The Astrophysical Journal Supplement Series, 222:4 (29pp), 2016 January doi:10.3847/0067-0049/222/1/4

IDENTIFICATION OF $z \gtrsim 2$ *Herschel* 500 μm SOURCES USING COLOR DECONFUSION

X. W. Shu[1,2], D. Elbaz[1], N. Bourne[3], C. Schreiber[1], T. Wang[1,4], J. S. Dunlop[3], A. Fontana[5], R. Leiton[6], M. Pannella[1,7], K. Okumura[1], M. J. Michałowski[3], P. Santini[5], E. Merlin[5], F. Buitrago[3], V. A. Bruce[3], R. Amorin[5], M. Castellano[5], S. Derriere[8], A. Comastri[9], N. Cappelluti[9], J. X. Wang[10], and H. C. Ferguson[11]

[1] Laboratoire AIM-Paris-Saclay, CEA/DSM/Irfu—CNRS—Université Paris Diderot, CEA-Saclay, pt courrier 131, F-91191 Gif-sur-Yvette, France; xwshu@mail.ustc.edu.cn
[2] Department of Physics, Anhui Normal University, Wuhu, Anhui, 241000, China
[3] SUPA, Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh, EH9 3HJ, UK
[4] School of Astronomy and Astrophysics, Nanjing University, Nanjing, 210093, China
[5] INAF—Osservatorio Astronomico di Roma, Via Frascati 33, I-00040 Monte Porzio Catone (RM), Italy
[6] Astronomy Department, Universidad de Concepción, Casilla 160-C, Concepcin, Chile
[7] Ludwig-Maximilians-Universitt, Department of Physics, Scheinerstr. 1, D-81679 München, Germany
[8] Observatoire astronomique de Strasbourg, Universit de Strasbourg, CNRS, UMR 7550, 11 rue de lUniversit, F-67000 Strasbourg, France
[9] INAF—Osservatorio Astronomico di Bologna, Via Ranzani 1, I-40127, Bologna, Italy
[10] Department of Astronomy, University of Science and Technology of China, Hefei, Anhui 230026, China
[11] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
Received 2015 July 6; accepted 2015 November 26; published 2016 January 8

ABSTRACT

We present a new method to search for candidate $z \gtrsim 2$ *Herschel* 500 μm sources in the Great Observatories Origins Deep Survey-North field using a $S_{500\,\mu m}/S_{24\,\mu m}$ "color deconfusion" technique. Potential high-z sources are selected against low-redshift ones from their large 500 to 24 μm flux density ratios. By effectively reducing the contribution from low-redshift populations to the observed 500 μm emission, we are able to identify counterparts to high-z 500 μm sources whose 24 μm fluxes are relatively faint. The recovery of known $z \gtrsim 4$ starbursts confirms the efficiency of this approach in selecting high-z *Herschel* sources. The resulting sample consists of 34 dusty star-forming galaxies at $z \gtrsim 2$. The inferred infrared luminosities are in the range 1.5×10^{12}–$1.8 \times 10^{13}\,L_{\odot}$, corresponding to dust-obscured star formation rates (SFRs) of ~260–3100 $M_{\odot}\,\mathrm{yr}^{-1}$ for a Salpeter initial mass function. Comparison with previous SCUBA 850 μm-selected galaxy samples shows that our method is more efficient at selecting high-z dusty galaxies, with a median redshift of $z = 3.07 \pm 0.83$ and with 10 of the sources at $z \gtrsim 4$. We find that at a fixed luminosity, the dust temperature is ~5 K cooler than that expected from the $T_d - L_{\mathrm{IR}}$ relation at $z \lesssim 1$, though different temperature selection effects should be taken into account. The radio-detected subsample (excluding three strong active galactic nucleus) follows the far-infrared (far-IR)/radio correlation at lower redshifts, and no evolution with redshift is observed out to $z \sim 5$, suggesting that the far-IR emission is star formation dominated. The contribution of the high-z *Herschel* 500 μm sources to the cosmic SFR density is comparable to that of (sub)millimeter galaxy populations at $z \sim 2.5$ and at least 40% of the extinction-corrected UV samples at $z \sim 4$. Further investigation into the nature of these high-z dusty galaxies will be crucial for our understanding of the star formation histories and the buildup of stellar mass at the earliest cosmic epochs.

Key words: galaxies: high-redshift – galaxies: starburst – infrared: galaxies – submillimeter: galaxies

1. INTRODUCTION

Ultraluminous infrared galaxies (ULIRGs; with rest-frame 8–1000 μm luminosities in excess of $10^{12}\,L_{\odot}$) are among the brightest far-infrared (far-IR) emitters in the universe (Sanders & Mirabel 1996; Casey et al. 2014 and references therein). The far-IR luminosities of ULIRGs are dominated by reprocessed thermal dust emission due to star formation activity, corresponding to a star formation rate (SFR) of $>170\,M_{\odot}\,\mathrm{yr}^{-1}$. Locally, these luminous, dusty star-forming galaxies are rare, but they become more abundant with increasing redshift and dominate the IR luminosity density of galaxies around $z \sim 2$ (Chapman et al. 2005; Lagache et al. 2005; Le Floc'h et al. 2009; Magnelli et al. 2009, 2013). The formation process of ULIRGs is currently in dispute. While mergers are believed to be a dominant process in local ULIRGs, studies have suggested that the driver of star formation in ULIRGs at high redshift ($z \gtrsim 2$) appears different (e.g., Dekel et al. 2009), with a significant fraction being pure disks (e.g., Kartaltepe et al. 2012; Targett et al. 2013). Reproducing the number counts and extreme far-IR luminosities for this population, especially those at the highest redshifts, can place tight constraints on galaxy formation models (Baugh et al. 2005; Swinbank et al. 2008; Coppin et al. 2009). Since a significant fraction of cosmic star formation is likely hidden by dust (e.g., Chapman et al. 2005; Lagache et al. 2005; Barger et al. 2012; Casey et al. 2012a, 2012b, 2014; Lutz 2014; Madau & Dickinson 2014), understanding high-z far-IR luminous galaxies is crucial in order to construct a complete picture of galaxy evolution.

Observations at submillimeter (submm) and millimeter (mm) wavelengths have identified a population of ULIRGs at high redshift (e.g., Borys et al. 2003; Pope et al. 2005; Perera et al. 2008; Michałowski et al. 2010a, 2010b; Wardlow et al. 2011; Yun et al. 2012; Casey et al. 2013; Roseboom et al. 2013), which benefits from a strong negative K-correction in the (sub)mm. Follow-up studies of submm-selected galaxies (SMGs) have shown that the bulk of SMGs is between $z \sim 2$–3 (e.g., Chapman et al. 2005), with less than ~30% being at $z > 3$ (Chapin et al. 2009; Biggs et al. 2011; Wardlow et al. 2011; Michałowski et al. 2012). While radio identification has been the common technique to precisely localize SMGs at

other wavelengths for follow-up observations, a significant number of them remain undetected even in the current deepest radio observations (e.g., Barger et al. 2014; Swinbank et al. 2014), indicating that they are possibly at higher redshifts. Interferometric submm imaging of SMGs has been used to unambiguously identify counterparts in the optical/near-IR in a relatively unbiased way, revealing a significant population of SMGs at $z \gtrsim 3$ (e.g., Younger et al. 2007, 2009; Smolčić et al. 2012a, 2012b; Miettinen et al. 2015; Smolčić et al. 2015), challenging the previously derived redshift distribution of SMGs (see also Vieira et al. 2013). On the other hand, the selection of SMGs is known to be severely biased against galaxies with warmer dust temperatures (e.g., Chapman et al. 2003, 2010; Casey et al. 2009; Magdis et al. 2010; Magnelli et al. 2010). Therefore, the properties of thewhole high-z ULIRG population, such as the redshift distribution, luminosity and number density, are still not well characterized.

The *Herschel Space Observatory* (hereafter *Herschel*) carried out observations in the far-IR bands up to 500 μm and mapped much larger sky areas down to the confusion limit than previous (sub)mm surveys. By sampling the peak of heated dust emission for the first time, *Herschel* enables a direct assessment of the far-IR spectral energy distribution (SED) of SMGs (e.g., Magnelli et al. 2010, 2012, 2014; Huang et al. 2014; Swinbank et al. 2014). It has provided unbiased measurements of obscured SFR, dust temperature and dust mass for a large number of star-forming galaxies up to $z \sim 2$ (e.g., Elbaz et al. 2010, 2011; Oliver et al. 2012; and references therein).

It has also been shown that *Herschel* is efficient at detecting galaxies at very high redshift when a well-characterized selection is used (e.g., Pope & Chary 2010). For instance, Riechers et al. (2013) pre-selected sources with red SPIRE colors (i.e., $S_{500\,\mu m}/S_{350\,\mu m}$ versus $S_{350\,\mu m}/S_{250\,\mu m}$), yielding a sample of potential $z > 4$ sources, including one at $z = 6.34$, the most distant known ULIRG to date. Nevertheless, the requirement of $S_{500\,\mu m} > S_{350\,\mu m} > S_{250\,\mu m}$ means that such selection is biased against some $z > 4$ dusty star-forming galaxies with warm dust temperatures (Daddi et al. 2009b; Capak et al. 2011; Smolčić et al. 2015), and limited to the brightest 500 μm sources ($\gtrsim$30 mJy, Dowell et al. 2014). Therefore, a more conservative and efficient way to select high-z ULIRGs is required to better understand the properties of this important population of dust obscured galaxies in the early universe.

A challenge in studying the properties of high-z *Herschel* sources has been the low resolution of the data, especially in the SPIRE passbands (FWHM ~18″, 25″, and 36″ at 250, 350, and 500 μm, respectively). Source confusion is a severe issue in the *Herschel* data, which makes it difficult to identify the correct counterparts and measure fluxes for individual *Herschel* sources. The most common technique to deconvolve the *Herschel* sources is the use of *Spitzer*-MIPS 24 μm data of better resolution (~5.″7) as the position priors to fit their fluxes (e.g., Magnelli et al. 2009, 2013; Roseboom et al. 2010; Elbaz et al. 2011; Leiton et al. 2015). This de-blending approach assumes that all *Herschel* sources are detectable in the deep 24 μm images, and thus introduces a bias against 24 μm faint galaxies which fall below the detection limit. Moreover, the extraction of 24 μm catalogs usually requires IRAC source positions as priors, which can introduce additional identification biases. Magdis et al. (2011) have demonstrated that in the Great Observatories Origins Deep Survey (GOODS) fields, about 2% of the PACS sources are missed in the MIPS-24 μm catalog, most of which being at $z \sim 1.3$ and $z \sim 0.4$ with strong silicate absorption features. There has been no systematic study to address this issue for SPIRE sources, though simulations suggest that statistically very few sources would be missed when requiring a 24 μm prior (Leiton et al. 2015).

