Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **466,** 861–883 (2017)
Advance Access publication 2016 November 30

doi:10.1093/mnras/stw3088

A deep ALMA image of the *Hubble Ultra Deep Field*

J. S. Dunlop,[1⋆] R. J. McLure,[1] A. D. Biggs,[2] J. E. Geach,[3] M. J. Michałowski,[1] R. J. Ivison,[1,2] W. Rujopakarn,[4] E. van Kampen,[2] A. Kirkpatrick,[5] A. Pope,[5] D. Scott,[6] A. M. Swinbank,[7] T. A. Targett,[8] I. Aretxaga,[9] J. E. Austermann,[10] P. N. Best,[1] V. A. Bruce,[1] E. L. Chapin,[11] S. Charlot,[12] M. Cirasuolo,[2] K. Coppin,[3] R. S. Ellis,[2] S. L. Finkelstein,[13] C. C. Hayward,[14] D. H. Hughes,[9] E. Ibar,[15] P. Jagannathan,[16] S. Khochfar,[1] M. P. Koprowski,[3] D. Narayanan,[17] K. Nyland,[18] C. Papovich,[19] J. A. Peacock,[1] G. H. Rieke,[20] B. Robertson,[21] T. Vernstrom,[22] P. P. van der Werf,[23] G. W. Wilson[5] and M. Yun[5]

Affiliations are listed at the end of the paper

Accepted 2016 November 25. Received 2016 November 25; in original form 2016 June 6

ABSTRACT

We present the results of the first, deep Atacama Large Millimeter Array (ALMA) imaging covering the full ≃4.5 arcmin^2 of the *Hubble Ultra Deep Field* (HUDF) imaged with Wide Field Camera 3/IR on *HST*. Using a 45-pointing mosaic, we have obtained a homogeneous 1.3-mm image reaching $\sigma_{1.3} \simeq 35$ μJy, at a resolution of ≃0.7 arcsec. From an initial list of ≃50 $> 3.5\sigma$ peaks, a rigorous analysis confirms 16 sources with $S_{1.3} > 120$ μJy. All of these have secure galaxy counterparts with robust redshifts ($\langle z \rangle = 2.15$). Due to the unparalleled supporting data, the physical properties of the ALMA sources are well constrained, including their stellar masses (M_*) and UV+FIR star formation rates (SFR). Our results show that stellar mass is the best predictor of SFR in the high-redshift Universe; indeed at $z \geq 2$ our ALMA sample contains seven of the nine galaxies in the HUDF with $M_* \geq 2 \times 10^{10}$ M$_\odot$, and we detect only one galaxy at $z > 3.5$, reflecting the rapid drop-off of high-mass galaxies with increasing redshift. The detections, coupled with stacking, allow us to probe the redshift/mass distribution of the 1.3-mm background down to $S_{1.3} \simeq 10$ μJy. We find strong evidence for a steep star-forming 'main sequence' at $z \simeq 2$, with SFR $\propto M_*$ and a mean specific SFR $\simeq 2.2$ Gyr^{-1}. Moreover, we find that ≃85 per cent of total star formation at $z \simeq 2$ is enshrouded in dust, with ≃65 per cent of all star formation at this epoch occurring in high-mass galaxies ($M_* > 2 \times 10^{10}$ M$_\odot$), for which the average obscured:unobscured SF ratio is ≃200. Finally, we revisit the cosmic evolution of SFR density; we find this peaks at $z \simeq 2.5$, and that the star-forming Universe transits from primarily unobscured to primarily obscured at $z \simeq 4$.

Key words: galaxies: evolution – galaxies: high-redshift – galaxies: starburst – cosmology: observations – submillimetre: galaxies.

1 INTRODUCTION

A complete understanding of cosmic star formation history, and the physical mechanisms that drive galaxy formation and evolution, requires that we connect our UV/optical and infrared/mm views of the Universe (e.g. Hopkins & Beacom 2006; Dunlop 2011; Burgarella et al. 2013; Madau & Dickinson 2014). Until the advent of Atacama Large Millimeter Array (ALMA), these two views have been largely disconnected, for both technical and physical reasons. Benefiting from low background and high angular resolution, deep UV/optical surveys have proved extremely effective at completing our inventory of *unobscured* star formation which, certainly at high redshift, is dominated by large numbers of low-mass galaxies with individual star formation rates (SFR) ≃1 M$_\odot$ yr^{-1} (e.g. McLure et al. 2013; Bouwens et al. 2015; Bowler et al. 2015; Finkelstein et al. 2015; McLeod et al. 2015).

However, UV/optical observations are unable to uncover the most extreme star-forming galaxies, which, following the breakthroughs in far-IR/sub-mm astronomy at the end of the last century, are

⋆ E-mail: jsd@roe.ac.uk

known to be enshrouded in dust (Smail, Ivison & Blain 1997; Barger et al. 1998; Hughes et al. 1998; Eales et al. 1999). Such objects have now been uncovered in significant numbers through far-IR/mm surveys with the James Clerk Maxwell Telescope (e.g. Scott et al. 2002; Coppin et al. 2006; Scott, Dunlop & Serjeant 2006; Austermann et al. 2010; Geach et al. 2013), IRAM (Dannerbauer et al. 2004; Greve et al. 2004; Lindner et al. 2011), APEX (Weiss et al. 2009; Smolcic et al. 2012), ASTE (Scott et al. 2008, 2010; Hatsukade et al. 2010, 2011), BLAST (Devlin et al. 2009; Dunlop et al. 2010) and *Herschel* (Eales et al. 2010; Oliver et al. 2010; Elbaz et al. 2011; Lutz et al. 2011), with 'sub-mm galaxies' now known out to redshifts $z > 6$ (Riechers et al. 2013).

While sub-mm surveys for high-redshift galaxies benefit from a strong negative *K*-correction [provided by a modified blackbody spectral energy distribution (SED); Blain & Longair 1993; Hughes et al. 1993; Dunlop et al. 1994], the high background and/or the relatively poor angular resolution provided by single-dish telescopes at these wavelengths means that they are only really effective at uncovering rare, extreme star-forming galaxies with SFR $> 300\,M_{\odot}\,yr^{-1}$ (albeit reaching down to SFR $> 100\,M_{\odot}\,yr^{-1}$ in the very deepest Submillimetre Common-User Bolometer Array-2 (SCUBA-2) 450/850 μm imaging; Geach et al. 2013; Roseboom et al. 2013; Koprowski et al. 2016). The existence of such objects presents an interesting and important challenge to theoretical models of galaxy formation (e.g. Baugh et al. 2005; Khochfar & Silk 2009; Davé et al. 2010; Hayward et al. 2011; Narayanan et al. 2015), but they provide only ≃10–15 per cent of the known far-infrared/mm background (Fixsen et al. 1998; Scott et al. 2012; Geach et al. 2013), and attempts to complete our inventory of obscured star formation have had to rely on stacking experiments (e.g. Peacock et al. 2000; Marsden et al. 2009; Geach et al. 2013; Coppin et al. 2015).

A key goal, therefore, of deep surveys with ALMA is to close the depth/resolution gap between UV/optical and far-infrared/mm studies of the high-redshift Universe, and hence enable a complete study of visible+obscured star formation within the overall galaxy population. Over the last 2 yr, ALMA has begun to make important contributions in this area. Most early ALMA programmes have focused (sensibly) on pointed observations of known objects (e.g. Hodge et al. 2013; Karim et al. 2013; Ouchi et al. 2013; Bussmann et al. 2015; Capak et al. 2015; Maiolino et al. 2015; Simpson et al. 2015a,b; Scoville et al. 2016), including gravitationally lensed sources (e.g. Weiss et al. 2013; Watson et al. 2015; Béthermin et al. 2016; Knudsen et al. 2016; Spilker et al. 2016). However, strenuous efforts have been made to exploit the resulting combined 'blank-field' survey by-product to improve our understanding of the deep mm source counts (e.g. Ono et al. 2014; Carniani et al. 2015; Fujimoto et al. 2016; Oteo et al. 2016) albeit with interestingly different results. More recently, time has been awarded to programmes that aim to deliver contiguous ALMA mosaic imaging of small regions of sky with excellent multiwavelength supporting data (e.g. Hatsukade et al. 2015; Umehata et al. 2015). Such programmes offer not only further improvements in our knowledge of the sub-mm/mm source counts, but also the ability to determine the nature and physical properties (redshifts, stellar masses, SFR) of the ALMA-detected galaxies. For example, ALMA 1.1-mm imaging of 1.5 arcmin^2 within the CANDELS/UDS field (PI: Kohno) has provided new results on the 1.1-mm counts, and enabled the study of several ALMA-detected galaxies (Tadaki et al. 2015; Hatsukade et al. 2016).

However, to date, no homogeneous ALMA imaging has been undertaken within the best-studied region of deep 'blank-field' sky, the *Hubble Ultra Deep Field* (HUDF). On scales of a few arcmin^2, the HUDF remains unarguably the key ultradeep extragalactic survey field and, lying within the GOODS-South field at RA 03^h, Dec. -28°, is ideally located for deep ALMA observations. While four of the six Hubble Frontier Fields[1] provide alternative target fields for deep ALMA observations, the quality of the optical-near-infrared data in these fields will never seriously rival that which has already been achieved in the HUDF. In part this is due to the huge investment in *HST* optical imaging in this field made prior to the degradation of the ACS camera (Beckwith et al. 2006). However, it is also a result of the more recent investment in imaging with Wide Field Camera 3 (WFC3)/IR on the *HST* since 2009. Specifically, the combination of the UDF09 campaign (Bouwens et al. 2010; Oesch et al. 2010; Illingworth et al. 2013) followed by the UDF12 programme (Dunlop et al. 2013; Ellis et al. 2013; Koekemoer et al. 2013), has delivered the deepest near-infrared imaging ever achieved (reaching 30 AB mag, 5σ) over an area of 4.5 arcmin^2. As a result of coupling this multiband *HST* imaging with the recently augmented ultradeep *Spitzer* data (Labbé et al. 2015), accurate photometric redshifts, stellar masses and UV SFR are now known for ≃3000 galaxies in this field (e.g. Parsa et al. 2016). For a field of this size, the HUDF is also uniquely rich in optical/infrared *spectroscopic* information, with a combination of ground-based optical spectroscopy and *HST* WFC3/IR near-infrared grism spectroscopy delivering redshifts and emission-line strengths for over 300 galaxies (see Section 2.2). Finally, the HUDF lies in the centre of the *Chandra* Deep Field South (CDFS) 4-ms X-ray imaging (Xue et al. 2011), and has recently been the focus of a new programme of ultradeep radio imaging with the JVLA (PI: Rujopakarn).

The aim of the work presented here was to exploit this unique data base by using ALMA to construct the deepest homogeneous mm-wavelength image obtained to date on the relevant scales. As described in detail in the next section, ≃20 h of ALMA observations were approved in Cycle 1, and rolled over into Cycle 2, to enable us to complete a 1.3-mm mosaic covering the full 4.5 arcmin^2 imaged with WFC3/IR, seeking to reach an rms depth of $\sigma_{1.3} \simeq 30$ μJy (PI: Dunlop). We chose to undertake this first deep ALMA image of the HUDF at 1.3 mm (rather than at shorter wavelengths) for three reasons. First, in practice it maximizes sensitivity to higher redshift dusty star-forming galaxies at $z > 3$. Second, it is at these longer wavelengths that the resolution of single-dish surveys is undoubtedly poorest, and hence the imaging most confused. Third, this decision aided the feasibility of the observations in early ALMA cycles, with only 45 pointings required to complete the mosaic, and both nighttime and daytime observations being acceptable. Astrophysically, we sought to reach detections 4–5 times deeper than can be achieved with the deepest single-dish surveys (corresponding to SFR ≃25 $M_{\odot}\,yr^{-1}$ out to the very highest redshifts), and to exploit the uniquely complete HUDF galaxy data base in deep stacking experiments.

Data taking for this project commenced in 2014, and was completed in summer 2015, and in this paper we present the first results. We present and discuss the properties of the ALMA map, the sources uncovered within it, and the implications for our understanding of cosmic star formation and galaxy evolution. The remainder of this paper is structured as follows. In Section 2, we describe the ALMA observations, explain how the data were reduced and provide a summary of all the key multiwavelength supporting data in the field. In Section 3, we explain how sources were extracted from the ALMA map, and then, in Section 4, describe how cross-identifications with

[1] http://www.stsci.edu/hst/campaigns/frontier-fields/

the *HST* sources in the field enabled us to clean the source list to a final sample of 16 robust ALMA sources. In Section 5, we consider the implications of the number of sources we have detected, aided by the results of source injection and retrieval simulations, and compare our results to other recent estimates of deep mm number counts. In Section 6, we derive the physical properties of the sources we have detected, and explore the implications for star formation in galaxies at $z = 1$–3. Then, in Section 7, we present the results of stacking the 1.3-mm signal on the positions of known galaxy populations in the HUDF, and consider the consequences for the mm-wavelength background and for the ratio of obscured/unobscured star formation over cosmic history. Finally, we discuss the astrophysical implications of our findings in Section 8, and summarize our conclusions in Section 9. Throughout, all magnitudes are quoted in the AB system (Oke 1974; Oke & Gunn 1983), and all cosmological calculations assume a flat cold dark matter cosmology with $\Omega_{\rm M} = 0.3$, $\Omega_{\Lambda} = 0.7$, and $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$.

2 DATA

2.1 ALMA observations and data reduction

The ALMA observations of the HUDF were taken during two separate observing seasons – the first nine execution blocks (EBs) in 2014 July–September and the remaining four in 2015 May. As the primary goal of these observations was to produce a continuum map of the HUDF, the correlator was configured to process the maximum 7.5-GHz bandwidth in the form of four 1875 MHz-wide spectral windows, each with 128 dual-polarization channels. However, the velocity resolution of 40–45 km s^{-1} is still sufficient to resolve spectral lines that are typically observed in high-redshift star-forming galaxies. The correlator averaging time was 2 s per sample.

The HUDF was observed using a 45-pointing mosaic, with each pointing separated by 0.8 times the antenna beamsize. This mosaic pattern was observed twice per EB, except for one which terminated after only 20 pointings of the first mosaic pass. However, no problems were found with these data and they were included in the final map. The amplitude and bandpass calibrator for each EB was the unresolved quasar J0334−401, this also serving as the phase calibrator during the second observing season. Although relatively far (12.°4) from the HUDF, the phase solutions varied smoothly over the course of each EB, and maps made from these data demonstrated that phase referencing had indeed been successful. In the first season, the phase calibrator was J0348−2749 that is only 3.°5 from the target. The array configuration varied greatly during the observations, with the first season generally using baselines twice as long as required to achieve the requested angular resolution. A summary of the observations is given in Table 1.

All data reduction was carried out using CASA and followed standard procedures. First, the data from the second season needed to be corrected for incorrect antenna positions that had been used during correlation. Other a priori calibrations included application of system temperature tables and water vapour radiometer phase corrections. The latter were particularly large and time variable for the second season, presumably as a result of observing during the day. Very little data needed to be flagged, a notable exception being the outer four channels of each spectral window which have very poor sensitivity. After the removal of the frequency response of each antenna using the bandpass calibrator, amplitude and phase corrections were calculated as a function of time for the flux and phase calibrators. The flux scale was then set with reference to the regularly monitored flux density of J0334−401 and the gain solutions interpolated on to the HUDF scans.

Table 1. Summary of the ALMA observations of the HUDF. The date of each EB is given along with the approximate maximum baseline length and the average amount of precipitable water vapour (PWV).

Observing date	Maximum baseline (m)	PWV (mm)
2014 July 18	650	0.43
2014 July 29	820	1.04
2014 August 17	1100	0.94
2014 August 18	1250	1.51
2014 August 18	1250	1.45
2014 August 27	1100	1.35
2014 August 28	1100	1.20
2014 August 28	1100	1.25
2014 September 1	1100	1.08
2015 May 16	550	0.65
2015 May 16	550	0.80
2015 May 17	550	1.00
2015 May 17	550	1.80

A continuum mosaicked image of the calibrated data was produced using the task CLEAN. To enhance mapping speed, the data were first averaged in both frequency and time to produce a data set with 10 frequency channels per spectral window and a time sampling of 10 s. The data were naturally weighted for maximum sensitivity, but the relatively large array configurations still produced a synthesized beam (589×503 mas^2) that was significantly smaller than the circular 0.7-arcsec beam that had been requested. As this would potentially lead to problems with detecting resolved sources, we experimented with various u, v tapers in order to find the best combination of angular resolution and mosaic sensitivity. A $\simeq 220 \times 180\,k\lambda$ taper, with position angle (PA) oriented to circularize the beam as much as possible, produced a beam close to that requested (707×672 mas^2) and a final mosaic sensitivity as measured over a large central area of the map of 34 μJy beam^{-1}. As the detected source flux densities were very weak, and the synthesized beam sidelobes very low, no deconvolution (cleaning) was performed. The resulting image is shown in Fig. 1. Finally, to aid checks on data quality, and source reality, we also constructed three alternative 50:50 splits of the ALMA 1.3-mm image, splitting the data in half by observing date, sideband and polarization.

