

Cosmic star formation probed via parametric stack-fitting of known sources to radio imaging

I. G. Roseboom[★] and P. N. Best

Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

Accepted 2013 December 16. Received 2013 November 28; in original form 2013 September 13

ABSTRACT

The promise of multiwavelength astronomy has been tempered by the large disparity in sensitivity and resolution between different wavelength regimes. Here, we present a statistical approach which attempts to overcome this by fitting parametric models directly to image data. Specifically, we fit a model for the radio luminosity function (LF) of star-forming galaxies to pixel intensity distributions at 1.4 GHz coincident with near-IR selected sources in COSMOS. Taking a mass-limited sample in redshift bins across the range $0 < z < 4$, we are able to fit the radio LF with ~ 0.2 dex precision in the key parameters (e.g. Φ^* , L^*). Good agreement is seen between our results and those using standard methods at radio and other wavelengths. Integrating our LFs to get the star formation rate density, we find that galaxies with $M_* > 10^{9.5} M_\odot$ contribute $\gtrsim 50$ per cent of cosmic star formation at $0 < z < 4$. The scalability of our approach is empirically estimated, with the precision in LF parameter estimates found to scale with the number of sources in the stack, N_s , as $\propto \sqrt{N_s}$. This type of approach will be invaluable in the multiwavelength analysis of upcoming surveys with the Square Kilometre Array pathfinder facilities: LOFAR, ASKAP and MeerKAT.

Key words: methods: data analysis – radio continuum: galaxies.

1 INTRODUCTION

The cosmic history of star formation, typically summarized as the redshift evolution of the star formation rate density (SFRd; Lilly et al. 1996; Madau et al. 1996; Hopkins & Beacom 2006; Behroozi, Wechsler & Conroy 2013), is widely regarded as one of the benchmark measurements of galaxy evolution; its precise measure is a key goal of modern observational astrophysics, and any model for galaxy formation must replicate these observations to be taken seriously.

The rapid evolution of observational facilities in the last decade, in particular upgrades to the wide-field optical (with the Advanced Camera for Surveys in 2002) and near-IR (with Wide Field Camera 3 in 2009) capabilities of the *Hubble Space Telescope* (*HST*), has resulted in robust estimates of the SFRd out to at least $z = 7$ (Bouwens et al. 2006, 2008, 2011; McLure et al. 2010, McLure et al. 2013). However, the *HST* view of the high- z Universe is effectively limited to rest-frame UV and optical wavelengths and hence to relatively unobscured galaxies. This limitation is worrisome, as populations of highly obscured star-forming ($100\text{--}1000 M_\odot \text{ yr}^{-1}$) galaxies are seen in the far-IR and submm (e.g. Hughes et al. 1998). Large-scale submm surveys have shown that the number density of these so-called ultraluminous IR galaxies (ULIRGs) are dwarfed by the more

moderate star-forming population at all redshifts (e.g. Reddy et al. 2008), although they do occur in large enough quantities to cause tension with numerical models of galaxy formation (e.g. Baugh et al. 2005; Lacey et al. 2010). Existing far-IR/submm facilities do not possess the sensitivity to probe even moderate SFRs (i.e. $\text{SFR} < 100 M_\odot \text{ yr}^{-1}$) at $z > 1$ and hence the overlap in UV/optical and far-IR/submm identified galaxies at high redshift is small.

This situation is clearly unsatisfactory; observing the full range in star formation rates (SFRs) at high redshift requires measurements at two distinct wavelengths which can only be cross-calibrated in the nearby Universe (i.e. $z < 0.2$). Deep, wide-area observations in the UV/optical would offer the observationally cheapest route to reconciling this disconnect. However, the strong relationship between SFR and dust obscuration means that the galaxies with the highest SFRs may be amongst the faintest at the UV/optical wavelengths (e.g. Meurer, Heckman & Calzetti 1999; Hopkins et al. 2001; Kewley et al. 2002; Reddy et al. 2010; Roseboom et al. 2012). Moreover, there is some evidence that UV/optical tracers of the level of dust obscuration (e.g. the UV slope, β_{UV} , the ratio of Balmer lines) may only be accurate at moderate SFRs (i.e. $< 100 M_\odot \text{ yr}^{-1}$; Buat et al. 2010; Reddy et al. 2010; Roseboom et al. 2012).

Thus, there is no alternative but to push the sensitivity limits of SFR tracers which are not affected by dust obscuration. In recent years, it has become de rigueur to build vast multiwavelength data sets on a number of well-known extragalactic survey fields (e.g. GOODS, Lockman Hole, COSMOS, etc.). While the disparity

[★]E-mail: igr@roe.ac.uk

in observational capabilities across the electromagnetic spectrum has limited the usefulness of multiwavelength studies of individual objects, the practice of stacking less-sensitive data sets (typically X-ray, far-IR/submm, radio) at the positions of the deepest catalogued sources (typically optical/near-IR) has become commonplace. For imaging afflicted by white noise, the stacked signal-to-noise should increase proportionally to the square-root of the number of sources in the stack. Hence, for extragalactic fields with tens of thousands of sources, the aggregate, dust-unbiased SFRs can be robustly estimated at a level ~ 100 times fainter than the detection limit for individual sources.

However, traditional stacking methods can only recover the mean (or median) SFR of any sample, and hence only the mean relationship between SFR and the binned variable can be established. To counter this, it is common to bin the sources in more than a single parameter, e.g. the stellar mass–redshift plane (Oliver et al. 2010, Karim et al. 2011; Roseboom et al. 2013). This approach has the significant drawback of reducing the number of sources, N , in each stack, which in turn reduces the stacked signal by a factor proportional to N^2 .

In this paper, we present a statistical framework that allows the underlying distributions, i.e. $P(\text{SFR}|M_*, z, \beta_{\text{UV}}, \dots)$ to be extracted in a parametric way from the stack. The clear advantage of this approach is that it allows the distribution in SFR to be constrained using all the available sources in a consistent way. To illustrate this point, Karim et al. (2011; hereafter K11) perform a stacking analysis with practically the same data set we will use here; optically/near-IR-selected galaxies stacked into Very Large Array (VLA) imaging in COSMOS. They split their sample into 9×7 redshift–mass bins and determine the mean SFR in each bin via median stacking of cutouts of the radio data around each source. If no binning is required, and the precision in the parametric fit scales as \sqrt{N} , then our method could provide estimates of the SFR– M_* – z relation with ~ 8 times more precision (~ 3 if we still bin in redshift). These types of gains are non-negligible considering the observational expense required to improve the raw sensitivity of large multiwavelength data sets like COSMOS by these factors.