Further problems arise when priors are too close for the *Herschel* flux to be deconvolved: when two 24 μm priors are closer than the FWHM/3 of any given *Herschel* band, the flux is usually attributed to one or the other. Previous studies have systematically favored the brightest of the two 24 μm counterparts (e.g., Magnelli et al. 2009, 2013). While such a choice is generally justified for sources at low redshift, it can misidentify the counterpart of a very distant source which is faint at 24 μm but much brighter at *Herschel* wavelengths, especially in the SPIRE 500 μm band. For the same *Herschel* 500 μm flux, a $z = 3.5$ galaxy can be a factor of ~40 fainter in the 24 μm band than a galaxy at $z = 0.8$ (e.g., Leiton et al. 2015). Indeed, the recent study of Yan et al. (2014) has shown that the brightest 24 μm source does not necessarily contribute predominately to the far-IR emission seen in the SPIRE bands. It is therefore important to investigate in detail the misassociated and/or missed population of luminous, high-z *Herschel* sources which could play a significant role in the star formation history in the early universe (Barger et al. 2012; Casey et al. 2012b, 2014; Dowell et al. 2014).

In this paper, we have undertaken a systematic search for high-z *Herschel* 500 μm sources in the GOODS-*North* field, which has the deepest SPIRE observations currently available, using a map-based "color deconfusion" technique. We show that our method, by constructing a $S_{500\,\mu m}/S_{24\,\mu m}$ ratio map, can effectively extract high-z sources contributing to the 500 μm emission. The resulting sample consists of 36 galaxies identified at $z \sim 2$–4.6. In Section 2 we present the data and our method. In Sections 3 and 4 we present and discuss the derived properties of the galaxies, and we summarize the results in Section 5. Throughout this paper, we adopt $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$, and we use a Salpeter initial-mass function (IMF).

2. DATA AND METHOD

2.1. Data

The data used in this paper come from the *Herschel* observations of GOODS-North field as a part of the GOODS-*Herschel* program. The observations cover a total area of $10' \times 17'$ down to depths of 1.1, 2.7, 5.7, 7.2, and 9 mJy (3σ) at 100, 160, 250, 350, and 500 μm , respectively. The SPIRE images of GOODS-North are the deepest observations undertaken by *Herschel* to date, while the PACS data are slightly shallower than the central region of GOODS-South obtained in GOODS-*Herschel* that was combined with the PEP survey in Magnelli et al. (2013). A detailed description of the observations is given in Elbaz et al. (2011). We take advantage of these deep *Herschel* observations, along with the deep *Spitzer*/MIPS 24 μm observation of the GOODS-North field (PI: M. Dickinson), to identify the best candidates of distant dusty star-forming galaxies seen by *Herschel*. Other ancillary data are used including *Spitzer*/IRAC at 3.6, 4.5, 5.8, and 8 μm, VLA at 1.4 GHz ($5\sigma \sim 20\mu$Jy, Morrison et al. 2010), SCUBA at

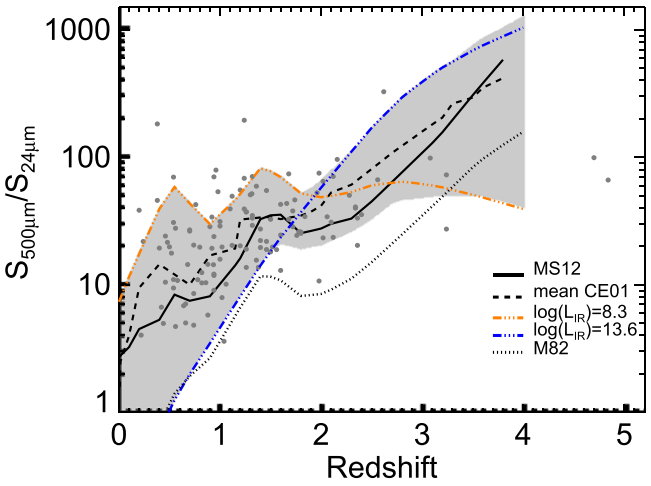

Figure 1. $S_{500\ \mu m}/S_{24\ \mu m}$ flux ratio as a function of redshift for the GOODS-*Herschel* (Elbaz et al. 2011) galaxies with "clean" 500 μm measurements (gray dots). Solid black line: main-sequence SED from Magdis et al. (2012). Dashed black line: mean trend of the Chary & Elbaz (2001) SEDs that are parameterized by the total IR luminosity. Shaded region: range of $S_{500\ \mu m}/S_{24\ \mu m}$ ratios for CE01 templates, with orange dotted–dashed line representing the SED of the minimum IR luminosity ($\log(L_{IR}/L_{\odot}) = 8.3$) and blue dotted–dashed line for the maximum IR luminosity ($\log(L_{IR}/L_{\odot}) = 13.6$). Black dotted line shows the SED from the local starburst galaxy M82 that is out of the range of CE01 libraries at $z \lesssim 3.2$.

850 μm (3.5σ catalog, Pope et al. 2006; Wall et al. 2008), SCUBA-2 at 850 μm (4σ catalog, Barger et al. 2014), and AzTEC at 1.1 mm (3.8σ catalog, Perera et al. 2008; Chapin et al. 2009).

Photometric redshifts used in this paper are drawn from a K_S-band selected multi-wavelength catalog in the GOODS-North field, spanning 20 photometric bands from *GALEX* NUV to IRAC 8 μm. We refer the reader to Pannella et al. (2015) for a more detailed description of the multi-wavelength data, catalog production and photometric redshift estimation. This catalog contains 14,828 galaxies, ~81% of which are brighter than the 5σ limiting magnitude of $K_s = 24.5$, and 3775 galaxies having spectroscopic redshifts. By comparing to the spectroscopic subsample, Pannella et al. found that the relative accuracy $[\Delta z = (z_{\rm phot} - z_{\rm spec})/(1 + z_{\rm spec})]$ of photometric redshifts reaches ~3%, with less than 3% catastrophic outliers (e.g., objects with $\Delta z > 0.2$). Based on the multi-wavelength photometric catalog, stellar masses were determined with *FAST* for each of K_s-detected sources (see Pannella et al. 2015 for more details).

2.2. Construction of the $S_{500\ \mu m}/S_{24\ \mu m}$ Ratio Map

In this section, we present our method to identify candidate high-z dusty galaxies. The observed far-IR/24 μm colors of individual galaxies are known to be correlated with redshift (Leiton et al. 2015). The trend is strongest for the 500 μm/24 μm color. In Figure 1, we show the evolution of the 500 μm/24 μm ratio as a function of redshift for the main-sequence SED from Magdis et al. (2012; solid black line) and the mean trend of the Chary & Elbaz (2001; CE01) templates (dashed black line). While the SEDs for galaxies at $z > 4$ are poorly constrained, distant galaxies are characterized by large 500 μm/24 μm ratios (e.g., $S_{500\ \mu m}/S_{24\ \mu m} > 30$ at $z > 2$). These SEDs are globally in agreement with the observed 500 μm/24 μm color distribution of the "clean" sources in the GOODS-North field (gray dots).[12] We note that there exists a population of low-luminosity galaxies below $z \sim 2$ with relatively high $S_{500\ \mu m}/S_{24\ \mu m}$ ratios, causing the large scatter of the distribution. They might be cold galaxies with abnormal low dust temperature of 10–20 K (~25 per degree^2, e.g., Rowan-Robinson et al. 2010). However, given the very small area covered by GOODS-North, we estimated that the number of such galaxies is low (~2) in the field. Some silicate-break galaxies at $z \sim 0.5$ and $z \sim 1.5$ may contribute the elevated $S_{500\ \mu m}/S_{24\ \mu m}$ ratios due to the shift of the silicate absorption features into the MIPS/24 μm band at these redshifts, but their fraction is relatively small (less than 10%, Magdis et al. 2011). Alternatively, the large scatter could be partly attributed to the catastrophic photometric redshift outliers, or chance associations of a high-z 500 μm galaxy with a lower redshift 24 μm prior. These require further investigations.

Therefore, with the assumption that redshift dominates over temperature evolution in producing the observed shift of the IR SEDs, we are able to select distant galaxies against the low-redshift ones based on their large 500–24 μm flux ratios. We choose this color over the shorter SPIRE wavelengths because the 500 μm band probes the peak of far-IR emission up to the highest redshifts. However, due to the source blending issues, it is difficult to unambiguously assign a measured 500 μm flux density to a specific object seen in higher-resolution images. A single 500 μm source is usually made of multiple 24 μm counterparts that are not necessarily associated. Simulations have shown that 500 μm flux densities can be systematically overestimated due to source blending (Leiton et al. 2015). Therefore, we do not simply search for targets using pre-existing 500 μm photometric catalogs where flux measurements might be highly uncertain, but use a map-based technique keeping the information of the flux distribution in the *observed* images.

Our search approach consists of building a 500 μm / 24 μm ratio map, which relies on the deep *Spitzer*/MIPS 24 μm data. For the GOODS-*North* field of interest here, the 24 μm data reach fainter far-IR luminosities than the *Herschel* bands up to redshift $z \sim 3$ (see Figure 4 in Elbaz et al. 2011). To build our $S_{500\ \mu m}/S_{24\ \mu m}$ ratio map, we use the real SPIRE 500 μm map that we divide by an image containing all 24 μm detections down to 21 μJy (at a level of 3σ, totally 2704 galaxies), which are distributed at their actual positions. We do not use the real *Spitzer*-MIPS 24 μm image because we found that even after smearing the image to the 500 μm beam, negative pixels due to noise fluctuations introduced fake sources in the $S_{500\ \mu m}/S_{24\ \mu m}$ ratio map. In brief, we added point sources using the 500 μm point-spread function (PSF) at the positions 24 μm sources under identical astrometric projection as the 500 μm map, scaling each PSF to the 24 μm flux, to build an image that matches the actual 500 μm map in beam size and pixel scale. The astrometric accuracy of the 500 μm map is verified by stacking at the positions of 241 24 μm sources with $S_{250\ \mu m}$ in the range 10–35 mJy in the field. The stacked map has an offset from the center of the 24 μm counterparts of ~0.8 arcsec, which is small compared to the 7.2 arcsec pixel size of the 500 μm map. In order to take into account the astrometry uncertainties of the *Herschel* data, a random astrometric error of ~0.″5 was introduced to the source

[12] Clean galaxies are defined as sources with $\leqslant$1 neighbor of $S_{\rm neighbor} > 0.5 S_{\rm central}$ within a distance of 20″ (e.g., Elbaz et al. 2011), for which blending effects are small.

positions. Note that in this beam-smeared image, 24 μm flux density for each source is kept, and the background is dominated by the smeared faint sources populating the whole image.

We then construct a 500 μm/24 μm ratio map by dividing the actual 500 μm image by the beam-smeared 24 μm map. Since low-z galaxies are expected to have much lower $S_{500\,\mu m}/S_{24\,\mu m}$ ratios compared to high-z ones (Figure 1), this procedure thus effectively reduces the contribution to the map from the bulk of the low-redshift galaxy population, while considerably boosting the signal from candidate distant sources.

