^{24.59}$ W Hz^{-1}, $\alpha_1 = 1.27$, and $\beta_1 = 0.49$. For the SFGs in Model C, we use the local SFG RLF (Φ_0^{SF}) from Novak et al. (2017) obtained by fitting a lognormal power law to combined data from Condon, Cotton & Broderick (2002), Best et al. (2005), and Mauch & Sadler (2007), which contain low-resolution and deep high-resolution information to constrain both the faint and bright ends of the SFG RLF. Using an analytical function in the form of a lognormal power law (second function of equation 8) their best-fitting parameters are $\Phi_2^* = 1.42 \times 10^{-3}$ Mpc^{-3} mag^{-1} (scaled to our binning), $L_2^* = 1.85 \times 10^{21}$ W Hz^{-1}, $\delta = 1.22$, and $\sigma_{LF} = 0.63$.

The most common ways to quantify evolution in the RLF are through density or luminosity evolution, although we note that the true evolution is probably a mixture of the two (e.g. Yuan et al. 2016). Density evolution causes a vertical shift in the RLF with redshift, and luminosity evolution causes a horizontal shift with redshift. The SFGs and AGN are known to evolve differently, hence we evolve these two populations separately. The combined density- and luminosity-evolution fit is known to have large degeneracies when the knee of the RLF for SFGs is not well constrained, and pure density evolution (PDE) can overestimate sources at low luminosities (e.g. Novak et al. 2017, 2018). Therefore, we only consider a pure luminosity evolution (PLE) of the form

$$\Phi(L, z) = \Phi_0^{SF}\left[\frac{L}{(1+z)^{\alpha_L^{SF} + z\beta_L^{SF}}}\right]$$
$$+ \Phi_0^{AGN}\left[\frac{L}{(1+z)^{\alpha_L^{AGN} + z\beta_L^{AGN}}}\right], \quad (9)$$

where $\alpha_L^{SF,AGN}$ and $\beta_L^{SF,AGN}$ are the evolution parameters.

As the likelihood (equation 4) is computed in flux-density space, we need to convert models describing the RLF, $\Phi(L, z)$, to the source-

Table 2. Assumed priors. $L_{5\sigma}$ is the luminosity corresponding to the $5\sigma_n$ flux-density cut for a given redshift.

Parameter	Prior
$\alpha_2, \beta_2, \delta$	Uniform $\in [-5, 5]$
α_1, β_1	Uniform $\in [-5, 5]$
σ_{LF}	Gaussian $\sim (\mu = 0.6, \sigma = 0.1)$
$\log_{10}[L_{\min_{\{1,2\}}}/(\text{W Hz}^{-1})]$	Uniform $\in [16, 23]$
$\log_{10}[L_{\max_{\{1,2\}}}/(\text{W Hz}^{-1})]$	Uniform $\in [25, 30]$
$\log_{10}[\phi_{\{1,2\}}^*/(\text{Mpc}^{-3}\,\text{mag}^{-1})]$	Uniform $\in [-12, -2]$
$\log_{10}[L_1^*/(\text{W Hz}^{-1})]$	Uniform $\in [23.5, 30]$
$\log_{10}[L_2^*/(\text{W Hz}^{-1})]$	Uniform $\in [18, 24]$
σ_n	$\delta(\sigma_{\text{survey}})$

counts models (dN/dS):

$$\frac{dN}{dS} = \frac{dN}{dL}\frac{dL}{dS}$$
$$= \rho(L)V_M 4\pi D_L^2 (1 + z_M)^{-\alpha - 1}$$
$$= \frac{\Phi(L)V_M}{L \ln(10^{0.4})} 4\pi D_L^2 (1 + z_M)^{-\alpha - 1}, \quad (10)$$

where z_M is the median (photometric) redshift of the redshift bin (Table 1), D_L is the comoving distance corresponding to z_M, and V_M is the comoving volume of the survey in the redshift bin. We use the median values because we are working with binned data and the individual source information is lost (in the binning process). This then introduces a small error in the $(1 + z)^{-\alpha - 1}$ factor and in D_L. As a consequence, this approach works well with small redshift bins and large samples.

3.4 Priors

We assign a uniform prior to the power-law slopes $\alpha_{1,2}$, $\beta_{1,2}$, and δ. The parameter σ_{LF} is assigned a Gaussian prior. We avoid degeneracies in the slopes for the double power law, by imposing $\alpha_{1,2} \geq \beta_{1,2}$. The parameters $L_{1,2}^*$, $L_{\min_{1,2}}$, $L_{\max_{1,2}}$, and $\phi_{1,2}^*$ all have uniform priors in log-space. We impose an additional prior on the AGN break (L_1^*), in that it must never be less than $\log_{10}[L_1^*/(\text{W Hz}^{-1})] \sim 23$. We adopt this 'tight' prior because the bright end of our RLF is not always well constrained by our data. Furthermore, the prior is justified because the AGN RLF is well explored in the literature and the break is found at luminosities well above the luminosity corresponding to 5σ at all redshifts considered (e.g. Smolčić et al. 2017c; Ceraj et al. 2018). The priors are shown in Table 2.

3.5 Sampling and Bayesian evidence

Sampling is generally computationally expensive, especially when the evidence is calculated to enable model selection. We use nested sampling (Skilling 2004) to compute the posterior distribution, which significantly reduces sampling computational time.

Additionally, nested sampling produces the Bayesian evidence as a by-product. The Bayesian evidence is useful in quantifying which model between A and B better fits the data. This is done using the Bayes factor, which is the difference of their log-evidence, $\ln(\mathcal{Z}_B) - \ln(\mathcal{Z}_A)$. The Bayes factor allows one to say that Model A is (i) 'not significantly' when $\Delta \ln \mathcal{Z} < 1$, (ii) 'significantly' when $1 < \Delta \ln \mathcal{Z} < 2.5$, (iii) 'strongly' ($2.5 < \Delta \ln \mathcal{Z} < 5$), and (iv) 'decisively' ($\Delta \ln \mathcal{Z} > 5$) better than Model B (Jeffreys 1961). We adopt this convention when comparing the different models RLFs. We run BAYESTACK through a PYTHON implementation of PYMULTINEST (Buchner et al. 2014).

Table 3. The relative evidence for the different models (Section 3.3) in each redshift bin of the NIR-selected radio data. In each redshift bin the reference evidence is from the model with the lowest log-evidence and the winning model is in bold. For Model C, we provide the evidence for the fits to the individual redshift bins.

Model	$\triangle \log_{10} \mathcal{Z}$	$\triangle \log_{10} \mathcal{Z}$	$\triangle \log_{10} \mathcal{Z}$	$\triangle \log_{10} \mathcal{Z}$	$\triangle \log_{10} \mathcal{Z}$
	$0.10 < z < 0.40$	$0.40 < z < 0.60$	$0.60 < z < 0.80$	$0.80 < z < 1.00$	$1.00 < z < 1.30$
A	45.5 ± 0.27	41.1 ± 0.27	$\mathbf{65.5 \pm 0.28}$	60.9 ± 0.25	29.8 ± 0.28
B	$\mathbf{47.7 \pm 0.26}$	$\mathbf{44.3 \pm 0.27}$	62.8 ± 0.28	$\mathbf{62.2 \pm 0.27}$	$\mathbf{32.4 \pm 0.28}$
C	0.0 ± 0.00	0.0 ± 0.00	0.0 ± 0.00	0.0 ± 0.00	0.0 ± 0.00
	$1.30 < z < 1.60$	$1.60 < z < 2.00$	$2.00 < z < 2.50$	$2.50 < z < 3.20$	$3.20 < z < 4.00$
A	50.9 ± 0.27	$\mathbf{129.1 \pm 0.27}$	100.8 ± 0.26	$\mathbf{95.1 \pm 0.27}$	15.3 ± 0.24
B	$\mathbf{61.3 \pm 0.27}$	128.9 ± 0.27	$\mathbf{102.6 \pm 0.26}$	94.0 ± 0.26	$\mathbf{17.3 \pm 0.23}$
C	0.0 ± 0.00	0.0 ± 0.00	0.0 ± 0.00	0.0 ± 0.00	0.0 ± 0.00

4 RESULTS

In this section, we provide a binned RLF for the radio-detected sources in our mass-selected sample, based on the $1/V_{max}$ statistic, and then present the results of our RLF modelling described in Section 3.3.