In Section 2, the method used to extract $P(\text{SFR}|z)$ from a stack of galaxy positions is presented, Section 3 describes the data sets used to test this method and Section 4 presents our results. Section 5.1 compares our results to the literature and discusses the implications of our results in the context of the cosmic SFRd, while Section 5.2 discusses the potential future applications for this method and possible obstacles. Finally, Section 6 presents our conclusions.

Throughout we assume a Λ cold dark matter cosmology with $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Where relevant, all quoted quantities assume a Chabrier (2003) Initial Mass Function (IMF).

2 PARAMETRIC STACKING METHOD

2.1 Basic framework

If we consider a pixelated sky image \mathbf{d} in units of flux density per telescope beam, then for a list of sources with known positions \mathbf{x} , the best estimate of the mean flux density will be given by

$$\frac{\sum \mathbf{d}_x / \sigma_x^2}{\sum 1 / \sigma_x^2},$$

where \mathbf{d}_x and σ_x are the pixel flux density and noise estimates for positions \mathbf{x} , respectively. Turning this around, if the true flux

densities at positions \mathbf{x} are f_x the observed, pixel intensities d_x will be simply

$$\mathbf{d}_x = f_x + \delta,$$

where δ is some noise value drawn from a Gaussian distribution; $N(\mu = 0, \sigma_x)$. If a model M can be constructed which predicts the probability $P(f)$ of a true flux density f for a known source, then the probability of observing a particular pixel intensity d at the location of that source is

$$P(d|M, \sigma) = \int_0^\infty \frac{1}{\sigma \sqrt{2\pi}} P(f) \exp\left(-\frac{(f-d)^2}{2\sigma^2}\right) df, \quad (1)$$

and so for a stack of sources \mathbf{x} ,

$$P(\mathbf{d}_x|\mathbf{x}, M, \sigma_x) = \prod_x P(d|M, \sigma). \quad (2)$$

Applying Bayes' theorem, we see that

$$P(M, \sigma_x|\mathbf{d}_x) \propto P(M)P(\mathbf{d}_x|\mathbf{x}, M, \sigma_x),$$

and so the best-fitting parameters for the model M can be found by maximizing the likelihood given in equation (2).

2.2 Building the flux density distribution model

Equations (1) and (2) give a generic framework to fit a stack (i.e. stack-fit) of flux densities for a set of known positions in the presence of noise. This framework could be used for parameter estimation in models with a wide range of purposes, from simple parametrizations of the number counts (e.g. Mitchell-Wynne et al. 2014) to full-blown galaxy formation simulations (e.g. Henriques et al. 2009). Here, we propose to model the rest-frame luminosity distribution function at a given redshift, z ; $\hat{\Phi}(l, z)$. In this scenario, we need to convert our proposed luminosities to the observed flux density in order to use equation (2). Considering both the luminosity distance and a k -correction as a function of redshift, we obtain

$$P(f) = \frac{\hat{\Phi}(4\pi D(z)^2 k(z) f, z)}{\int_{l=0}^\infty \hat{\Phi}(l, z) dl dz}, \quad (3)$$

where $D(z)$ is the luminosity distance and $k(z)$ the k -correction at a redshift z . It is worth emphasizing that $P(f)$ here is the probability of a *particular* galaxy at a known position having a true flux density f . If for a given sample $\hat{\Phi}$ is independent of redshift, then we can simply propose $\hat{\Phi}(l)$ and remove the integral over dz in the denominator of equation (3). Similarly, if we can split our sample into narrow redshift bins then the simplest non-parametric estimate of the redshift evolution can be obtained by measuring $\hat{\Phi}(l)$ in each redshift bin independently. This is the approach we will take later in Section 3.2.

Equation (3) is analogous to the formula for $P(M_{\text{abs}}, z)$, where M_{abs} is the absolute magnitude, which appears in the Sandage, Tamman & Yahil (1979; hereafter STY) maximum likelihood luminosity function (LF) estimator. However, in our scenario the sample is not selected at the wavelength at which $\hat{\Phi}$ is determined, but by a selection at some other wavelength (or physical galaxy property). In the STY formalism, the integral in the denominator of equation (3) would have a lower bound equal to the minimum observable luminosity at the redshift of the source. However, in our formalism, the sources in the stack can have no such limit and so the lower bound of this integral is zero. This means that whatever parametric form we choose for $\hat{\Phi}$ must have a finite integral from zero to infinity, otherwise the ratio in equation (3) will become undefined.

If the redshifts for each source are not known accurately, or only the redshift distribution for the population of sources is known (as

is the case for e.g. BzK galaxies), then the probability density of the redshift, $P(z)$, can be integrated over, i.e.

$$P(f) = \frac{1}{\int_{l=0}^{\infty} \hat{\Phi}(l, z) dl dz} \int_{z=0}^{\infty} P(z) \hat{\Phi}(4\pi D(z)^2 k(z) f, z) dz. \quad (4)$$

Similarly, if the k -correction cannot be described as a simple function of redshift then whatever other parameters, η , that are required to describe the k -correction must be added to the integral, i.e. $\int \int k(z, \eta) dz d\eta$.

As is the case with the STY LF estimator, equation (3) is insensitive to the normalization of the LF model $\hat{\Phi}$. The value of the normalization can be recovered by considering the total number of observed sources, their visibility and various incompletenesses, as we will detail in Section 2.4.

2.3 Dealing with incompleteness

Armed with equations (1) and (4), we have the statistical framework needed to fit a model specified in terms of rest-frame luminosities directly to the pixel stack. It is worth noting that this setup assumes implicitly that the input catalogue to the stack is complete. This is not an onerous requirement, the depth of optical/near-IR surveys is such that it is possible to construct complete volume- or mass-limited samples out to high redshift. However, it would be preferable to generalize our method to accommodate known incompleteness in the input catalogues. To accomplish this, we need to calculate $P(f|\theta)$, where θ are the properties of a source in the catalogue providing the positions, e.g. apparent magnitude, stellar mass, photometric redshift, optical colour, etc. If $P(f)$ and $P(\theta)$ are independent, then

$$P(f|\theta) = P(f)P(\theta), \quad (5)$$

i.e. we simply multiply $P(f)$ by the estimated completeness for sources with properties θ .