In Figure 2, we show an example of a spectroscopically confirmed $z = 4.05$ SMG (GN10, Daddi et al. 2009a) identified in the $S_{500\,\mu m}/S_{24\,\mu m}$ ratio map (see Section 2.3.1), whose 500 μm flux was initially wrongly associated with a nearby source at $z_{\rm phot} = 1.44$ due to the strong source blending. It can be seen from Figure 2 that GN10 is faint at 24 μm, and its SPIRE emission particularly at 500 μm is heavily blended with nearby sources, making it difficult to identify directly. In contrast, the galaxy is unambiguously uncovered in the $S_{500\,\mu m}/S_{24\,\mu m}$ ratio map at a position coincident with its 24 μm counterpart. Note that the measured SPIRE 350 and 500 μm flux in the public catalog (Elbaz et al. 2011) is associated with the nearby $z_{\rm phot} = 1.44$ galaxy. However, as shown in Figure 2 (lower panel), SED fittings using the whole library of CE01 SED templates to the 24 μm to far-IR photometry suggest that this source contributes to <20% of the measured 500 μm flux. The use of the starburst SED derived by Magdis et al. (2012) didn't improve the fittings and yielded similar results, strongly suggesting that the source is likely incorrectly associated as the counterpart. Conversely, the best-fitting SED model at the redshift of GN10 yields a much better result ($\Delta\chi^2 = 31.9$), indicating that it should dominate the measured 500 μm emission. This demonstrates the effectiveness of our method for uncovering and selecting high-z 500 μm sources. The selection efficiency of our method will be investigated in detail on simulated data (Section 2.3.1), and a validation of the method using SCUBA2 data is presented in Section 2.3.3.

2.3. Object Identification in the $S_{500\,\mu m}/S_{24\,\mu m}$ Map

Given that the $S_{500\,\mu m}/S_{24\,\mu m}$ ratio is very sensitive to high-z sources, we then perform a systematic search for candidates on the ratio map using SExtractor (Bertin & Arnouts 1996). Since the noise distribution is highly non-Gaussian in the final ratio map, it is difficult to define a minimum signal-to-noise ratio (S/N) requirement for source extraction. We therefore used simulations to determine the minimum S/N threshold in order to optimize the "purity" and completeness of the resulting sample. Based on such simulations (see Section 2.3.1), we adopted a minimum S/N of 2, corresponding to a minimum $S_{500\,\mu m}/S_{24\,\mu m}$ ratio of 20.4, where the noise fluctuation, defined as "effective" noise, was estimated from the dispersion of pixel distribution in the ratio map, P(D) plot, as shown in Figure 3 (black curve). Note that this S/N requirement based on simulations cannot be translated to a nominal flux detection limit in a trivial way, but it is sufficient for our purpose of identifying candidate high-z galaxies. A more detailed description on the simulations and the detection efficiency will be given in the next section.

2.3.1. Selection Efficiency

For our sample to be robust and meaningful, we need to determine the efficiency of our selection of $z \gtrsim 2$ infrared-luminous galaxies. To do this we perform Monte-Carlo simulations by injecting fake sources into the actual maps. The positions of injected sources are randomly distributed but are chosen to avoid overlapping within known sources detected in the ratio map (Section 2.3). In the simulations, we assumed a mean $S_{500\,\mu m}/S_{24\,\mu m}$ color evolution with redshift, as shown by the dashed curve in Figure 1. The evolutionary trend is derived by averaging the $S_{500\,\mu m}/S_{24\,\mu m}$ color over a suite of CE01 templates for a given redshift bin. As we will describe later, we used mainly the CE01 library to fit the mid-to-far-IR photometry and derive the integrated far-IR luminosities. The CE01 library comprises a total of 105 SEDs with a range of dust temperatures and is widely used to characterize the far-IR properties of ULIRGs (e.g., Magnelli et al. 2012; Swinbank et al. 2014). The simulation samples a $S_{500\,\mu m}$ range from 5 mJy to 30 mJy in nine steps and a redshift range $1.5 < z < 5$ in six steps. We estimated the efficiency of detection at a given flux and redshift as the fraction of the injected sources that were recovered in the ratio map, using the same selection procedure as above.[13]

We carried out $N_{\rm sim} = 200$ independent simulations for each of the flux densities, each of which has $N_{\rm src} = 20$ sources injected. $N_{\rm src}$ was chosen to be small enough to avoid overlapping between the simulated sources. The averaged detection efficiency over the $N_{\rm sim}$ simulations forms the final detection efficiency, which is displayed as a function of redshift and $S_{500\,\mu m}$ in Figure 4. This figure shows that, as expected, the recovery efficiency increases at high flux densities and redshifts. For instance, for a $S_{500\,\mu m} = 15$ mJy (typical 3σ confusion limit, Nguyen et al. 2010) source at redshift $z = 2$, the selection efficiency is ~45%, while it increases to $\gtrsim$80% for $z > 3$ galaxies with the same flux. Above $z = 3$, the curves of the detection rate as a function of flux evolve very similarly, suggesting that our method is effective at selecting high-redshift sources with a completeness of $\gtrsim$50% for a $S_{500\,\mu m} \gtrsim 10$ mJy source. Note that the detection efficiency for each real source depends on the intrinsic SED, as well as the possible presence of source blending.

Due to the contamination of bright nearby sources, which is particularly strong in the SPIRE 500 μm band, the selection could introduce a bias against the areas where the projected number density of 24 μm sources is higher. Our random distribution of sources in the real 500 μm map allows for checking the effect of the "local" confusion noise on source detections. We find that the recovery rate decreases with the increase of projected 24 μm source density ($N_{\rm neib}$).[14] As shown in Figure 4 (right), for a galaxy at $z = 3$ with $S_{500\,\mu m} = 10$ mJy, the probability that the source with $N_{\rm neib} \lesssim 2$ is detected in the ratio map is ~60%, while it decreases to less than 40% in regions where there are more than six 24 μm sources within a radius of 20″. Correcting for such bias, the detection efficiency would be higher. Visual inspection of source spatial distribution suggests that sources appear to be found in areas where the projected 24 μm source density is relatively low. Note that

[13] Sources are considered recovered if they are found within 2 pixels of the injected position with a peak S/N above 2σ.

[14] For a given galaxy of interest, the source density, $N_{\rm neib}$, is defined as the number of 24 μm sources within a 20″ radius.

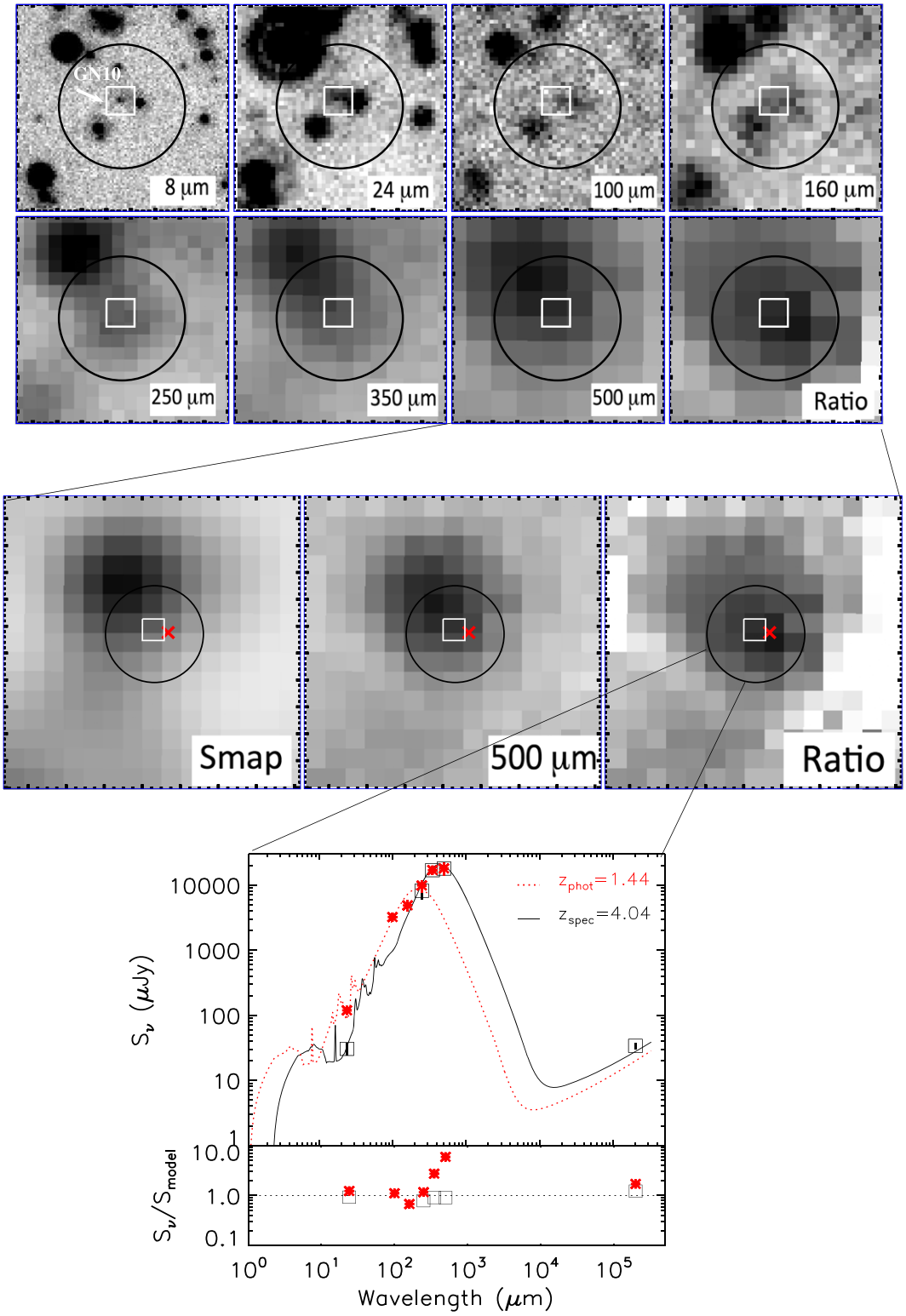

Figure 2. Upper panel: multiwavelength image cutouts of a high-z source, GN10 (Daddi et al. 2009a), in our sample. Each cutout is 30″ on a side. The white square denotes the position of the 8 μm (24 μm) source and the black circle indicates the beam FWHM (36″) at 500 μm. In the ratio map it is evident that there is an excess emission centered on GN10, which is detected with greater significance in the ratio map than in the 500 μm map. Middle panel: zoomed-out view of the 500 μm and ratio image, with a size of 1.8 × 1.8 arcmin^2. On the left, we also show the "beam-smoothed" 24 μm image (Section 2.2) for comparison. Red cross marks the position for a $z_{\rm phot}$ = 1.44 galaxy close to GN10 (~5 arcsec separation), for which the SPIRE 350 and 500 μm flux is associated in the official *Herschel* catalog (Elbaz et al. 2011). Lower panel: 24 μm and the *Herschel* photometry (red stars) for the $z_{\rm phot}$ = 1.44 galaxy. Red dotted curve shows the best-fitting SED at $z_{\rm phot}$ = 1.4. Black curve is the best-fitting SED to the far-IR photometry at the redshift of GN10 ($z_{\rm spec}$ = 4.04). The SED fittings suggest that GN10 (instead of the nearby source) dominates the far-IR emission at 350 and 500 μm. Details of the counterpart identifications to the *Herschel* 500 μm sources detected in the ratio map will be presented in Section 2.3.