2.2 Supporting multifrequency data

2.2.1 Optical/near-infrared imaging

The key data set that defined the area that we aimed to cover with the ALMA 1.3-mm mosaic is the ultradeep near-infrared imaging of the HUDF obtained with WFC3/IR on *HST* via the UDF09 (e.g. Bouwens et al. 2010; Bunker et al. 2010; Finkelstein et al. 2010, 2012; McLure et al. 2010; Oesch et al. 2010) and UDF12 (e.g. Dunlop et al. 2013; Ellis et al. 2013; Illingworth et al. 2013; McLure et al. 2013; Finkelstein et al. 2015) programmes. As described in Koekemoer et al. (2013), the final UDF12 WFC3/IR imaging reaches a 5σ detection depth of 29.7 mag in the Y_{105} filter, and 29.2 mag in the J_{125}, J_{140} and H_{160} filters (total magnitudes, as derived from small-aperture magnitudes assuming point-source corrections). These unparalleled near-infrared data, covering an area $\simeq$4.5 arcmin^2, are complemented by what remains the deepest ever

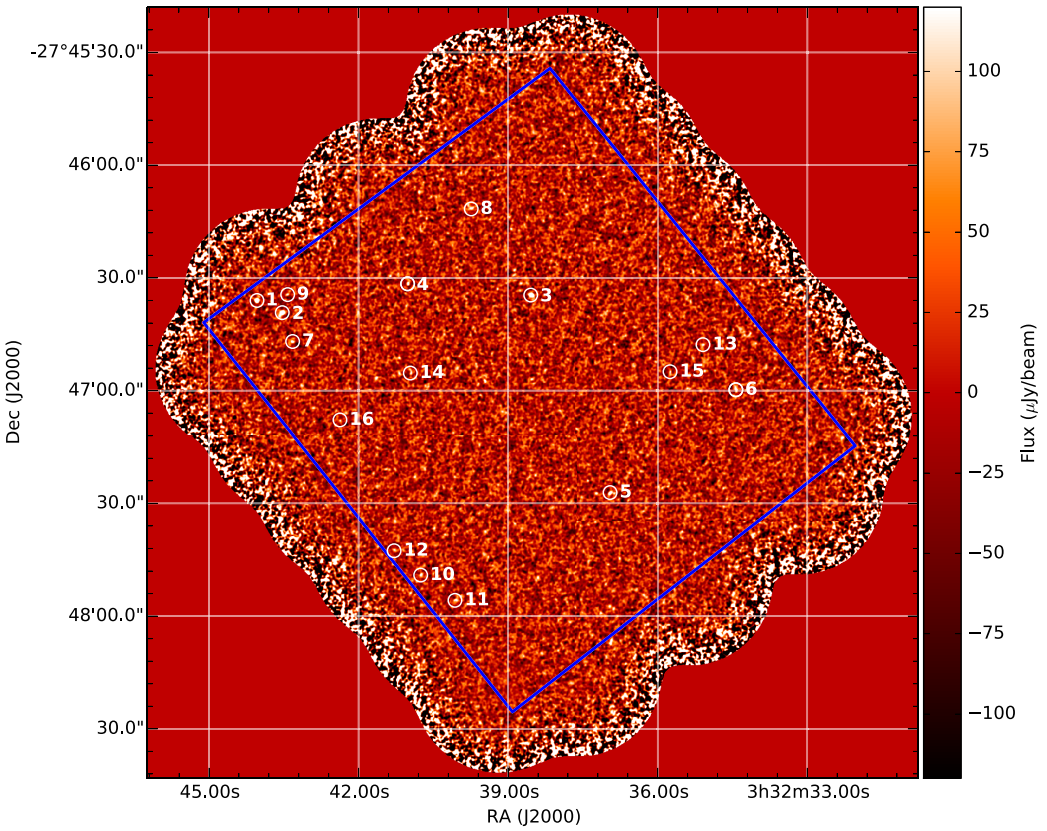

Figure 1. The ALMA 1.3-mm map of the HUDF, with the positions of the 16 sources listed in Table 2 marked by 3.6-arcsec diameter circles. The border of the homogenously deep region of near-infrared WFC3/IR imaging obtained through the UDF09 and UDF12 *HST* programmes is indicated by the dark-blue rectangle. The ALMA image, constructed from a mosaic of 45 individual pointings, provides homogeneous 1.3-mm coverage of this region, with a typical noise per beam of $\sigma_{1.3} \simeq 35$ μJy.

optical imaging obtained with ACS on *HST* (Beckwith et al. 2006). This provides imaging in the B_{435}, v_{606}, i_{775} and z_{850} filters, reaching 5σ detection depths of 29.7, 30.2, 29.9 and 29.8 mag, respectively. More recently, the CANDELS programme (Grogin et al. 2011) has provided deep i_{814} data across the HUDF (reaching 29.8 mag, 5σ) as part of the CANDELS-DEEP imaging of GOODS-South (Koekemoer et al. 2011; see also Guo et al. 2013).

The core *HST* imaging data set is extended to shorter wavelengths by the inclusion of deep VLT VIMOS imaging in the *U*-band (reaching 28 mag, 5σ; Nonino et al. 2009), and to longer wavelengths by the deepest ever K_s-band imaging obtained with HAWK-I on the VLT through the HUGS survey (Fontana et al. 2014), which reaches $K_s = 26.5$ mag (5σ). Imaging longward of 2.2 μm has been obtained with *Spitzer*, with new ultradeep IRAC imaging of the HUDF at 3.6 and 4.5 μm being provided by our own stack of the available public data described by Labbé et al. (2015) (see also Ashby et al. 2013, 2015). This reaches deconfused 5σ detection depths of ≃26.5 mag at 3.6 μm and ≃26.3 mag at 4.5 μm.

Galaxy detection and photometry in the deep *HST* imaging data set was undertaken using SExtractor v2.8.6 (Bertin & Arnouts 1996) in dual image mode with H_{160} as the detection image, and the *HST* photometry homogenized through appropriately scaled apertures at shorter wavelengths. The ground-based (*U* and K_s) and *Spitzer* photometry was extracted by deconfusing the data using *HST* positional priors both using the method described in McLure et al. (2011, 2013), and independently using tphot (Merlin et al. 2015).

The resulting optical–near-infrared catalogue contains ≃2900 objects with 12-band photometry (see for example Parsa et al. 2016).

2.2.2 *Mid/far-infrared/sub-mm imaging*

Longward of 4.5 μm, the original GOODS *Spitzer* imaging (PID 104; PI Dickinson) provides the deepest available data at 5.6, 8.0 (from IRAC: Fazio et al. 2004) and 24 μm (from MIPS). The 24 μm imaging has been augmented and incorporated within the *Spitzer* Far-Infrared Deep Extragalactic Legacy[2] survey (Magnelli et al. 2009) and reaches a 5σ detection limit of $S_{24} \simeq 30$ μJy.

Data at longer far-infrared wavelengths are provided by *Herschel* (Pilbratt et al. 2010), and we utilize here the final public image products from three major guaranteed-time surveys. PACS (Poglitsch et al. 2010) images at 100 and 160 μm, reaching rms depths of 0.17 and 0.42 mJy respectively are provided by a combination of the data obtained through the GOODS-*Herschel* (Elbaz et al. 2011) and the PACS Evolutionary Probe (Lutz et al. 2011) surveys, while SPIRE (Griffin et al. 2010) images at 250, 350 and 500 μm, reaching 5.86, 6.34 and 6.88 mJy respectively (including confusion noise) are provided by the *Herschel* Multitiered Extragalactic Survey (Oliver et al. 2010, 2012).

Because the *Herschel* (and especially the SPIRE) imaging has such low angular resolution compared to the ALMA imaging, care

[2] PI M. Dickinson, see http://www.noao.edu/noao/fidel/

must be taken to attempt to deconfuse the *Herschel* images in order to avoid obtaining biased, or artificially accurate far-infrared photometry for the ALMA sources (see Appendix A, Fig. A1). We therefore fitted the *Herschel* maps of the HUDF region with appropriate beams centred at the positional priors of all the ALMA and 24 μm sources in the field. The best-fitting beam normalizations, and associated covariance errors then allowed us to extract *Herschel* fluxes/non-detections.

2.2.3 *Radio imaging*

Until recently, the deepest available radio imaging in the HUDF was provided by the 1.4 GHz observations of the extended CDFS as described by Kellermann et al. (2008) and Miller et al. (2008). This produced imaging with a 2.8 arcsec × 1.6 arcsec beam reaching a typical rms sensitivity of $\sigma_{1.4} \simeq 7.5$ μJy. However, recently (2014 March to 2015 September) a new, ultradeep, JVLA 4–8 GHz survey has been undertaken within GOODS-South, with a single pointing (7.2 arcmin primary beam at 6 GHz) centred on the HUDF (at RA $03^{\rm h}\ 32^{\rm m}\ 38.6^{\rm s}$, Dec. $-27^{\circ}46'59\farcs83$). This new imaging, which we utilize here and in a companion paper on mm/radio source sizes within the HUDF (Rujopakarn et al. 2016), comprises 149, 17 and 11 h of imaging in the A, B and C configurations, respectively. The result is an image with a synthesized beam of 0.31 arcsec × 0.61 arcsec (PA $= -3\fdg6$), reaching an rms sensitivity at 6 GHz of $\sigma_6 \simeq 0.32$ μJy per beam at the phase centre, and $\sigma_6 \simeq 0.35$ μJy per beam at the edge of the HUDF. This imaging, which in effect (for a power-law radio spectral slope of $\alpha = 0.8$, $f_\nu \propto \nu^{-\alpha}$) is ≃10 times deeper than the pre-existing 1.4 GHz radio map, reveals 27 radio sources with peak S/N > 5 within the 4.5 arcmin^2 area of the HUDF marked in Fig. 1.

2.2.4 *X-ray imaging*

The deepest X-ray imaging in the HUDF is provided by the 4 ms imaging with *Chandra* of the CDFS that reaches an on-axis flux-density detection limit of $\simeq 3 \times 10^{-17}$ erg cm^{-2} s^{-1} across the full soft+hard band (0.5–8 keV) (Xue et al. 2011). Various authors have studied the galaxy counterparts of the X-ray sources within the wider GOODS-South field (e.g. Rangel et al. 2013; Hsu et al. 2014; Giallongo et al. 2015) but for the present study focused on the HUDF we work with the original X-ray positions, and establish our own galaxy identifications and redshift information as required.

2.2.5 *Optical/near-infrared spectroscopy*

New spectroscopic observations of the HUDF were taken with MUSE as part of the guaranteed-time programme between 2014 September and 2015 January. The MUSE IFU provides full spectral coverage spanning 4770–9300 Å, and a contiguous field of view of 60 arcsec × 60 arcsec, with a spatial sampling of 0.2 arcsec pixel^{-1}, and a spectral resolution of $R = 3500$ at $\lambda = 7000$ Å. The publicly available MUSE data in the HUDF comprises a 3 × 3 mosaic of ≃ 18.2 ks integrations, plus a single deep ≃ 65 ks exposure in the centre of the field.

We downloaded the public data set and reduced it using the ESOREX pipeline. This pipeline identifies the location of the data on the CCD using the flat-field image, and then extracts the flat-field, arc and science data. It then wavelength calibrates and flat-fields each slice and constructs the data cube. Each science exposure was interspersed with a flat-field to improve the slice-by-slice flat-field (illumination). Residual slice-to-slice variations were then modelled and removed using a set of custom routines that attempt to model the (wavelength-dependent) offsets. Sky subtraction was performed on each subexposure by identifying and subtracting the sky emission using blank areas of sky at each wavelength slice (after masking continuum sources), and the final mosaics were constructed using an average with a 3σ-clip to reject cosmic rays, using point sources in each (wavelength collapsed) image to register the cubes. The final cube was then registered to the *HST* / WFC3/IR J_{125} image using point sources in both frames. Flux calibration was carried out using observations of known standard stars at similar airmass that were taken immediately before or after the science observations (and in each case we confirmed the flux calibration by measuring the flux density of stars with known photometry in the MUSE science field). More details on the MUSE HUDF project will be provided in Bacon et al. (in preparation).

To search for redshifts from each ALMA-identified source, we extracted one- and two-dimensional spectra from a 1.5 arcsec × 1.5 arcsec region centred at the ALMA position and searched for emission and absorption lines. This yielded spectroscopic redshifts for 6 of the 16 sources listed in Table 2, of which 4 are new, with the other 2 confirming previous ground-based redshifts derived using FORS2 on the VLT (Vanzella et al. 2008; Kurk et al. 2013).

The new redshifts being provided by MUSE add to an already impressive data base of spectroscopic redshifts in the GOODS-South field, and in the HUDF in particular. The various pre-existing ground-based spectroscopic campaigns are summarized in Parsa et al. (2016), but in recent years *HST* WFC3/IR near-infrared grism spectroscopy has also made an important contribution, with a combination of the 3D-HST programme, and CANDELS supernovae grism follow-up observations delivering ≃1000 redshifts in the GOODS-South field (Brammer et al. 2012; Skelton et al. 2014; Morris et al. 2015; Momcheva et al. 2016). While many of the *HST* redshifts simply provide (useful) confirmation of the results of earlier ground-based spectroscopic observations, this grism work has been particularly helpful in filling in the traditional 'redshift desert' between $z \simeq 1.2$ and $z \simeq 2$, where relatively few strong emission lines are accessible in the optical regime.

In total, these multiple efforts (extending over the last ≃15 yr) have yielded spectroscopic redshifts for nearly 3000 galaxies in the GOODS-South field, with over 200 robust spectroscopic redshifts now available within the subregion defined by the HUDF. As a consequence of this uniquely rich/dense spectroscopic data base, we are able to provide spectroscopic redshifts for 13 of the 16 galaxies in the final ALMA-selected sample (the selection of which is described below in Sections 3 and 4). These redshifts, along with the appropriate references, are given in the final column of Table 2.

The redshift of one ALMA-identified source (UDF3) is confirmed independently from our ALMA observations, via the detection of three spectral lines from CO, CI and H_2O in our ALMA data cube (see later, and Ivison et al., in preparation).

3 ALMA SOURCE EXTRACTION

To detect sources in the ALMA image, we first constructed a noise map that provides an estimate of the local pixel-to-pixel variance on scales comparable to the beam. For every pixel, we evaluated the standard deviation of flux-density values within a window of size $10 \times \theta$ where $\theta = \sqrt{(1.331 \times a \times b)}$, where a and b are the semimajor and semiminor axes of the synthesized beam. To mitigate the contribution from bright sources, we applied local 4σ clipping before evaluating the standard deviation. This noise map

Table 2. Details of the final sample of 16 ALMA-detected sources in the HUDF, selected and refined as described in Sections 3 and 4. Column 1 gives source numbers as also used in Fig. 1, while columns 2 and 3 give the positions of the ALMA sources as determined from the 1.3-mm map. Estimated total flux densities (see Section 4.3 for details on corrections to point-source flux densities) and peak S/N at 1.3 mm are given in columns 4 and 5. Then in columns 6 and 7, we give the co-ordinates of the adopted galaxy counterpart as determined from the H_{160} WFC3/IR *HST* image of the HUDF. Columns 8 and 9 give the positional offsets between the ALMA and *HST* positions, before (Δ_1) and after (Δ_2) moving the *HST* positions south by 0.25 arcsec (see Section 4.2 for a discussion of this astrometric shift, and its calculation/motivation; the cumulative distributions produced by these two alternative sets of positional offsets are shown in Fig. 2). Column 10 gives the total H_{160} magnitude of each *HST* galaxy counterpart, while column 11 lists the redshift for each source. The 13 spectroscopic redshifts are given to three decimal places, with the three photometric redshifts given to two decimal places. The sources of the spectroscopic redshifts are indicated by the flag in column 12, and are as follows: (1) Brammer (private communication); 2) MUSE (this work); (3) Momcheva et al. (2016); (4) Kurk et al. (2013); (5) Hathi, Malhotra & Rhoads (2008); (6) Vanzella et al. (2008).

ID	RA (ALMA) (deg)	Dec. (ALMA) (deg)	$S_{1.3mm}$ (μJy)	S/N (1.3 mm)	RA (*HST*) (deg)	Dec. (*HST*) (deg)	Δ_1 (arcsec)	Δ_2 (arcsec)	H_{160} (AB mag)	z	Ref
UDF1	53.183 48	−27.776 67	924 ± 76	18.37	53.183 45	−27.776 58	0.33	0.13	24.75	3.00	
UDF2	53.181 37	−27.777 57	996 ± 87	16.82	53.181 40	−27.777 46	0.38	0.15	24.70	2.794	1
UDF3	53.160 62	−27.776 27	863 ± 84	13.99	53.160 60	−27.776 13	0.51	0.27	23.41	2.541	2
UDF4	53.170 90	−27.775 44	303 ± 46	6.63	53.170 90	−27.775 39	0.18	0.06	24.85	2.43	
UDF5	53.153 98	−27.790 87	311 ± 49	6.33	53.154 05	−27.790 91	0.24	0.42	23.30	1.759	3
UDF6	53.143 47	−27.783 27	239 ± 49	4.93	53.143 47	−27.783 21	0.22	0.03	22.27	1.411	2
UDF7	53.180 51	−27.779 70	231 ± 48	4.92	53.180 52	−27.779 65	0.21	0.06	24.17	2.59	
UDF8	53.165 59	−27.769 90	208 ± 46	4.50	53.165 55	−27.769 79	0.43	0.22	21.75	1.552	4
UDF9	53.180 92	−27.776 24	198 ± 39	4.26	53.181 05	−27.776 17	0.46	0.40	21.41	0.667	2
UDF10	53.169 81	−27.796 97	184 ± 46	4.02	53.169 69	−27.797 02	0.42	0.56	23.32	2.086	3
UDF11	53.166 95	−27.798 84	186 ± 46	4.02	53.166 90	−27.798 69	0.54	0.31	21.62	1.996	2, 4
UDF12	53.172 03	−27.795 17	154 ± 40	3.86	53.172 12	−27.795 09	0.39	0.28	27.00	5.000	5
UDF13	53.146 22	−27.779 94	174 ± 45	3.85	53.146 15	−27.779 88	0.31	0.24	23.27	2.497	3
UDF14	53.170 67	−27.782 04	160 ± 44	3.67	53.170 69	−27.781 97	0.24	0.06	22.76	0.769	2
UDF15	53.148 97	−27.781 94	166 ± 46	3.56	53.149 02	−27.781 96	0.18	0.36	23.37	1.721	3
UDF16	53.176 55	−27.785 50	155 ± 44	3.51	53.176 58	−27.785 45	0.22	0.09	21.42	1.314	2, 6

then allowed us to construct a signal-to-noise map that we used as the detection image. A simple peak-finding algorithm was adopted: first we identified significant ($>5\sigma$) peaks, and co-added these to construct a model point spread function (PSF). This PSF was then used to subtract sources from the map as they were identified, starting from the most significant peak and moving down until a threshold floor significance was reached.

We limited source detection to map regions with $\sigma_{1.3} < 40\,\mu$Jy, which yielded an effective survey area of 4.4 arcmin^2. Within this area, we detected 47 candidate sources with peak S/N > 3.5 and a point-source flux density $S_{1.3} \geq 120\,\mu$Jy. However, running an identical source extraction on the negative map (i.e. the real map multiplied by -1) yielded 29 sources with S/N > 3.5 and $S_{1.3} \geq 120\,\mu$Jy.

It is interesting to consider whether this is as expected. Adopting a beam angular radius of 0.35 arcsec, the map contains $\simeq$42 000 beams, and thus, based on Gaussian statistics, we would expect $\simeq$10 spurious peaks with S/N > 3.5. However, if, as pointed out by Condon (1997) and Condon et al. (1998), there are effectively twice as many statistically independent noise samples as naively expected, then these numbers rise to $\simeq$20 spurious peaks with S/N > 3.5, in much better accord with what is actually found from source extraction on the negative image. The noise level only then needs to be altered by <5 per cent to bring the numbers into essentially exact agreement. This suggests that there is no serious issue with the noise in the map, and indeed a full simulation of the image involving beam filtering of white noise confirms that the numbers and S/N distribution of the spurious sources as derived from the negative map are as expected (Peacock et al., in preparation).

The implication is that only $\simeq$15–20 of the 'sources' extracted from the positive image are real, and the challenge is to identify which these are.

4 GALAXY COUNTERPARTS AND SOURCE LIST REFINEMENT

4.1 Galaxy identifications

Refining the source list is not as straightforward as, for example, confining attention to sources with S/N > 4, given that there are seven such 'sources' in the negative map. A clean source list can be produced by limiting the selection to S/N > 6, but this leaves only five sources, and clearly does not make optimal use of the new ALMA data. Fortunately, we are able to use the excellent positional accuracy of the ALMA sources, along with the wealth of supporting multifrequency data, to identify which of the $>3.5\sigma$ peaks extracted as described in the previous section correspond to real ALMA sources.

First, it was very evident that the brightest sources in the ALMA source list had obvious galaxy counterparts in the *HST* imaging, with positions coincident to within <0.5 arcsec. Excellent positional correspondence is certainly expected since, even for a 3.5σ source, the predicted 1σ uncertainty in RA and Dec. given a beam-size of 0.7 arcsec (full width at half-maximum) is 0.085 arcsec, and the corresponding conservative 3σ search radius is $\simeq$0.25 arcsec (see Ivison et al. 2007). This level of positional accuracy is approached by the positional offsets between the ALMA and radio sources (albeit increased by a factor $\simeq$2 by image pixelization for a 10σ source), but ambiguity over the true centroid of some of the *HST* counterparts, astrometric uncertainties and potentially even optical-mm physical offsets combine to make the positional correspondence between the ALMA sources and their *HST* counterparts not quite as precise as theoretically predicted. None the less, for the obviously secure galaxy identifications confirmed by radio detections we found that $\sigma_{pos} = 0.2$ arcsec, and so adopted a search radius of 0.6 arcsec. This very small search radius makes

chance ALMA–*HST* coincidences very unlikely for all but the very faintest galaxies, and indeed, applied to the negative ALMA source list, yields only three random galaxy identifications.