2.4 Determining the normalization

In converting our model $\hat{\Phi}(l)$ to a probability distribution, the absolute normalization, $\hat{\Phi}^*$, is lost. To recover this, we need to determine which value of $\hat{\Phi}^*$ will give the correct number of observed galaxies, N , i.e.

$$\hat{\Phi}^* = \frac{\bar{n}}{\int_{l=0}^{\infty} \hat{\Phi}(l, z) dl dz}, \quad (6)$$

where \bar{n} is the mean density for objects in the stack. For a complete, volume-limited sample of sources, it is clear that

$$\bar{n} = \frac{N}{V},$$

where V is the observable volume for sources in the stack list. For non-complete stack lists with variable completeness, C , and observable volumes, V_{\max} , this becomes

$$\bar{n} = \sum_x [C_x V_{\max}]^{-1} \quad (7)$$

3 DATA AND MODEL SETUP

3.1 ULTRAVISTA and VLA data in COSMOS

To use our method effectively, we need two overlapping data sets; an input catalogue which is effectively deeper and provides redshift

estimates and an image which has pseudo-white noise and similar angular resolution to the input catalogue (to avoid issues with confusion). The ULTRAVISTA photometric redshift catalogue of Muzzin et al. (2013a) and the VLA-COSMOS 1.4 GHz radio image presented by Schinnerer et al. (2007) satisfy these two requirements, respectively. The Muzzin et al. (2013a) catalogue provides photometric redshifts and stellar mass estimates for over 260 000 galaxies to a limiting magnitude of $K_s < 24.35$ (3σ) from the combination of the 1.63 deg² ULTRAVISTA Y, J, H, K_s imaging (McCracken et al. 2012) with overlapping ground-based and *HST* optical imaging of the COSMOS field (see Scoville et al. 2007). Only reliable sources from the ULTRAVISTA catalogue are considered here by requiring the metrics `STAR = 0`, `K_FLAG < 4`, `CONTAMINATION! = 1` and `NAN_CONTAM < 3`, and a magnitude of $K_s < 23.4$, as recommended by Muzzin et al. (2013a). These flags allow the removal of stars as well as sources with corrupted photometry due to bad pixels and nearby bright sources. This magnitude limit is chosen as it represents the 90 per cent completeness level of the catalogue.

Meanwhile, the VLA imaging of the COSMOS field consists of a mosaic of 23 tiled pointings of the VLA at 1.4 GHz in the A and C configuration (Schinnerer et al. 2007). The image has an rms of $\sim 10 \mu\text{Jy beam}^{-1}$ at the field centre and a beam full width at half-maximum of 1.4 arcsec \times 1.5 arcsec. In this work, we only make use of the 1.6 deg² of the mosaic that overlap with ULTRAVISTA coverage. The publicly available image has been ‘cleaned’, i.e. point sources above 40 μJy responding to the synthesized (‘dirty’) beam have been removed and replaced with a Gaussian beam conserving the peak flux density. For our peak pixel stacks this has no effect, as the peak flux density is conserved in the cleaning process; however, this does have some impact on the integrated-to-peak flux density ratio, as we will discuss in Section 3.3.

3.2 Modelling the mass-limited radio LF

To use equation (3), we need to define some parametric form for $\Phi(\hat{l}, z)$. As we have high-quality photometric redshifts for all sources in the input catalogue, we prefer to divide our sample into redshift bins across which the evolution of the LF is negligible and fit for $\Phi(\hat{l})$ in each of these redshift bins independently. This allows the redshift dependence of the LF to be probed in a non-parametric way. Typically, the radio LF of star-forming galaxies is assumed to follow a modified Schechter function (Saunders et al. 1990):

$$\hat{\Phi}(l) = \hat{\Phi}^* \left(\frac{l}{l^*}\right)^{1-\alpha} \exp\left[-\frac{1}{2\sigma_{\text{LF}}^2} \log_{10}^2\left(1 + \frac{l}{l^*}\right)\right] \frac{dl}{l}. \quad (8)$$

Motivated by this, and to allow direct comparison with literature estimates, we also assume that the luminosity distribution for our mass-limited sample in each redshift bin is described by equation (8). As the radio emission in these star-forming galaxies will be dominated by the synchrotron radiation, we assume that the radio Spectral Energy Distribution (SED) can be described as $S_\nu \propto \nu^{-\alpha_s}$. Thus, the k -correction is given by $k(z) = (1+z)^{1-\alpha_s}$. Here, we assume $\alpha = 0.8$ (Condon 1992).

Finally, as we are working with a mass-limited sample, we require that the integral of $\hat{\Phi}(l, z)$ be finite, i.e. $\alpha < 1$.

3.3 Redshift binning and model fitting

We build samples for input to our stack-fitting method by slicing the ULTRAVISTA catalogue into nine redshift bins across the range $0.1 < z < 4$. In order to restrict the sample to only star-forming galaxies, i.e. excluding AGN and passive galaxies, we use

the *UVJ* colour selection first described by Williams et al. (2009). Specifically, we require rest-frame $[U - V] \leq 0.88[V - J] + c$ where $c = 0.69$ at $z < 1$ and $c = 0.59$ at $z > 1$. Table 2 details the redshifts and numbers of galaxies in our redshift bins. As an additional restriction, we consider only sources with stellar mass $M_* > 10^{9.5} M_\odot$, the $\gtrsim 50$ per cent completeness limit for the ULTRAVISTA sample at the highest redshifts we consider.

For each source in the ULTRAVISTA catalogue, we calculate a completeness to use in equation (7). The completeness correction is estimated by comparing the number of observed sources in a mass bin (with $\Delta M_* = 0.5$) to the predicted number taking into account the geometry of the ULTRAVISTA survey, the redshift and the parametric fit to the mass function for star-forming galaxies given by Muzzin et al. (2013b).