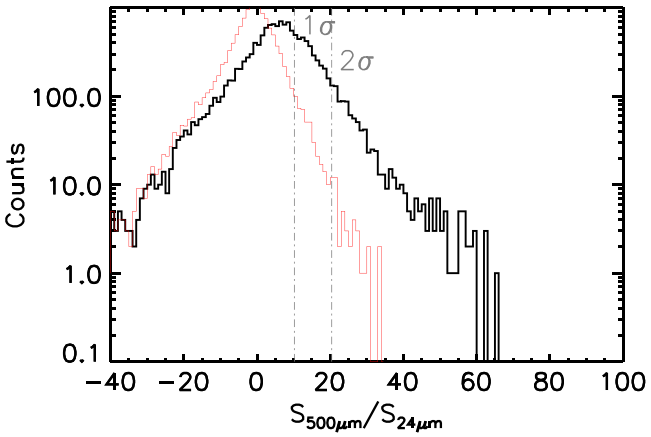

Figure 3. P(D) plot for the $S_{500\,\mu m}/S_{24\,\mu m}$ ratio map. The vertical dotted–dashed lines correspond 1σ and 2σ of the dispersion. The red curve represents the $S_{500\,\mu m}/S_{24\,\mu m}$ ratio after removing all individually detected 500 μm sources in the *Herschel* catalog used in Elbaz et al. (2011). This illustrates that the false detections caused by map artifacts that would pass our selection threshold are negligible (see Section 2.5).

sources which are significantly fainter than their close neighbors in the SPIRE/500 μm would not pass our S/N selection. There are (rare) $z \gtrsim 4$ ULIRGs detected at submm and/or mm that are too faint to be detected in the SPIRE data, such as the one at $z = 5.2$ (HDF850.1, Walter et al. 2012). Such sources would not be selected by our method and hence would be missing from our sample. In addition, our minimum S/N requirement in detecting sources in the ratio map discriminates against galaxies with M82-like SED unless they are at $z > 3$ (Figure 1).

2.3.2. Counterpart Identification

We describe in this section the procedure used to attribute shorter-wavelength counterparts to the *Herschel* 500 μm sources detected in the $S_{500\,\mu m}/S_{24\,\mu m}$ ratio map. It is important to note that the use of this ratio map, constructed from a list of 24 μm sources, does not preclude the identification of 500 μm sources with no 24 μm counterparts. Such sources will naturally emerge as objects with a high ratio. In the following, we show how probabilities of association are determined for potential 24 μm counterparts. We also define a quantitative threshold below which no 24 μm counterpart is considered reliable, hence identifying 24 μm dropouts. Once a source was detected in the ratio map, we searched for potential counterparts in the higher-resolution 24 μm image out to a radius of 15 arcsec from the $S_{500\,\mu m}/S_{24\,\mu m}$ centroid. The search radius was determined from our simulations (Section 2.3.1) where we found that for more than 90% of sources, the extracted positions are less than 15 arcsec from the input positions, i.e., the probability of missing potential counterparts is less than 10%.[15] Note that at the extreme depth reached by the 24 μm imaging reported here, the cumulative surface density of 24 μm sources yields 1–8 sources per search area with a median of 3 sources. For larger radii, one will find more counterparts but the probability of chance associations will also increase.

The primary method for identifying the most likely counterparts to each 500 μm source is "*P*-statistics" (Downes et al. 1986), by calculating the Poissonian probability of sources in a high-resolution catalog (24 μm and/or radio) that lie within the search radius, θ. Given a potential counterpart with flux density S in the high-resolution imaging, the probability of finding at least one object within θ of at least that flux density is $P = 1 - \exp(-\pi n_{(>S)}\theta^2)$, where $n_{(>S)}$ is the surface density of sources above flux density S. P gives the likelihood that a 500 μm source is associated with a counterpart in a higher-resolution map by chance. The lower the P-value, the less likely it is that the candidate is associated with a 500 μm source by chance, and the higher the probability of a genuine match. However, as has been pointed out in the literature (e.g., Yun et al. 2012), the often adopted high-resolution 24 μm or radio continuum data for the counterpart identification may suffer from a systematic bias against high-z sources. Since the main goal of the ratio map is to favor higher-z sources which have statistically higher $S_{500\,\mu m}/S_{24\,\mu m}$ ratios, we replace the surface density of $n_{(>S)}$ in the computation of P-value by $n_{(>z,>S)}$ to take into account the redshift information. For a given counterpart, $n_{(>z,>S)}$ is the surface density of objects brighter than S which have redshifts above z. This means that in a catalog that is cut at flux density S for a given counterpart, we considered only those with redshifts above z in calculating the probability of it being a chance association.

Unlike what happens at shorter wavelengths, the best counterpart for a 500 μm detection is not the brightest 24 μm prior but instead results from a combination of source luminosity and redshift. To account for this, For a given potential 24 μm counterpart, we do not use the number density of sources above flux S at the 24 μm in calculating the Poissonian probability of chance associations, but of 24 μm counterparts brighter than a "predicted" 500 μm flux density, $S^{\rm pred}_{500\,\mu m}$. This is obviously only possible in a field such as GOODS-*North* with high-redshift completeness. To estimate $S^{\rm pred}_{500\,\mu m}$ for each 24 μm counterpart and then determine a density criterion based on it, we assume a characteristic 500/24 μm flux density ratio that depends on redshift following the trend shown with a dashed line in Figure 1. Such a redshift dependence of the 500/24 μm color represents the evolutionary trend of mean $S_{500\,\mu m}/S_{24\,\mu m}$ ratio of CE01 SEDs. Since high-redshift 500 μm sources are expected to present a larger $S_{500\,\mu m}/S_{24\,\mu m}$ ratio and hence a higher probability of being detected in the ratio map, this procedure will naturally select the corresponding 24 μm sources as potential counterparts and reduce the probability of chance associations with lower-redshift sources. However, one has to treat this statistic with caution, as we assumed that lower-z galaxies with abnormal IR SEDs that would have similarly high $S_{500\,\mu m}/S_{24\,\mu m}$ ratios are rare (e.g., Rowan-Robinson et al. 2010). We will comment on later in this paper, ~17% of identifications might be due to the contaminations from some of the lower-z populations. Note that using the observed 500 μm flux to predict 24 μm flux is an alternative way to identify counterparts, but maybe problematic (regardless of the uncertainty in choosing IR SEDs), since without a perfect knowledge of priors, the precise measurements of the 500 μm flux are impossible due to the effect of blending, resulting in the predictions of 24 μm flux with high uncertainty.

In order to define a threshold for our modified P-value calculated for the 24 μm catalog, $P^C_{24\,\mu m}$, that will correspond to a realistic association probability, we use the realistic mock *Herschel* data that are described in detail in Schreiber et al.

[15] This search radius also represents a 2σ positional uncertainty, with $\sigma \sim 0.6 \times$ FWHM/(S/N) (Ivison et al. 2007), where FWHM = 36 arcsec, for a detection in the 500 μm map with S/N = 3.

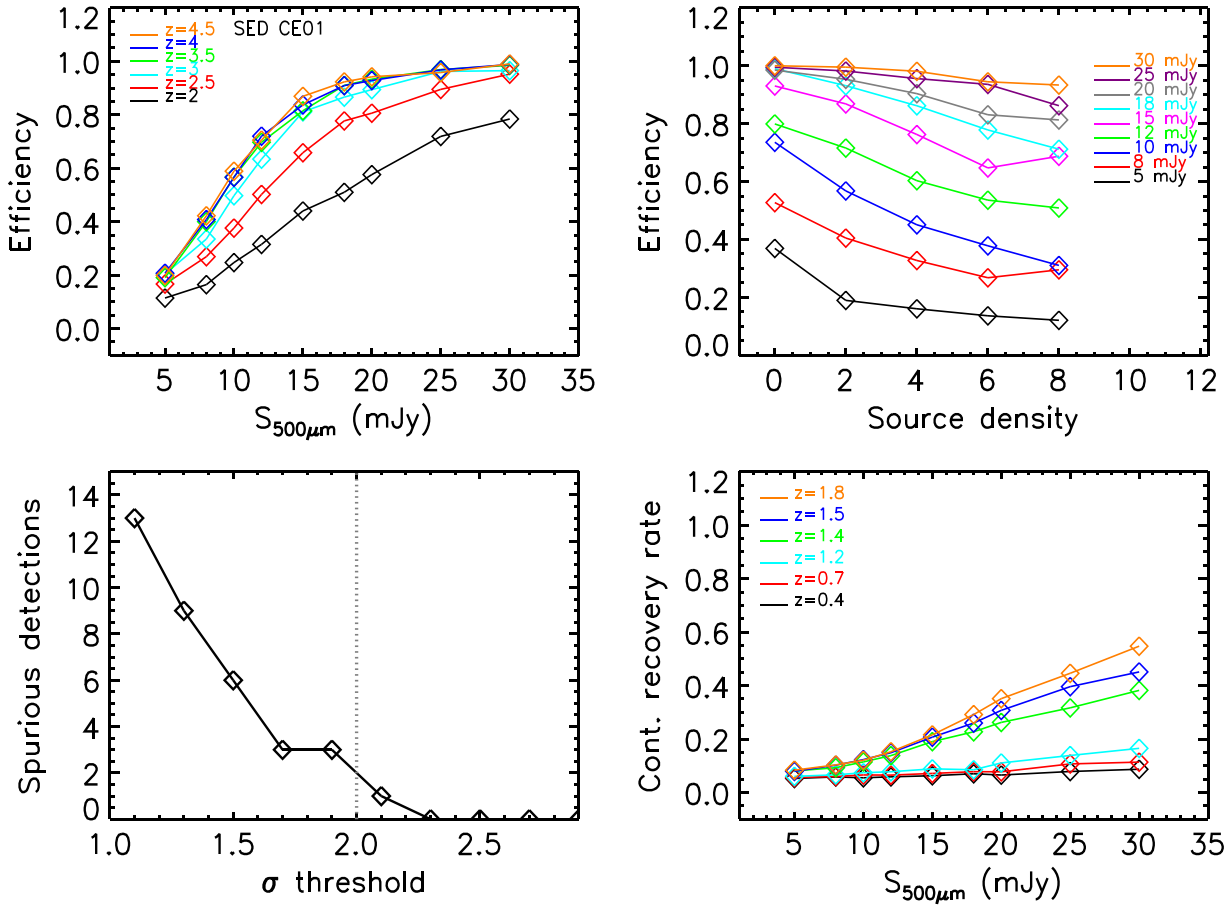

Figure 4. Upper left: selection efficiency as a function of 500 μm flux density and redshift, assuming a mean CE01 SED (Chary & Elbaz 2001). Note that using the main-sequence SED from Magdis et al. (2012) in simulations yields similar results. For sources with $z > 3$, the selection functions appear very similar and the recovery rate is $\gtrsim$60% above 10 mJy. Upper right: selection efficiency as a function of source density for simulated galaxies at $z = 3$. The source density is defined as the number of 24 μm sources within 20″ radius. Lower left: number of expected false detections above a given detection threshold. Lower right: the detection rate of contamination sources at $z < 2$ as a function of 500 μm flux for different redshift bins.