Applied to the positive source sample, searching for near-infrared galaxy candidates within a radius of 0.6 arcsec (which obviously assumes that real ALMA sources have an *HST* counterpart in the UDF09+UDF12 imaging; see below) reduced the potential source sample to 21 sources, 12 of which are independently confirmed as real sources in the new ultradeep JVLA 6 GHz imaging (see Rujopakarn et al. 2016).

4.2 Astrometry issues

While performing the galaxy counterpart identification, we noticed a systematic positional offset, primarily in declination, between the ALMA positions and the positions of their (often obvious) galaxy counterparts. We checked this using the stacking results described below (see Section 7), and deduced that the ALMA positions lie systematically two ALMA pixels (i.e. $\simeq$0.24 arcsec) south of the *HST* positions.

At the same time we found that, for the 12 radio-detected ALMA sources, there was no significant systematic offset in position between the JVLA and ALMA positions, with the mean offset being $<$35 mas in both RA and Dec. (despite the use of different phase calibrators).

We explored this further, using the radio data that extend over somewhat more of GOODS-South than just the HUDF region, and found that the *HST* positions (based on the H_{160} HUDF/GOODS-South imaging) are systematically offset from the radio positions by +0.279 arcsec in Dec., and −0.076 arcsec in RA, and are systematically offset from the 2MASS positions by +0.247 arcsec in Dec. and +0.035 arcsec in RA (see Rujopakarn et al. 2016).

Given the apparent consistency of the offset in Dec., we experimented with simply shifting all *HST* positions south by 0.25 arcsec, and repeating the identification process. The impact of this change is documented in Table 2 (which gives the ALMA-*HST* offset in arcsec both before and after modifying the *HST* coordinate system), and in Fig. 2 , which shows the tightening of the distribution of positional offsets after applying this shift.

One might reasonably ask which coordinate system is correct? However, the agreement between the JVLA, ALMA and 2MASS positions strongly suggests that it is the *HST* coordinate system that is wrong. In fact, there is good evidence that this is the case. The HUDF and CANDELS astrometry has been tied to the GOODS ACS astrometric solution. Referring back to the documentation accompanying the GOODS 2008 data release, it transpires that, for v2.0 of the GOODS *HST* data, the GOODS team decided to shift the GOODS-North coordinate system south by 0.3 arcsec in Dec. However, a similar shift was not applied to the GOODS-South v2.0 image mosaic.[3] The stated rationale was the lack of available comparison data of the necessarily quality in the GOODS-South field at the time, although it was claimed that 'an analysis of Chandra Deep Field South astrometry by the MUSYC team using the Yale Southern Observatory Double Astrograph telescope suggests that the mean GOODS-S world coordinate system is absolutely accurate at a level better than 0.1 arcsec'. It seems clear, now, in the light of the new ALMA and JVLA data, that this is not the case, and that the GOODS-South world coordinate system should, as was done for GOODS-North v2.0, be moved south by $\simeq$0.25–0.3 arcsec. This could be implemented for future CANDELS/HUDF releases, but for now we continue to give *HST* co-ordinates in Table 2 in the existing GOODS-South/CANDELS/HUDF system (to ease object identification in existing *HST*-based catalogues), and simply note the improved positional correspondence achieved for our galaxy identifications when this astrometric shift is systematically applied.

[3] https://archive.stsci.edu/pub/hlsp/goods/v2/h_goods_v2.0_rdm.html#4.0

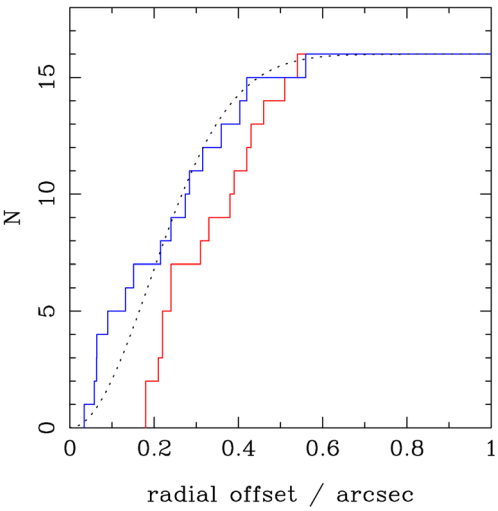

Figure 2. The cumulative distribution of radial separations between the 16 ALMA 1.3-mm sources and their adopted *HST* near-infrared galaxy counterparts, as detailed in Table 1. The red line shows the distribution prior to moving the *HST* astrometric reference frame, while the blue line shows the significantly tighter distribution which arises after moving all *HST* positions south by 0.25 arcsec (see Section 4.2). The dashed black line shows the cumulative distribution expected assuming a Gaussian distribution of positional errors, with $\sigma_{\rm RA} = \sigma_{\rm Dec.} = 0.2$ arcsec (see Section 4 for further details).

4.3 Final ALMA HUDF source sample

Application of this astrometric shift, as well as tightening the positional agreement for solid identifications, also led to the rejection of three others, and finally we also rejected the two sources for which the only available galaxy counterpart had $H_{160} > 28.5$. This latter decision was made on continuity grounds (no other remaining ALMA source has $H_{160} > 27$), and because, as evidenced from searching for galaxy counterparts to the negative pseudo-sources, the *HST* source density at these extreme depths is expected to yield $\simeq$2 chance coincidences within a search radius $r < 0.55$ arcsec.

This leaves the final sample of 16 sources whose positions in the ALMA map are indicated in Fig. 1. Accurate positions (both for the ALMA sources and their *HST* counterparts) and flux densities are given in Table 2 (along with ALMA-*HST* positional offsets before and after the afore-mentioned astrometric shift and, for completeness and ease of reference, redshifts; see Section 5 below). Reassuringly, this final 16-source sample, culled on the basis of the search for near-infrared counterparts in the deep *HST* imaging, still contains all 12 radio-detected sources from the original $\simeq$50-source sample of 1.3-mm peaks (see Table 3).

Our final sample of 16 sources is thus very similar in size to what would be expected on the basis of comparing the numbers of positive and negative $>3.5\sigma$ peaks as explained above in Section 3. Nevertheless, one might be concerned that, by culling the ALMA source list on the basis of secure galaxy identifications, we are effectively excluding the possibility that the ALMA map

Table 3. The radio (6 GHz, JVLA) and X-ray (0.5–8 keV) detections of the 16 ALMA sources in the HUDF. Radio flux densities and associated uncertainties are from the new ultradeep JVLA image of the HUDF region obtained by Rujopakarn et al. (2016). We do not report radio-source positions here, simply because they are coincident with the ALMA positions within 50 milliarcsec. The X-ray flux densities, and derived luminosities are the total (i.e. soft+hard) values derived from the *Chandra* 4-ms imaging (Xue et al. 2011; Hsu et al. 2014). The X-ray positions for all five detected sources differ from the ALMA positions by <0.5 arcsec.

ID	S_{6GHz} (μJy)	$S_X/10^{-17}$ (erg cm^{-2} s^{-1})	$L_X/10^{42}$ (erg s^{-1})
UDF1	9.02 ± 0.57	150 ± 7	25.1 ± 1.2
UDF2	6.21 ± 0.57		
UDF3	12.65 ± 0.55	6 ± 3	0.8 ± 0.4
UDF4	3.11 ± 0.62		
UDF5	6.25 ± 0.46		
UDF6	8.22 ± 0.51		
UDF7	18.69 ± 0.60	8 ± 3	1.0 ± 0.4
UDF8	7.21 ± 0.47	330 ± 15	20.0 ± 0.9
UDF9	2.92 ± 0.58		
UDF10	<0.70		
UDF11	9.34 ± 0.74	11 ± 4	0.8 ± 0.4
UDF12	<0.70		
UDF13	4.67 ± 0.53		
UDF14	<0.68		
UDF15	<0.68		
UDF16	5.49 ± 0.46		

might reveal sources that are not visible in the deep *HST* imaging. In fact, we believe this is not a concern for three reasons. First, it must be remembered that, as a result of the UDF12 programme (Dunlop et al. 2013; Ellis et al. 2013; Koekemoer et al. 2013; McLure et al. 2013), the near-infrared imaging in this field is the deepest ever obtained, and we completed our search for galaxy counterparts in a stack of Y_{105}+J_{125}+J_{140}+H_{160} imaging reaching a detection threshold of >30 mag. Secondly, continuity arguments imply no significant number of near-infrared non-detections of the ALMA sources in our sample; as can be seen from Table 2, even though the sample has been culled of objects that lack galaxy counterparts at $H_{160} < 28.5$, in practice all the objects have $H_{160} \leq 27$, and indeed 15/16 have $H_{160} < 25$. Thus, in the context of the extremely deep WFC3/IR imaging available here, the galaxy counterparts of the secure ALMA sources are relatively bright, and it is extremely hard to argue that only slightly fainter ALMA sources should suddenly have galaxy counterparts that are two order-of-magnitude fainter in the near-infrared. Thirdly, we initially uncovered one reasonably significant source (S/N = 4.9, originally source number 9 in the master sample) for which we could not find any galaxy counterpart down to the limit of the WFC3/IR imaging. Notwithstanding the knowledge that there are two >4.5σ pseudo-sources in the negative image, we still explored this source in detail, in case it represented an extremely unusual (perhaps very distant) dusty object. As part of this exploration, we interrogated the data splits described earlier in Section 2.1, and found that this source featured at $\simeq$6σ in one half of the time-stream, but at less than 3σ in the other half. This is not the behaviour expected for a genuine 5σ source, and confirmed our suspicion that this was indeed our brightest false single-band detection.

We conclude, to the best of our ability (i.e. using all available supporting information, utilizing the negative ALMA 'sample' as a control, and examining carefully various 50:50 splits of the ALMA data) that the final sample presented in Table 2 and Fig. 1 represents all the robust ALMA sources detected in our map with peak S/N > 3.5, and point-source flux density $S_{1.3} > 120\,\mu$Jy.

While the sources listed in Table 2 were all selected on the basis of peak S/N > 3.5, and point-source flux density $S_{1.3} > 120\,\mu$Jy, subsequent fitting to the images showed that at least the first three sources are clearly resolved. For UDF1, UDF2 and UDF3, the ratio of total to point-source flux density was found to be 1.26, 1.56 and 1.50, respectively, and it is the total integrated flux density that is given in Table 2. Thereafter, however, we found that the fainter sources in the robust list were too faint for accurate total fluxes to be estimated by individual source fitting. While it may well be the case that the fainter sources are smaller, we decided it was unreasonable to assume they were simply point sources, an assumption that would clearly bias there estimated fluxes systematically low (albeit a subset will be flux-boosted).

We therefore decided to make a systematic correction to the point-source flux densities of the remaining 13 sources, to provide a best estimate of their true total 1.3-mm flux densities. We created a stack of the brightest five sources, and found that fitting to this yielded a total to point-source flux density ratio of $\simeq$1.3. We therefore decided to make a conservative, systematic correction to the point-source flux densities of sources UDF4 through UDF16, by simply multiplying their point-source flux densities (and associated errors) by a factor 1.25. It is these estimated total flux densities that are tabulated in column 4 of Table 2, but in column 5 we also give the original peak S/N ratio for each source (as derived at the detection stage).

Finally, we note that, for source UDF3, we have detected molecular emission lines from H_2O, CO and CI that, as well as confirming its spectroscopic redshift at mm wavelengths, also in this case make a significant contribution to the total flux density given in Table 2. Our best estimate is that removal of the line contribution reduces the flux density of UDF3 from $S_{1.3} = 863 \pm 84\,\mu$Jy to $S_{1.3} = 717 \pm 134\,\mu$Jy, and we use this latter value as appropriate for SFR estimates later in this paper. Sources UDF8 and UDF11 also appear to have emission lines within our sampled band-pass, but not at a level that seriously impacts on the estimated continuum flux (see Ivison et al., in preparation).

5 NUMBER COUNTS

5.1 Simulations, completeness and flux boosting

To quantify incompleteness and the impact of flux boosting, we performed a series of source injection and retrieval simulations. To make this as realistic as possible, we drew random samples from current best estimates of the source counts at the depths of interest. The results shown in Fig. 3(a) were based on 1000 realizations of an HUDF-size image, with the sources drawn randomly from the Schechter-function fits to the 1.1-mm source counts given by Hatsukade et al. (2016), after scaling the 1.1-mm flux densities to 1.3-mm values by dividing by 1.65 (Fujimoto et al. 2016). The scaled Hatsukade et al. (2016) differential number-count model fit is plotted as a magenta line in Fig. 3(a), with the resulting input to our simulations shown in navy blue (1000 HUDF samples are not sufficient to sample the 1.3-mm number counts brighter than $\simeq$1 mJy, but this is not important here).

We created a fake sky map by randomly placing single-pixel point sources into an equivalent pixel grid as the real map (with no clustering), convolved this with the ALMA PSF, and added this

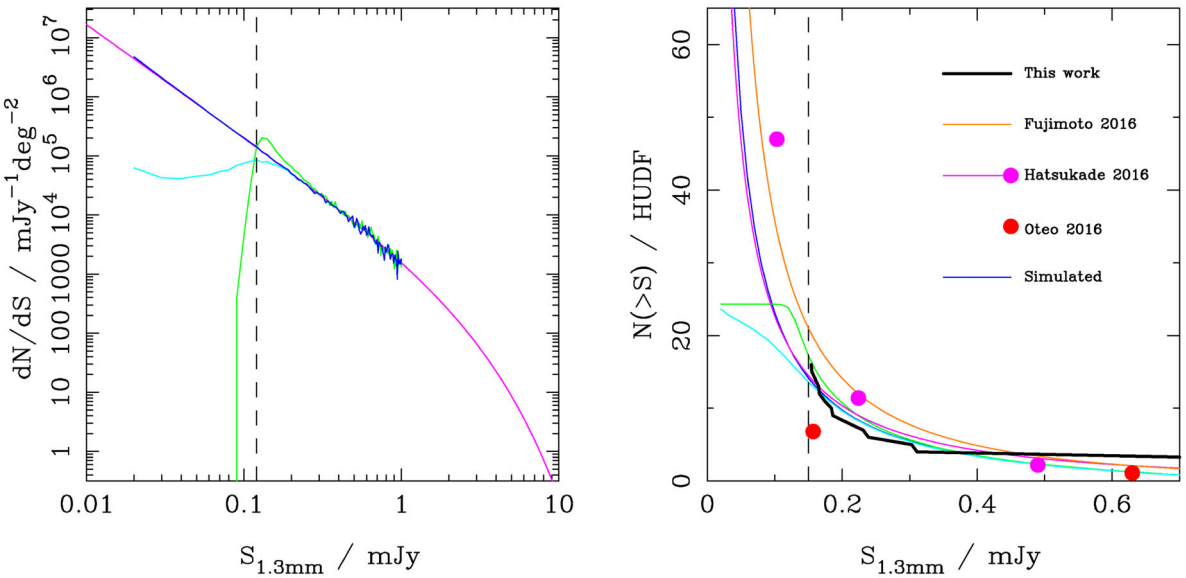

Figure 3. Left: Completeness and flux boosting in the HUDF ALMA 1.3-mm image. The magenta curve shows the differential number counts as given by the Schechter-function fit of Hatsukade et al. (2016), shifted from 1.1 to 1.3 mm by scaling flux density down by a factor of 1.65 (Fujimoto et al. 2016). The navy blue line shows the results of randomly drawing 1000 HUDF-size samples from this number-count model, and so represents the input for our source-injection simulations. The green curve shows the number counts as retrieved from the 1000 ALMA HUDF images using the same source extraction technique as applied to uncover the real sources, and insisting on S/N > 3.5. The cyan curve shows the same retrieved sources, but at their input rather than retrieved flux densities. The vertical dashed line indicates our chosen point-source detection threshold of $S_{1.3} = 0.120$ mJy, where it can be seen that flux boosting almost perfectly offsets the effect of incompleteness. Right: Cumulative number counts scaled to the size of the HUDF. The results from the present study are plotted as the thick black line (after scaling by a factor of 1.25 to account for average resolved flux density), and compared with the results of other recent studies (scaled to 1.3-mm) as indicated in the legend. We found $\simeq$1.5 times sources than anticipated given the number counts from Fujimoto et al. (2016) (orange curve), and also found fewer sources than indicated by the binned cumulative number counts reported by Hatsukade et al. (2016) (magenta points). However, our ALMA HUDF counts are in good agreement with integration of the Schechter fit to the differential counts provided by Hatsukade et al. (2016) (magenta line), and lie above the recent number-count results reported by Oteo et al. (2016) (red points), which were based on higher significance cuts, and arguably less biased pointings than the afore-mentioned studies. As in the left-hand panel, our simulated number counts are indicated by the navy blue line, and the impact of flux boosting+incompleteness is indicated by the green curve, with incompleteness indicated in cyan. Extrapolation of this model to still fainter flux densities suggests that to uncover $\simeq$100 1.3-mm sources in the HUDF would require reaching a detection limit of $S_{1.3} = 30$ μJy (i.e. $\simeq$4 times deeper than achieved here).

model to the real map. These simulated sources were then located and extracted from the image in exactly the same way as for the real sources (see Section 3). We are then able to quantify incompleteness by plotting the number of sources retrieved (at S/N > 3.5σ) as a function of their input flux density (cyan curve), and the combined impact of incompleteness and flux boosting by plotting the number of sources retrieved (at S/N > 3.5σ) as a function of output flux density (green curve).

From this plot, it can be seen that the differential number counts at our chosen point-source selection limit of S/N > 3.5σ, $S_{1.3}$ > 0.12 mJy should be basically correct, with the effect of incompleteness almost exactly compensated by the flux boosting of intrinsically fainter sources.

5.2 Observed and predicted number counts

Investigating the 1.3-mm number counts is not the main focus of this study, as the area of sky imaged here is small, and the flux-density range limited. Indeed, several recent studies have explored the 1.1–1.3-mm number counts at comparable depths by combining existing ALMA surveys and single pointings that together sample a significantly larger sky area, and include subregions of imaging reaching significantly deeper than achieved here (e.g. Ono et al. 2014; Carniani et al. 2015; Fujimoto et al. 2016; Oteo et al. 2016). Moreover, such studies, by including multiple sightlines, can potentially mitigate the impact of cosmic variance. None the less, given the homogeneity of our data set, and the unbiased nature of the HUDF (as compared to pointings centred on known extragalactic objects of specific interest for ALMA follow-up), it is of interest to check how the number of sources detected here compares with expectations based on recent number-count studies.

This comparison is presented in Fig. 3(b), which shows our observed cumulative counts, compared with appropriately scaled results from several of the aforementioned recent studies (integrated Schechter functions, and also binned counts). This shows that, down to a flux density $S_{1.3} = 0.15$ mJy (our effective flux-density limit after scaling the flux densities by a factor $\simeq$1.25), the 1.3-mm number counts in the HUDF are lower by a factor of $\simeq$1.5 than predicted by the Schechter-function source-count fit provided by Fujimoto et al. (2016), and also lower than the *binned* cumulative counts given by Hatsukade et al. (2016). However, our number counts are in very good agreement with integration of the Schechter-function fit to the differential counts provided by Hatsukade et al. (2016) (plotted in Fig. 3a as the magenta line), and somewhat higher than the binned cumulative counts published by Oteo et al. (2016) (who noted that their derived counts are a factor $\simeq$2 lower than those reported in most other recent studies).