4.1 The binned RLF

The RLF for sources with high signal-to-noise ratio (described as 'detected') can be calculated by converting flux density to luminosity (using equation 6) and binning the sources in luminosity. We use the $1/V_{max}$ method (Schmidt 1968),

$$\Phi(L_\nu) = \frac{1}{\Delta m} \sum_{i=1}^{N} \left(\frac{1}{V_{max}} \right)_i , \qquad (11)$$

with an uncertainty

$$\sigma(\Phi) = \frac{1}{\Delta m} \left[\sum_{i=1}^{N} \left(\frac{1}{V_{max}} \right)_i^2 \right]^{1/2} , \qquad (12)$$

where V_{max} is the maximum comoving volume at which the source can be detected given the depth of the data. We assume that the radio sources have a detection in the NIR data, a photometric redshift, and that the value of V_{max} is determined either by the upper limit of the particular redshift bin or by the radio luminosity of the source.

Fig. 4 includes the $1/V_{max}$ measurements for our stellar-mass-selected sources above the nominal 5σ detection threshold. Because of our stellar-mass selection our RLFs are not expected to be exactly the same as the RLF determined using a purely radio-selected sample. However, we note that the VLA-COSMOS 3-GHz sources all have optical/NIR counterparts up to $z \sim 1.5$, and ~95 per cent completeness at $z \sim 4$ (Smolčić et al. 2017b). Given that the main goal of this paper is to measure the RLF for the fainter population of SFGs and how they evolve, this does not affect our results.

Our $1/V_{max}$ data points (dark-blue data points in Fig. 4) are in good agreement with McAlpine et al. (2013) and Novak et al. (2018) measurements for $z < 2$. At $z > 2$ our $1/V_{max}$ points lie below the comoving volume density found in these studies at the intermediate luminosities in the range $24 < \log_{10}[L_{1.4}/\mathrm{W\,Hz}^{-1}] < 26$, where AGN begin to dominate the RLF. This is mainly due to our mass selection, rather than using the full optical/NIR data and the associated photometric redshifts. However, as we note above, this has little effect on our main results.

4.2 The free RLF models

We use BAYESTACK to determine the best-fitting parameters for the RLF of our mass-selected sample using Models A and B (Section 3.3) in each redshift bin. For each redshift bin we record the Bayesian evidence, posterior distributions for each parameter (supplementary material) along with the median, maximum likelihood, and maximum a posteriori (MAP) values for each parameter (shown in the supplementary material). The Bayes factors for each redshift bin are shown in Table 3, where the reference evidence is for the model with the lowest evidence and the model with the highest evidence is in bold text. We find that the data mostly prefer Model B, the model with a lognormal power law describing the faint sources (dominated by SFGs) and a double power law describing the bright-end sources.

In Fig. 4, we show the stellar-mass-selected RLF, reconstructed using the MAP parameters from Model B along with the 95 per cent confidence interval. The 95 per cent region is calculated by reconstructing the RLF at all luminosities using all the models in the posterior, and determining the 95 per cent limits at each luminosity independently. The MAP reconstruction follows the $1/V_{max}$ data points very well above 5σ.

However, the MAP reconstruction lies above the Novak et al. (2017, 2018) extrapolated evolution fits around and below the detection threshold (faint end), with a maximum difference of 0.5 Mpc^{-3} mag^{-1}, at $z \lesssim 1.6$. This is likely due to the fact that the Novak et al. (2017, 2018) RLFs are extrapolated to these low luminosities and not a direct measurement. Conversely, if our mass-selected sample is complete to SFGs at these luminosities, our measurement will be more accurate. We are complete to a stellar mass limit of $M \sim 10^{9.5}$ $\mathrm{M_\odot}$ to $z \sim 1.6$ (see Fig. 1), which is well below the knee in the stellar mass function (e.g. Adams et al. 2021). Using the star-forming main sequence (Johnston et al. 2015) this corresponds to a SFR $\lesssim 10$ $\mathrm{M_\odot}$ yr^{-1}, or a radio luminosity of $L_{1.4\,\mathrm{GHz}} \approx 3 \times 10^{22}$ $\mathrm{W\,Hz}^{-1}$ using the relation from Delhaize et al. (2017), which is indeed below the luminosities sampled by Novak et al. (2018) and where we observe an excess in the RLF in comparison to this previous work. We also note that BAYESTACK naturally accounts for Eddington bias, where the more numerous low-luminosity sources are boosted to higher luminosities due to the noise in the image. To model this accurately, one needs to know the shape of the luminosity function below the detection threshold to account for the number of sources that are boosted above the flux-density limit. BAYESTACK on the other hand fits the luminosity function while fully accounting for the noise, and as such Eddington bias is fully accounted for in the modelling. This could also cause the observed discrepancy between our RLF and that of Novak et al. (2017).

At higher redshifts ($z > 1.6$) the faint end of the reconstructed RLF lies below the extrapolated evolution fit from Novak et al. (2017, 2018). This is due to Novak et al. (2017, 2018) adopting a fixed faint-end slope that extends below their detection threshold (and is essentially fixed by the low-redshift data). Instead, we are using a mass-selected sample with the aim of probing this regime, and so in Models A and B, we allow the faint-end slope of the SFG RLF to vary freely. Thus, as our mass-selected sample becomes less complete to galaxies at or below the knee in the stellar mass function, and hence at or below the knee in the RLF through the relation between stellar mass and SFR, we observe a downturn in our reconstructed RLF. We emphasize that this is not due to a real downturn in the RLF, but due to the stellar-mass selection of our sample; a deeper NIR survey, probing to lower stellar masses would not produce such a downturn.

We also see that the reconstructed RLF, at much lower luminosities (two or more orders of magnitude below the detection threshold, noticeably above $z \sim 0.4$), also drops off steeply. This is due to the mass selection, in that we are approaching the point in the RLF where there are not any galaxies at low stellar mass to populate this part of the RLF, due to the relationship between galaxy mass and SFR (Daddi et al. 2007; Noeske et al. 2007; Whitaker et al. 2014; Johnston et al. 2015). We have checked this by including all NIR-detected sources (rather than using the mass-limited sample) and find that the luminosity where the drop-off occurs moves to lower luminosities, as expected. This shows that it is not a feature of the RLF, but a feature of the parent sample, due to the lack of low-stellar-mass sources in our sample, and this stellar-mass limit obviously increases with redshift due to the flux limit of the NIR data.

We also note that some of the difference at these higher redshifts could be due to high-mass galaxies that are missing from our observations. Recent studies have found that ~ 10–30 per cent of luminous and dusty galaxies (sub-mm galaxies, SMG), which peak in the redshift range $1.8 < z < 3.4$ (Chapman et al. 2005; Simpson et al. 2014, 2020; Dudzevičiūtė et al. 2020), are missing from deep NIR surveys (Simpson et al. 2014; Brisbin et al. 2017; Cowie et al. 2018; Dudzevičiūtė et al. 2020), even though they are mainly high stellar mass galaxies that primarily reside at the bright end of the SFG RLFs (see Fig. 8). These galaxies may also therefore play a role in our underestimation of Φ at higher redshifts.

If one were to conduct a simple mean-stacking experiment, we would expect to obtain constraints on the mean flux down to $\sim \sigma/\sqrt{N}$. We are fitting a parametric form, and as such, the specific depth for which we are able to obtain good constraints on the RLF is less well defined. This is the principal reason we fit for L_{min}. However, we note that there is a limit imposed by our chosen parametric forms and priors. Once the faint-end slope is set by the large number of sources just below the knee in the RLF, then it is limited in how much it can vary by the total number of sources in the parent sample (modulo Poisson uncertainties), and by the lower luminosity limit that the model is allowed to fit to (i.e. the prior on L_{min}). We checked the former by introducing a third parametric function in our model to ensure that given enough flexibility the model would produce larger confidence regions at these faint luminosities, we found that it did and that the luminosity below which the confidence region expands significantly is around $\sim \sigma/\sqrt{N}$, as one would expect. In Fig. 4, we therefore show a line that denotes the $L \sim L[S = 5\sigma/\sqrt{N}, z]$ limit in each redshift bin to inform the reader of the luminosity range above which the model is robust. We note that we found that setting the lower limit on the prior range at $L_{min} = 0$ W Hz^{-1} led to a prohibitive increase in the run time to sample the posteriors (essentially expanding the prior range by 16 dex in luminosity), thus we retained our original priors but note this limitation.

4.3 The fixed RLF model

Our main goal is to measure the RLF to low radio luminosities to obtain a measurement of the cosmic SFRD from a stellar-mass-selected sample. Through the BAYESTACK technique we are able to constrain the RLF to luminosities below the nominal 5σ threshold. However, as shown in Section 4.2, our mass selection causes the free-fitting models to drop off towards lower radio luminosities. This is not an underlying feature of the SFG RLF, and will therefore affect the cosmic SFRD estimation. To address this we follow the work of McAlpine et al. (2013) and Novak et al. (2017, 2018) in fixing the shape of the RLF to that of the local RLF (Section 3.3).