For each redshift bin, we build a stack of radio flux densities by taking the value of the pixel in the VLA-COSMOS image that corresponds to each position. In addition to the single peak pixel stack, we also build a 2D image stack. This 2D stack is used to assess the integrated-to-peak flux density conversion ratio for each set of objects. For point sources, the integrated-to-peak flux density ratio should be unity; however, some of the sources in our stacks may be extended (especially at low z), and bandwidth smearing at the edge of the VLA pointings has the effect of making sources marginally extended. To determine the correct integrated-to-peak flux density ratio for each of our stacks, we produce 501×501 pixel stacks for each of our redshift bins. The median image is produced across these stacks and then de-convolved with the dirty beam. The peak flux density is then compared to the integrated value across the median image. The integrated-to-peak flux density ratios ($S_{\text{int}}/S_{\text{peak}}$) for each redshift bin assessed in this way are given in Table 2. It is interesting to compare our values to those of K11 who performed a similar stacking analysis of the VLA-COSMOS data. While they use a slightly different map (the COSMOS internal Deep map rather than the public Large Project map) and redshift binning, the trend of large $S_{\text{int}}/S_{\text{peak}}$ at low- z decreasing to a typical value of ~ 1.7 – 2 at $z > 0.5$ is also seen.

After correcting the pixel stack values by $S_{\text{int}}/S_{\text{peak}}$, we fit the model specified by equations (1) and (3) using Monte Carlo Markov Chain (MCMC) methods. Specifically, we use an implementation of the affine-invariant sampling proposed by Goodman & Weare (2010). A detailed discussion of this MCMC approach and its application to astronomy, can be found in Foreman-Mackey et al. (2013); we briefly summarize the main points here. The affine-invariant sampling method makes use of a set of test positions, or ‘walkers’, on the posterior probability. In each iteration, the proposed ‘step’ for a walker is generated by defining a vector in parameter space between it and another randomly chosen ‘walker’ and moving along this vector by a random fraction between $1/a$ and a , where a is a tuning parameter typically set to 2. Proposed steps are then accepted with a probability such that MCMC chain is ‘balanced’ (i.e. it is equally likely to step from $x \rightarrow x'$ as its inverse $x' \rightarrow x$). Here, we use our own IDL implementation of the affine-invariant ensemble sampler; for each stack, we utilize 200 walkers and the tuning parameter $a = 2$. The correlation length (i.e. the number of steps required for the chain to retain no knowledge of its starting point) for each parameter is measured using the method of Goodman & Weare (2010) and the MCMC chain run until the number of steps exceeds 10 times the maximum correlation length amongst all of the free parameters.

The model described in Sections 2 and 3 contains five free parameters: the Gaussian noise in the radio image, σ_N ; the LF normalization, Φ^* ; the characteristic luminosity, L^* ; the faint end slope, α ;

Table 1. Details of parameters used to fit pixel stacks.

Parameter	Min	Max	μ	σ
σ_N^a	0	50	13.6	1
Φ^*	–10	12	–	–
L^*	18	25	–	–
α	0	1	–	–
σ_{LF}^a	0	2	0.6	0.1

^a Gaussian prior.

and the Gaussian term in the LF, σ_{LF} . For the parameters σ_N and σ_{LF} , we assume a Gaussian prior, while for the others we assume a ‘top-hat’ prior with an arbitrarily chosen range. A summary of the model parameters and their priors is given in Table 1. For the image noise, σ_N , we set the prior to the value obtained from the standard deviation of all the pixels in radio image within the ULTRAVISTA coverage, giving $\sigma_N = 13.6 \pm 1$. Meanwhile, the prior for the LF parameter, σ_{LF} , is set to the best estimate and uncertainty from fitting equation (8) to the measurements of the local radio LF for star-forming galaxies from Condon et al. (2002) and Mauch & Sadler (2007).

4 RESULTS

The median stack-fit estimates and 68 per cent confidence intervals for the model parameters are given in Table 2. Fig. 1 shows the distribution of flux density in stack pixels for the nine redshift bins compared to both the median stack-fit and a simple Gaussian fit to the pixel stack. Also shown is the median of the flux density distribution, which is always found to take a positive value. In each redshift bin, there is an excess of pixels with positive values over the Gaussian model, validating our basic model assumption.

While the degeneracies between the model parameters make direct interpretation of the median stack-fit parameters problematic, there are some noteworthy trends. Reassuringly, the estimates of the noise in the map are consistent across the nine redshift bins. Some real variation in this quantity is to be expected due to the variation in $S_{\text{int}}/S_{\text{peak}}$ from both extended sources and bandwidth smearing. Apart from the lowest redshift bin (where we would expect many sources to be resolved), the variation in σ_N is within the measurement errors.

The relationship between $\Phi^*(z)$ and $L^*(z)$ is interesting. At $z \lesssim 1$, $\Phi^*(z)$ remains roughly constant, while $L^*(z)$ increases steadily, i.e. pure luminosity evolution. This picture is consistent with simple models for the evolution of star-forming galaxies in the radio from previous work (e.g. Haarsma et al. 2000; Smolčić et al. 2009). At $z \gtrsim 1$, $\Phi^*(z)$ decreases rapidly while $L^*(z)$ continues to increase strongly to our limit of $z \sim 4$. While radio LF measurements for star-forming galaxies do not currently exist above $z = 1$, this evolution in the LF parameters is in agreement with the known trend for star-forming galaxies from other tracers at high redshift (e.g. Sobral et al. 2012; Gruppioni et al. 2013).