In summary, Fig. 3(b) shows that significant work still needs to be done to clarify the faint-end of the mm number counts with ALMA, and that our results from the HUDF are consistent with current constraints.

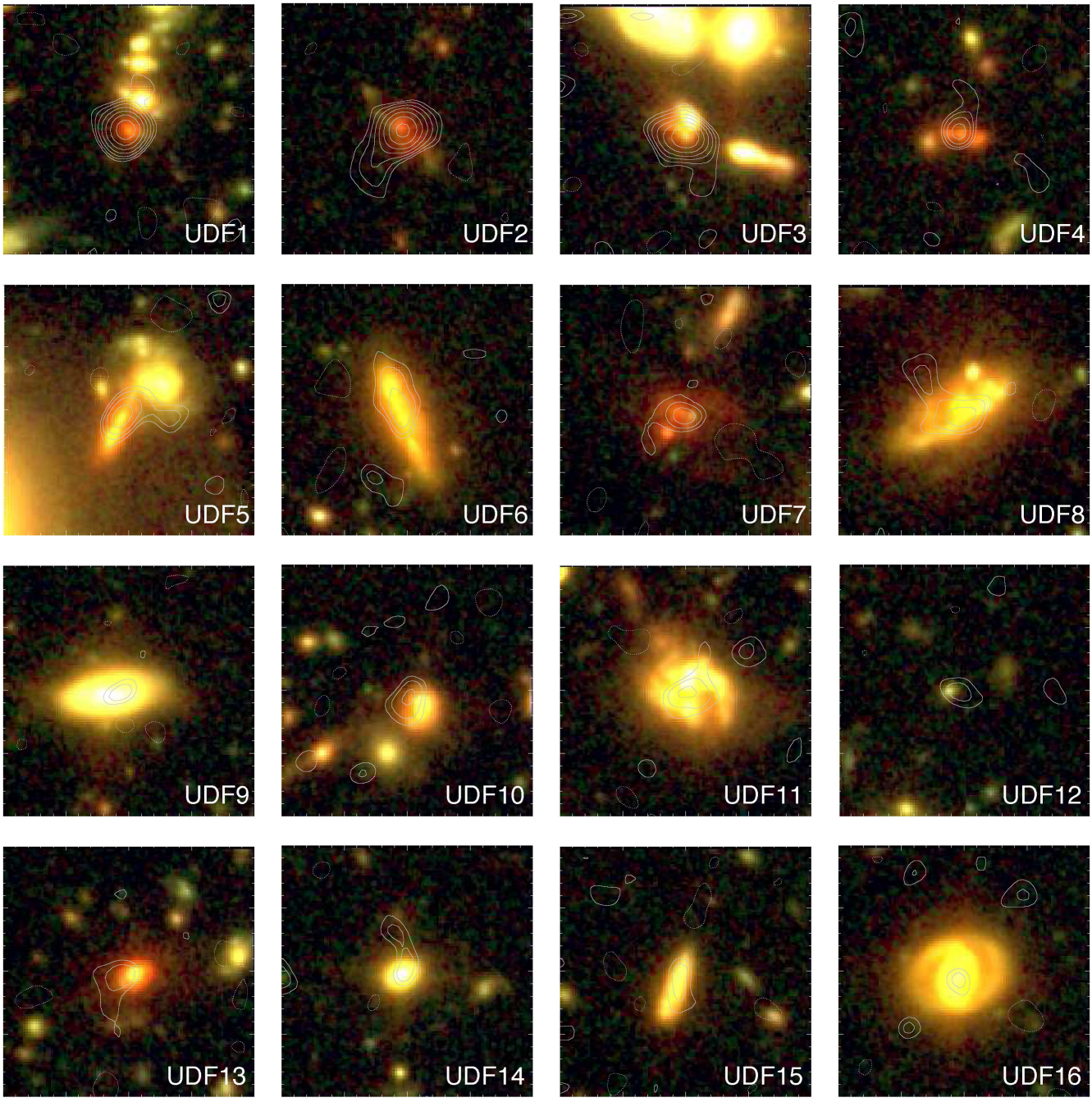

Figure 4. Colour (i_{775}+Y_{105}+H_{160}) *HST* postage-stamp images of the 16 ALMA-detected galaxies listed in Table 1, with the contours from the ALMA 1.3-mm image overlaid (in pale grey). Each stamp is 6 × 6 arcsec in size, with north to the top, and east to the left. ALMA contours are at -2σ, 2σ, $2^{1.5}\sigma$, $2^{2.0}\sigma$ and $2^{2.5}\sigma$.

Our ALMA image of the HUDF can also be used to explore the number counts significantly faintward of $S_{1.3} = 0.15$ mJy both via a $P(D)$ analysis (which will be presented elsewhere) and by stacking on the positions of the known galaxies in the HUDF at all redshifts. The results of such stacking experiments are presented in Section 7, along with the implications for the 1.3-mm background.

6 SOURCE PROPERTIES

The galaxies revealed by our ALMA imaging have several interesting properties. First, as is clear from the colour (i+Y+H) postage-stamp images presented in Fig. 4, the vast majority are noticeably red. Indeed, certainly the $z > 1.5$ ALMA sources can essentially be spotted by eye as the reddest objects in this particular colour representation of the HUDF. Secondly, because these objects are actually quite bright (15/16 have $H_{160} < 25$ mag), and because of the wealth of supporting spectroscopy and photometry in the HUDF, we have complete, high-quality redshift information for essentially the whole sample (see Table 2, and associated caption). Thirdly, armed with this redshift information and multiwavelength photometry (e.g. see Appendix A, Fig. A1) we can derive relatively robust stellar masses and SFR for these galaxies, as we now discuss (see Table 4).

Table 4. Stellar masses, estimated dust extinction, unobscured and obscured SFR, and specific SFR for the 16 ALMA-detected galaxies in the HUDF. Column 1 gives source numbers as also used in Figs 1 and 3, and Tables 2 and 3. Column 2 gives the stellar mass of each galaxy, determined as described in Section 6.2, while column 3 then gives the raw UV SFR (SFR_{UV}) based on the uncorrected UV absolute magnitude. Column 4 then gives the best estimate of extinction, A_v, as derived from the optical-infrared SED fitting (see Section 6.2). Then, in columns 5–8, we give alternative measures of SFR derived as described in Section 6.3, namely: (i) the dust-corrected SFR derived from the optical-infrared SED fitting undertaken to determine the stellar mass (SFR_{SED}); (ii) the SFR derived from fitting the long-wavelength *Spitzer*+*Herschel*+ALMA (24 μm to 1.3 mm) photometry (see Appendix A, Fig. A1) with the star formation template plotted in Fig. 8 (SFR_{FIR1}); (iii) the SFR derived from the same long-wavelength photometry, but adopting the best fitting of three alternative long-wavelength star-forming SEDs (SFR_{FIR2}); (iv) the SFR inferred from the radio detections (SFR_{Rad}). For the reasons described in Section 6.3, we adopt SFR_{FIR1} as the best/simplest estimate of true SFR, and use this to calculate the ratio of obscured:unobscured SFR given in column 9, and finally the estimates of specific SFR given in column 10. All values given here assume a Chabrier (2003) IMF. Finally we note that, for source UDF3, marked by a †, the derived values of SFR_{FIR1} and SFR_{FIR2} (and hence also the ratio of obscured:unobscured SFR and sSFR) are given after correcting for the unusually large contribution made to the 1.3-mm flux density in this object by molecular line emission (see Section 4.3 for details, and Ivison et al., in preparation).

ID	$\log_{10}(M_*/M_\odot)$	SFR_{UV} ($M_\odot$ yr^{-1})	A_V (mag)	SFR_{SED} ($M_\odot$ yr^{-1})	SFR_{FIR1} ($M_\odot$ yr^{-1})	SFR_{FIR2} ($M_\odot$ yr^{-1})	SFR_{Rad} ($M_\odot$ yr^{-1})	SFR_{obs}/ SFR_{UV}	sSFR (Gyr^{-1})
UDF1	10.7 ± 0.10	0.31 ± 0.05	3.1	399.4	326 ± 83	364 ± 82	439 ± 28	1052 ± 317	6.50 ± 2.24
UDF2	11.1 ± 0.15	0.32 ± 0.10	2.2	50.2	247 ± 76	194 ± 64	242 ± 22	772 ± 339	1.96 ± 0.92
UDF3†	10.3 ± 0.15	4.70 ± 0.30	0.9	42.0	195 ± 69	173 ± 1	400 ± 17	41 ± 15	9.77 ± 4.88
UDF4	10.5 ± 0.15	0.43 ± 0.20	1.6	20.0	94 ± 4	58 ± 5	89 ± 17	219 ± 102	2.97 ± 1.05
UDF5	10.4 ± 0.15	0.20 ± 0.05	2.4	36.1	102 ± 7	67 ± 25	86 ± 6	510 ± 132	4.06 ± 1.46
UDF6	10.5 ± 0.10	0.10 ± 0.02	2.8	78.0	87 ± 11	66 ± 5	68 ± 5	870 ± 205	2.75 ± 0.73
UDF7	10.6 ± 0.10	0.50 ± 0.03	1.5	16.5	56 ± 22	77 ± 42	617 ± 20	112 ± 45	1.41 ± 0.64
UDF8	11.2 ± 0.15	0.98 ± 0.02	1.6	35.8	149 ± 90	94 ± 37	73 ± 5	152 ± 92	0.94 ± 0.66
UDF9	10.0 ± 0.10	0.06 ± 0.01	0.9	0.5	23 ± 25	5 ± 2	5 ± 1	383 ± 421	2.30 ± 2.56
UDF10	10.2 ± 0.15	1.14 ± 0.10	1.5	37.0	45 ± 22	34 ± 7	<35	39 ± 20	2.84 ± 1.71
UDF11	10.8 ± 0.10	6.29 ± 0.20	1.4	162.8	162 ± 94	232 ± 10	172 ± 14	26 ± 15	2.57 ± 1.60
UDF12	9.6 ± 0.15	1.55 ± 0.10	0.2	2.6	37 ± 14	21 ± 7	<100	24 ± 10	9.29 ± 4.80
UDF13	10.8 ± 0.10	0.95 ± 0.05	1.3	18.0	68 ± 18	60 ± 19	142 ± 17	72 ± 19	1.08 ± 0.38
UDF14	9.7 ± 0.10	0.05 ± 0.01	1.3	1.0	44 ± 17	3 ± 2	<4	880 ± 383	8.78 ± 3.96
UDF15	9.9 ± 0.15	1.14 ± 0.02	1.1	15.5	38 ± 27	25 ± 8	<20	33 ± 24	4.78 ± 3.79
UDF16	10.9 ± 0.10	0.10 ± 0.05	0.6	0.5	40 ± 18	25 ± 4	38 ± 3	400 ± 269	0.50 ± 0.26

6.1 Redshift distribution

As explained in Section 2.2.5, and tabulated in Table 2, the wealth of deep spectroscopy in the HUDF field results in 13 out of our 16 ALMA-detected galaxies having optical–near-infrared spectroscopic redshifts. To complete the redshift content of the sample, we estimated the redshifts of the remaining sources by adopting the median value from five alternative determinations of z_{phot} based on different SED fitting codes. The range of values returned by these codes, coupled with tests of z_{phot} versus z_{spec} (for the 13 sources in Table 2 with spectroscopic redshifts), indicates that the three photometric redshifts listed in Table 2 carry a typical rms uncertainty of $\delta z \simeq 0.1$.

The final redshift distribution of the ALMA-detected HUDF galaxy sample is shown in Fig. 5. Although the present study probes an order-of-magnitude deeper (in terms of dust-enshrouded SFR) than the brighter sub-mm/mm selected samples produced by SCUBA, MAMBO, LABOCA, AzTEC and SCUBA-2 over the last 15–20 yr, the redshift distribution of detected sources is very little changed. The mean redshift of the ALMA HUDF sources is $\langle z \rangle = 2.15$, with 13/16 sources ($\simeq$80 per cent) in the redshift range $1.5 < z < 3$. This is very similar to the redshift distribution displayed by, for example, the much brighter AzTEC sources uncovered in the SHADES fields, which have a median redshift of $z \simeq 2.2$, with $\simeq$75 per cent of sources lying in the redshift range $1.5 < z < 3.5$ (Michałowski et al. 2012a; see also, e.g. Chapman et al. 2005; Aretxaga et al. 2007; Chapin et al. 2009), although there is some evidence that the most extreme sources are confined to somewhat higher redshifts (e.g. Ivison et al. 2007; Smolcic et al. 2012; Koprowski et al. 2014; Michałowski et al. 2016).

In Fig. 6, we plot the galaxies in the HUDF on the M_{1500}-redshift plane, highlighting the locations of the ALMA-detected galaxies.

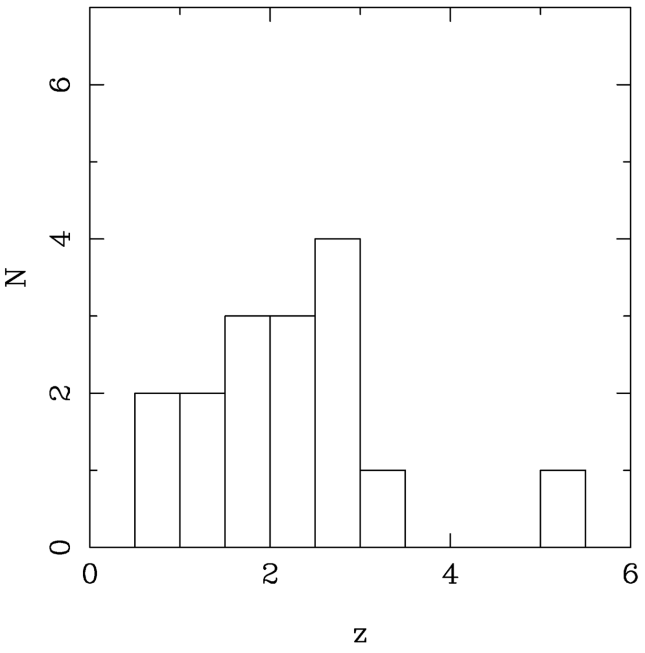

Figure 5. The redshift distribution of the 16 ALMA-detected galaxies in the HUDF. The redshift information consists of 13 spectroscopic redshifts, and three accurate ($\delta z \simeq 0.1$) photometric redshifts, as listed in Table 2.

This shows that the ALMA-detected galaxies span a wide range of (observed) UV luminosities.

However, a different picture emerges when stellar mass is plotted versus redshift, as shown in Fig. 7. Here, it can be seen that essentially all of the ALMA-detected galaxies have high stellar masses.

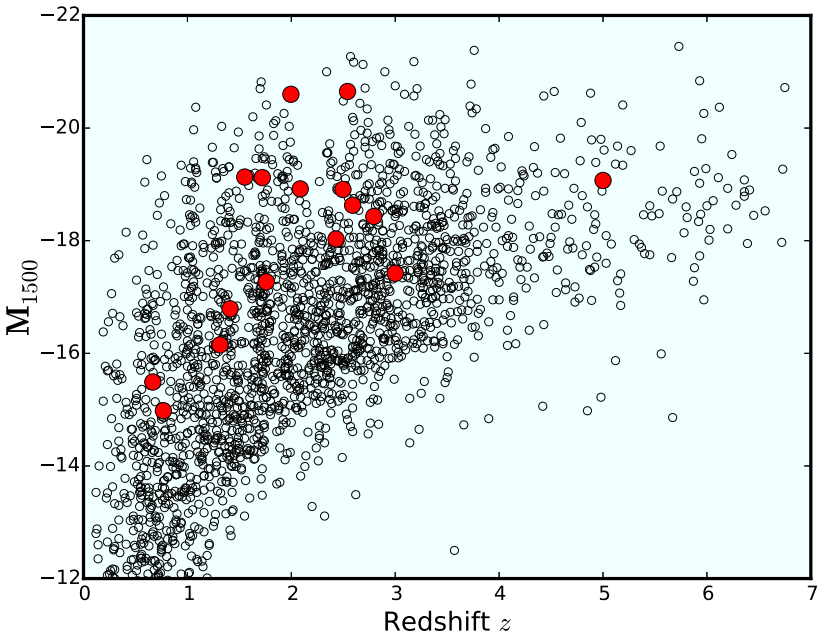

Figure 6. UV absolute magnitude versus redshift for the galaxies in the HUDF, highlighting (in red) those detected in our ALMA 1.3-mm image. It can be seen that the ALMA-detected galaxies span a wide range of raw unobscured UV luminosity, and appear unexceptional on the M_{1500}–z plane. However, if corrected for dust obscuration, they would be the brightest galaxies in the field (see Section 6).

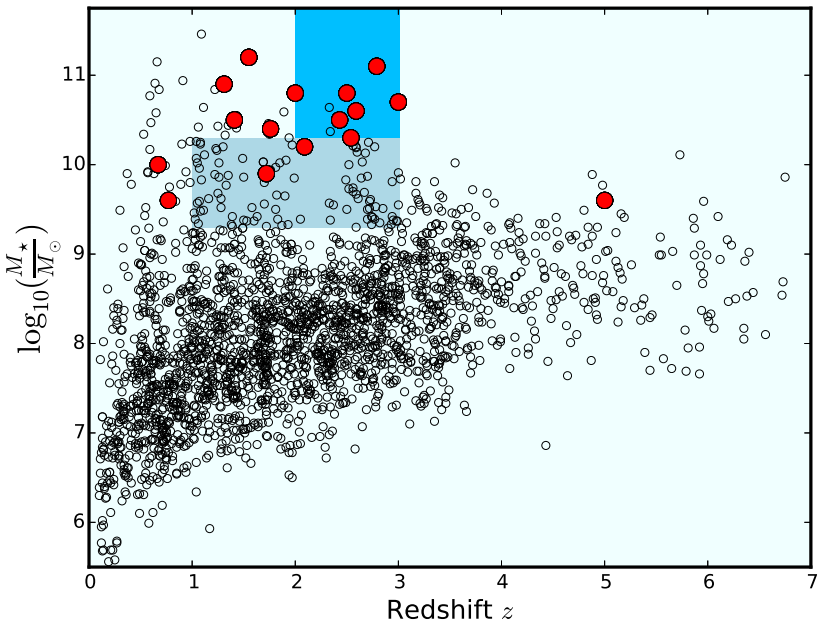

Figure 7. Stellar mass versus redshift for the galaxies in the HUDF, highlighting (in red) those detected in our ALMA 1.3-mm image. The connection between dust-enshrouded SFR and M_* is clear, and indeed, as emphasized by the bright-blue box at the top of the plot, we detect seven of the nine galaxies (i.e. $\simeq$80 per cent) with $M_* \geq 2 \times 10^{10}\,\mathrm{M}_\odot$ at $z > 2$. Also apparent is the emergence of a significant population of quenched high-mass galaxies at $1 < z < 2$, where the ALMA detection rate for galaxies with $M_* \geq 2 \times 10^{10}\,\mathrm{M}_\odot$ drops to 5/19 (i.e. $\simeq$25 per cent). It can be also seen that the lack of ALMA detections beyond $z > 3$ in our sample simply reflects the dearth of comparably massive galaxies in the HUDF at these redshifts (due to the evolution of the galaxy stellar mass function). The blue-grey rectangle indicates the sample of $\simeq$90 galaxies in the redshift range $1 < z < 3$ and mass range $9.3 < \log_{10}(M_*/\mathrm{M}_\odot) < 10.3$ used to produce the stacked ALMA image shown in Fig. 8, and discussed in Section 7.