We start by modelling the individual redshift bins using the fixed model (Model C), with $\beta_{LF}^{SF,AGN} = 0$ (i.e. only allowing a renormalization of the RLF in each redshift bin, with a single luminosity evolution term). The resulting posterior plots are provided in the supplementary material and the RLF is shown in Fig. 5. Table 3 shows that Model C is (almost) always the least preferred model (having the lowest log-evidence). This is because Model C *forces* a fixed faint-end slope. For our mass-selected sample the fall-off in the number of sources at the low-mass end, and therefore at low radio luminosity, means that this fixed slope will struggle to produce a fit as good as the models with more freedom, as it assumes that the lower mass galaxies are in the sample. Thus, formally it is the worst fitting model, even though it may accurately represent the underlying RLF.

We also run BAYESTACK simultaneously over all of the redshift bins using the PLE model. The PLE for our AGN and SFG galaxies is given by the following MAP values and 95 per cent confidence limits,

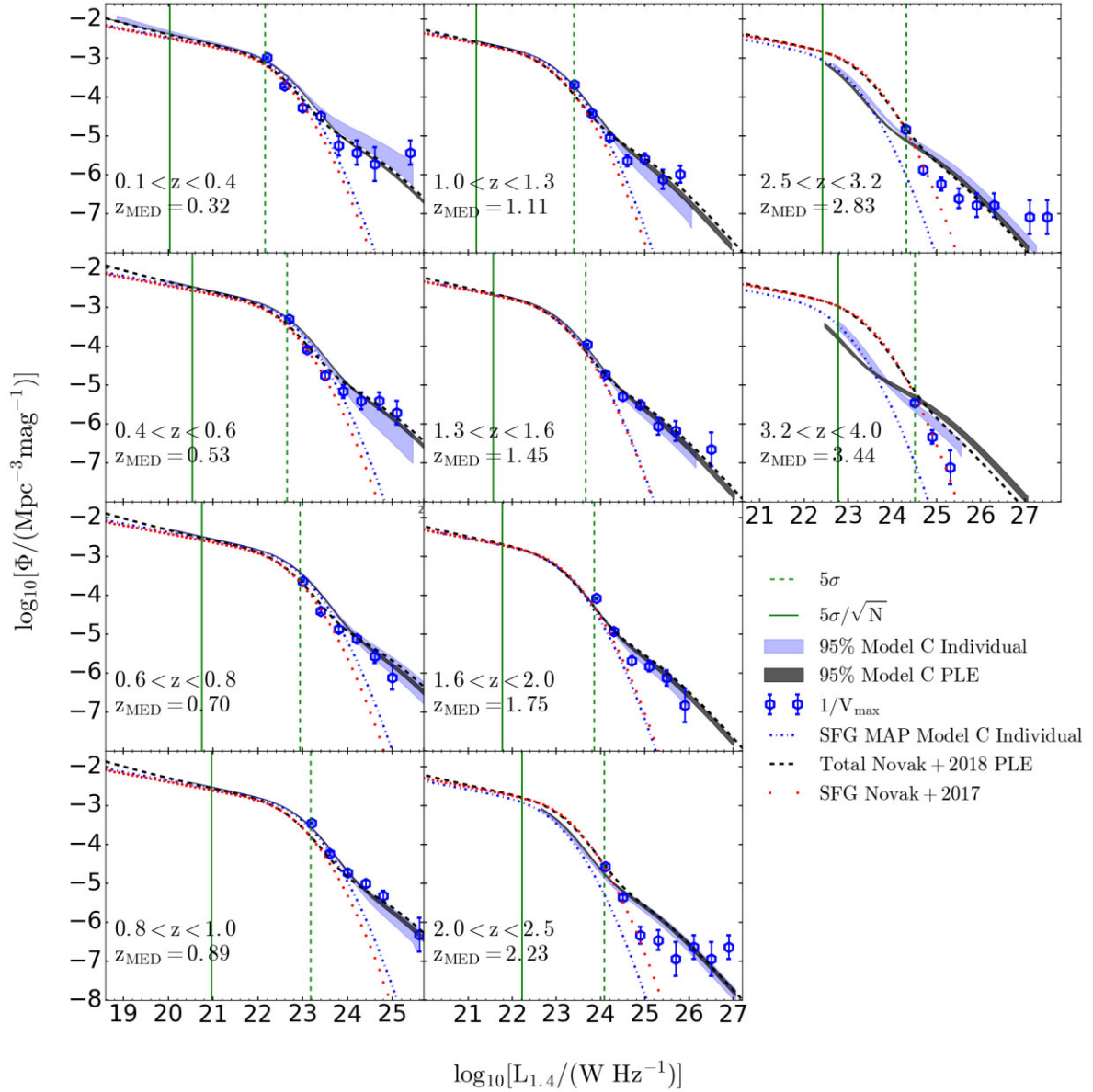$$L_{AGN} \propto (1 + z)^{2.03 \pm 0.15 - (0.35 \pm 0.08)z}$$

and

$$L_{SFG} \propto (1 + z)^{4.57 \pm 0.04 - (1.19 \pm 0.02)z}.$$

In Fig. 5, we show the RLF fits and 95 per cent confidence intervals for the individual redshift bins, alongside the PLE RLF model fits, both with the fixed RLF shape. The PLE RLF model agrees with the $1/V_{max}$ data points across all redshifts up to $z \sim 2$. At the highest redshifts ($z > 2$) we find some differences between our model (plus our binned data) and models of Novak et al. (2017, 2018). At the high radio luminosities, small number statistics, coupled with slight differences in the photometric redshifts used by us and Novak et al. (2017), offers some explanation as to why our RLF lies below theirs. Furthermore, we note that we are also becoming increasingly incomplete at these redshifts, and sources that are relatively bright at radio wavelengths could have lower mass/faint host galaxies (e.g. Jarvis et al. 2009; Norris et al. 2011; Collier et al. 2014). However, more relevant to the focus of this work are the differences in the evolution of the lower luminosity component of the RLF, which we assume to be dominated by SFGs.

The SFG component of our RLF model evolves with a similar strength to that of Novak et al. (2017) up to $z \sim 1.6$, with the degeneracy between α_L^{SF} and β_L^{SF} across this redshift range explaining the apparent difference in the evolutionary parameters (Table 4). However, beyond $z \sim 2$, the Novak et al. (2017) RLF continues to evolve, whereas we find that the SFG RLF from our model reaches a steady state and then begins to decline (for the Novak et al. 2017 PLE model, this decline does not take effect until $z > 3.5$).

We note that this decline coincides with the decrease in low stellar-mass sources in our stellar-mass-limited sample at these redshifts. However, it is also worth mentioning that the total RLF (AGN+SFGs) does continue to be a reasonable fit to the binned
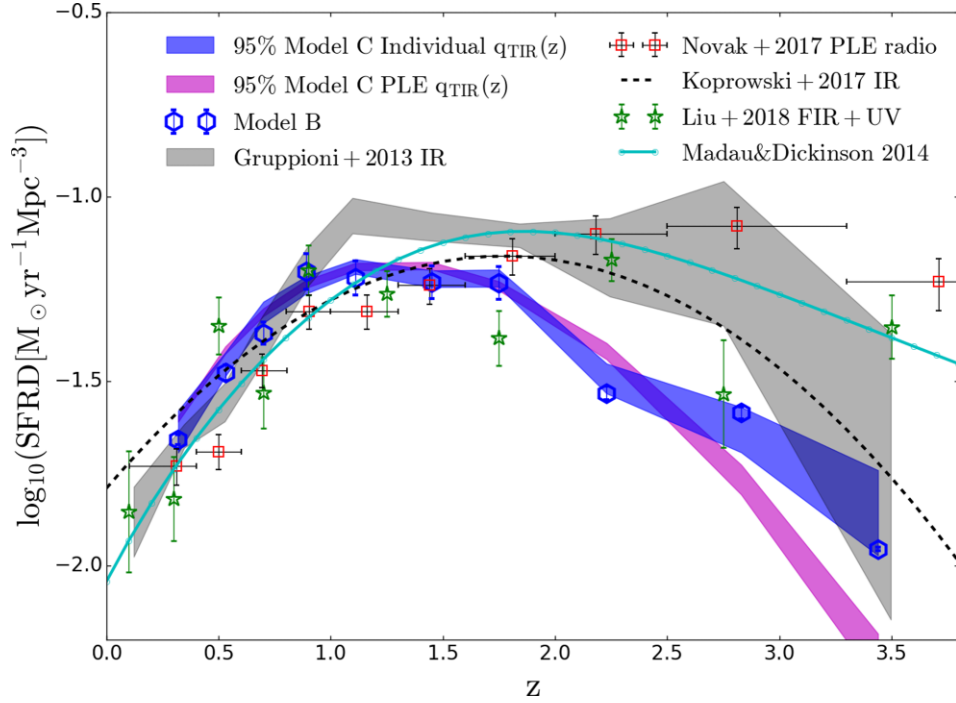
Figure 5. The rest-frame RLF of both AGN and SFGs in the COSMOS field. The blue hexagons represent $1/V_{max}$ estimations for the UltraVISTA sources with VLA-COSMOS 3-GHz detected sources. The dark grey and blue regions represent the 95 per cent confidence interval of the distribution of reconstructions for Model C pure luminosity evolution (PLE) fit and Model C individual fit to each redshift bin, respectively. The green and blue dashed–dotted lines represent the SFG components of the total RLF of the Model C PLE fit and Model C individual fit to each redshift bin, respectively. The curved, black dashed line representing the PLE fit to the radio-selected RLF from Novak et al. (2018). The red dots are a PLE fit to the SFGs from Novak et al. (2017). The vertical, green dashed lines correspond to the detection threshold (5σ) computed using the median redshift for each redshift bin. The vertical green solid lines represent the $5\sigma/\sqrt{N}$, which is an indicator of the luminosity below which the 95 per cent region is expected to be larger than shown. Both 5σ and $5\sigma/\sqrt{N}$ are calculated using the median redshift for each redshift bin.