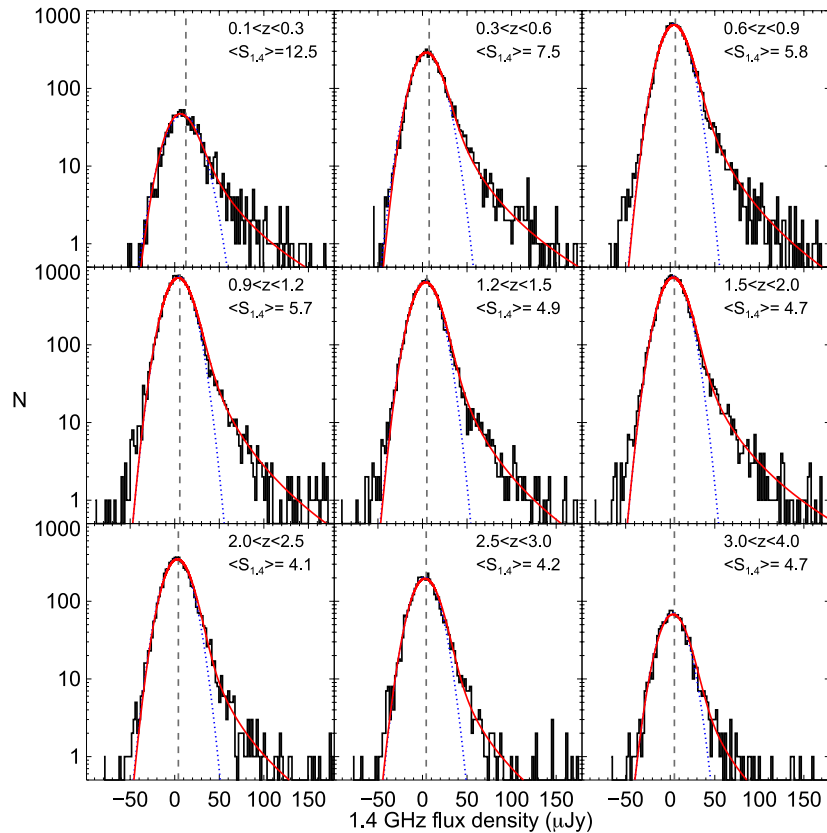
5 DISCUSSION

5.1 LFs and the cosmic SFRd

The stack-fit mass-limited radio LFs for star-forming galaxies are shown in Fig. 2. The 68 per cent confidence interval on our LF estimates (and subsequently SFRd estimates) are constructed by

Table 2. Details of the stacking bins and median fit values to the stacks achieved via MCMC. In each case, the quoted errors represent the 68 per cent confidence interval.

z	N	$\langle S_{1.4} \rangle$ (μJy)	$S_{\text{int}}/S_{\text{peak}}$	σ_N (μJy)	Φ^* (Mpc^{-3})	L^* (W Hz^{-1})	α	σ_{LF} (W Hz^{-1})
0.1–0.3	1083	12.5	3.6	$13.8^{+0.5}_{-0.5}$	$-2.72^{+0.10}_{-0.19}$	$20.4^{+0.4}_{-0.3}$	$0.36^{+0.21}_{-0.25}$	$0.63^{+0.05}_{-0.06}$
0.3–0.6	5585	7.5	2.65	$13.0^{+0.2}_{-0.2}$	$-2.86^{+0.14}_{-0.13}$	$20.7^{+0.3}_{-0.2}$	$0.19^{+0.20}_{-0.14}$	$0.63^{+0.02}_{-0.03}$
0.6–0.9	12 095	5.8	1.88	$13.0^{+0.1}_{-0.1}$	$-2.79^{+0.08}_{-0.06}$	$21.2^{+0.1}_{-0.1}$	$0.08^{+0.12}_{-0.06}$	$0.58^{+0.02}_{-0.02}$
0.9–1.2	13 319	5.7	1.70	$13.0^{+0.1}_{-0.1}$	$-2.92^{+0.06}_{-0.05}$	$21.4^{+0.1}_{-0.1}$	$0.06^{+0.08}_{-0.04}$	$0.58^{+0.02}_{-0.02}$
1.2–1.5	11 517	4.9	1.80	$13.0^{+0.1}_{-0.1}$	$-3.05^{+0.10}_{-0.08}$	$21.6^{+0.2}_{-0.1}$	$0.12^{+0.15}_{-0.09}$	$0.60^{+0.02}_{-0.02}$
1.5–2	13 376	4.7	1.76	$13.1^{+0.1}_{-0.1}$	$-3.26^{+0.14}_{-0.12}$	$21.7^{+0.3}_{-0.2}$	$0.16^{+0.18}_{-0.12}$	$0.67^{+0.02}_{-0.02}$
2–2.5	6255	4.1	1.90	$13.2^{+0.2}_{-0.2}$	$-3.41^{+0.10}_{-0.18}$	$22.1^{+0.4}_{-0.3}$	$0.37^{+0.23}_{-0.23}$	$0.65^{+0.03}_{-0.04}$
2.5–3	3606	4.2	1.98	$13.4^{+0.2}_{-0.2}$	$-3.60^{+0.08}_{-0.22}$	$22.3^{+0.5}_{-0.4}$	$0.38^{+0.25}_{-0.27}$	$0.67^{+0.04}_{-0.05}$
3–4	1257	4.7	1.92	$13.3^{+0.4}_{-0.4}$	$-4.29^{+0.08}_{-0.22}$	$22.6^{+0.5}_{-0.5}$	$0.41^{+0.26}_{-0.29}$	$0.67^{+0.06}_{-0.07}$

**Figure 1.** Pixel flux density distribution compared to both the median-fit model (red solid line) and a simple Gaussian fit (blue dotted line). The median of the flux density distribution is shown as a grey dashed line and quoted in the top-right corner. In each case, it can be seen that there is both a non-zero median of the flux density distribution and an excess of pixels in the stack with positive values over the simple Gaussian fit, validating the basic assumptions of our method.

generating Monte Carlo realizations of our best-fitting parameters, using the covariance matrix output from the MCMC chains. In each realization, the normalization of the LF is forced to equal the completeness-corrected number of sources in the field, with an additional noise term taking into account both Poisson noise and cosmic variance. The effect of cosmic variance is estimated using the results of Moster et al. (2011), and found to be 6 per cent for the lowest redshift bin, increasing to 18 per cent for the highest redshift bin.

Directly comparing our LF stack-fits to literature values is complicated by the fact that our sample is stellar mass limited, rather than radio luminosity limited. Nonetheless, we compare our results to to-

tal LF's in the radio from Best et al. (2005) and Smolčić et al. (2009), the latter of which makes use of the same VLA-COSMOS imaging as this work. The two sets of radio LFs show good agreement with our estimates, although the two highest redshift bins probed by Smolčić et al. (2009) suggest an upturn at $L_{1.4} \gtrsim 10^{24.5} \text{ W Hz}^{-1}$ that is not possible for our parametric fits to replicate.