Indeed, from Fig. 7 it can then be seen that the most obvious physical reason for the lack of very high-redshift galaxies in our sample (i.e. only one detection beyond $z = 3.1$) is the absence of high-mass galaxies at these redshifts within the relatively small cosmological volumes sampled by the 4.5 arcmin^2 field. This is discussed further in the next subsection. By contrast, the decline in ALMA detections at $z < 1.5$ is driven by the quenching of star formation activity in high-mass galaxies, as is again evident from Fig. 7.

6.2 Stellar masses

To derive the stellar masses of the galaxies, we fitted a range of single-component, and then also two-component Bruzual & Charlot (2003) evolutionary synthesis models to the optical-infrared photometry of the sources (at the redshifts listed in Table 2). For the single-component models, the minimum age was set to 50 Myr, with τ allowed to vary from 300 Myr up to an essentially constant SFR (see Wuyts et al. 2009). We applied reddening assuming the dust attenuation law of Calzetti et al. (2000), with extinction allowed to vary up to $A_v \simeq 6$ (see Dunlop, Cirasuolo & McLure 2007), and the impact of IGM absorption was modelled following Madau (1995).

The derived galaxy masses, given in Table 4, proved to be (perhaps surprisingly) robust to the range of star formation histories that provided acceptable fits to the photometry (possibly as a result of the high signal:noise of the available optical-infrared data in the HUDF). The uncertainties in stellar mass given in Table 4 reflect the range in stellar masses delivered by acceptable star formation histories.

We assumed the initial mass function (IMF) of Chabrier (2003), and that the stellar masses given in Table 2 are M_* 'total', which means the mass of living stars plus stellar remnants. One advantage of this choice is that the conversion from M_* assuming Chabrier (2003) to M_* with a Salpeter (1955) IMF is relatively immune to age, involving multiplication by a factor $\simeq$1.65. This M_*-total is typically $\simeq$0.05 dex larger than M_*-living, and $\simeq$0.15 dex smaller than $M_{\rm gal}$, which includes recycled gas, although these conversions are functions of galaxy age and star formation history.

It can be seen that the stellar masses are high, with 13 out of the 16 sources having $M_* > 10^{10}\,{\rm M}_\odot$ (with the adopted Chabrier IMF). Such objects are rare at these epochs in the relatively small volumes probed by the HUDF. In particular, from Fig. 7 (which shows mass versus redshift for all galaxies in the HUDF), it can be seen that the HUDF contains only nine galaxies with $M_* > 2 \times 10^{10}\,{\rm M}_\odot$ at $z \geq 2$, and that our ALMA observations have detected seven of them (i.e. $\simeq$80 per cent). This provides compelling evidence that stellar mass is the best predictor of SFR at these epochs (rather than, for example rest-frame UV luminosity), as expected from the 'main sequence' of star-forming galaxies (e.g. Daddi et al. 2007; Noeske et al. 2007; Elbaz et al. 2010; Michałowski et al. 2012b; Roseboom et al. 2013; Koprowski et al. 2014, 2016; Speagle et al. 2014; Renzini & Peng 2015; Schreiber et al. 2015). The location of our ALMA-detected galaxies relative to the 'main-sequence' is discussed further below in Section 8.1.

As is also apparent from Fig. 7, at $1 < z < 2$ the fraction of high-mass ($M_* > 2 \times 10^{10}\,{\rm M}_\odot$) galaxies that we detect with ALMA drops to 5/19 (i.e. $\simeq$25 per cent) reflecting the emergence of a significant population of quenched high-mass galaxies over this redshift range.

At redshifts $z > 3.5$ we have detected only one object, UDF12, which lies at $z = 5.0$. Among our ALMA detections, this is in fact the galaxy with the lowest estimated stellar mass, but from Fig. 7 it can be seen that it is one of the most massive galaxies in the HUDF for its redshift (i.e. it is one of the very few galaxies in this field at $z > 4$ that has $M_* > 10^{9.5}\,{\rm M}_\odot$). We can therefore speculate that this detection may reflect a modest increase in typical specific SFR between $z \simeq 2$ and $z \simeq 5$ (e.g. Steinhardt et al. 2014; Marmel-Queralto et al. 2016), combined with the sensitivity of 1.3-mm observations to extreme redshift dusty star-forming galaxies, and that moderately deeper sub-mm/mm observations of the field may yield significantly more detections at $z > 3$.

6.3 Star formation rates

The completeness and quality of the redshift information, and the available high-quality multiwavelength photometry allows us to derive various alternative estimates of SFR in the 16 ALMA-detected galaxies, which we present in Table 4.

First, we convert the rest-frame UV ($\lambda = 1500$ Å) absolute magnitude, M_{1500} of each source into an estimate of the unobscured SFR. The adopted calibration means that an absolute magnitude of $M_{1500} = -18.75$ corresponds to a SFR of 1 ${\rm M}_\odot\,{\rm yr}^{-1}$, consistent with the conversion given by Kennicutt & Evans (2012). The resulting values are given in column 3 of Table 4, and are unspectacular (with SFR $< 1\,{\rm M}_\odot\,{\rm yr}^{-1}$ for the majority of the sources).

Secondly, we use the SED fitting, as described above (to derive the stellar masses), to estimate the extinction-corrected SFR from the UV–near-infrared SED. These estimates are given in column 5 of Table 4, with the corresponding best-fitting values of A_v given in column 4 (extinction was allowed to range up to $A_v = 6$ mag in the SED fitting, but the best-fitting values lie in the range 0.2–3.1 mag).

Thirdly, we estimate dust-enshrouded SFR from the long-wavelength photometry, utilizing the ALMA data, the deconfused *Herschel* photometry, and testing the impact of including or excluding *Spitzer* 24-μm photometry. Given the redshifts of the ALMA-detected galaxies, the key ALMA 1.3-mm data point samples the rest-frame SEDs of the sources significantly longward of the peak of any feasible far-infrared SED. This is good for the estimation of robust dust masses, but means that the inferred SFR obviously depends on the form of the adopted far-infrared SED template.

We therefore investigated the long-wavelength SEDs of the individually detected sources, and also amalgamated the 16-source photometry into a combined SED (after de-redshifting and normalizing at $\lambda_{\rm rest} = 1.3$ mm) that we show in Fig. 8. This combined SED (which, due to the accuracy of the available redshift information, displays the rest-frame 8-μm feature) can be used to establish the best template SED to use for the estimating of SFR for each source. The best-fitting model SED shown in Fig. 8 is the active galactic nucleus (AGN)+star-forming composite model of Kirkpatrick et al. (2015), and yields a best fit with a 20 per cent AGN contribution to bolometric 8–1000 μm infrared luminosity. We have therefore used this template to estimate the far-infrared luminosity of each of the ALMA sources, exploring the impact of both including and excluding the 24-μm photometry to yield realistic uncertainties, and then converting to SFR using the conversion of Murphy et al. (2011) with a final minor scaling applied to convert from a Kroupa to a Chabrier IMF. The resulting values are tabulated in column 6 of Table 4, and are adopted hereafter as our best estimates of dust-enshrouded SFR. We note that, although this calculation has been performed using the composite template shown in Fig. 8, and then reducing the far-infrared luminosity by 20 per cent to compensate for the typical AGN contribution, in practice near identical values are obtained by simply fitting the star-forming component to the ALMA data point. The star-forming component is essentially the sub-mm galaxy template of Pope et al. (2008),[4] and at the redshifts of interest here ($z \simeq 2$) produces a conversion between observed 1.3-mm flux density and SFR that can be approximated by SFR (in ${\rm M}_\odot\,{\rm yr}^{-1}$) $\simeq 0.30 \times S_{1.3}$ (in μJy). With this template, and adopting a Chabrier (2003) IMF, the flux-density limit of the current survey thus corresponds to a limiting SFR sensitivity of $\simeq 40\,{\rm M}_\odot\,{\rm yr}^{-1}$.

[4] http://www.alexandrapope.com/#downloads/t0u6h

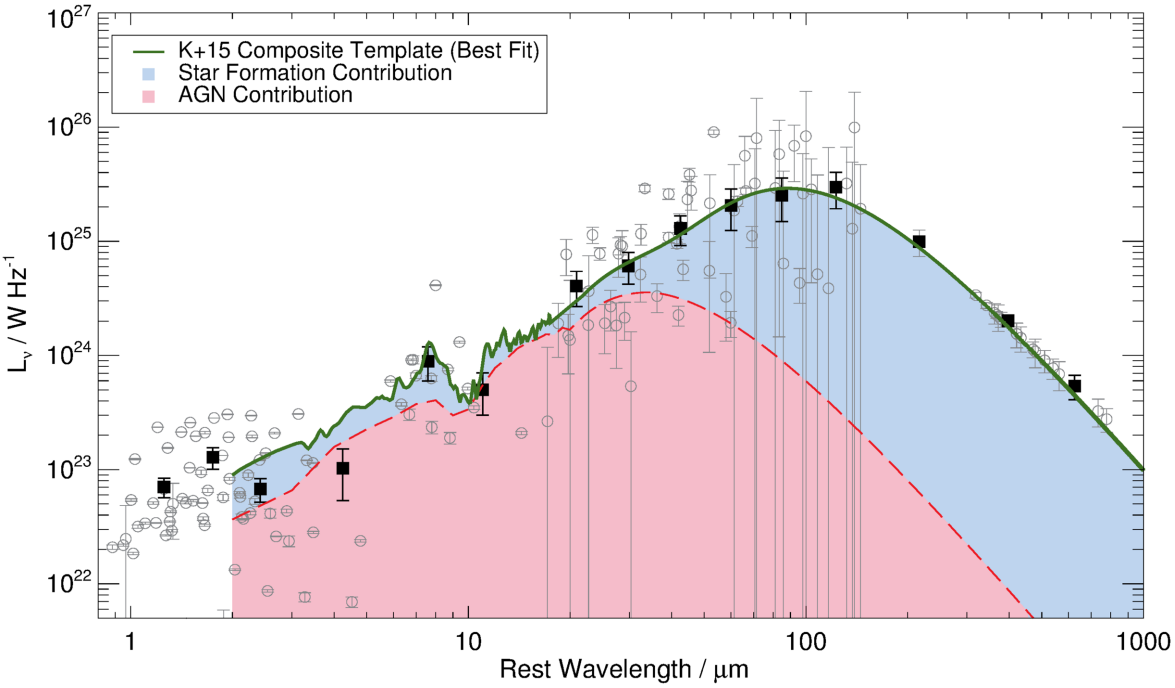

Figure 8. The combined *Spitzer*+*Herschel*+ALMA photometry of the 16 ALMA sources (after de-redshifting and scaling to the same rest-frame 1.3-mm luminosity), fitted by the composite star-forming+AGN template of Kirkpatrick et al. (2015). The solid black squares indicate the weighted mean of the scaled multisource photometry within a given wavelength bin. The accuracy of the redshift information results in the 8-μm feature being clearly visible in the observed combined rest-frame SED. The best-fitting composite template has an AGN component which contributes only ≃20 per cent to the bolometric rest-frame far-infrared (8–1000 μm) luminosity, and is completely negligible at the wavelengths sampled by the ALMA imaging.

For completeness, we also explored the impact of attempting to determine the best-fitting far-infrared SED template for each individual source, rather than adopting a single template for all sources. This approach makes more use of the deconfused *Herschel* photometry and limits, but the decision between alternative SEDs is inevitably rather uncertain on a source-by-source basis. We fitted each source with an M82, Arp220 or sub-mm galaxy template (Silva et al. 1998; Michalowski, Hjorth & Watson 2010), and again explored the impact of including and excluding the 24 μm data. The results are given in column 7 of Table 4; we give the average of the derived SFR with and without including the 24 μm data, with the adopted error being the larger of the statistical error in the fitting or the range of results dictated by the impact of the 24 μm data. The most important information to be gleaned from these results is that the derived SFR is generally reassuringly similar to, or slightly lower than the values estimated from the single template (shown in Fig. 8) as described above. The somewhat lower values for some sources simply reflect the fact that, when a source is better fit with the Michałowski et al. (2010) sub-mm galaxy SED than either M82 or Arp220, the bolometric luminosity of the Pope et al. (2008) sub-mm galaxy template is ≃1.4 larger than that of the Michałowski et al. (2010) sub-mm template when normalized at rest-frame wavelengths $\lambda_{\rm rest}$ ≃500 μm (when normalized at $\lambda_{\rm rest}$ ≃150 μm, the ratio is only ≃1.1).

Finally, we provide an estimate of SFR based on the new JVLA 6 GHz photometry of the ALMA sources (Table 3). Given the potential for AGN contamination at radio wavelengths, uncertainty in the precise radio-SFR calibration, and the need to adopt a consistent estimator for SFR for all 16 ALMA sources, we do not make further use of the radio-based estimates of SFR in this paper. Nevertheless, the radio-inferred estimates of SFR given in column 8 of Table 4 provide reassurance that our far-infrared derived values are not seriously overestimated.

In this context, we note from Tables 3 and 4 that there appear to be two radio-loud AGNs in our sample, UDF3 and UDF7. There are also two obvious X-ray AGN, UDF1 and UDF8. As listed in Table 3, the catalogue produced by Xue et al. (2011) also yields detections of three more galaxies in our sample, but these detections are >20 times fainter than the obvious X-ray AGN, and, depending on the adopted extinction correction, may in fact be explained by the star formation activity in these galaxies (although two of these weaker X-ray detections do correspond to the two radio-loud AGN).

7 STACKING AND THE MM BACKGROUND

A key advantage of the HUDF, not shared by many other deep ALMA pointings, is the quality, depth and completeness of the galaxy catalogue in the field. This enables stacking of chosen galaxy subpopulations within the ALMA image, allowing us to explore source properties to significantly fainter flux densities than achieved in the selection of robustly detected individual sources. Thus, not only can we attempt to derive an estimate of the total 1.3-mm flux density present in the field, but to the extent allowed by population statistics, we can explore how this (and hence dust-obscured star formation activity) is distributed as a function of redshift, galaxy stellar mass and/or UV luminosity.

We therefore explored stacking of galaxy subpopulations selected from a range of bins defined by redshift z, stellar mass M_* and UV absolute magnitude $M_{\rm UV}$. Perhaps unsurprisingly, given the clear link between stellar mass and ALMA flux density revealed

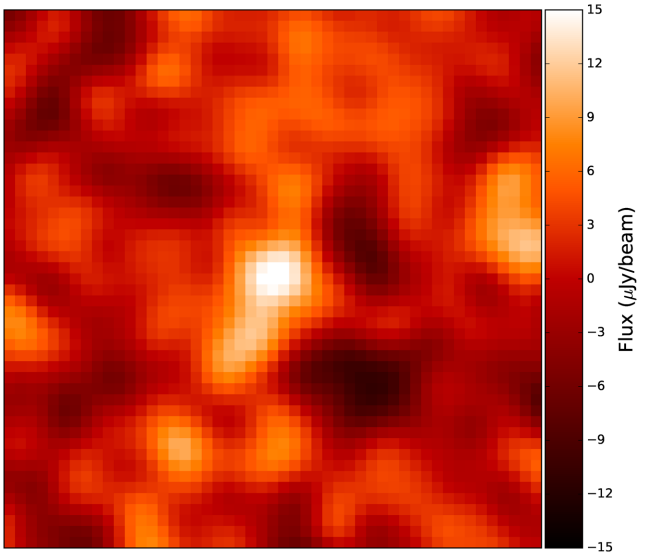

Figure 9. The result of stacking the ALMA imaging on the positions of galaxies in the redshift range $1 < z < 3$ and the stellar mass range $9.3 < \log_{10}(M_*/\mathrm{M}_\odot) < 10.3$ (excluding the sources already detected and listed in Table 4). The image shown is 5.9 × 5.9 arcsec. The stacked 'source' includes 89 galaxies, and has a mean flux density (point-source +25 per cent) of $S_{1.3} = 20.1 \pm 4.6$ μJy, corresponding to a mean SFR of $6.0 \pm 1.4\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$. This means that galaxies in this mass and redshift range contribute a total dust-enshrouded SFR $\simeq 530 \pm 130\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$. The same sources have an average UV luminosity corresponding to an absolute magnitude of $M_{\mathrm{UV}} = -19.38$, and hence contribute a total unobscured (raw UV) SFR $\simeq 160\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$. These results imply an average ratio of obscured:unobscured SFR of ≃3.3 (or total-SFR/unobscured-SFR ≃4.3), and a sSFR $\simeq 1.95\,\mathrm{Gyr}^{-1}$ at a median mass of $\log_{10}(M_*/\mathrm{M}_\odot) = 9.6$ and median redshift $z = 1.92$.

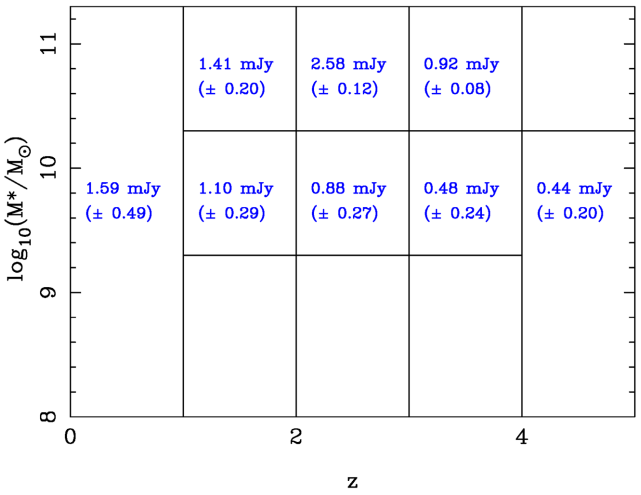

Figure 10. The distribution of 1.3-mm flux density found within the HUDF across the galaxy redshift–star-mass plane. The figures given here combine the flux densities of the detected sources and the results of stacking in the appropriate redshift–mass bins (excluding detected sources to avoid double counting). Results are only given for bins within which stacking yielded better than a 2σ detection, typically corresponding to an average flux density $S_{1.3} \geq 0.010$ mJy. At $z > 4$ the HUDF contains no galaxies more massive than $\log_{10}(M_*/\mathrm{M}_\odot) = 10.3$, and so the right-hand cell in the top row is genuinely blank. The cells below $\log_{10}(M_*/\mathrm{M}_\odot) = 9.3$ at $1 < z < 3$ may contain additional 1.3-mm flux density, but not at a level that could be securely detected via stacking in the current image. At $z < 1$ and $z > 4$, low sample size precludes bin subdivision by mass, and so we simply give the total flux detected via stacking of all sources in the $z < 1$ and $z > 4$ redshift bins. Addition of all the numbers in this figure, combined with an estimated contribution of ≃1 mJy per HUDF from sources brighter than $S_{1.3} \simeq 1$ mJy (equivalent to $0.8\,\mathrm{Jy\,deg}^{-2}$) gives an estimate of 10.6 ± 1.0 mJy for the 1.3-mm flux density in the HUDF down to $S_{1.3} \simeq 0.01$ mJy (equivalent to $8.5 \pm 0.9\,\mathrm{Jy\,deg}^{-2}$; see Fig. 10). Of this total estimated background, ≃70 per cent is provided by sources with $\log_{10}(M_*/\mathrm{M}_\odot) > 9.3$ in the redshift range $1 < z < 3$. Within our background estimate, ≃45 per cent of the flux density is contributed by the 16 individually detected sources in our ALMA image of the HUDF, with only ≃35 per cent added by our attempts to extend flux retrieval down to ≃0.01 mJy sources through stacking.

by the location of the ALMA-detected sources in Fig. 7, the most significant stack detections were achieved in the next mass bin, which we defined as $9.3 < \log_{10}(M_*/\mathrm{M}_\odot) < 10.3$. The 1.3-mm image stack of all galaxies in this mass range, and in the redshift range $1 < z < 3$, is shown in Fig. 9, and (excluding the two detected galaxies in this bin) produces a 4.4σ detection, with a mean flux density (point source +25 per cent) $S_{1.3} = 20.1 \pm 4.6$ μJy (corresponding to a mean SFR of $6.0 \pm 1.4\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$; reassuringly a stack of the galaxies in this same bin into the JVLA 6-GHz image yielded a detection with mean flux density $S_{6\,\mathrm{GHz}} = 262 \pm 24$ nJy, corresponding to a mean SFR of $6.4 \pm 1.0\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$).