Table 4. Comparison with determinations in the literature of the pure luminosity evolution (PLE) of the radio luminosity function.

Reference	Description	α_L^{SF}	β_L^{SF}	α_L^{AGN}	β_L^{AGN}
This work	Total RLF fit	4.57 ± 0.04	-1.19 ± 0.02	2.03 ± 0.15	-0.35 ± 0.08
Novak et al. (2018)	Total RLF fit	2.95 ± 0.04	-0.29 ± 0.02	2.86 ± 0.16	-0.70 ± 0.06
Novak et al. (2017) and Smolčić et al. (2017c)	Individual SF and AGN fit	3.16 ± 0.04	-0.32 ± 0.02	2.88 ± 0.17	-0.84 ± 0.07
McAlpine et al. (2013)	Total RLF fit	2.47 ± 0.12^a		1.18 ± 0.21^a	

[a]McAlpine et al. (2013) fit only one evolutionary term each for the SF and AGN.

Figure 6. The cosmic star formation rate density (SFRD). The blue hexagons are generated from the SFG MAP values for the Model B fit to each individual redshift bin. The blue shading corresponds to the 95 per cent confidence region of the SFG component of the Model C fit to each redshift bin (individually). The magenta shading corresponds to the 95 per cent confidence region of the SFG component of the Model C PLE fit to the combined redshift bins. The green stars are the combined uncorrected IR and UV data from Liu et al. (2018) and the red squares are from the SFG RLF of Novak et al. (2017), using a PLE fit. The black dashed line is from Koprowski et al. (2017) and the connected cyan curve represents a fit by Madau & Dickinson (2014) to various cosmic SFRD measurements in the literature. The grey shaded region is the cosmic SFRD generated from the IR LF of Gruppioni et al. (2013).

$1/V_{max}$ points of Novak et al. (2018) out to $z \sim 2.5$, suggesting that some of the deficit in the low-luminosity RLF is compensated for by the evolving high-luminosity RLF that we associate with AGN.

5 COSMIC HISTORY OF STAR FORMATION

The RLFs for SFGs derived from our stellar-mass-selected sample provide a means to obtain the SFRD by integrating under them, i.e.

$$\text{SFRD} = \int_{L_{min}}^{L_{max}} \text{SFR}(L_{1.4}) \Phi(L_{1.4}) \, dL, \tag{13}$$

where $\Phi(L_{1.4})$ is our RLF for SFGs and $\text{SFR}(L_{1.4})$ is the SFR associated with 1.4-GHz radio luminosities. Using the Kennicutt (1998) calibration, the total IR luminosity (L_{TIR}) is related to the SFR by

$$\frac{\text{SFR}}{\text{M}_\odot \, \text{yr}^{-1}} = 4.5 \times 10^{-37} \frac{L_{TIR}}{\text{W}}, \tag{14}$$

where L_{TIR} is the total IR luminosity. The radio luminosity can be converted into the total IR luminosity and linked to SFR using the infrared–radio correlation (IRRC; e.g. Delhaize et al. 2017),

$$\frac{\text{SFR}}{\text{M}_\odot \, \text{yr}^{-1}} = f_{IMF} \times 10^{q_{TIR}-24} \frac{L_{1.4\,\text{GHz}}}{\text{W Hz}^{-1}}, \tag{15}$$

where f_{IMF} is the IMF (equal to 1 for a Chabrier IMF; Chabrier 2003) and q_{TIR} is a parameter that quantifies the IRRC given by

$$q_{TIR} = \log\left(\frac{L_{TIR}}{3.75 \times 10^{12}\,\text{W}}\right) - \log\left(\frac{L_{1.4\,\text{GHz}}}{\text{W Hz}^{-1}}\right). \tag{16}$$

We adopt a q_{TIR} value that evolves with redshift, given by $q_{TIR}(z) = 2.78 \pm 0.02(1 + z)^{-0.14 \pm 0.01}$ (Novak et al. 2017). Although we note that the evolution may be due to a mass dependence of the IRRC (e.g. Gürkan et al. 2018; Delvecchio et al. 2021; Smith et al. 2021), we note that for the purposes of our study, adopting the relation of Delhaize et al. (2017) effectively accounts for the mass dependence in our sample, as that evolving FIRC was derived using similar data. We then obtain the SFRD by numerically integrating the product of the RLF and the SFR over 1.4-GHz radio luminosities (equation 13). For this, the integral should cover all radio luminosities and not just the range dictated by our fitted values of L_{min} and L_{max}. Although we note that this makes little difference in the derived cosmic SFRD, due to the shallowness of the faint-end slope for low SFRs, and the steep exponential decline at high SFRs. With this in mind we use $L_{min} = 10^{21}$ W Hz^{-1} in all redshift bins and the SFG RLF models.

In Fig. 6, we present the cosmic SFRDs obtained using Model B (blue data points) and Model C (blue shading), both from fitting individual redshift bins, and for the PLE fit to Model C (magenta shading).[3] As would be expected, the different models used in fitting the RLF result in slightly different determinations of the cosmic SFRD.

In all our models, the SFRD steadily increases with redshift out to $z \sim 1.2$, flattens to $z \sim 1.7$, and then steadily decreases towards higher redshifts. This shape is due to the fact that the stellar-mass

---

[3] The 95 per cent region is calculated using the 95 per cent confidence interval from the RLF. The conversion error associated with the q value (Novak et al. 2017) has not being accounted for in this calculation.

selection still enables the knee in the SFG RLF to be well constrained to $z \sim 1.5$.

At $z > 2.5$ the stellar-mass limit starts imposing on our ability to constrain the position of the knee in the SFG RLF for all models, and the best-fitting evolutionary terms force the position of the knee to lower radio luminosities in order to fit the incomplete parent sample. All our SFRDs estimates start to steadily decline in the range $1.5 < z < 2$ because of the rising stellar-mass limit with redshift. The SFRD based on the Model C PLE RLF behaves similar to Model B and the individual Model C RLFs but drops off more quickly at the highest redshifts ($z > 2.5$). This is due to the restricted form of the evolution model, once the downturn sets in at $z \sim 1.7$, then the PLE model ensures that the strength of the downturn continues to high redshift.

5.1 Comparison to the literature

In this subsection, we compare our determination of the cosmic SFRD, based on stellar-mass-selected galaxies (constrained below the nominal detection threshold using the BAYESTACK technique), to literature measurements of the cosmic SFRD using a variety of SFR tracers.

5.1.1 Comparison with the radio-selected cosmic SFRD

We first compare our cosmic SFRDs to the cosmic SFRD determined by Novak et al. (2017), which is based on COSMOS2015 photometry and SED fits (Laigle et al. 2016; Delvecchio et al. 2017) and VLA-COSMOS 3-GHz data. Our results for the SFRD generally lie above the Novak et al. (2017) at $z < 1.2$. This deviation is directly due to our RLFs having a higher Φ towards and below the detection threshold in these redshift range compared to those from Novak et al. (2017). As highlighted in Section 4.2, this is because we are able to better constrain the RLF using sources that lie below 5σ in the radio data. This also results in our uncertainties being much tighter in this regime.