It is also interesting to compare our results to the median radio luminosities found via traditional stacking by K11. For each redshift bin in Fig. 2, we show the median radio luminosity for $M_* > 10^{9.5} M_\odot$ from K11 as a purple star, while the median luminosity of our parametric fit is shown as a red dashed line. It can be seen that the K11 values agree well with the median luminosity of our LFs,

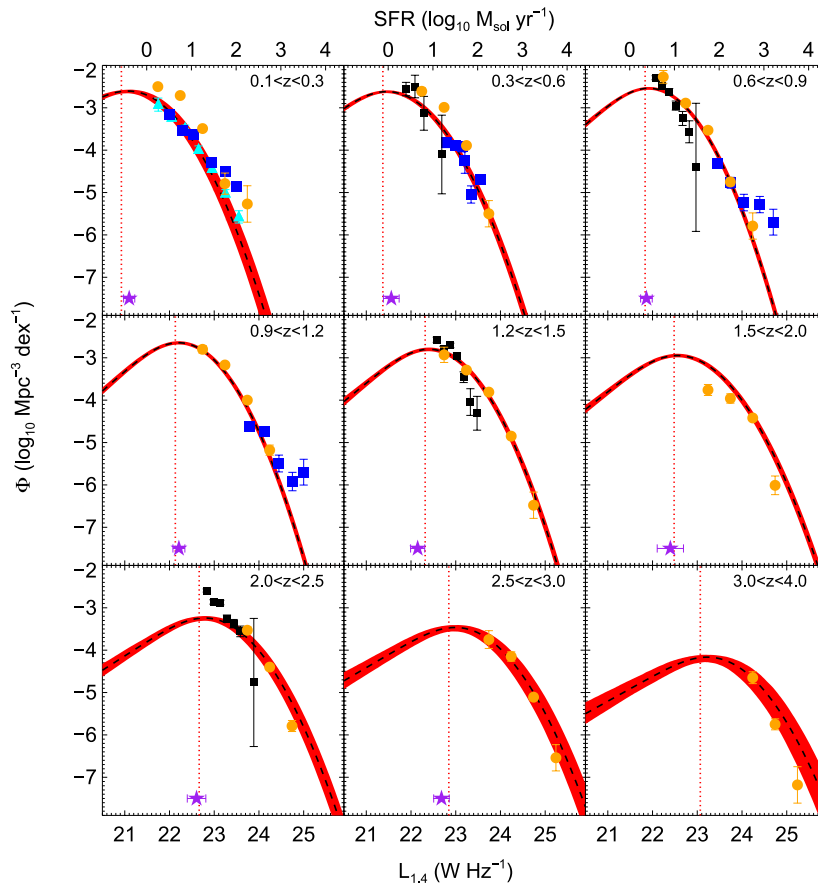


Figure 2. LFs from stack-fitting to the radio pixel intensity stacks. The solid dashed line is the LF from the median fit parameters given in Table 2 while the red shaded region is the 68 per cent interval in Φ at each $L_{1.4}$. Literature values of the radio LF from other works are also shown as: cyan triangles (NVSS–SDSS; Best et al. 2005) and blue squares (VLA-COSMOS; Smolčić et al. 2009). Far-IR and mass-limited H α estimates, converted to radio luminosity assuming star formation as the common origin, are shown as orange circles (*Herschel* Far-IR; Gruppioni et al. 2013) and black squares (H α ; Sobral et al. 2014). The purple star is the median $L_{1.4}$ found in equivalent mass and redshift bins by K11, while the dotted line represents the median of our parametric fit.

both validating our approach and also demonstrating the power of our parametric stack-fitting over traditional stacking techniques.

Finally, we compare our results to the total far-IR LF (Gruppioni et al. 2013) and the mass-limited H α LF (Sobral et al. 2014). Assuming in all cases that star formation dominates the luminosity, we can convert the far-IR and H α measurements to 1.4 GHz luminosity by combining the SFR calibrations of Kennicutt et al. (1998) and Bell (2003); $L_{1.4} = L_{\text{IR}} + 11.49$ and $L_{1.4} = L_{\text{H}\alpha} - 19.87$, assuming a Chabrier (2003) IMF. Excellent agreement is found between our estimate and the far-IR measurements from Gruppioni et al. (2013) at $z > 0.6$, while lower redshift bins appear to show an ‘excess’ of far-IR sources over the radio estimates. This result is not unique to our work and has been noted by previous authors when comparing radio and far-IR LFs for star-forming galaxies (e.g. Bell 2003; Smolčić et al. 2009).

The comparison with the H α estimates is especially interesting as these measurements have been made with the same mass-limit as our sample. For the $0.3 < z < 0.6$ bin, the agreement between the H α and radio mass-limited LF is very good. At higher redshifts, the overall normalization of the H α and radio LFs agree well, but the H α shows a steeper cut-off towards high luminosities (i.e. lower L^*). This could point to insufficient dust corrections to the H α fluxes for galaxies with the highest SFRs (e.g. Roseboom et al. 2012). Alternatively, our radio measurements could be polluted with emission not related to star formation (e.g. AGN). However, the

agreement seen between the radio and far-IR at high luminosities means that the far-IR estimates also need to be influenced in the same way. The physical origin of far-IR emission in ULIRGs (and by proxy the radio) is an area of some debate, although most studies agree that it is very difficult to power luminous IR emission at long wavelengths (i.e. $\gtrsim 100 \mu\text{m}$) with a process other than star formation (e.g. Lutz et al. 2008; Hatziminaoglou et al. 2010).

Armed with our LFs for $M_* > 10^{9.5} M_{\odot}$ galaxies out to $z = 4$ and assuming that the radio luminosity is generated by star formation, we can ask the question: What contribution do these galaxies make to the cosmic SFRd? In Fig. 3, we show the cosmic SFRd estimated by integrating the LFs from Fig. 2 and assuming the SFR calibration of Bell (2003). Also shown in Fig. 3 is the recent compilation of total SFRd estimates by Behroozi et al. (2013; hereafter B13), as well as a mass-limited ($M_* > 10^{9.8} M_{\odot}$) estimate constructed from the results of Santini et al. (2009). While the error in our estimate and the scatter in the B13 literature values are significant, it is clear that galaxies with $M_* > 10^{9.5} M_{\odot}$ make up $\gtrsim 50$ per cent of the SFRd up to $z = 4$. This is in good agreement with both the K11 and Santini et al. (2009) results up to $z \sim 3$.