With 89 galaxies contributing to the stacked detection shown in Fig. 9, the resulting inferred total 1.3-mm flux density in this bin is 1788 ± 410 μJy. We subdivided this bin by redshift (at $z = 2$), and proceeded in an analogous way to seek significant ($>2\sigma$) stacked 1.3-mm detections in other regions of the mass–redshift plane. In practice, given the available number of galaxies, and the depth of the ALMA imaging, such detections correspond to bins that yield an average flux density $S_{1.3} > 10$ μJy. In Fig. 10, we show the final result of this process, where, within each bin, we have also added back the contribution of the individually detected sources. The coverage of the redshift–mass plane is limited by the fact that (i) there are no galaxies in the HUDF with $\log_{10}(M_*/\mathrm{M}_\odot) > 10.3$ to stack at $z > 4$ (see Fig. 7); (ii) no significant stacked detections were achieved in mass bins confined to $\log_{10}(M_*/\mathrm{M}_\odot) < 9.3$ and (iii) the limited number of galaxies in the field at $z < 1$ and $z > 4$ prohibited significant stacked detections subdivided by mass at these redshifts. Nevertheless, given the evidence for the steep dependence of dust-obscured star formation on stellar mass explored further below, the sum of the figures given in Fig. 10 can be expected to yield a reasonably complete estimate of the 1.3-mm background, and we believe this represents the best estimate to date of the distribution of this background as a function of redshift and galaxy stellar mass.

We note that, at $z > 1$, the numbers given in Fig. 10 can be reasonably converted into estimates of total dust-enshrouded SFR per bin, by multiplying by ≃300 (see Section 6.3). However, at $z < 1$, this will yield a serious overestimate of SFR, because of the lack of a negative *K*-correction in this regime (i.e. a significant amount of the flux density in the $z < 1$ bin is contributed by relatively low-redshift, but intrinsically not very luminous sources). For this reason, and because at $z \simeq 1$, observations at 1.3-mm sample the far-infrared SEDs of the sources rather far from the peak of emission, we do not use our stacked results at $z < 1$ in the discussion of SFR in the remainder of this paper. However, the contribution of sources at $z < 1$ is still of interest when considering the implications for faint source counts, and the 1.3-mm background.

In Fig. 11, we explore our resulting estimate of integrated flux density at $\lambda = 1.3$ mm, as a function of source flux density, utilizing our detections down to $S_{1.3} \simeq 150$ μJy, and the contribution produced by the sum of stacked fluxes (excluding the individual

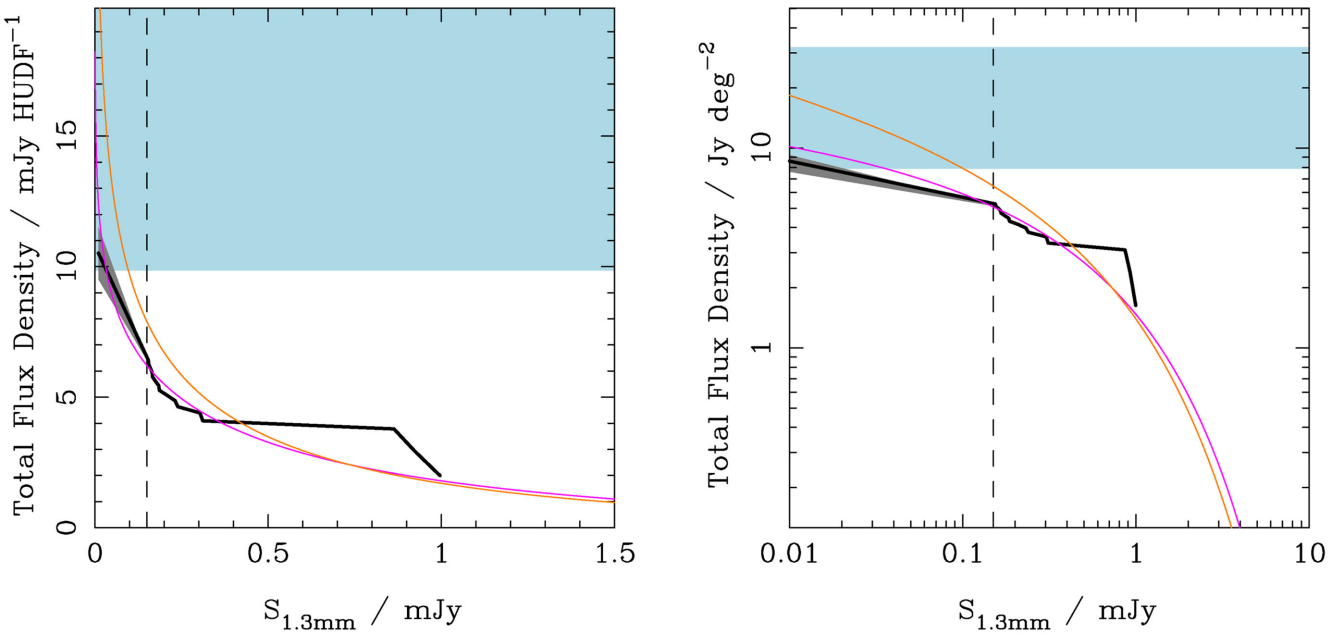

Figure 11. Total integrated flux density at $\lambda = 1.3$ mm. The left-hand plot shows results per HUDF (i.e. per 4.5 arcmin^2) in linear units, while the right-hand plot shows the more standard logarithmic representation in units of Jy deg^{-2}. The solid black line in each plot shows simply the running sum of the flux densities of the 16 ALMA HUDF sources, starting at the total contributed by sources brighter that $S_{1.3} = 1$ mJy (i.e. $\simeq$0.8 Jy deg^{-2}; Scott et al. 2012) and summing down to our (effective) detection threshold of $S_{1.3} = 0.15$ mJy (marked by the vertical dashed line in both plots). Below that flux density, we extrapolate the black line (with grey shading to indicate the uncertainty) to account for the contribution estimated from our stacking analysis, which samples the fainter population down to $S_{1.3} \simeq 0.01$ mJy (and adds an additional 4.1 ± 1.0 mJy / HUDF). The blue-grey shaded region in both panels indicates the (highly uncertain) estimate of the 1.3-mm background as measured by COBE (i.e. 17^{+16}_{-9} Jy deg^{-2}; Puget et al. 1996; Fixsen et al. 1998). It can be seen that our HUDF-derived flux-density total is (just) consistent with the 1σ lower bound on the background estimate. The magenta and orange curves give the predicted background as a function of 1.3-mm flux density for the scaled Schechter number-count models of Hatsukade et al. (2016) and Fujimoto et al. (2016) already utilized in Fig. 3, and discussed in Section 5.2. Our results can be plausibly reconciled with those of Hatsukade et al. (2016), but to achieve the much higher estimated background reported by Fujimoto et al. (2016) down to $S_{1.3} = 0.01$ Jy we would require an approximate doubling of the flux density we have actually been able to uncover in the HUDF through source detection and stacking.

source contributions from Fig. 10) reaching down to $S_{1.3} \simeq 10\,\mu$Jy. The left-hand panel of Fig. 11 shows results per HUDF (i.e. per 4.5 arcmin^2) in linear units, while the right-hand panel shows the more standard logarithmic representation in units of Jy deg^{-2}. The solid black line in each plot shows simply the running sum of the flux densities of the 16 ALMA HUDF sources, starting at the total contributed by sources brighter that $S_{1.3} = 1$ mJy (i.e. $\simeq$0.8 Jy deg^{-2}; Scott et al. 2012) down to our (effective) detection threshold of $S_{1.3} = 0.15$ mJy. Below this flux density, we extrapolate the line down to $S_{1.3} \simeq 10\,\mu$Jy to account for the contribution estimated from our stacking analysis (which adds an additional 4.1 ± 1.0 mJy / HUDF). Our HUDF-derived estimate of total flux density can be seen to be (just) consistent with the 1σ lower bound on the estimated background as measured by COBE (i.e. 17^{+16}_{-9} Jy deg^{-2}; Puget et al. 1996; Fixsen et al. 1998). Our results can be plausibly reconciled with those of Hatsukade et al. (2016), but to achieve the much higher estimated backgrounds reported by Fujimoto et al. (2016) or Carniani et al. (2015) down to $S_{1.3} = 0.01$ mJy, we would require to approximately double the flux density we have actually been able to uncover in the HUDF through source detection and stacking.

8 DISCUSSION

8.1 The star-forming main sequence

Armed with the physical knowledge of the properties of the individual ALMA-detected galaxies (Table 4), the stacking results discussed in the previous section, and our knowledge of the redshifts, rest-frame UV luminosities and stellar masses of all galaxies in the HUDF, we now investigate the implications for the relationship between star formation and stellar mass at $z \simeq 2$.

We confine our attention to the 13 ALMA-detected sources with $1.0 < z \leq 3.0$, derive the average properties of this sample, and consider also the average results from stacking in the same redshift range within the two stellar mass bins defined by $9.3 < \log_{10}(M_*/\mathrm{M}_\odot) < 10.3$ and $8.3 < \log_{10}(M_*/\mathrm{M}_\odot) < 9.3$. The median redshift of the detected sample is $z = 2.086$, while for the two stacks the median redshifts are $z \simeq 1.92$ and $z = 2.09$. The corresponding median stellar masses are $\log_{10}(M_*/\mathrm{M}_\odot) = 10.70$, 9.63 and 8.62.

The results are plotted in Fig. 12. This shows (as solid red circles) the positions on the SFR–M_* plane of the 13 ALMA-detected sources, with associated uncertainties (see Table 4), with the average (median) values of total (FIR+UV) SFR plotted for the three mass bins as solid black squares (with standard errors). For the lowest of the stellar mass bins, no detection was achieved in the ALMA stack of the 391 galaxies in the redshift range $1.0 < z \leq 3.0$, and so we plot an upper limit derived from the rms of $S_{1.3} = 1.5\,\mu$Jy achieved in this stack (corresponding to a 1σ limit on average obscured SFR $< 0.45\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$, and hence a limit of total SFR $< 0.87\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$).

The solid black line in Fig. 12 shows a simple relation of the form SFR $\propto M_*^{1.0}$, with sSFR $= 2.2\,\mathrm{Gyr}^{-1}$. This is clearly an excellent fit to our data. Also shown is the original 'main-sequence' (MS) of

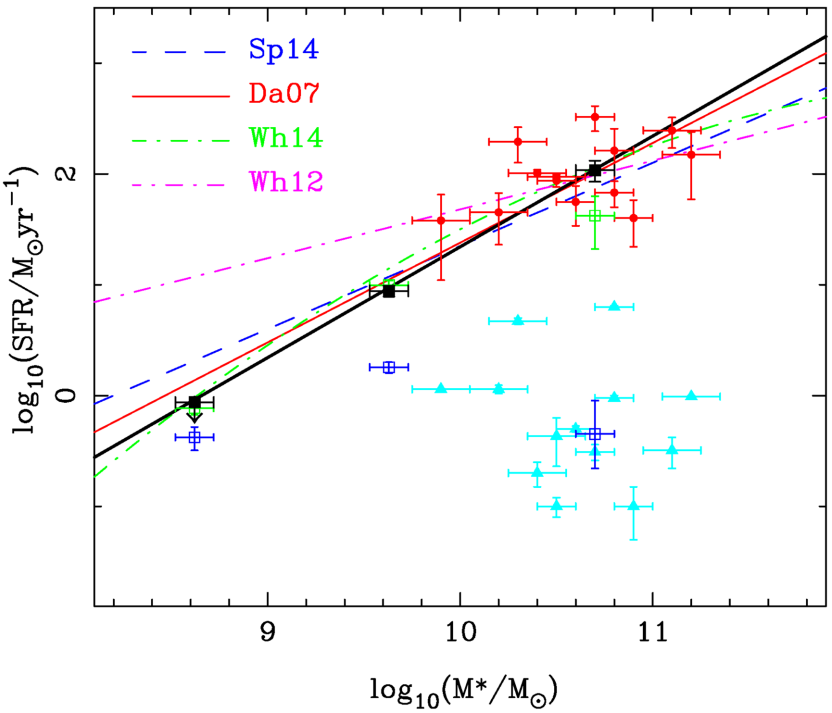

Figure 12. SFR versus galaxy stellar mass at $z \simeq 2$. The 13 ALMA sources that lie in the redshift range $1 < z < 3$ (median redshift $z = 2.09$) are shown twice, first adopting the total SFR (i.e. UV+FIR SFR; red points) and then adopting only the raw UV SFR (cyan triangles). Also shown is the average value in three stellar-mass bins at $1 < z \leq 3$ (derived from stacking as described in the text) for total UV+FIR SFR (black solid squares), raw UV SFR (navy-blue open squares) and dust-corrected SFR (from the optical–near-infrared SED fitting; green open squares). The binned points are plotted at the median stellar mass of each bin. Reassuringly, the green and black values agree well in the two lower mass bins, but the SED A_v-corrected values fall about a factor of 2 short of the true ALMA-derived average in the highest mass bin. The solid black line shows a simple relation of the form SFR $\propto M_*^{1.0}$, with sSFR $= 2.2$ Gyr^{-1}. The other curves are proposed fits to the 'main-sequence' of star-forming galaxies, at $z \simeq 2$, published by Daddi et al. (2007; solid red line, after conversion to a Chabrier IMF), Whitaker et al. (2012; dot–dashed magenta line), Whitaker et al. (2014; dot–dashed green curve) and Speagle et al. (2014; dashed blue line).

star-forming galaxies at $z \simeq 2$ derived by Daddi et al. (2007) (SFR $\propto M_*^{0.9}$, adjusted in mass normalization to account for the change from Salpeter IMF to Chabrier IMF), the shallower $z \simeq 2$ MS presented by Whitaker et al. (2012) (SFR $\propto M_*^{0.6}$), the revised steeper polynomial form presented by Whitaker et al. (2014), and the result of the meta analysis undertaken by Speagle et al. (2014) (calculated at $z \simeq 2$). All of these (and many more) published relations are perfectly consistent with our data at $\log_{10}(M_*/\mathrm{M}_\odot) \simeq 10.7$, proving beyond doubt that the ALMA-detected galaxies lie on the MS at $z \simeq 2$. However, over the dynamic range probed here, none fit any better than (or indeed as well as) the simply constant sSFR relation plotted in black (although both the original Daddi et al. and the revised Whitaker et al. relations are also clearly acceptable).

In Fig. 12, we also plot the corresponding results for the raw UV SFR for the individual sources (cyan triangles), and median values in each stellar mass bin (open navy-blue squares). Finally, in each mass bin we also plot the median values of dust-corrected UV SFR derived from the optical–near-infrared SED fitting (open green squares), as would be obtained in the absence of any direct mm/far-IR information (i.e. based on the values given in column 5 of Table 4, and analogous results for all galaxies in the two lower mass bins). In the two lower mass bins the positions of these green points are reassuringly close to the black points, indicating that dust-corrected SFR from optical–near-infrared SED fitting works well for moderately obscured lower mass galaxies, and also confirming that the steepness of the MS at low masses is not an artefact of the ALMA stacking procedure. However, in the highest mass bin, the SED-estimated median SFR falls short of the true ALMA-derived values by a factor $\simeq 2$. This is perhaps not surprising, as the ratio of median total SFR to median raw UV-estimated SFR at $\log_{10}(M_*/\mathrm{M}_\odot) = 10.70$ can be seen to be $\simeq 300$. The ALMA-derived results argue against any flattening of the MS, at least up to stellar masses $\log_{10}(M_*/\mathrm{M}_\odot) \simeq 11$ (provided quenched galaxies are excluded, although the lack of quenched high-mass galaxies in the field makes this distinction academic at $z > 2$).

8.2 Mass dependence of obscuration

The results discussed in the previous subsection, and presented in Fig. 12, clearly imply a strong mass dependence for the ratio of obscured to unobscured star formation. We explore this explicitly, again focusing on $z \simeq 2$ (i.e. $1 < z \leq 3$), in Fig. 13. Here we plot the ratio of total SFR (i.e. UV+FIR) to UV–visible SFR as a function of mass. Again we plot the individual ratios for the 13 ALMA-detected sources (see column 9 in able 4), and we also plot the ratios for each of our three mass bins (for the lowest mass bin, an upper limit, as in Fig. 12), summing all SFR (detected-FIR + stacked-FIR + total-UV) in each mass bin, and dividing by all the raw UV SFR in that bin.

It can be seen that the median value of this ratio for the detected sources is 152, or in fact 218 if the two detected galaxies in this redshift range with $\log_{10}(M_*/\mathrm{M}_\odot) < 10.3$ are disregarded (for consistency with the high mass bin for which $10.3 < \log_{10}(M_*/\mathrm{M}_\odot) < 11.3$). However, the stacked (true total ratio)

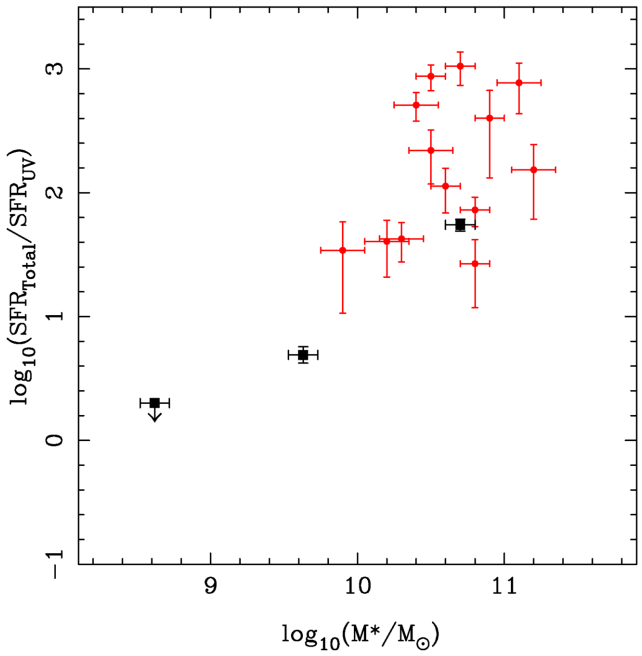

Figure 13. The ratio of total SFR (i.e. UV+FIR) to UV–visible SFR as a function of galaxy stellar mass at $z \simeq 2$. The individual ratios for the 13 ALMA-detected sources at $1 < z \leq 3.0$ are indicated by the red points, and the median value at $\log_{10}(M_*/\mathrm{M}_\odot) > 10.3$ is $\simeq$200 (see column 9 in Table 4). The black squares indicate the ratios for each of our three mass bins, summing all SFR (detected-FIR + stacked-FIR + total-UV) in each bin, and dividing by all the raw UV SFR in that mass bin. The stacked (true total ratio) value at median mass $\log_{10}(M_*/\mathrm{M}_\odot) = 10.7$ is 55 ± 6 (significantly lower than the median ratio of ALMA-detected sources). The average ratio at $\log_{10}(M_*/\mathrm{M}_\odot) = 9.6$ is 4.9 ± 0.8, while at $\log_{10}(M_*/\mathrm{M}_\odot) = 8.6$ our analysis yields a limit on this ratio of <2.5 (see Section 8.2).

value at median mass $\log_{10}(M_*/\mathrm{M}_\odot) = 10.7$ is significantly lower, at 55 ± 6. The average ratio at $\log_{10}(M_*/\mathrm{M}_\odot) = 9.6$ is 4.9 ± 0.8, while at $\log_{10}(M_*/\mathrm{M}_\odot) = 8.6$ our analysis yields <2.5.