(2015). We construct a $S_{500\,\mu m}/S_{24\,\mu m}$ map for the mock *Herschel* data and detect candidate high-z sources using the same procedure as we described above. In Figure 5 (upper), we plot the P-statistics calculated using the formulation above, against the redshifts for all input 24 μm sources within the search radius (filled squares). We found that at a threshold of $P \lesssim 0.1$, ~90% sources detected in the ratio map would have at least one 24 μm counterpart at $z \gtrsim 2$ (red circles), and no source below $z \sim 1.5$ has a P-value greater than 0.1. Figure 5 (lower) shows the distribution of the ratio of 500 μm flux of candidate counterparts ($P \lesssim 0.1$) to the total 500 μm flux of all 24 μm sources within the search radius. This ratio provides a measure of the importance, in terms of 500 μm flux, of the identifications. A ratio close to 1 indicates that the 500 μm emission is dominated by a single (high-z) 24 μm counterpart. Only ~14% (6/42) 500 μm sources detected in the ratio map have more than one 24 μm galaxy with additional source(s) contributing up to 50% of the total 500 μm flux, and five of them in fact have multiple $z \gtrsim 2$ counterparts. Therefore ~86% (36/42) of the 500 μm sources have been associated with a single $z \gtrsim 2$ source, which is main contribution to the 500 μm emission (>50%), suggesting a high efficiency of the identifications.

In Figure 6, we plot the cumulative distribution of our modified P-values when taking into account the redshift information for all 24 μm counterparts within the search radius (solid line). Also shown is the distribution of the traditional P-values which were calculated using only the 24 μm flux, as often adopted in literature (dotted line). Based on the resulted Poisson probabilities, almost none of the associations would be judged to be significant at $\geqslant$90% confidence (i.e., $P \leqslant 0.1$). Using the predicted 500 μm flux instead (dashed line) would result in slightly more sources with $P \leqslant 0.1$, but still not statistically high enough to quantify the formal significance of the identifications.

2.3.3. Validation of the Method with SCUBA2

In order to further test our method and its efficiency in uncovering candidates of high-z *Herschel* 500 μm sources, we build a $S_{500\,\mu m}/S_{24\,\mu m}$ map using the data from the COSMOS field for which we have a SCUBA-2 450 μm image for direct comparison. The higher resolution of SCUBA-2 (7.″5 FWHM, Geach et al. 2013) at 450 μm, compared to 36″ at 500 μm, offers a unique opportunity to assess the reliability of the high-z 500 μm sources identified using our $S_{500\,\mu m}/S_{24\,\mu m}$ method. Figure 7 shows the SCUBA-2 450 μm image of the COSMOS field (left). The corresponding *Herschel* 500 μm image of similar size is shown in the middle. On the right, we show the *Herschel* 500 μm to *Spitzer* 24 μm flux ratio map, overlaid with detected sources (red circles) using the same search

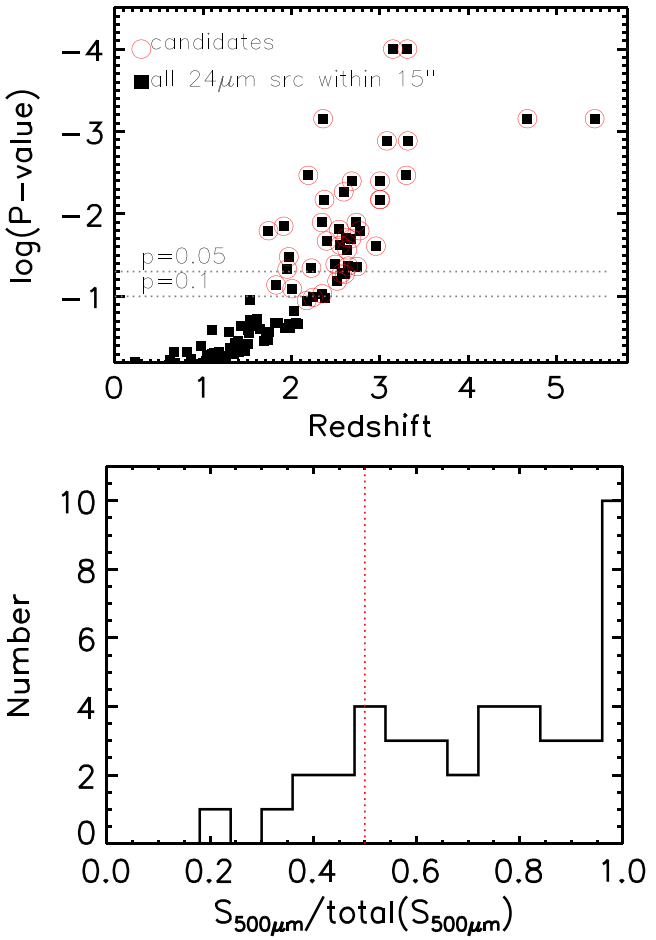

Figure 5. Upper: *P*-values calculated as $P = 1 - \exp(-\pi n\theta^2)$, where n is $n(> S^{\rm Pred}_{500\,\mu{\rm m}}, > z)$ and θ is the search radius of 15″, are plotted against redshifts for all 24 μm sources within the search radius (filled squares). Candidate counterparts identified with $P \lesssim 0.1$ are shown in red circles. For sources with multiple candidate counterparts (11/42), the one with the least *P*-value is plotted. The $S^{\rm Pred}_{500\,\mu{\rm m}}$ is predicted 500 μm flux for each source (see Section 2.3.2). Lower: the distribution of $S_{500\,\mu{\rm m}}/{\rm total}(S_{500\,\mu{\rm m}})$, where $S_{500\,\mu{\rm m}}$ is the 500 μm flux density of the counterpart identified with $P \lesssim 0.1$, and ${\rm total}(S_{500\,\mu{\rm m}})$ is the summed 500 μm flux of all *input* 24 μm sources within the search radius. Here we used the *input* $S_{500\,\mu{\rm m}}$ of each source in the simulated map. Note that 36 out of 42 sources have one 24 μm galaxy contributing more than 50% of the total 500 μm flux (a majority contribution to the 500 μm emission), suggesting a high efficiency of the *P*-value in identifying the *best* counterpart. For the remaining six objects with $S_{500\,\mu{\rm m}}/{\rm total}(S_{500\,\mu{\rm m}}) < 0.5$, five have multiple counterparts (*P*-values between 0.0027 and 0.0864). The summed 500 μm flux density of those counterparts to the ${\rm total}(S_{500\,\mu{\rm m}})$ is greater than 0.5. The red dashed vertical line represents $S_{500\,\mu{\rm m}}/{\rm total}(S_{500\,\mu{\rm m}}) = 0.5$.

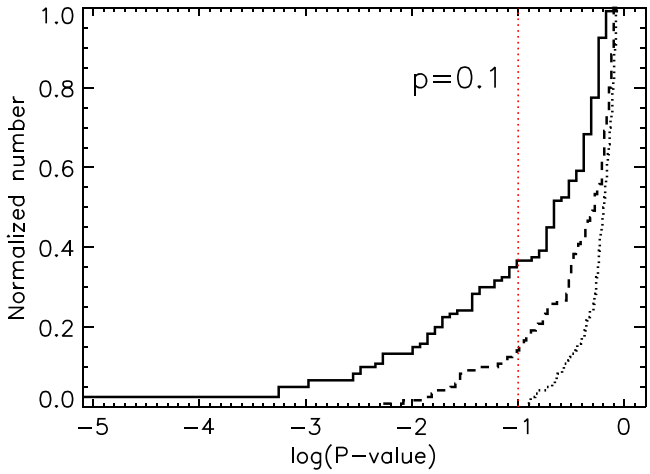

Figure 6. Cumulative distribution of *P*-values for all 24 μm counterparts within the search radius. The solid lines shows the distribution of the modified *P*-values when taking into account the redshift information. Also shown is the distribution of *P*-values which were calculated using only the 24 μm flux, as often adopted in literature (dotted line), and those using only the predicted 500 μm flux (dashed line).

criteria as described in Section 2.1. It can be seen that many bright 500 μm sources in the actual map (middle) disappear in the $S_{500\,\mu{\rm m}}/S_{24\,\mu{\rm m}}$ map. This is likely due to the fact that many of them are dominated by low-z galaxies where the $S_{500\,\mu{\rm m}}/S_{24\,\mu{\rm m}}$ ratios are low. Conversely, while some sources appear relatively faint and undetectable in the actual 500 μm map (close to the confusion limit ~6 mJy), they are more unambiguously seen in the ratio map (e.g., the last two panels in Figure 9). These fainter 500 μm sources are in fact clearly seen in the deeper SCUBA-2 450 μm imaging (rms ~ 2.5–4.0 mJy/beam) of COSMOS.

When compared to the field covered by the SCUBA-2 450 μm map, a total of 22 sources are detected in the ratio map, and 18 of them have 450 μm counterparts within 15″. For the remaining four sources without SCUBA-2 counterparts, three are located close to the edge of the SCUBA-2 map where the local noise level is relatively high (white circles in Figure 7). We also checked the positional offsets between the sources identified in the ratio map and the SCUBA-2 450 μm sources. We found that the positional offsets for most sources are less than 10″. Considering the SCUBA-2 beam size of ~7″ which is close to the pixel size of *Herschel* 500 μm map (7.″2), this strongly suggests a physical association between the sources detected in the ratio map and those in the SCUBA-2 map.

With the same *P*-statistics and cut ($P \lesssim 0.1$), we present in Figure 8 the counterpart identifications to high-z 500 μm sources in the COSMOS field, where the higher resolution SCUBA-2 450 μm imaging allows for more precisely localizing the 24 μm counterparts. We found a similar trend of *P*-value against redshift as in Figure 5. At a *P*-value cut of $\lesssim$0.1, most, if not all, 500 μm sources detected in the ratio map have at least one 24 μm counterpart at $z \gtrsim 2$. Most high-z 500 μm sources that we identified are also detected in the SCUBA-2 450 μm image with positions coincident with the 24 μm counterpart, suggesting a low probability for chance associations. These results further support the reliability of our approach for counterpart identification. We therefore adopt a cut on the Poisson probabilities of $P \lesssim 0.1$ for the counterpart identifications to our sample.

In Figure 9, we show examples of the identifications of two relatively clean sources from our catalog where the blending effect is small, as well as four sources in the COSMOS field. Though they could also be identified in the original 500 μm map, this demonstrates that the detection in the ratio map is without any doubt coincident with a high-z source: one is spectroscopically confirmed at $z = 3.19$ and another has a photometric redshift at $z = 4.15$. Further support comes from the comparison with the SCUBA2 data at the 450 μm, showing that candidate high-z 500 μm sources have been unambiguously identified.