5.2 Future prospects for parametric stack-fitting

From the results above, it is clear that stack-fitting has substantial benefits if source catalogues at one wavelength are significantly

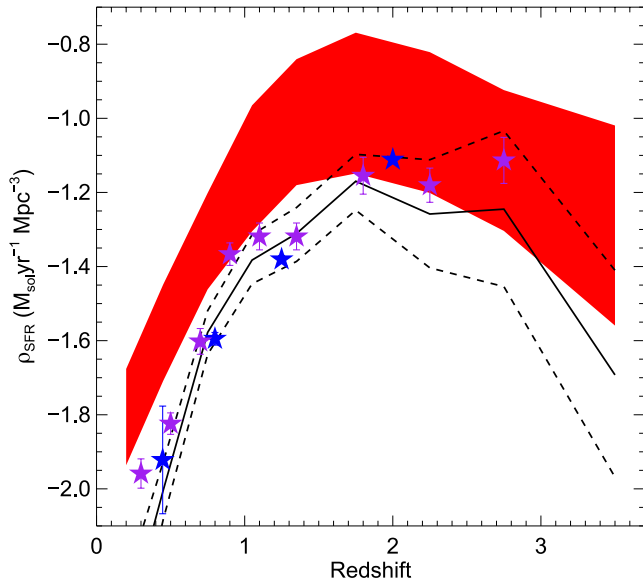


Figure 3. SFRd for galaxies with $M_* > 10^{9.5} M_\odot$ as a function of redshift implied by our LFs (solid line). The dashed lines represent the 1σ uncertainty on our parametric fit. For comparison is shown the best estimate of the total SFRd(z) from the compilation of B13 (red shaded) as well as mass-limited estimates from both the UV (Santini et al. 2009; blue circles) and stacked radio (K11; purple stars). Both our parametric fits and the literature results imply that $M_* \gtrsim 10^{9.5} M_\odot$ galaxies contribute $\gtrsim 50$ per cent of the total SFRd up to $z < 4$.

deeper compared to overlapping data and a flexible parametric description of the data can be provided. As we have shown, this technique is well suited to interferometric radio data, as the disparity in angular resolution compared to optical/near-IR imaging means that source confusion is not an issue.

Several large surveys of the extragalactic sky at radio wavelengths are either underway (e.g. with LOFAR; van Haarlem et al. 2013) or will commence in the near future (e.g. MIGHTEE, Jarvis et al. 2012; EMU, Norris et al. 2011). These surveys will provide all-sky observations at comparable depth to the COSMOS data as well as significantly deeper regions of comparable area. However, none of these surveys will provide radio imaging that is effectively deeper than ancillary optical/near-IR data sets. Thus, techniques such as the one presented here may be invaluable in terms of interpreting these surveys.

At other wavelengths, large areas of the sky have already been imaged at IR/submm wavelengths by *Herschel*, *WISE* and SCUBA-2. Again, overlapping optical/near-IR data sets are significantly deeper and so the stack-fitting approach may offer significant benefits, with the caveat that the mismatch in the beam size (roughly a factor of 10 between ground-based near-IR and submm) means that confusion noise must be taken into account, something which is not currently accommodated by our method.

While we leave a detailed treatment of implementing our method in the presence of confusion noise to future work, we briefly consider here how confusion noise could be accounted for. The simplest solution to the problem of source confusion is to assume that the confusion introduces some noise term, similar to the instrumental noise in equation (1), which is a function of both the instrument beam and the intrinsic flux density distribution,

$$P^{\text{conf}}(f) = P(f) + \int_{s=0}^{\infty} \int_{r=0}^{\infty} B(r)sP(s) dr ds, \quad (9)$$

Table 3. Effect of varying N_s on the precision of the LF parameters, L^* and α , and the area 95 per cent confidence ellipse between these parameters, A_{95} .

N_s	σ_{L^*}	σ_α	A_{95}
13 376	0.21	0.14	0.19
9515	0.25	0.16	0.26
4577	0.28	0.18	0.36
1828	0.31	0.18	0.56
935	0.39	0.21	0.90
468	0.36	0.20	1.09
182	0.72	0.27	1.55

where $B(r)$ is the one-dimensional beam profile and $P(s)$ is the number density of sources with intrinsic flux density s . An implicit assumption of equation (9) is that the sources are randomly distributed on the sky. In the presence of clustering, if the sources in the stack can be described by a simple model (e.g. a power law, or halo occupation model), this can be added to the integral such that the probability of source occurring within the beam is also a function of radius from the source, i.e. $P(s)$ becomes $P(s, r)$.

Finally, it is of interest to understand how our method scales with the number of sources (N_s) in the stacks. For traditional stacking the signal-to-noise in the stacked signal increases as $\sqrt{N_s}$. However, it is not obvious how the precision in our model LF parameters should vary as a function of N_s . To test this, we take the redshift bin with the most sources, $1.5 < z < 2$, from our analysis and sparsely sample it to simulate the impact of decreasing the number of sources. Details of this process are given in Table 3.

One difficulty in the analysis of this simulation is that the parameters in the model are degenerate, so the individual error estimates do not fairly represent the true uncertainty. To mitigate this, we consider the area of the 1σ error ellipse defined by the covariance matrix for each model fit. As σ_N and σ_{LF} are constrained by Gaussian priors (and only weakly degenerate with L^*) and Φ^* is determined from the absolute normalization, we need only consider the ellipse formed by the uncertainty in L^* and α . Thus, we also list the area of the ellipse which contains 95 per cent of the uncertainty in L^* and α , A_{95} , in Table 3.

For traditional stacking techniques, the precision in the stacked estimate increases $\propto N^{1/2}$. To allow quantitative comparison with these techniques, we fit the data from Table 3 with a model assuming that the precision, Δ , improves with the number of sources to some fractional exponent, i.e. $\Delta \propto N_s^{-1/\tau}$. This gives $\tau = 4.2 \pm 0.7$ for $\sigma_{L^*}(N_s)$, $\tau = 8 \pm 2$ for $\sigma_\alpha(N_s)$ and $\tau = 2 \pm 0.1$ for the 95 per cent confidence ellipse area (A_{95}). It is clear that the large degeneracies between the parameters significantly inflate the exponents for the marginalized parameter uncertainties. However, the area of the error ellipse between L^* and α is seen to scale as $N_s^{-1/2}$, i.e. as the square-root of the number of sources in the stack, similar to simple stacking methods.

6 CONCLUSIONS

In the last decade, there has been an increasing focus on multiwavelength approaches to outstanding problems in galaxy formation and evolution. While multiwavelength tracers of galaxy parameters (e.g. SFR) are the key in reducing systematic uncertainties, the disparity in observing capabilities across wavelength regimes, in terms of both sensitivity and angular resolution, severely limits the

ability to build cross-matched samples of individual galaxies. In this work, we have presented an alternative method designed to maximize the information gleaned from multiwavelength data sets by using a statistical approach to constrain a parametric description of the LF. Taking a mass-limited sample in the COSMOS field and stack-fitting pixel intensity distributions from the VLA-COSMOS imaging, we achieve the following results.