At $z < 1.6$ our determination of the SFRD lies below that found by Novak et al. (2017). This is due to the fact that our RLF lies above that of the RLF derived by Novak et al. (2017) due to the reasoning provided in Section 4.2. This suggests that even with the deepest available radio continuum data over $\sim$degree-scale fields, our ability to measure the knee and the faint-end slope of the RLF is severely limited. This clearly demonstrates some of the advantages of using BAYESTACK, where we are able to better constrain these key parameters that describe the RLF to much higher redshift, and are more limited by the depth of the optical and NIR data.

Our SFRDs also deviate from Novak et al. (2017) at $z > 1.6$ as a result of our stellar-mass selection, along with the possibility of missing SMGs in our sample. The impact of this is most apparent in Fig. 5, where the discrepancy between the SF component(s) of our 'Model C' RLF(s) and that of the Novak et al. (2017) RLF increases with redshift. However, although not shown in Fig. 6, at $z > 1.6$ our results are in broad agreement with other radio-based estimates of the cosmic SFRD in the literature (e.g. Smolčić et al. 2009; Karim et al. 2011; Ocran et al. 2020; Lesile et al. 2020). Thus, our results should be considered as complementary to those that use radio selection to measure the RLF. In our case, the incompleteness arises from the stellar-mass selection only, but we are able to directly constrain the faint-end slope to higher redshifts than the pure radio selection. Whereas completeness corrections for radio-selected samples are required for both the radio data (e.g. Eddington and Malmquist bias) and in terms of the ability to identify a host galaxy and measure a
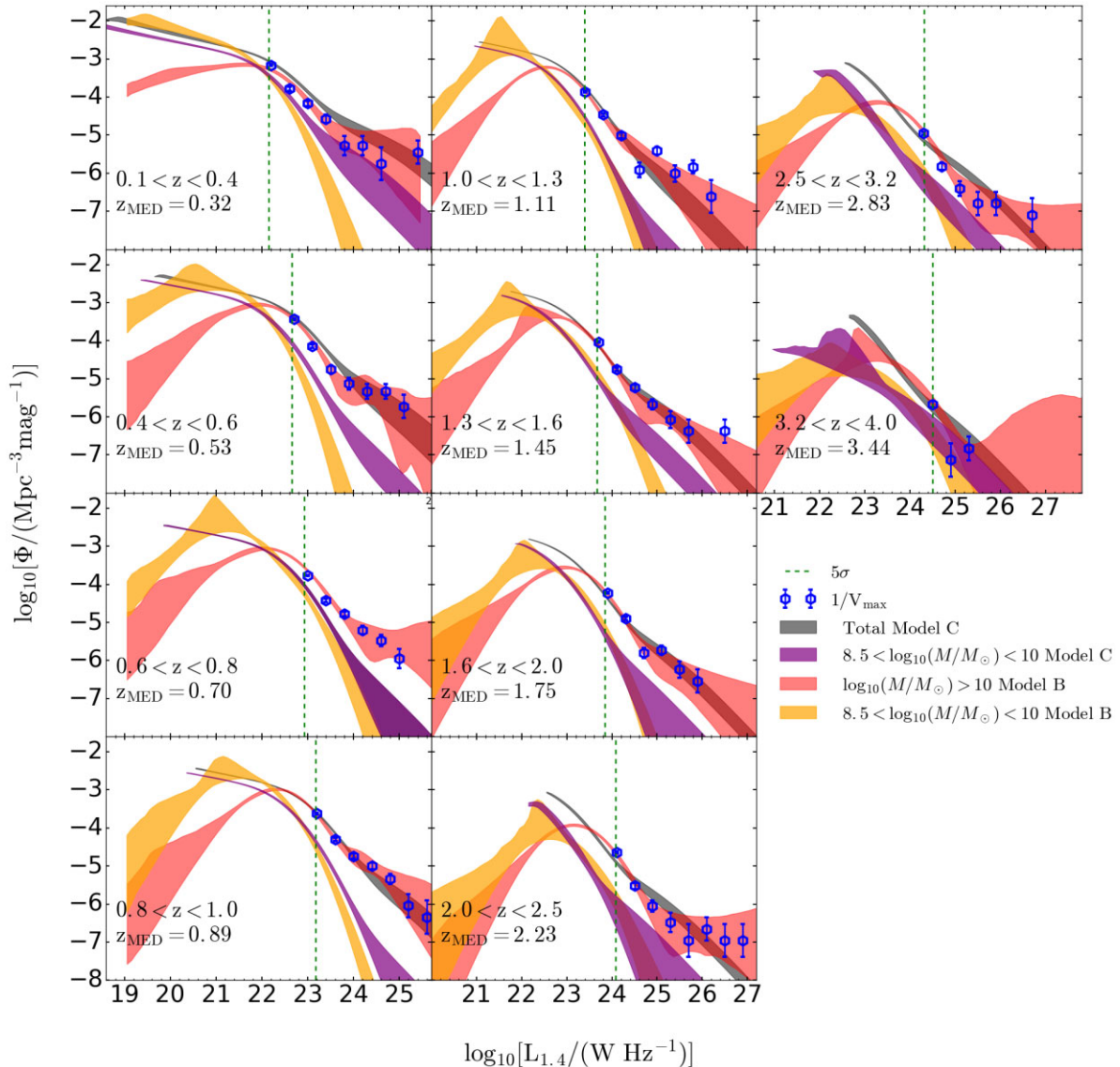
redshift (which is less of a problem for fields with excellent ancillary data, such as COSMOS).

5.1.2 Comparison to other studies

Next, we compare our cosmic SFRDs to the dust-obscured cosmic SFRD from Gruppioni et al. (2013). For this they use a total IR LF based on deep *Herschel* data, from the PACS Evolutionary Probe (PEP; Lutz et al. 2011) and the complementary *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al. 2012), out to $z \sim 4$. We convert their $L_{\rm TIR}$ density (where the $L_{\rm TIR}$ is obtained from an integral over the whole thermal IR spectrum) to the SFRD using equation (14). In Fig. 6, we also plot the IR-based cosmic SFRD from Koprowski et al. (2017), who used *Herschel* FIR flux densities for their LFs, extracted at the positions of sub-mm sources identified using the James Clerk Maxwell Telescope's SCUBA-2 Cosmology Legacy Survey (S2CLS; Geach et al. 2017) and the ALMA (Dunlop et al. 2017) in the COSMOS and UKIRT Infrared Deep Sky Survey (UKIDSS)-UltraDeep Survey (UDS) fields. Our results are in good agreement with Gruppioni et al. (2013) below $z \sim 1$, but then deviate at higher redshifts where the IR SFRD continues to increase (before flattening and then falling around $z \sim 3$). It should be noted that there are a lot of uncertainties in measuring $L_{\rm TIR}$ from a few data points. There are also *k*-correction effects, since, as one goes to higher redshifts we move away from the peak of the thermal emission at $\sim$100 μm in the rest frame. At these high redshifts the $L_{\rm TIR}$ becomes dominated by hotter dust systems, which are more likely to have AGN contributions. This implies that converting from $L_{\rm TIR}$ to SFRD for these systems may lead to an overestimate of the cosmic SFRD. The SFRD by Koprowski et al. (2017) is higher than both of our SFRD determinations at most redshifts, except around $z \sim 1$ where our results overlap. We note that Gruppioni & Pozzi (2019) attributed the discrepancies between the two IR SFRD functions (Gruppioni et al. 2013; Koprowski et al. 2017) to selection bias, incompleteness effects, and the choice of SED in the SCUBA-selected data from Koprowski et al. (2017), which reinforces some of the issues we mention above.

We also compare our results with the cosmic SFRD from Liu et al. (2018), which is derived using a combination of the dust-obscured and unobscured cosmic SFRD measurements. Specifically, Liu et al. (2018) determined their SFRD using superdeblended FIR to sub-mm *Herschel* photometry from confused galaxies in the northern field of the Great Observatories Origins Deep Survey (GOODS). The FIR/sub-mm photometry is extracted based on fitting SEDs to sources selected from deep *Spitzer Space Telescope* Multiband Imaging Photometer (MIPS; Rieke et al. 2004) and 1.4-GHz VLA (Morrison et al. 2010; Owen 2018) data. The derived SFRD from Liu et al. (2018) is in good agreement with our results up to $z \sim 3$ with minor deviations. They agree with our decline above $z \sim 1.6$, this is most likely because their sample is also limited by stellar mass, due to their optical/NIR selection.