In order to considering alternative counterparts for the few sources with no 24 μm counterpart, which may be an IRAC source with a predicted 500 μm flux density greater than the one associated with the 24 μm counterparts, we in fact

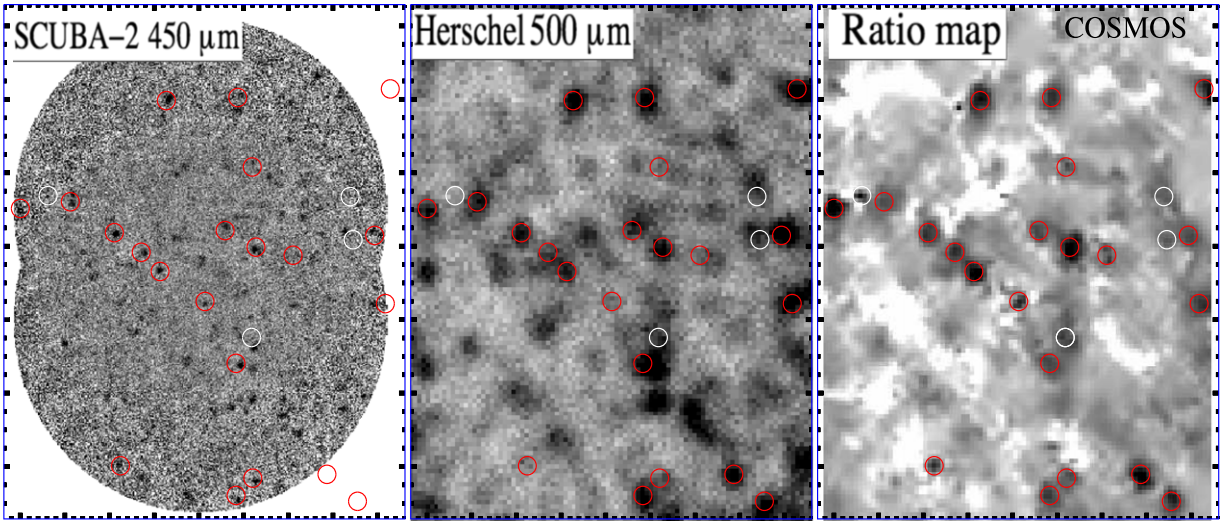

Figure 7. SCUBA-2 450 μm image (left), *Herschel* 500 μm image (middle), and the *Herschel* 500 μm to *Spitzer* 24 μm ratio map (right). The data are taken from observations of the COSMOS field. Sources that are detected in the ratio map are shown in red circles with a radius of 15″, while white shows those without counterparts in the 450 μm map. Note that while many bright 500 μm sources disappear in the ratio map, there are few faint sources being more unambiguously revealed in the ratio map. The SCUBA-2 450 μm data are taken from the SCUBA-2 Cosmology Legacy Survey (Geach et al. 2013).

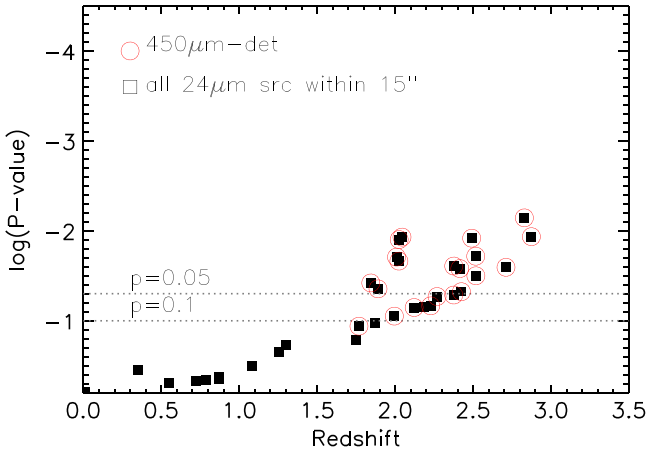

Figure 8. Plot of redshift vs. *P*-value for all 24 μm sources (filled squares) within the search radius of 15″, and counterparts to 500 μm sources (red circles). Most high-z candidate 500 μm sources (red circles) identified with $P \lesssim 0.1$ (at $\gtrsim$90% confidence) are also detected in the SCUBA-2 450 μm imaging with higher resolution, suggesting that the rate of chance associations is low.

considered two extrapolations of $S_{500\,\mu\mathrm{m}}$, the one derived from the observed 24 μm and the one derived from star formation main-sequence relation (Column 8 in Table 1). The latter approach is based on redshift and stellar mass measurements, associating an SFR to each IRAC source using the main-sequence relation (Schreiber et al. 2015). The SFR was converted into a total IR luminosity using the Kennicutt (1998) relation.[16] We then assigned a favorable CE01 SED to each galaxy based on the IR luminosity to predict $S_{500\,\mu\mathrm{m}}$. We used the UVJ selection technique (Whitaker et al. 2012) to isolated passive galaxies and excluded them from the counterpart catalog. Although the main-sequence approach provides only a lower limit to SFR and hence estimate of $S_{500\,\mu\mathrm{m}}$, it will help to discriminate against candidates of high-z passive galaxies which are undetected at 24 μm. Note that 500 μm detections that do not present 24 μm counterparts fulfilling the *P*-value criterion will therefore be considered as 24 μm dropouts.

Using the above procedure, we have found at least one 24 μm counterpart for 28 (out of 36) sources detected in the ratio map (Table 1). In eight cases where no such counterparts can be found, we first proceed to search in the IRAC 3.6 μm catalog using the same radius and *P*-value cut. This results in two further identifications in the IRAC band (GH500.01 and GH500.15). For the remaining six sources, we used the radio data for counterpart identifications. Out of a total of 36 sources, seven have more than one potential 24 μm counterpart, and seven are found to have additional counterpart(s) from the IRAC identifications. Such multiple statistical associations could be partly due to the source clustering at similar redshifts (e.g., GH500.26, GH500.27). Therefore, a unique counterpart is identifiable in about 60% of the cases, while two or more candidates are present in others. For the latter case, we choose the one(s) detected in the radio (if available) as the most probable counterpart for further analysis (Table 2).[17] Candidate *Spitzer* and radio counterparts and their computed *P*-statistics[18] are given in Table 1.

We note that some bright silicate absorption galaxies at $z \sim 0.4$ and $z \sim 1.3$ which have depressed 24 μm emission (Magdis et al. 2011), may have elevated $S_{500\,\mu\mathrm{m}}/S_{24\,\mu\mathrm{m}}$ ratios meeting our selection criteria. By inspecting the photometric redshift and far-IR SED, we found one such lower-z galaxy (GH500.06) that may coincidently have a large $S_{500\,\mu\mathrm{m}}/S_{24\,\mu\mathrm{m}}$ ratio and we excluded it from our sample. In addition, there are five sources at redshifts less than 2 (GH500.03, GH500.11, GH500.22, GH500.24, GH500.36), among which three are spectroscopically confirmed at $z_{\mathrm{spec}} = 1.6$, $z_{\mathrm{spec}} = 1.76$, and

[16] We used a mass-dependent dust extinction, i.e., IR excess $\mathrm{IRX} = \log_{10}(L_{\mathrm{IR}}/L_{\mathrm{UV}})$, to decompose the SFR into a dust-obscured component seen in the FIR and a dust-free component which emerges in the UV (e.g., Heinis et al. 2014).

[17] For GH500.08, we choose the one of not being detected in the radio as best counterpart (see Appendix A).

[18] In computing the *P*-statistics for the radio catalog, we used the nominal method without taking into account the redshift information (e.g., Pope et al. 2006).

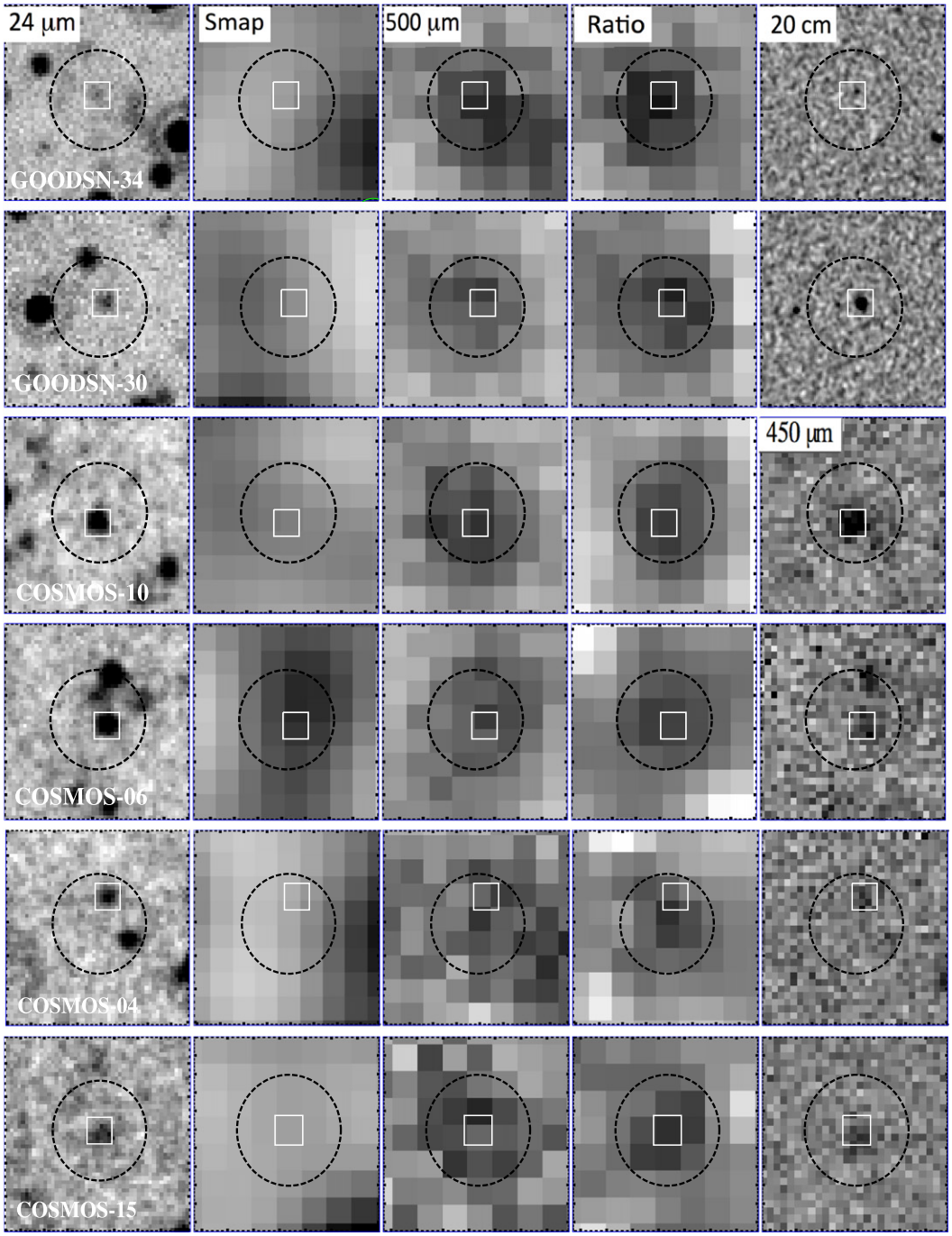

Figure 9. Examples of "clean" sources detected in the ratio map. Image cutouts from left to right: the 24 μm, the "beam-smoothed" 24 μm (Section 2.2), the 500 μm, the $S_{500\ \mu m}/S_{24\ \mu m}$ ratio, and the radio 1.4 GHz map. In the lower four panels, we also show example sources in the COSMOS field where the SCUBA2-450 μm image is available for direct comparison (last column).