(i) We constrain the parameters of the mass-limited radio LF for star-forming galaxies out to $z \sim 4$, with ~ 0.2 dex error in L^* .

(ii) Our stack-fit LFs show good agreement with comparable literature estimates. Compared to direct estimates using the same radio data (VLA-COSMOS), we can constrain the LF 1–2 dex below what is possible with individually detected sources at $z > 0.3$.

(iii) Converting our radio LFs to SFRd, we find that $M_* > 10^{9.5} M_\odot$ galaxies make up $\gtrsim 50$ per cent of the total SFRd in the interval $0.1 < z < 4$, in good agreement with other estimates from the optical and mid-IR.

Given the wealth of existing multiwavelength data at far-IR/submm wavelengths and the promise of large-scale radio continuum surveys in the near future leading to the Square Kilometre Array (SKA), it is clear that statistical methods such as the one presented here will be invaluable in determining the cosmic history of star-forming galaxies. As a final result, we show that the precision of our method scales like $\sqrt{N_s}$, where N_s is the number of sources used in the pixel stack. Thus, it is clear that even with large area, shallow radio surveys such as those proposed for the SKA pathfinders ASKAP and MeerKAT (EMU, Norris et al. 2011; MIGHTEE, Jarvis et al. 2012) it will be possible to obtain interesting results about the very high-redshift Universe well below the nominal detection limits for individual sources.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their constructive comments which enhanced this work. We would also like to thank J. Peacock, A. Karim and J. Dunlop for useful conversations which helped this work, E. Schinnerer for providing the VLA-COSMOS dirty beam map, A. Muzzin and D. Marchesini for both making the UL-TRAVISTA catalogue public and providing the catalogue completeness curves, and D. Sobral for providing the mass-limited $H\alpha$ LF estimates.

REFERENCES

Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, *MNRAS*, 356, 1191
 Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57 (B13)
 Bell E. F., 2003, *ApJ*, 586, 794
 Best P. N., Kauffmann G., Heckman T. M., Ivezić Ž., 2005, *MNRAS*, 362, 9
 Bouwens R. J., Illingworth G. D., Blakeslee J. P., Franx M., 2006, *ApJ*, 653, 53
 Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2008, *ApJ*, 686, 230
 Bouwens R. J. et al., 2011, *ApJ*, 737, 9
 Buat V. et al., 2010, *MNRAS*, 409, L1
 Chabrier G., 2003, *PASP*, 115, 763

Condon J. J., 1992, *ARA&A*, 30, 575
 Condon J. J., Cotton W. D., Broderick J. J., 2002, *AJ*, 124, 675
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
 Goodman J., Weare J., 2010, *Comm. Appl. Math. Comput. Sci.*, 5, 65
 Gruppioni C. et al., 2013, *MNRAS*, 432, 23
 Haarsma D. B., Partridge R. B., Windhorst R. A., Richards E. A., 2000, *ApJ*, 544, 641
 Hatziminaoglou E. et al., 2010, *A&A*, 518, L33
 Henriques B. M. B., Thomas P. A., Oliver S., Roseboom I., 2009, *MNRAS*, 396, 535
 Hopkins A. M., Beacom J. F., 2006, *ApJ*, 651, 142
 Hopkins A. M., Connolly A. J., Haarsma D. B., Cram L. E., 2001, *AJ*, 122, 288
 Hughes D. H. et al., 1998, *Nature*, 394, 241
 Jarvis M. J., 2012, *Afr. Skies*, 16, 44
 Karim A. et al., 2011, *ApJ*, 730, 61 (K11)
 Kennicutt R. C., Jr, 1998, *ARA&A*, 36, 189
 Kewley L. J., Geller M. J., Jansen R. A., Dopita M. A., 2002, *AJ*, 124, 3135
 Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., Orsi A., Silva L., Granato G. L., Bressan A., 2010, *MNRAS*, 405, 2
 Lilly S. J., Le Fevre O., Hammer F., Crampton D., 1996, *ApJ*, 460, L1
 Lutz D. et al., 2008, *ApJ*, 684, 853
 McCracken H. J. et al., 2012, *A&A*, 544, A156
 McLure R. J., Dunlop J. S., Cirasuolo M., Koekemoer A. M., Sabbi E., Stark D. P., Targett T. A., Ellis R. S., 2010, *MNRAS*, 403, 960
 McLure R. J. et al., 2013, *MNRAS*, 432, 2696
 Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, *MNRAS*, 283, 1388
 Mauch T., Sadler E. M., 2007, *MNRAS*, 375, 931
 Meurer G. R., Heckman T. M., Calzetti D., 1999, *ApJ*, 521, 64
 Mitchell-Wynne K., Santos M. G., Afonso J., Jarvis M. J., 2014, *MNRAS*, 437, 2270
 Moster B. P., Somerville R. S., Newman J. A., Rix H.-W., 2011, *ApJ*, 731, 113
 Muzzin A. et al., 2013a, *ApJS*, 206, 8
 Muzzin A. et al., 2013b, *ApJ*, 777, 18
 Norris R. P. et al., 2011, *Publ. Astron. Soc. Aust.*, 28, 215
 Oliver S. et al., 2010, *MNRAS*, 405, 2279
 Reddy N. A., Steidel C. C., Pettini M., Adelberger K. L., Shapley A. E., Erb D. K., Dickinson M., 2008, *ApJS*, 175, 48
 Reddy N. A., Erb D. K., Pettini M., Steidel C. C., Shapley A. E., 2010, *ApJ*, 712, 1070
 Roseboom I. G. et al., 2012, *MNRAS*, 426, 1782
 Roseboom I. G. et al., 2013, *MNRAS*, 436, 430
 Sandage A., Tammann G. A., Yahil A., 1979, *ApJ*, 232, 352 (STY)
 Santini P. et al., 2009, *A&A*, 504, 751
 Saunders W., Rowan-Robinson M., Lawrence A., Efstathiou G., Kaiser N., Ellis R. S., Frenk C. S., 1990, *MNRAS*, 242, 318
 Schinnerer E. et al., 2007, *ApJS*, 172, 46
 Scoville N. et al., 2007, *ApJS*, 172, 1
 Smolčić V. et al., 2009, *ApJ*, 690, 610
 Sobral D., Best P. N., Matsuda Y., 2012, *MNRAS*, 420, 1926
 Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, *MNRAS*, 437, 3516
 van Haarlem M. P. et al., 2013, *A&A*, 556, A2
 Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, *ApJ*, 691, 1879

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