Clearly, the ratio of total SFR to UV–visible SFR is a very steep function of mass above $\log_{10}(M_*/\mathrm{M}_\odot) \simeq 9.5$; over the next decade in mass, this ratio also increases by a factor $\simeq$10, indicating that essentially all the increase in SFR with stellar mass on the main sequence at higher masses is delivered in a dust-obscured form.

This clarifies why sub-mm observations are so effective at detecting high-mass galaxies at $z \geq 2$. Because the galaxies lie on the star-forming main sequence, a galaxy with $\log_{10}(M_*/\mathrm{M}_\odot) = 10.7$ will have 10 times the intrinsic SFR of a galaxy of stellar-mass $\log_{10}(M_*/\mathrm{M}_\odot) = 9.7$, and the obscured:unobscured ratio is also typically 10 times greater.

8.3 Cosmic star formation history

Finally, we use our new ALMA results, combined with our existing knowledge of the rest-frame UV properties of the galaxies in the HUDF, to explore the evolution of cosmic SFR density, ρ_{SFR}. In Fig. 14, we plot the dust-obscured values of ρ_{SFR} in unit redshift bins from $z = 1$ to $z = 5$, combining our ALMA detections (Table 4) and the results of the stacking summarized in Section 7. In essence, these results (indicated by the red data points in both panels of Fig. 14) are the result of collapsing the 1.3-mm flux distribution shown in Fig. 10 along the stellar-mass axis, converting to ρ_{SFR}(obscured) using the SED plotted in Fig. 8 (see Section 6.3) and dividing by the comoving cosmological volume sampled by the 4.5 arcmin^2 of the HUDF in each redshift bin. This thus represents a direct sum of the observed dust-obscured star formation activity in the field, and does not rely on, for example, assumptions regarding the poorly constrained faint-end slope of the far-infrared luminosity function at these redshifts.

In a similar manner, we have summed up all the individual values of raw UV SFR for all the $\simeq$2000 galaxies uncovered through the *HST* imaging of the HUDF, to construct the evolution of ρ_{SFR}(visible) over the same redshift range (indicated by the solid blue points in both panels of Fig. 14). These two values are then simply added to produce an estimate of ρ_{SFR}(total), which is plotted as the black points in the left-hand panel of Fig. 14, and compared with two recently published parametric fits to the evolution of ρ_{SFR} as derived from reviews of the existing literature results (Behroozi et al. 2013; Madau & Dickinson 2014).

From the left-hand panel of Fig. 14, it can be seen that the results of this simple calculation are remarkably consistent with the published curves at $z \simeq 1$–3 (which, where required, have been scaled to the Chabrier IMF), but our derived value appears somewhat low at $z \simeq 4.5$. However, by these redshifts the census of UV SFR produced by this simple summing of galaxy contributions is inevitably incomplete at the faint end, with the median value of M_{UV} changing from -16.8 at $z \simeq 1.5$ to -18.1 at $z \simeq 4.5$.

Therefore, to enable a proper and consistent accounting of the UV contribution, we integrated the evolving galaxy UV luminosity function (weighted by UV luminosity) down to a consistent luminosity limit corresponding to $M_{\mathrm{UV}} = -15$. To do this we used the latest UV luminosity functions produced by Parsa et al. (2016) and Bouwens et al. (2015), and the results are shown by the open blue squares in the right-hand panel of Fig. 14, where they are also then added to the (unchanged) values of ρ_{SFR}(visible) to produce a revised estimate for ρ_{SFR}(total). Reassuringly, it can be seen that the two alternative values of ρ_{SFR}(visible) are indistinguishable at $z \simeq 1.5$, where the HUDF data are deep enough to sample the UV luminosity function to $M_{\mathrm{UV}} < -15$ (Parsa et al. 2016). However, at $z < 1$ the HUDF is too small to properly reflect the contribution made to UV luminosity density by the brighter galaxies, while at $z > 2$ the direct census becomes increasingly incomplete, and the integration of the UV luminosity function yields a systematically increasing upward correction.

At $z \simeq 1$–3, the overwhelming dominance of dust-obscured star formation means that completing the UV census in this way makes little impact on the estimated value of ρ_{SFR}(total). However, at $z > 3$ the impact is more pronounced, and by $z \simeq 4$–5 the effect is dramatic enough to result in ρ_{SFR}(visible) being larger than ρ_{SFR}(obscured), and ρ_{SFR}(total) being lifted up to values that are consistently higher than predicted by the Madau & Dickinson (2014) fit, but in excellent agreement with (at least the high-redshift end of) the parametric fit obtained by Behroozi et al. (2013).

We stress that this is the first time a direct census of ρ_{SFR}(obscured) has been performed at these redshifts, and that the integration of the UV LF has not always been performed to a consistently deep luminosity limit at all redshifts. It is therefore reassuring to see such good agreement with the results of existing literature reviews, but also interesting to note how the balance of power shifts from unobscured to obscured star formation with cosmic time. It appears that, at redshift $z > 4$, most of the star formation in the Universe is unobscured, and relatively modest corrections are therefore required to infer ρ_{SFR}(total) from rest-frame UV observations. By contrast, at $z < 4$ the obscured mode of star formation becomes increasingly dominant until it is primarily responsible for producing the peak in ρ_{SFR}(total) at $z \simeq 2.5$. Both Behroozi et al. (2013) and

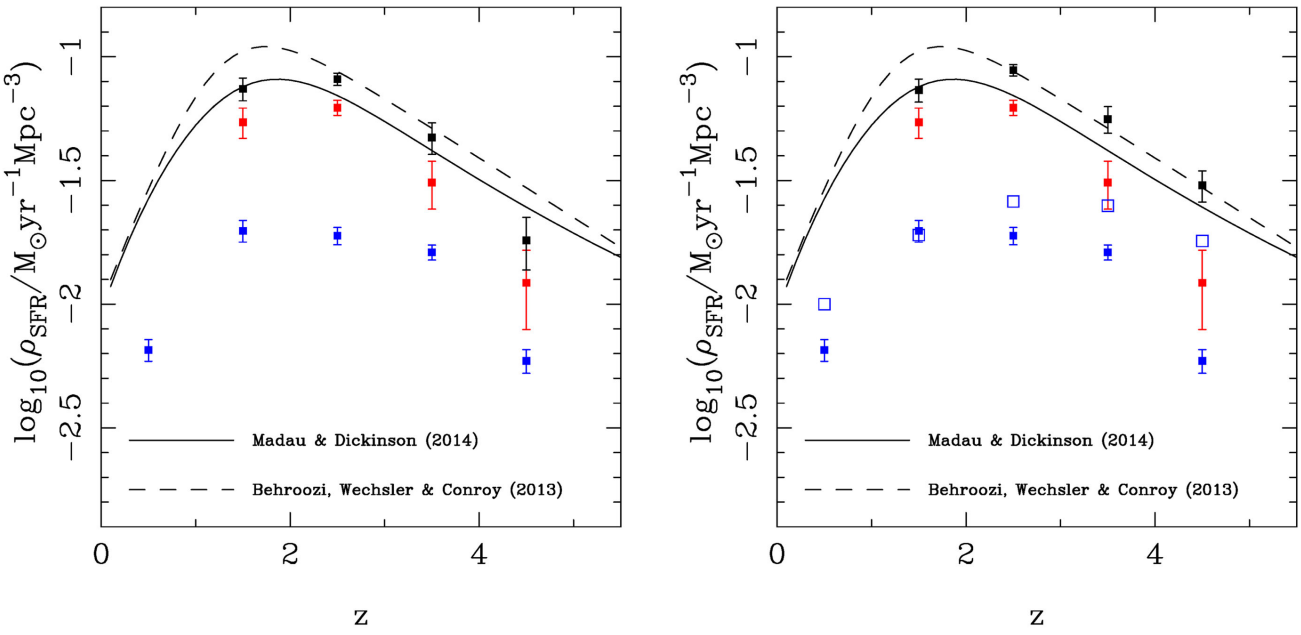

Figure 14. The cosmic evolution of comoving SFR density, $\rho_{\rm SFR}$. In the left-hand panel, we plot our new ALMA-derived estimates of $\rho_{\rm SFR}$(obscured) (red points), estimated values of $\rho_{\rm SFR}$(visible) produced by summing the raw UV-derived SFR of the galaxies in the HUDF (blue points) and the values of $\rho_{\rm SFR}$(total) that result from adding these two contributions at each redshift (black points). The solid and dashed curves show two recent parametric descriptions of the evolution of $\rho_{\rm SFR}$ as derived from the literature by Madau & Dickinson (2014) and Behroozi, Wechsler & Conroy (2013), respectively (scaled to the IMF of Chabrier 2003). In the right-hand panel, we show the impact of including a more complete census of $\rho_{\rm SFR}$(visible), based on the luminosity-weighted integral of the evolving UV galaxy luminosity function, down to $M_{\rm UV} = -15$ (Parsa et al. 2016; Bouwens et al. 2015). In this panel, the values of $\rho_{\rm SFR}$(obscured) are unchanged, but the black points indicating $\rho_{\rm SFR}$(total) now include the more complete (and, at high redshift, higher) census of $\rho_{\rm SFR}$(visible), as shown by the open blue squares. At $z \simeq 2$ the agreement with the parametric fit provided by Behroozi et al. (2013) is excellent, and our results indicate that there is a transition from $\rho_{\rm SFR}$(visible)/ $\rho_{\rm SFR}$(obscured) > 1, to $\rho_{\rm SFR}$(visible)/ $\rho_{\rm SFR}$(obscured) < 1 at $z \simeq 4$.

Madau & Dickinson (2014) favoured a peak in $\rho_{\rm SFR}$(total) at $z \leq 2$, and it may be the case that the very brightest sub-mm sources not sampled by the small volume of the HUDF contribute sufficiently around this redshift to both boost $\rho_{\rm SFR}$ by a few per cent, and perhaps shift the peak to slightly lower redshifts. However, recent studies utilizing *Herschel* and SCUBA-2 data to probe the dust-obscured contribution out to $z \simeq 3$ also appear to favour a peak redshift in the range $z \simeq 2$–3 (e.g. Burgarella et al. 2013; Bourne et al. 2016).

While the average metallicity of the Universe must obviously increase with cosmic time, the apparent transition at $z \simeq 4$ from predominantly visible star formation at higher redshifts, to primarily dust-obscured star formation at $z < 4$ is not necessarily driven by an increase in the prevalence of dust in galaxies of a given mass. Rather it appears to be largely due to the rapid growth in the number density of the high-mass galaxies that contain most of the dust-obscured star formation at $z \simeq 1$–3 (see Figs. 7 and 13); indeed there is at most only weak evidence for any evolution in the ratio of obscured:unobscured SFR with redshift for galaxies selected at constant stellar mass M_* (Bourne et al. 2016). As is clear from Fig. 7, deeper sub-mm imaging with ALMA has the potential to resolve this issue, by charting the evolution of dust-obscured star formation activity at constant stellar mass from $z \simeq 7$ to the present day.

9 CONCLUSIONS

We have constructed the first deep, contiguous and homogeneous ALMA image of the *Hubble Ultra Deep Field*, using a mosaic of 45 ALMA pointings at $\lambda = 1.3$ mm to map the full $\simeq 4.5$ arcmin^2 area previously imaged with WFC3/IR on the *HST*. The resulting image reaches an rms sensitivity $\sigma_{1.3} \simeq 35$ μJy, at a resolution of $\simeq 0.7$ arcsec. A search for sources in this image yielded an initial list of $\simeq 50$ $>3.5\sigma$ peaks, but an analogous analysis of the negative image showed that, as expected from the size and noise level of the map, 30–35 of these peaks were likely to be spurious. We then exploited the unparalleled optical/near-infrared data in the field to isolate the real sources, via the identification of robust galaxy counterparts within a search radius of $\simeq 0.5$ arcsec (in the process uncovering the need for a $\simeq 0.25$ arcsec shift in the *HST* coordinate system).

The result is a final sample of 16 ALMA sources with point-source flux densities $S_{1.3} > 120$ μJy. The brightest three of these sources were clearly resolved, and so we measured their total flux densities from image fitting. For the fainter sources, we estimated total flux densities by applying a 25 per cent boost to their point-source flux densities (a correction based on a stack of the brightest five sources).

All of the ALMA sources have secure galaxy counterparts with accurate redshifts (13 spectroscopic, 3 photometric), yielding a mean redshift $\langle z \rangle = 2.15$. Within our sample, 12 galaxies are also detected at 6 GHz in new ultradeep JVLA imaging. Due to the wealth of supporting data in this unique field, the physical properties of the ALMA-detected galaxies are well constrained, including their stellar masses and UV–visible SFR. To estimate the dust-obscured SFR for the sources, we established a template far-infrared SED by fitting their combined ALMA and (deconfused) *Spitzer*+*Herschel* photometry.

Our results confirm previous indications that stellar mass is the best predictor of SFR in the high-redshift Universe, with our ALMA sample containing seven of the nine galaxies in the HUDF with $M_* \geq 2 \times 10^{10}$ M$_\odot$ at $z \geq 2$. We detect only one galaxy at $z > 3.5$,

and show that the lack of high-redshift detections simply reflects the rapid drop-off of high-mass galaxies in the field at $z > 3$.

The detected sources, coupled with results of stacking in bins of redshift and mass, allow us to probe the redshift/mass distribution of the 1.3-mm background down to $S_{1.3} \simeq 10\,\mu$Jy, and we find that our estimate of the total 1.3-mm background provided by detected and stacked sources is (just) consistent with the background measurement made by COBE.

We find strong evidence for a steep 'main sequence' for star-forming galaxies at $z \simeq 2$, with SFR $\propto M_*$ and a mean specific SFR $\simeq 2.2$ Gyr^{-1}. Moreover, we find that $\simeq$85 per cent of total star formation at $z = 1$–3 is enshrouded in dust, with $\simeq$65 per cent of all star formation at this epoch occurring in high-mass galaxies ($M_* > 2 \times 10^{10}\,M_\odot$), for which the average obscured:unobscured SF ratio is $\simeq$200. Averaged over cosmic volume we find that, at $z \simeq 2$, the ratio of obscured to unobscured star formation activity rises roughly proportional to stellar mass, from a factor $\simeq$5 at $M_* \simeq 5 \times 10^9\,M_\odot$, to a factor $\simeq$50 at $M_* \simeq 5 \times 10^{10}\,M_\odot$.

Finally, we combine our new ALMA results with the existing *HST* data to attempt a complete census of obscured and visible star formation in the field, and hence revisit the cosmic evolution of SFR density ($\rho_{\rm SFR}$). We find reassuringly good agreement with recent estimates of the evolution of $\rho_{\rm SFR}$ with redshift, and our results indicate that, while most star formation in the young Universe is visible at rest-frame UV wavelengths, dust-obscured star formation becomes dominant at $z < 4$, due primarily to the rise in the number density of high-mass star-forming galaxies.

ACKNOWLEDGEMENTS

JSD acknowledges the support of the European Research Council via the award of an Advanced Grant (PI J. Dunlop), and the contribution of the EC FP7 SPACE project ASTRODEEP (Ref.No: 312725). RJM acknowledges the support of the European Research Council via the award of a Consolidator Grant (PI R. McLure). JEG thanks the Royal Society for support. MJM acknowledges the support of the UK Science and Technology Facilities Council, and the FWO Pegasus Marie Curie Fellowship. RJI acknowledges support from the European Research Council through the Advanced Grant COSMICISM 321302. WR acknowledges support from JSPS KAKENHI Grant Number JP15K17604 and Chulalongkorn University's CUniverse (CUAASC). RSE acknowledges support from the European Research Council through the Advanced Grant 669253. JAP acknowledges support from the European Research Council through the Advanced Grant 670193. PJ would like to thank NRAO for assistance in the form of a Reber fellowship.

This paper makes use of the following ALMA data: ADS/JAO.ALMA#2012.1.00173.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), NSC and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc; VLA data are from project ID VLA/14A-360.

This work is based in part on observations made with the NASA/ESA *Hubble Space Telescope*, which is operated by the Association of Universities for Research in Astronomy, Inc, under NASA contract NAS5-26555. This work is also based in part on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under NASA contract 1407. *Herschel* is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA.