$z_{\rm spec} = 0.8$, and another two have photometric redshifts of $z_{\rm phot} = 1.61$ and $z_{\rm phot} = 0.60$, respectively. As we are mainly interested in the $z \gtrsim 2$ population of *Herschel* 500 μm sources, we exclude them from the following analysis.

For the remaining 35 sources (including four with more than one candidate counterpart), we provide the coordinates, photometric redshift (or spectroscopic redshift if available), mid-IR to radio flux densities, and the derived far-IR 8–1000 μm luminosity and dust temperature in Table 2. Their distribution in the rest-frame $(U - V)$ versus $(V - J)$ color digram is shown in Figure 10. It can be seen that most objects (32/36, ~90%) can be classified as dusty starburst galaxies (Whitaker et al. 2012), consistent with their high dust attenuation of $A_V \sim 2$ mag inferred from the UV-MIR SED fittings. We note that there are two sources (GH500.10a and GH500.25) falling into the regime for quiescent galaxies. However, both are significantly detected in the *Herschel* - PACS 160 μm (~4–6 mJy, Table 2), suggesting the presence of ongoing obscured star formation. The possibility of wrong identifications seems low, as both are relatively isolated and well detected in the 24 μm and radio. Since the two sources have relatively high photometric redshift at $z \gtrsim 4$ where the

Table 1
Identifications of *Herschel* 500 μm Sources in the GOODS-North Field

Name	R.A. (J2000)	decl. (J2000)	$z_{\rm phot}$	$z_{\rm spec}$	Dist. (arcsec)	$P^C_{24\,\mu{\rm m}}$	$P^C_{3.6\,\mu{\rm m}}$	$P_{1.4\,{\rm GHz}}$	$S_{3.6\,\mu{\rm m}}$ (μJy)	$S_{5.8\,\mu{\rm m}}$ (μJy)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
GH500.01[a]	12:35:39.50	62:13:11.9	4.24	...	2.9	...	0.0098	...	6.6 ± 0.3	19.9 ± 1.4
GH500.02	12:35:39.54	62:12:43.8	3.03	...	0.9	0.057	0.047	...	7.5 ± 0.2	14.4 ± 1.1
GH500.03	12:35:53.12	62:10:37.2	1.53	1.60	9.7	0.081	0.079	0.13	91.5 ± 0.3	100.8 ± 1.0
GH500.04	12:35:54.28	62:13:43.5	3.2	...	6.7	0.018	0.022	0.109	9.6 ± 0.1	18.0 ± 0.6
	12:35:53.24	62:13:37.5	2.8	...	3.4	0.093	0.069	0.068	6.6 ± 0.2	9.5 ± 0.5
GH500.05	12:36:05.68	62:08:38.4	3.21	...	8.6	0.053	0.085	...	4.5 ± 0.1	6.3 ± 0.5
GH500.06[b]	12:36:08.81	62:11:43.7	1.31	1.33	7.1	0.173	0.167	0.056	79.7 ± 0.4	64.2 ± 0.6
GH500.07[b]	12:36:08.60	62:14:35.37	2.15	...	6.1	0.228	0.212	0.086	5.1 ± 0.1	10.6 ± 0.4
GH500.08	12:36:08.84	62:08:04.1	3.56	...	2.4	0.0087	0.0088	...	10.0 ± 0.2	23.8 ± 0.6
	12:36:10.59	62:08:10.7	4.54	...	11.8	0.0033	0.033	0.04	2.5 ± 0.2	2.4 ± 0.6
GH500.09	12:36:16.11	62:15:13.7	2.70	2.57	1.4	0.028	0.118	0.05	10.5 ± 0.2	23.4 ± 0.6
GH500.10	12:36:16.51	62:07:03.0	3.07	...	6.7	0.0076	0.014	0.13	24.6 ± 0.6	44.1 ± 0.9
	12:36:15.61	62:06:43.1	4.09	...	14.1	0.016	0.023	0.07	2.9 ± 0.2	5.2 ± 0.6
GH500.11	12:36:18.49	62:09:03.3	1.61	...	0.15	0.061	0.061	...	57.7 ± 0.2	56.8 ± 0.7
GH500.12	12:36:18.39	62:15:50.6	1.90	2.0	2.3	0.079	0.056	0.011	16.2 ± 0.2	31.3 ± 0.6
GH500.13	12:36:20.98	62:17:09.8	1.49	1.99	2.5	0.067	0.151	0.055	14.9 ± 0.1	23.5 ± 0.3
	12:36:21.27	62:17:08.2	2.32	...	0.91	0.205	0.075	0.010	13.8 ± 0.1	24.6 ± 0.4
GH500.14	12:36:22.66	62:16:29.7	1.92	2.47	7.2	0.024	0.266	0.03	19.9 ± 0.3	35.7 ± 0.9
	12:36:22.07	62:16:15.9	4.63	...	7.4	0.0033	0.005	0.12	3.4 ± 0.2	11.2 ± 0.6
GH500.15[a]	12:36:27.19	62:06:05.4	4.4	...	7.2	...	0.014	0.11	1.9 ± 0.2	5.1 ± 0.7
GH500.16[b]	12:36:31.29	62:09:58.0	2.02	...	1.7	0.11	0.219	0.014	13.2 ± 0.2	20.9 ± 0.5
GH500.17[b]	12:36:32.65	62:06:21.1	2.91	...	9.1	...	0.226	0.13	3.6 ± 0.1	9.5 ± 0.4
	12:36:34.53	62:06:13.1	4.31	...	11.3	...	0.055	...	2.1 ± 0.2	2.5 ± 0.6
GH500.18	12:36:33.40	62:14:08.5	...	4.04	1.7	0.022	0.042	0.089	0.5 ± 0.1	1.4 ± 0.5
GH500.19	12:36:44.03	62:19:38.4	3.41	...	6.9	0.035	0.029	0.123	2.3 ± 0.09	6.9 ± 1.1
GH500.20	12:36:46.07	62:14:48.9	2.67	...	2.6	0.062	0.121	0.02	7.9 ± 0.2	14.1 ± 0.5
GH500.21	12:37:00.26	62:09:09.9	2.13	...	12.5	0.077	0.075	0.004	22.1 ± 0.2	40.8 ± 0.7
GH500.22	12:37:00.96	62:11:45.6	2.01	1.76	3.92	0.036	0.058	0.022	61.5 ± 0.2	79.5 ± 0.5
GH500.23	12:37:01.52	62:20:24.8	3.00	...	0.39	0.060	0.038	0.049	8.7 ± 0.2	17.4 ± 0.6
	12:37:02.12	62:20:24.9	5.92	...	3.81	...	0.001	...	1.1 ± 0.2	1.2 ± 0.4
GH500.24[a]	12:37:01.10	62:21:09.6	0.79	0.8	7.9	0.87	0.316	0.005	239.8 ± 0.6	118.0 ± 1.2
GH500.25	12:37:02.55	62:13:02.3	4.44	...	2.2	0.003	0.007	0.12	3.9 ± 0.08	8.3 ± 0.3
	12:37:02.90	62:13:07.0	4.49	...	7.5	...	0.039	...	1.5 ± 0.08	2.4 ± 0.3
GH500.26	12:37:07.21	62:14:08.2	2.45	2.48	1.1	0.048	0.059	0.12	17.6 ± 0.1	28.4 ± 0.6
	12:37:07.58	62:14:09.5	5.1	2.48	3.9	0.149	0.018	0.10	24.1 ± 0.3	51.5 ± 0.8
GH500.27	12:37:11.81	62:22:12.3	4.16	4.04	8.6	0.0087	0.015	0.046	8.3 ± 0.2	19.7 ± 1.0
	12:37:09.73	62:22:02.3	4.10	4.04	9.1	0.022	0.014	0.139	7.5 ± 0.3	16.8 ± 0.9
	12:37:08.73	62:22:02.5	4.39	...	13.4	0.015	0.028	...	4.4 ± 0.2	3.5 ± 0.4
	12:37:08.78	62:22:01.8	4.19	4.04	14.7	0.021	0.025	0.014	4.0 ± 0.09	4.2 ± 0.4
	12:37:09.42	62:22:14.5	1.65	1.63	11.4	0.187	0.049	0.127	57.2 ± 0.2	68.2 ± 0.9
GH500.28	12:37:11.37	62:13:31.0	1.79	1.99	1.6	0.043	0.056	0.02	54.8 ± 0.1	72.9 ± 0.6
GH500.29	12:37:12.05	62:12:12.0	2.86	2.91	7.3	0.076	0.046	0.10	9.2 ± 0.09	17.9 ± 0.5
	12:37:14.29	62:12:08.5	2.2	3.15	9.3	0.013	0.063	...	17.5 ± 0.1	26.3 ± 0.5
GH500.30	12:37:13.86	62:18:26.3	4.15	...	2.1	0.010	0.011	0.003	4.4 ± 0.09	10.6 ± 0.6
GH500.31	12:37:28.08	62:19:20.2	4.32	...	2.3	0.0087	0.0033	0.11	7.7 ± 0.4	16.2 ± 1.6
	12:37:29.24	62:19:28.6	4.65	...	12.89	...	0.025	...	2.9 ± 0.2	4.2 ± 0.3
GH500.32	12:37:28.11	62:14:21.82	3.32	...	5.06	0.047	0.071	...	1.8 ± 0.08	4.7 ± 0.6
GH500.33	12:37:30.78	62:12:58.9	2.41	...	0.9	0.071	0.102	0.017	17.1 ± 0.9	17.7 ± 2.3
GH500.34	12:37:38.30	62:17:36.2	3.13	3.19	1.5	0.059	0.193	0.13	4.3 ± 0.2	8.6 ± 0.8
	12:37:39.37	62:17:35.0	4.33	...	7.69	...	0.023	...	4.1 ± 0.6	3.6 ± 0.9
GH500.35	12:38:00.27	62:16:21.2	2.4	...	4.6	0.085	0.045	...	14.4 ± 0.3	28.2 ± 0.7
	12:38:00.84	62:16:11.8	2.19	...	6.5	0.094	0.051	...	40.0 ± 0.2	48.3 ± 0.8
	12:37:58.82	62:16:11.7	1.98	...	10.6	0.096	0.119	...	10.2 ± 0.1	19.7 ± 0.5
GH500.36[b]	12:38:12.54	62:14:54.3	0.60	...	3.8	0.307	0.638	0.059	126.5 ± 0.6	116.4 ± 2.7

Notes. Column (1): Name; (2) and (3): coordinates for the 24 μm counterparts; (4): Photometric redshift; (5): Spectroscopic redshift from Barger et al. (2008), Daddi et al. (2009a, 2009b), (6): Distance of the 24 μm counterparts from the center of source in the ratio map; (7); Probability P for random associations of 24 μm counterparts is calculated as $P = 1 - \exp[-\pi n\theta^2]$, where n is the surface density of sources brighter than the predicted 500 μm flux, $S^{\rm Pred}_{500\,\mu m}$, and with redshifts above z for a given counterpart of interest (see Section 2.3.2). (8): P-values for candidate counterparts from the IRAC 3.6 μm catalog, which are calculated similar to the 24 μm catalog, but using the main-sequence relation to predict 500 μm flux from stellar masses. (9): P-values for the radio 1.4 GHz catalog, but are calculated using $n_{(>S)}$ (the surface density of sources brighter than radio flux $S_{1.4\,{\rm GHz}}$). Traditionally, a radio counterpart is considered to be robust with $P < 0.05$.
[a] Counterpart(s) identified from the IRAC 3.6 μm.
[b] Counterpart(s) identified based on the radio 1.4 GHz detection.