We also show the Madau & Dickinson (2014) cosmic SFRD, which is a fit to various cosmic SFRDs in the literature. Our results are again slightly above the Madau & Dickinson (2014) below $z \sim 1$. Our results further deviate at $z > 1$, which is certainly influenced by our stellar-mass selection. However, it should also be noted that we have assumed the IRRC form that evolves negatively with increasing redshift, meaning that for a given radio luminosity, the SFR would be lower at high redshift than at low redshift. This could obviously result in a false decrease in the SFRD if the real IRRC did not evolve with redshift, which other studies have suggested, depending

Figure 7. The contribution to the total rest-frame 1.4-GHz RLF from sources with different stellar masses in the COSMOS field. Low stellar mass ($10^{8.5} < M/M_\odot < 10^{10}$) is represented by the purple and orange shaded regions, corresponding to the 95 per cent confidence interval of the distribution of reconstructions of models in the posteriors. The contribution from sources with high stellar mass ($M > 10^{10} M_\odot$) is represented by the 95 per cent region. The total RLFs based on Model C fit to each redshift bin are represented by the grey shading that corresponds to the 95 per cent region. The blue hexagons represent $1/V_{max}$ estimations for our detected sources. The vertical green dotted lines correspond to the detection threshold (5σ) computed using the median redshift for each redshift bin.

on how the galaxies have been selected (e.g. Molnár et al. 2018). For example, one aspect of this is that Gürkan et al. (2018), Delvecchio et al. (2021), and Smith et al. (2021) all find that the IRRC has a dependence on the stellar mass of the galaxy, and this may be responsible for the observed evolution of the IRRC, as higher redshift samples are inevitably dominated by more massive galaxies due to the nature of flux-limited samples. However, mass is unlikely to be the only extra parameter that needs to be considered when using the IRRC to convert a radio luminosity to SFR, with Smith et al. (2014) and Read et al. (2018) showing that dust temperature, and how you include sensible k-corrections for a range of dust temperatures at different redshifts, can be crucial to measure the SFR.
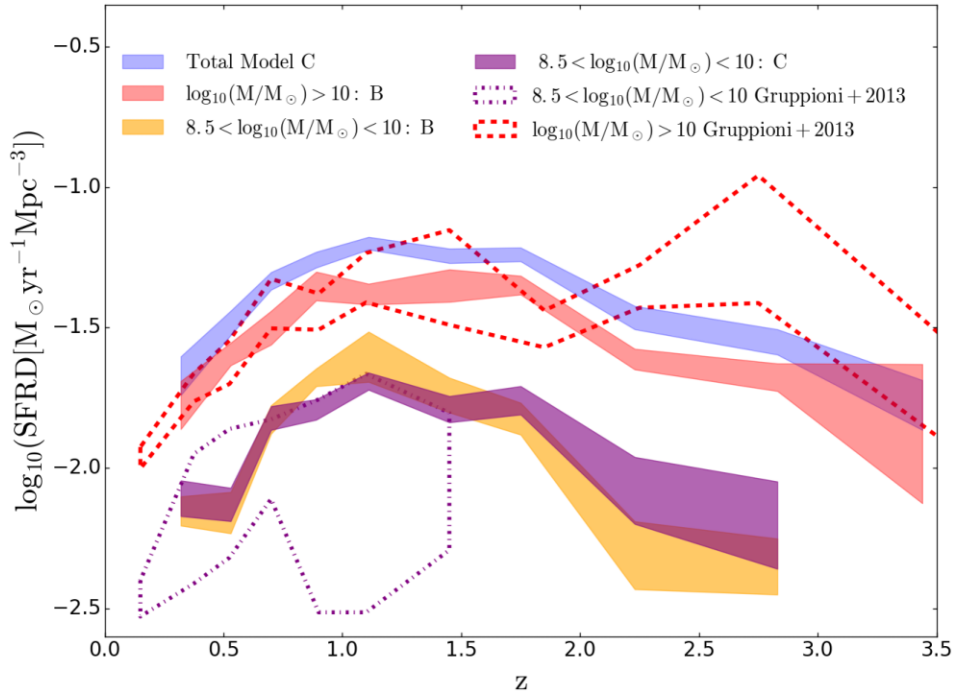
Furthermore, as we move beyond $z \sim 1$, inverse Compton scattering of the cosmic microwave background (CMB) photons may reduce the level of radio emission from SFGs observed at a given (relatively high) frequency (e.g. Murphy 2009). All of these

issues result in our understanding of any evolution in q_{TIR} being uncertain.

5.2 Contribution from different stellar mass populations

As noted previously, the stellar-mass selection we have applied to our sample (Section 2.1.3) means that we miss low stellar-mass ($M < 10^9 M_\odot$) sources at high redshift ($z \gtrsim 1.5$).

To further investigate the effects of stellar mass on the total RLF we divide our sources into low ($10^{8.5} \leq M \leq 10^{10} M_\odot$) and high ($M > 10^{10} M_\odot$) stellar-mass galaxies, shown in Fig. 7 for Models B and C. It should be noted that we are splitting at a stellar mass of $M = 10^{10} M_\odot$ due to the fact that our sample is complete to $M \sim 10^{10} M_\odot$ in our highest redshift bin (Fig. 1). Galaxies with high stellar mass typically have high radio luminosities, as expected, and host a large proportion of the detected radio sources.

Figure 8. Contribution from different stellar-mass populations to the cosmic star formation rate density (SFRD). The blue shading represents 95 per cent confidence region associated with the total cosmic SFRD from Model C fit to each redshift bin (individually). The orange and purple shading are the 95 per cent confidence region from sources with low stellar masses ($10^{8.5} \leq M \leq 10^{10}$ $M_\odot$) from Model C and Model B, respectively. The red shading is the 95 per cent confidence region from sources with high stellar masses ($M > 10^{10}$ $M_\odot$). The red dashed and purple dashed–dotted line are the high ($10^{10} \leq M \leq 10^{12}$ $M_\odot$) and low ($M < 10^{10}$ $M_\odot$) stellar mass contributions from Gruppioni et al. (2013).

The low stellar-mass galaxies typically have low radio luminosities and dominate the RLF below the 5σ detection threshold, for $z < 1.5$. Above $z > 1.5$, the contribution from the low stellar-mass sources decreases due to our stellar-mass completeness limit (see Fig. 1). It is clear that the bulk of the RLF that we are able to measure at $z > 0.4$ is dominated by galaxies with stellar mass $M > 10^{10}$ $M_\odot$.

In Fig. 8, we show the contribution from the low ($10^{8.5} \leq M/M_\odot \leq 10^{10}$) and high ($M > 10^{10}$ $M_\odot$) stellar-mass sources to the total cosmic SFRD. The RLF for low stellar-mass sources (orange and purple shades in Fig. 7) shows that they are not the dominant population contributing to the SFRD at any redshift. However, they are important to include as they are crucial to determine the position of the knee in the RLF (L^*), where the bulk of the SFRD is concentrated, and the steepness of the faint-end slope (with a steep slope resulting in a higher contribution to the SFRD from these faint sources). The fact that they are missing in our sample at high redshifts, means that this can affect our cosmic SFRD estimate.

The contribution to the SFRD from low stellar-mass galaxies increases with redshift up to $z \sim 1$ (for Model B and $z \sim 1.5$ for Model C) where it peaks and drops towards higher redshifts (due to the NIR flux-density limit). The shape of the SFRD derived from just the low stellar-mass galaxies is similar to the total SFRD below $z \sim 1$. However, it is the contribution from high stellar-mass galaxies that dominates the total SFRD at all redshifts, and hence has a shape almost identical to the total (with minor differences).

We compare our results to the cosmic SFRDs from Gruppioni et al. (2013), who divided their sources into three stellar mass (low, mid, and high) bins. Their low stellar mass contribution, $M < 10^{10}$ $M_\odot$, shows larger error bars but are fully in agreement with our results. We compare our high stellar mass contribution with the

combination of their mid ($10^{10} \leq M/M_\odot \leq 10^{11}$) and high stellar mass ($10^{11} \leq M/M_\odot \leq 10^{12}$), which are largely in agreement below $z \sim 2$. Above $z \sim 2$ the Gruppioni et al. (2013) (high stellar mass) sources result in a higher SFRD, with an increasing contribution from starburst galaxies and possibly AGN. This might imply that we are missing these sources in our stellar-mass selection, or it is possible that they instead contribute to the total RLF through the high-luminosity part that we do not use to determine the SFRD.