REFERENCES

Aretxaga I. et al., 2007, MNRAS, 379, 1571
Ashby M. L. N. et al., 2013, ApJ, 769, 80
Ashby M. L. N. et al., 2015, ApJS, 218, 33
Austermann J. E. et al., 2010, MNRAS, 401, 160
Barger A. J., Cowie L. L., Sanders D. B., Fulton E., Taniguchi Y., Sato Y., Kawara K., Okuda H., 1998, Nature, 394, 248
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
Beckwith S. V. W. et al., 2006, AJ, 132, 1729
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Béthermin M. et al., 2016, A&A, 586, L7
Blain A. W., Longair M. S., 1993, MNRAS, 264, 509
Bourne N. et al., 2016, preprint (arXiv:1607.04283)
Bouwens R. J. et al., 2010, ApJ, 709, L133
Bouwens R. J. et al., 2015, ApJ, 803, 34
Bowler R. A. A. et al., 2015, MNRAS, 452, 1817
Brammer G. B. et al., 2012, ApJS, 200, 13
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Bunker A. J. et al., 2010, MNRAS, 409, 855
Burgarella D. et al., 2013, A&A, 554, 70
Bussmann R. S. et al., 2015, ApJ, 812, 43
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koorneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Capak P. L. et al., 2015, Nature, 522, 455
Carniani S. et al., 2015, A&A, 584, 78
Chabrier G., 2003, PASP, 115, 763
Chapin E. L. et al., 2009, MNRAS, 398, 1793
Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772
Condon J. J., 1997, PASP, 109, 166
Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693
Coppin K. E. K. et al., 2006, MNRAS, 372, 1621
Coppin K. E. K. et al., 2015, MNRAS, 446, 1293
Daddi E. et al., 2007, ApJ, 670, 156
Dannerbauer H., Lehnert M. D., Lutz D., Tacconi L., Bertoldi F., Carilli C., Genzel R., Menten K. M., 2004, ApJ, 606, 664
Davé R., Finlator K., Oppenheimer B. D., Fardal M., Katz N., Keres D., Weinberg D. H., 2010, MNRAS, 404, 1355
Devlin M. J. et al., 2009, Nature, 458, 737
Dunlop J. S., 2011, Science, 333, 178
Dunlop J. S. et al., 1994, MNRAS, 350, 769
Dunlop J. S., Cirasuolo M., McLure R. J., 2007, MNRAS, 376, 1054
Dunlop J. S. et al., 2010, MNRAS, 408, 2022
Dunlop J. S. et al., 2013, MNRAS, 432, 3520
Eales S., Lilly S., Gear W., Dunne L., Bond J. R., Hammer F., Le Févre O., Crampton D., 1999, ApJ, 515, 518
Eales S. et al., 2010, PASP, 122, 499
Elbaz D. et al., 2010, A&A, 518, L29
Elbaz D. et al., 2011, A&A, 533, 119
Ellis R. S. et al., 2013, ApJ, 763, L7
Fazio G. G. et al., 2004, ApJS, 154, 10
Finkelstein S. L. et al., 2010, ApJ, 719, 1250
Finkelstein S. L. et al., 2012, ApJ, 758, 93
Finkelstein S. L. et al., 2015, ApJ, 810, 71
Fixsen D. J., Dwek E., Mather J. C., Bennett C. L., Shafer R. A., 1998, ApJ, 508, 123
Fontana A. et al., 2014, A&A, 570, 11
Fujimoto S., Ouchi M., Ono Y., Shibuya T., Ishigaki M., Nagai H., Momose R., 2016, ApJS, 222, 1

Geach J. E. et al., 2013, MNRAS, 432, 53
Giallongo E. et al., 2015, A&A, 578, 83
Greve T. R., Ivison R. J., Bertoldi F., Stevens J. A., Dunlop J. S., Lutz D., Carilli C. L., 2004, MNRAS, 354, 779
Griffin M. J. et al., 2010, A&A, 518, L3
Grogin N. A. et al., 2011, ApJS, 197, 35
Guo Y. et al., 2013, ApJS, 207, 24
Hathi N. P., Malhotra S., Rhoads J. E., 2008, ApJ, 673, 686
Hatsukade B. et al., 2010, AIPC, 1279, 324
Hatsukade B. et al., 2011, MNRAS, 411, 102
Hatsukade B., Ohta K., Yabe K., Seko A., Makiya R., Akiyama M., 2015, ApJ, 810, 91
Hatsukade B. et al., 2016, PASJ, 68, 36
Hayward C. C., Kereš D., Jonsson P., Narayanan D., Cox T. J., Hernquist L., 2011, ApJ, 743, 159
Hodge J. A. et al., 2013, ApJ, 768, 91
Hopkins A. M., Beacom J. F., 2006, ApJ, 651, 142
Hsu L.-T. et al., 2014, ApJ, 796, 60
Hughes D. H., Robson E. I., Dunlop J. S., Gear W. K., 1993, MNRAS, 263, 607
Hughes D. H. et al., 1998, Nature, 394, 241
Illingworth G. D. et al., 2013, ApJS, 209, 6
Ivison R. J. et al., 2007, MNRAS, 380, 199
Karim A. et al., 2013, MNRAS, 432, 2
Kellermann K. I., Fomalont E. B., Mainieri V., Padovani P., Rosati P., Shaver P., Tozzi P., Miller N., 2008, ApJS, 179, 71
Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531
Khochfar S., Silk J., 2009, ApJ, 700, L21
Kirkpatrick A., Pope A., Sajina A., Roebuck E., Yan L., Armus L., Díaz-Santos T., Stierwalt S., 2015, ApJ, 814, 9
Knudsen K. K., Watson D., Frayer D., Christensen L., Gallazzi A., Michałowski M. J., Richard J., Zavala J., 2016, preprint (arXiv:1603.03222)
Koekemoer A. M. et al., 2011, ApJS, 297, 36
Koekemoer A. M. et al., 2013, ApJS, 209, 3
Koprowski M. P., Dunlop J. S., Michałowski M. J., Cirasuolo M., Bowler R. A. A., 2014, MNRAS, 444, 117
Koprowski M. P. et al., 2016, MNRAS, 458, 4321
Kurk J. et al., 2013, A&A, 549, 63
Labbé I. et al., 2015, ApJS, 221, 23
Lindner R. R. et al., 2011, ApJ, 737, 83
Lutz D. et al., 2011, A&A, 532, 90
Madau P., 1995, ApJ, 441, 18
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2009, A&A, 496, 57
Maiolino R. et al., 2015, MNRAS, 452, 54
Marmel-Queralto E., McLure R. J., Cullen F., Dunlop J. S., Fontana A., McLeod D. J., 2016, MNRAS, 460, 3587
Marsden G. et al., 2009, ApJ, 707, 1729
McLeod D. J., McLure R. J., Dunlop J. S., Robertson B. E., Ellis R. S., Targett T. A., 2015, MNRAS, 450, 3032
McLure R. J., Dunlop J. S., Cirasuolo M., Koekemoer A. M., Sabbi E., Stark D., Targett T., Ellis R. S., 2010, MNRAS, 403, 960
McLure R. J. et al., 2011, MNRAS, 418, 2074
McLure R. J. et al., 2013, MNRAS, 432, 2696
Merlin E. et al., 2015, A&A, 582, 15
Michałowski M. J., Hjorth J., Watson D., 2010, A&A, 514, 67
Michałowski M. J. et al., 2012a, MNRAS, 426, 1845
Michałowski M. J., Dunlop J. S., Cirasuolo M., Hjorth J., Hayward C. C., Watson D., 2012b, A&A, 541, 85
Michałowski M. J. et al., 2016, preprint (arXiv:1610.02409)
Miller N. A., Fomalont E. B., Kellermann K. I., Mainieri V., Norman C., Padovani P., Rosati P., Tozzi P., 2008, ApJS, 179, 114
Momcheva I. G. et al., 2016, ApJS, 225, 27
Morris A. M. et al., 2015, AJ, 149, 178
Murphy E. J. et al., 2011, ApJ, 737, 67
Narayanan D. et al., 2015, Nature, 525, 496
Noeske K. G. et al., 2007, ApJ, 660, L43
Nonino M. et al., 2009, 183, 244
Oesch P. A. et al., 2010, ApJ, 709, L16
Oke J. B., 1974, ApJS, 27, 21
Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
Oliver S. J. et al., 2010, MNRAS, 405, 2279
Oliver S. J. et al., 2012, MNRAS, 424, 1614
Ono Y., Ouchi M., Kurono Y., Momose R., 2014, ApJ, 795, 50
Oteo I., Zwaan M. A., Ivison R. J., Smail I., Biggs A. D., 2016, ApJ, 822, 360
Ouchi M. et al., 2013, ApJ, 778, 1020
Parsa S., Dunlop J. S., McLure R. J., Mortlock A., 2016, MNRAS, 456, 3194
Peacock J. A. et al., 2000, MNRAS, 318, 535
Pilbratt G. L. et al., 2010, A&A, 518, L1
Poglitsch A. et al., 2010, a&A, 518, L2
Pope A. et al., 2008, ApJ, 675, 1171
Puget J.-L., Abergel A., Bernard J.-P., Boulanger F., Burton W. B., Desert F.-X., Hartmann D., 1996, A&A, 308, L5
Rangel C., Nandra K., Laird E. S., Orange P., 2013, MNRAS, 428, 3089
Renzini A., Peng Y., 2015, ApJ, 801, L29
Riechers D. A. et al., 2013, Nature, 502, 459
Roseboom I. G. et al., 2013, MNRAS, 436, 430
Rujopakarn W. et al., 2016, ApJ, 833, 12
Salpeter E. E., 1955, ApJ, 121, 161
Schreiber C. et al., 2015, A&A, 575 74
Scott S. E. et al., 2002, MNRAS, 331, 817
Scott S. E., Dunlop J. S., Serjeant S., 2006, MNRAS, 370, 1057
Scott K. S. et al., 2008, MNRAS, 385, 2225
Scott K. S. et al., 2010, MNRAS, 405, 2260
Scott K. S. et al., 2012, MNRAS, 423, 575
Scoville N. et al., 2016, ApJ, 820, 83
Silva L., Granato G. L., Bressan A., Denese L., 1998, ApJ, 509, 103
Simpson J. M. et al., 2015a, ApJ, 807, 128
Simpson J. M. et al., 2015b, ApJ, 799, 81
Skelton R. E. et al., 2014, ApJS, 214, 24
Smail I., Ivison R. J., Blain A. W., 1997, ApJ, 490, L5
Smolcic V. et al., 2012, A&A, 548, 4
Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15
Spilker J. et al., 2016, ApJ, 826, 112
Steinhardt C. L. et al., 2014, ApJ, 791, L25
Tadaki K. et al., 2015, ApJ, 811, L3
Umehata H. et al., 2015, ApJ, 815, L8
Vanzella E. et al., 2008, A&A, 478, 83
Watson D., Christensen L., Knudsen K. K., Richard J., Gallazi A., Michałowski M. J., 2015, Nature, 519, 327
Weiss A. et al., 2009, ApJ, 707, 1201
Weiss A. et al., 2013, ApJ, 767, 88
Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, ApJ, 754, L29
Whitaker K. E. et al., 2014, ApJ, 795, 104
Wuyts S., Franx M., Cox T. J., Hernquist L., Hopkins P. F., Robertson B. E., van Dokkum P. G., 2009, ApJ, 696, 348
Xue Y. Q. et al., 2011, ApJS, 195, 10

APPENDIX A: MID/FAR-INFRARED IMAGES

In this appendix, we provide postage-stamp *Spitzer* MIPS (24 μm), *Herschel* PACS (70, 100 and 160 μm) and *Herschel* SPIRE (250, 350 and 500 μm) images centred on each of the 16 ALMA sources listed in Tables 2 and 4 (for which colour *HST* images are provided in Fig. 4). The contrast in imaging resolution between ALMA and *Herschel* is readily apparent, particularly at the longer wavelengths accessed with the SPIRE instrument, but useful photometry was nevertheless obtained for many of the sources using the deconfusion techniques described in Section 2.2.2.

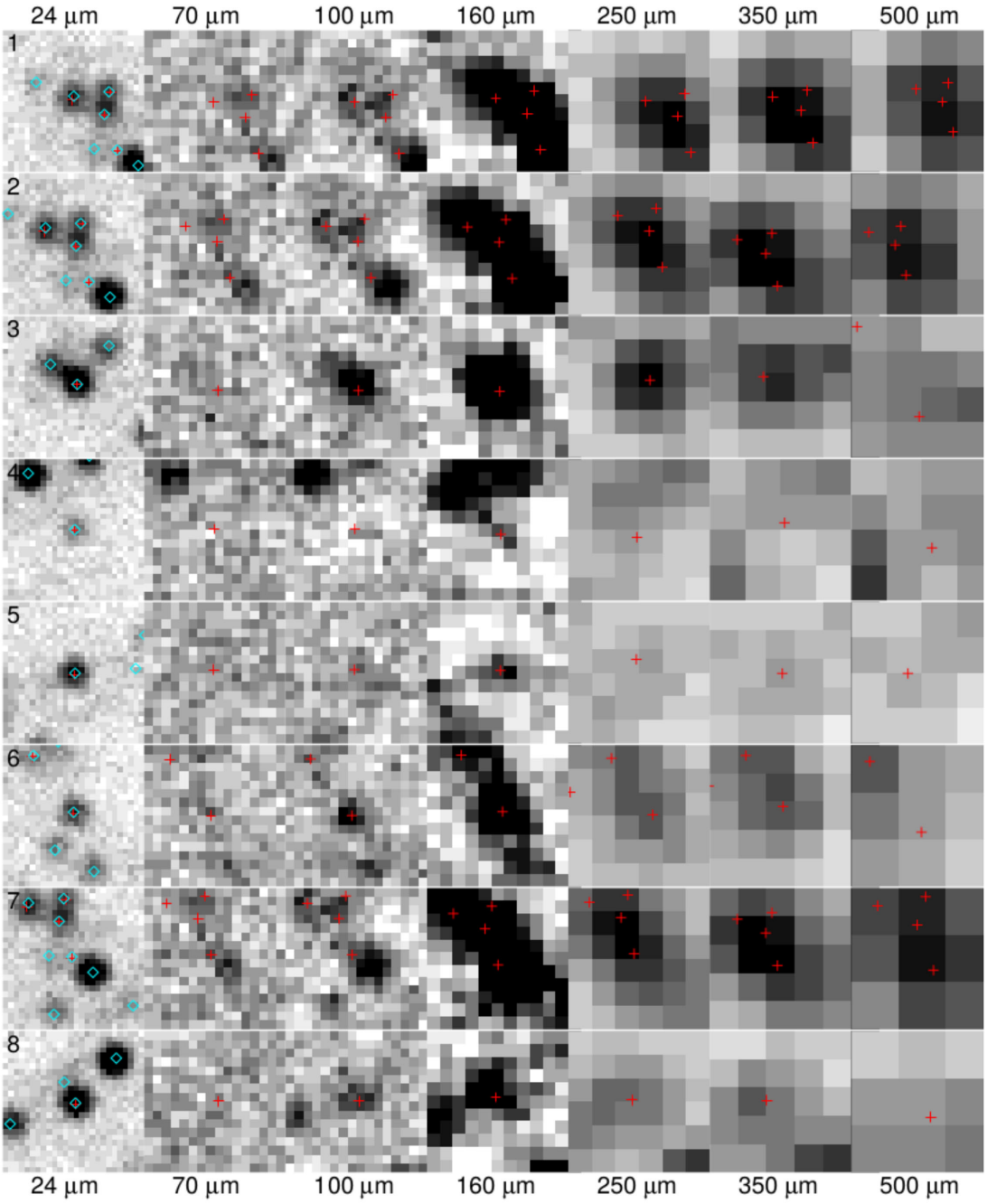

Figure A1. Postage-stamp grey-scale plots showing 30 arcsec × 30 arcsec *Spitzer* MIPS (24 μm), *Herschel* PACS (70, 100 and 160 μm) and *Herschel* SPIRE (250, 350 and 500 μm) images centred on each of the 16 ALMA sources. The ALMA positions are marked by the red crosses, while the blue diamonds mark 24 μm catalogue positions.

[1]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK*
[2]*European Southern Observatory, Karl-Schwarzchild-Str 2, D-95748 Garching b. Munchen, Germany*
[3]*Centre for Astrophysics Research, Science & Technology Research Institute, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK*
[4]*Department of Physics, Faculty of Science, Chulalongkorn University, 254 Phayathai Road, Pathumwan, Bangkok 10330, Thailand*
[5]*Department of Astronomy, University of Massachusetts, Amherst, MA 01002, USA*
[6]*Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T1Z1, Canada*
[7]*Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[8]*Department of Physics and Astronomy, Sonoma State University, 1801 East Cotati Avenue, Rohnert Park, CA 94928-3609, USA*
[9]*Instituto Nacional de Astrofisica, Optica y Electronica (INAOE), Aptdo. Postal 51 y 216, 72000 Puebla, Pue., Mexico*
[10]*NIST Quantum Devices Group, 325 Broadway Mailcode 817.03, Boulder, CO 80305, USA*
[11]*Herzberg Astronomy and Astrophysics, National Research Council Canada, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada*
[12]*Sorbonne Universites, UPMC-CNRS, UMR7095, Institut dAstrophysique de Paris, F-75014 Paris, France*

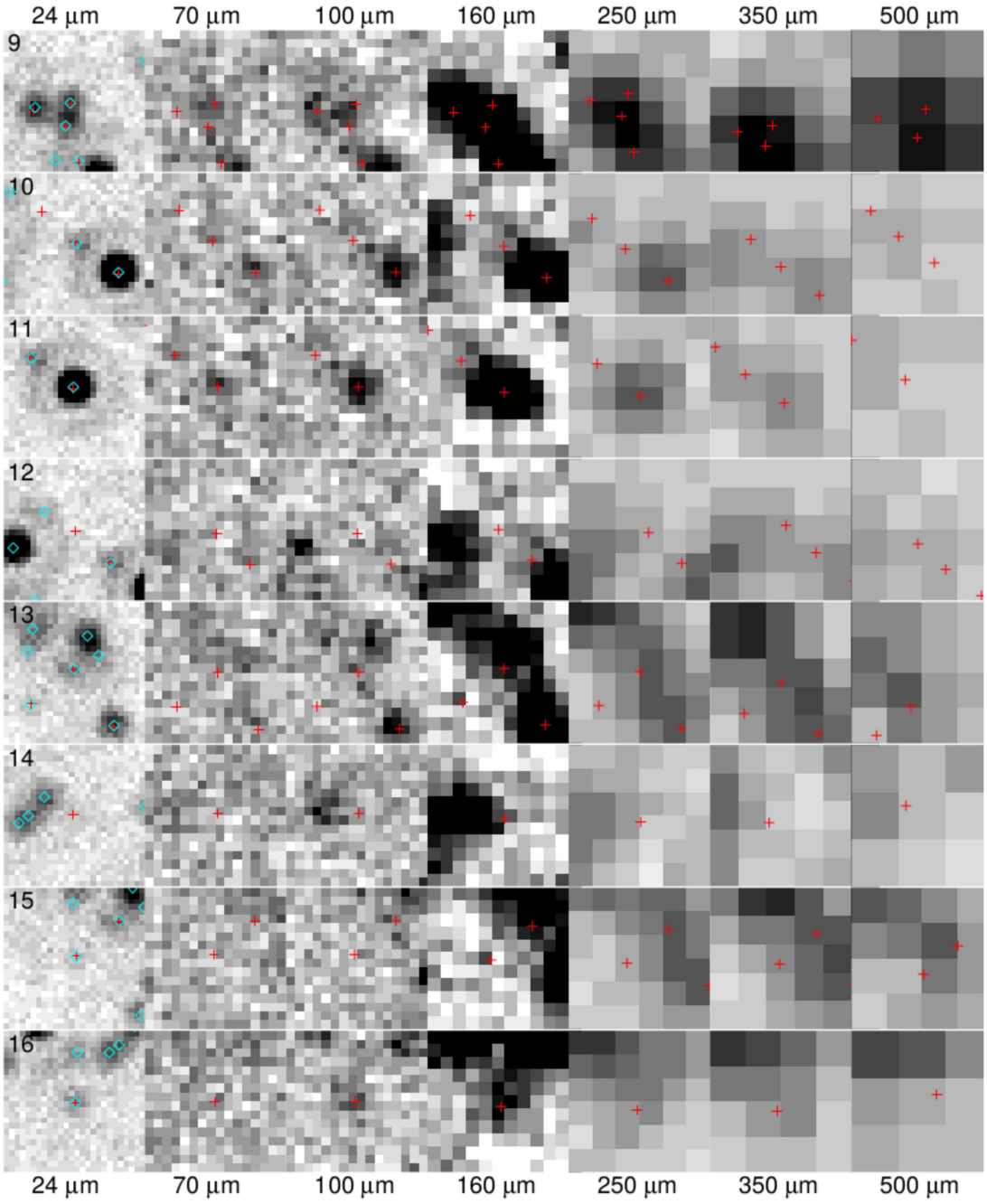

Figure A1 – *Continued.*

[13]*Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA*
[14]*Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA*
[15]*Instituto de Fisica y Astronomia, Universidad de Valparaiso, Avda. Gran Bretana 1111, Valparaiso, Chile*
[16]*National Radio Astronomy Observatory, 1003 Lopezville Road, Socorro, NM 87801, USA*
[17]*Department of Physics and Astronomy, Haverford College, Haverford, PA 19041, USA*
[18]*National Radio Astronomy Observatory, Charlottesville, VA 22903, USA*
[19]*Department of Physics and Astronomy, Texas A & M University, College Station, TX 77843-4242, USA*
[20]*Steward Observatory, University of Arizona, Tucson, AZ 85721, USA*
[21]*Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA*
[22]*Dunlap Institute for Astronomy and Astrophysics, University of Toronto, Toronto, ON M5S 3H4, Canada*
[23]*Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, the Netherlands*

This paper has been typeset from a TEX/LATEX file prepared by the author.