6 CONCLUSIONS

In this paper, we measure the RLF down to very low radio luminosities and obtain a measurement of the cosmic SFRD from our stellar-mass-limited sample. Using BAYESTACK we probe the stellar-mass-selected RLF orders of magnitude below the nominal 5σ detection threshold by fitting parametric models to the RLF for both SFGs (low-luminosity radio sources) and AGN (high-luminosity radio sources). The reconstructed RLFs follow the $1/V_{max}$ points very well above the detection limit. However, we find that our fitted models for the RLF lie above the Novak et al. (2018) extrapolated PLE around and below the detection threshold, up to $z \sim 1.6$. This is most likely due to the fact that we are able to directly measure the RLF at these low luminosities with our stacking technique, whereas their model is based on detected radio sources only, which means that their RLF below the break is based on an extrapolation using very limited data.

We see that in our best-fitting model (Model B – the unconstrained model), the faint-end slope of our SFG RLF falls-off towards low radio luminosities (below the detection threshold), particularly at the higher redshifts ($z > 1.5$). However, this fall-off is most probably not an underlying feature of the RLF of SFGs but instead the result

of a lack of fainter radio sources in our parent stellar-mass-selected sample. This is due to the known relation between stellar mass and SFRs (e.g. Daddi et al. 2007; Noeske et al. 2007; Johnston et al. 2015), where our stellar-mass-selected sample imposes a natural limit on the level of SF in galaxies we are able to probe. As our stellar-mass limit increases with redshift, due to the flux limit of the optical/NIR data, this means we do not probe as deeply the radio-faint SFGs in our sample.

We try to address this effect from the stellar-mass limit by fixing the shape of the RLF to that of the local RLF and allow it to evolve with redshift, since the local sample should not be affected by this stellar-mass limit. In order to use this local RLF shape, we start by obtaining the RLF in each individual redshift bin, by fitting for the knee in both the SFG and AGN RLFs. We next use a PLE fit (with two luminosity evolution terms) to fit the RLF with the prescribed functional form over all the redshift bins. We find that the best-fitting PLE model gives $L_{AGN} \propto (1 + z)^{2.03 \pm 0.15 - (0.35 \pm 0.08)z}$ and $L_{SFG} \propto (1 + z)^{4.57 \pm 0.04 - (1.19 \pm 0.02)z}$. As expected, this model does not fit the data as well since it does not take into account the drop at low radio luminosities due to the possible stellar-mass cut-off discussed above. However, this is exactly what we want to avoid the spurious effect of such mass limitation. The evolution strength is similar to that of Novak et al. (2017) up to $z \sim 1.6$. However, beyond $z \sim 2$, the Novak et al. (2017) RLF continues to evolve, whereas we find that the RLF does not evolve as strongly. The lack of strong evolution coincides with the decrease of low stellar-mass sources in our stellar-mass-limited sample at these redshifts. This results in the position of the knee in the RLF moving to lower luminosities for the SFG population, at $z > 2.5$. Another plausible reason our results deviate from Novak et al. (2017) is that we are missing heavily obscured SFRs from our stellar-mass-selected sample that still emit a significant amount of radio emission due to ongoing SF.

We then use our RLF models (the 'free' one and the 'fixed shape' one) to determine the radio-derived SFRD by numerically integrating the product of the 1.4 GHz RLF of SFGs and the SFR associated with the 1.4 GHz luminosity based on the IRRC. For both models, we found that our SFRD is consistent with the established behaviour using FIR data at $z < 1$, where it increases strongly with redshift (e.g. Gruppioni et al. 2013; Madau & Dickinson 2014; Koprowski et al. 2017). However, it lies above the previous determination from Novak et al. (2017) based on the same radio data set but using only direct detections. This is because using only those galaxies above the nominal detection threshold precludes a robust determination of the position of the break and the steepness of the faint-end slope in the RLF, even at $z < 1.5$. Whereas, our study is only limited by the depth of the optical and NIR data, thereby providing complementary and more complete information at and below the knee in the RLF at $z < 1.5$. This suggests that assuming a fixed functional form and fitting over all redshifts leads to an underprediction in the SFRD, due to the fact that the position of the knee in the RLF cannot be accurately determined at these higher redshifts.

We investigate the effects of stellar mass on the total RLF by splitting our sample into low ($10^{8.5} \leq M/M_{\odot} \leq 10^{10}$) and high ($M > 10^{10}$ M$_{\odot}$) stellar mass. We find that the low stellar-mass sources dominate the faint end of the RLF and the high stellar-mass sources are generally associated with the radio-detected sources, as expected given the relationship between stellar mass and SFR. We find that the SFRD is dominated by sources with high stellar masses ($>10^{10}$ M$_{\odot}$) at all redshifts.

Clearly, there is much more work to be done to understand the evolution of the SFRD, with various wavelengths suffering from different selection effects. Here, we have used a new method to determine the evolution of the RLF based on the radio emission from a stellar-mass-selected sample in the COSMOS field. We have shown that it is complementary to those studies that are based purely on radio selection, particular at $z < 1.5$ where the depth of the optical and NIR data ensure more robust constraints on the faint-end slope of the RLF. As such, future works could unite the two methods and provide a much more complete view of the complete RLF and the evolution of the SFRD to high redshifts.

Extending this study to other radio frequency would also be beneficial in overcoming some of the issues of using radio data to determine the SFRD at high redshifts. For example, the level of inverse Compton scattering of CMB photons and the contribution from free–free emission from H II regions will impact on the higher frequency emission more than at low frequency. This obviously is more of an issue at high redshifts, where the rest-frame frequency is >9 GHz for sources at $z > 2$ in the 3 GHz data we use here. Thus, undertaking a similar study as we have done here over the deep fields observed by the LOw-Frequency ARray (LOFAR; Tasse et al. 2021; Sabater et al. 2021) at an observed frequency of 150 MHz, and the MeerKAT International GHz Tiered Extragalactic Exploration (MIGHTEE; Jarvis et al. 2018; Heywood et al. 2021) survey at frequencies covering 544–1088 MHz, 856–1712 MHz, and 1750–3500 MHz, will provide crucial information necessary to advance our understanding of the cosmic SFRD further.

DATA AVAILABILITY

The radio data used in this paper were accessed from the VizieR service at https://doi.org//10.26093/cds/vizier.36020001. The optical/NIR data used in this paper were provided by Nathan Adams. Data are not yet published and will be shared on request to the corresponding author with permission of Nathan Adams. The code used to obtain the results is available in Zenodo at https://doi.org/10.5281/zenodo.4382263.

REFERENCES

Adams N. J., Bowler R. A. A., Jarvis M. J., Häußler B., Lagos C. D. P., 2021, MNRAS, 506, 4933
Adams N. J., Bowler R. A. A., Jarvis M. J., Häußler B., McLure R. J., Bunker A., Dunlop J. S., Verma A., 2020, MNRAS, 494, 1771
Aihara H. et al., 2018a, PASJ, 70, S4
Aihara H. et al., 2018b, PASJ, 70, S8
Aird J., Coil A. L., Georgakakis A., 2017, MNRAS, 465, 3390
Alam S. et al., 2015, ApJS, 219, 12
Arnouts S. et al., 2002, MNRAS, 329, 355
Ashby M. L. N. et al., 2018, ApJS, 237, 39
Belli S. et al., 2017, ApJ, 841, L6
Berta S. et al., 2013, A&A, 551, A100

Best P. N., Kauffmann G., Heckman T. M., Ivezić Ž., 2005, MNRAS, 362, 9
Bondi M. et al., 2018, A&A, 618, L8
Boogaard L. A. et al., 2019, ApJ, 882, 140
Bouwens R. J. et al., 2015, ApJ, 803, 34
Bowler R. A. A., Jarvis M. J., Dunlop J. S., McLure R. J., McLeod D. J., Adams N. J., Milvang-Jensen B., McCracken H. J., 2020, MNRAS, 493, 2059
Brisbin D. et al., 2017, A&A, 608, A15
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Buchner J. et al., 2014, A&A, 564, A125
Burgarella D., Buat V., Iglesias-Páramo J., 2005, MNRAS, 360, 1413
Ceraj L. et al., 2018, A&A, 620, A192
Chabrier G., 2003, PASP, 115, 763
Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772
Coil A. L. et al., 2011, ApJ, 741, 8
Collier J. D. et al., 2014, MNRAS, 439, 545
Condon J. J., Anderson M. L., Helou G., 1991, ApJ, 376, 95
Condon J. J., Cotton W. D., Broderick J. J., 2002, AJ, 124, 675
Cool R. J. et al., 2013, ApJ, 767, 118
Cotton W. D. et al., 2018, ApJ, 856, 67
Cowie L. L., González-López J., Barger A. J., Bauer F. E., Hsu L. Y., Wang W. H., 2018, ApJ, 865, 106
da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595
Daddi E. et al., 2007, ApJ, 670, 156
Daddi E., Cimatti A., Renzini A., Fontana A., Mignoli M., Pozzetti L., Tozzi P., Zamorani G., 2004, ApJ, 617, 746
de Jong T., Klein U., Wielebinski R., Wunderlich E., 1985, A&A, 147, L6
Delhaize J. et al., 2017, A&A, 602, A4
Delvecchio I. et al., 2017, A&A, 602, A3
Delvecchio I. et al., 2021, A&A, 647, A123
Drake A. B. et al., 2013, MNRAS, 433, 796
Dudzevičiūtė U. et al., 2020, MNRAS, 494, 3828
Dunlop J. S. et al., 2017, MNRAS, 466, 861
Emerson J., Sutherland W., 2010, Messenger, 139, 2
Franco M. et al., 2020, A&A, 643, A30
Geach J. E. et al., 2017, MNRAS, 465, 1789
Gruppioni C. et al., 2013, MNRAS, 432, 23
Gruppioni C. et al., 2020, A&A, 643, A8
Gruppioni C., Pozzi F., 2019, MNRAS, 483, 1993
Guidetti D. et al., 2017, MNRAS, 471, 210
Gürkan G. et al., 2018, MNRAS, 475, 3010
Hale C. L., Robotham A. S. G., Davies L. J. M., Jarvis M. J., Driver S. P., Heywood I., 2019, MNRAS, 487, 3971
Hales C. A., Murphy T., Curran J. R., Middelberg E., Gaensler B. M., Norris R. P., 2012, MNRAS, 425, 979
Hasinger G. et al., 2018, ApJ, 858, 77
Heywood I. et al., 2021, MNRAS, in press (arXiv:2110.00347)
Ilbert O. et al., 2009, ApJ, 690, 1236
Ilbert O. et al., 2013, A&A, 556, A55
Jarvis M. et al., 2018, PoS, MeerKAT2016, 006
Jarvis M. J. et al., 2010, MNRAS, 409, 92
Jarvis M. J., Teimourian H., Simpson C., Smith D. J. B., Rawlings S., Bonfield D., 2009, MNRAS, 398, L83
Jeffreys H., 1961, Theory of Probability, 3rd edn. Clarendon Press, Oxford
Jiménez-Andrade E. F. et al., 2019, A&A, 625, A114
Johnston R., Vaccari M., Jarvis M., Smith M., Giovannoli E., Häußler B., Prescott M., 2015, MNRAS, 453, 2540
Karim A. et al., 2011, ApJ, 730, 61
Kennicutt Robert C. J., 1998, ARA&A, 36, 189
Koprowski M. P., Dunlop J. S., Michałowski M. J., Coppin K. E. K., Geach J. E., McLure R. J., Scott D., van der Werf P. P., 2017, MNRAS, 471, 4155
Laigle C. et al., 2016, ApJS, 224, 24
Le Fèvre O. et al., 2013, A&A, 559, A14
Leja J., Tacchella S., Conroy C., 2019, ApJ, 880, L9
Leslie S. K., 2020, ApJ, 899, 58
Lilly S. J. et al., 2009, ApJS, 184, 218
Liu D. et al., 2018, ApJ, 853, 172
Lutz D. et al., 2011, A&A, 532, A90

Madau P., Dickinson M., 2014, ARA&A, 52, 415
Malefahlo E., Santos M. G., Jarvis M. J., White S. V., Zwart J. T. L., 2020, MNRAS, 492, 5297
Mauch T., Sadler E. M., 2007, MNRAS, 375, 931
McAlpine K., Jarvis M. J., Bonfield D. G., 2013, MNRAS, 436, 1084
McCracken H. J. et al., 2012, A&A, 544, A156
McLure R. J. et al., 2013, MNRAS, 432, 2696
Merlin E. et al., 2018, MNRAS, 473, 2098
Mitchell-Wynne K., Santos M. G., Afonso J., Jarvis M. J., 2014, MNRAS, 437, 2270
Mohan N., Rafferty D., 2015, Astrophysics Source Code Library, record ascl:1502.007
Molnár D. C. et al., 2018, MNRAS, 475, 827
Morrison G. E., Owen F. N., Dickinson M., Ivison R. J., Ibar E., 2010, ApJS, 188, 178
Murphy E. J., 2009, ApJ, 706, 482
Murphy E. J., Momjian E., Condon J. J., Chary R.-R., Dickinson M., Inami H., Taylor A. R., Weiner B. J., 2017, ApJ, 839, 35
Noeske K. G. et al., 2007, ApJ, 660, L43
Norris R. P. et al., 2011, ApJ, 736, 55
Novak M. et al., 2017, A&A, 602, A5
Novak M., Smolčić V., Schinnerer E., Zamorani G., Delvecchio I., Bondi M., Delhaize J., 2018, A&A, 614, A47
Ocran E. F., Taylor A. R., Vaccari M., Ishwara-Chand ra C. H., Prandoni I., Prescott M., Mancuso C., 2020, MNRAS, 491, 5911
Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
Oliver S. J. et al., 2012, MNRAS, 424, 1614
Ouchi M. et al., 2010, ApJ, 723, 869
Owen F. N., 2018, ApJS, 235, 34
Pan H., Jarvis M. J., Allison J. R., Heywood I., Santos M. G., Maddox N., Frank B. S., Kang X., 2020, MNRAS, 491, 1227
Polletta M. et al., 2007, ApJ, 663, 81
Prescott M. et al., 2016, MNRAS, 457, 730
Read S. C. et al., 2018, MNRAS, 480, 5625
Rieke G. H. et al., 2004, ApJS, 154, 25
Robotham A. S. G., Davies L. J. M., Driver S. P., Koushan S., Taranu D. S., Casura S., Liske J., 2018, MNRAS, 476, 3137
Sabater J. et al., 2021, A&A, 648, A2
Sanders D. B. et al., 2007, ApJS, 172, 86
Schmidt M., 1968, ApJ, 151, 393
Schober J., Schleicher D. R. G., Klessen R. S., 2015, MNRAS, 446, 2
Scoville N. et al., 2007, ApJS, 172, 1
Simpson J. M. et al., 2014, ApJ, 788, 125
Simpson J. M. et al., 2020, MNRAS, 495, 3409
Skilling J., 2004, in Fischer R., Preuss R., von Toussaint U., eds, AIP Conf. Proc. Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering. Am. Inst. Phys., Melville, NY, p. 395
Smith D. J. B. et al., 2014, MNRAS, 445, 2232
Smith D. J. B. et al., 2021, A&A, 648, A6
Smith D. J. B., Hayward C. C., 2018, MNRAS, 476, 1705
Smolčić V. et al., 2009, ApJ, 690, 610
Smolčić V. et al., 2017a, A&A, 602, A1
Smolčić V. et al., 2017b, A&A, 602, A2
Smolčić V. et al., 2017c, A&A, 602, A6
Steinhardt C. L. et al., 2014, ApJ, 791, L25
Tammann G. A., Yahil A., Sandage A., 1979, ApJ, 234, 775
Tasse C. et al., 2021, A&A, 648, A1
van der Kruit P. C., 1971, A&A, 15, 110
Whitaker K. E. et al., 2014, ApJ, 795, 104
Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, ApJ, 691, 1879
Willott C. J., Rawlings S., Blundell K. M., Lacy M., Eales S. A., 2001, MNRAS, 322, 536
Wuyts S. et al., 2007, ApJ, 655, 51
Yamaguchi Y. et al., 2020, PASJ, 72, 69
Yuan Z., Wang J., Zhou M., Mao J., 2016, ApJ, 820, 65
Zavala J. A. et al., 2019, ApJ, 887, 183
Zwart J. T. L., Santos M., Jarvis M. J., 2015, MNRAS, 453, 1740

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Figure S1. The triangle plot of Model C pure luminosity evolution to all the redshift bins.
Figure S2. The triangle plots for model B fit to the individual redshift bins.
Figure S3. The triangle plots for model C fit to the individual redshift bins.

Table S1. The MAP posterior parameters of Models A, B, and C for the NIR-selected RLF, in each of the redshift bins and their σ. The units of the parameters are as shown in Table 2.
Table S2. The MAP posterior parameters of the low (using Model B and Model C) and high (using Model B) stellar mass contribution to NIR-selected RLF, in each of the redshift bins and their σ. The units of the parameters are as shown in Table 2.
Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TeX/LaTeX file prepared by the author.