Table 2
Observed Photometry for *Herschel*-selected Dust Obscured Star-forming Galaxies

Name	R.A.	decl.	z	24 μm (μJy)	100 μm (mJy)	160 μm (mJy)	250 μm (mJy)	350 μm (mJy)	500 μm (mJy)	850 μm (mJy)	1.1 mm (mJy)	1.4 GHz (μJy)	$L_{\rm FIR}$ $10^{12}L_{\odot}$	$T_{\rm dust}$ (K)
GH500.1	12:35:39.49	62:13:11.7	4.24	...	...	...	5.6 ± 2.3	9.3 ± 2.1	14.9 ± 3.9	...	...	...	6.1	27.8 ± 8.8
GH500.2	12:35:39.60	62:12:44.6	3.03	54.2 ± 6.8	...	4.2 ± 1.0	7.8 ± 2.1	10.7 ± 2.7	10.6 ± 4.1	...	2.0	...	3.5	31.5 ± 4.5
GH500.4a	12:35:54.28	62:13:43.5	3.2	149.9 ± 8.1	1.1 ± 0.3	2.8 ± 0.9	12.7 ± 2.2	17.3 ± 2.3	12.6 ± 3.9	<13.8	3.1 ± 1.1	28.9 ± 12.1	5.4	30.7 ± 4.8
GH500.4b	12:35:53.28	62:13:37.6	2.81	32.5 ± 6.6	...	3.3 ± 0.9	20.6 ± 2.2	23.2 ± 2.3	15.6 ± 3.9	<6.4	...	39.8 ± 4.5	6.5	34.5 ± 8.1
GH500.5	12:36:05.68	62:08:38.5	3.2	37.47 ± 6.6	...	...	6.7 ± 2.1	10.2 ± 2.6	11.0 ± 3.9	<10.5	...	...	3.3	28.4 ± 9.9
GH500.7[a]	12:36:08.57	62:14:35.3	2.15	55.1 ± 6.7	...	3.1 ± 0.9	10.1 ± 2.2	11.9 ± 2.4	9.7 ± 3.4[a,b,c]	4.9 ± 2.3	2.4 ± 1.2	22.1 ± 7.1	1.51	23.8 ± 2.9
GH500.8	12:36:08.84	62:08:04.0	3.56	151.2 ± 8.2	...	3.5 ± 1.0	9.3 ± 2.1	10.2 ± 2.7	11.0 ± 4.0	<9.3	...	...	6.1	35.2 ± 7.3
GH500.9	12:36:16.116	62:15:13.7	2.57[c]	322.5 ± 10.	5.7 ± 4.1	12.0 ± 4.1	28.4 ± 2.1	29.0 ± 2.4	16.2 ± 4.2	4.9 ± 0.7	2.9 ± 1.1	35.8 ± 4.9	8.4	37.7 ± 1.6
GH500.10a	12:36:15.7	62:06:43.1	4.09	45.3 ± 6.7	...	4.2 ± 1.0	9.3 ± 2.0	9.8 ± 2.8	11.5 ± 4.1	<8.7	...	43.5 ± 5.8	9.0	41.0 ± 8.3
GH500.10b	12:36:16.51	62:07:03.0	3.07	302.4 ± 10	3.1 ± 4.1	...	21.6 ± 2.0	27.5 ± 2.8	17.2 ± 4.1	<12	...	26 ± 4.7	8.5	33.8 ± 4.9
GH500.12	12:36:18.31	62:15:50.5	2.0[c]	314.5 ± 10.	4.3 ± 0.4	23.4 ± 1.1	42.8 ± 2.3	47.1 ± 2.1	23.6 ± 4.0	7.5 ± 0.9	1.9 ± 1.1	164.4 ± 6.9	4.7	29.8 ± 1.2
GH500.13	12:36:20.96	62:17:10.0	1.99[c]	361.3 ± 10	3.98 ± 0.4	15.3 ± 1.0	21.6 ± 2.0	26.4 ± 2.6	13.7 ± 4.1	3.5 ± 0.7	...	44.9 ± 2.7	3.6	31.0 ± 1.7
GH500.14a	12:36:22.66	62:16:29.70	2.47[c]	412.8 ± 11	3.9 ± 0.4	12.3 ± 0.9	23.7 ± 2.0	21.2 ± 2.5	13.9 ± 4.2	<3.4	...	91.6 ± 9.9	5.8	33.3 ± 2.5
GH500.14b	12:36:22.10	62:16:15.9	4.63	50.4 ± 6.8	...	...	8.3 ± 2.0	16.0 ± 2.5	14.3 ± 4.2	9.3 ± 1.4	...	26.1 ± 6.1	11.2	33.7 ± 3.5
GH500.15	12:36:27.23	62:06:05.7	4.4[b]	...	...	...	7.3 ± 2.0	14.5 ± 2.8	13.7 ± 4.2	9.5 ± 2.3	4.3 ± 1.1	34.3 ± 5.9	10.0	35.2 ± 5.2
GH500.16	12:36:31.23	62:09:57.7	2.02	209.4 ± 8.3	4.9 ± 3.9	16.1 ± 1.0	25.6 ± 2.0	23.5 ± 2.7	14.8 ± 4.2	3.7 ± 0.8	...	141 ± 6.1	3.7	32.6 ± 1.6
GH500.17	12:36:32.65	62:06:21.1	2.91	...	...	...	13.6 ± 2.0	19.5 ± 2.7	17.5 ± 3.2[a,b,c]	<25.5	3.1 ± 1.1	23.8 ± 4.3	5.1	30.5 ± 5.4
GH500.18	12:36:33.42	62:14:08.7	4.04[c]	30.4 ± 6.6	...	...	8.2 ± 2.2	14.3 ± 2.7	13.2 ± 4.2	10.5 ± 1.6	5.4 ± 0.9	34.0 ± 4.0	8.1	29.1 ± 2.8
GH500.19	12:36:44.02	62:19:38.4	3.41	34.9 ± 6.5	...	...	11.9 ± 2.0	14.7 ± 2.8	13.1 ± 4.1	5.66 ± 0.97	2.3 ± 1.1	25.4 ± 2.9	6.5	31.0 ± 3.3
GH500.20	12:36:46.04	62:14:48.6	2.67	129.8 ± 8.4	3.9 ± 0.4	16.6 ± 1.0	25.9 ± 2.0	24.9 ± 2.6	21.1 ± 4.1	5.4 ± 0.5	...	109.6 ± 5.3	7.6	37.4 ± 1.5
GH500.21	12:37:00.26	62:09:09.9	2.13	246.9 ± 9.1	2.0 ± 0.4	6.6 ± 0.9	16.8 ± 1.9	16.6 ± 2.8	13.0 ± 4.2	3.5 ± 0.8	...	337.5 ± 4.3	2.7	34.4 ± 2.2
GH500.23	12:37:01.57	62:20:24.9	3.0	55.7 ± 6.6	...	3.1 ± 0.9	10.4 ± 2.0	13.8 ± 2.8	7.7 ± 3.9[a,b,c]	<8.1	...	48.1 ± 5.1	4.2	29.8 ± 6.3
GH500.25	12:37:02.55	62:13:02.3	4.44	79.0 ± 7.6	0.9 ± 0.3	6.0 ± 0.9	11.9 ± 2.0	11.6 ± 2.7	5.2 ± 2.7	3.2 ± 0.6	...	25.3 ± 4.0	10.7	50.8 ± 5.8
GH500.26	12:37:07.18	62:14:08.2	2.49[c]	248.9 ± 9.1	1.6 ± 0.4	8.6 ± 0.9	23.9 ± 2.0	29.0 ± 2.4	24.7 ± 4.1	7.6 ± 0.6	...	28.3 ± 4.2	5.2	29.7 ± 1.4
GH500.27a	12:37:11.89	62:22:11.8	4.05[c]	84.3 ± 7.8	...	6.1 ± 0.9	17.3 ± 1.9	30.1 ± 2.3	46.6 ± 4.1	20.3 ± 2.1	10.7 ± 9.4	79.3 ± 13.	18.0	32.3 ± 1.6
GH500.27b	12:37:09.73	62:22:02.5	4.055[c]	30.2 ± 6.6	...	2.9 ± 0.9	12.4 ± 2.0	15.2 ± 2.4	18.2 ± 4.0	...	...	19.9 ± 4.0	8.5	33.9 ± 5.5
GH500.27c	12:37:08.77	62:22:01.7	4.055[c]	38.8 ± 6.5	...	...	14.5 ± 2.0	18.0 ± 2.4	18.2 ± 4.0	9.9 ± 2.3	...	172.9 ± 9.0	10.2	36.5 ± 5.5
GH500.28	12:37:11.37	62:13:31.0	1.99[c]	748.0±15.9	11.3 ± 0.6	34.9 ± 1.5	54.1 ± 2.8	49.3 ± 2.1	29.2 ± 3.9	7.7 ± 0.6	4.0 ± 1.0	178.9±12.1	8.7	33.1 ± 0.9
GH500.29	12:37:12.04	62:12:11.7	2.91[c]	51.5 ± 6.7	1.0 ± 0.3	3.6 ± 1.0	11.8 ± 2.0	10.0 ± 2.6	6.5 ± 4.0[a,b,c]	4.6 ± 0.7	2.0 ± 1.0	31.7 ± 4.3	3.6	31.9 ± 3.3
GH500.30	12:37:13.86	62:18:26.2	4.15	63.4 ± 7.0	1.08 ± 0.4	3.5 ± 1.0	10.6 ± 2.0	12.5 ± 2.8	7.3 ± 3.2[a,b,c]	4.5 ± 0.8	2.4 ± 1.1	627 ± 8	7.8	41.4 ± 5.1
GH500.31	12:37:28.09	62:19:20.1	4.32	44.1 ± 6.8	...	2.2 ± 0.8	12.3 ± 2.0	13.8 ± 2.9	13.4 ± 4.2	<6.9	...	33.6 ± 5.4	10.1	36.7 ± 7.8
GH500.32	12:37:28.11	62:14:21.82	3.32	34.3 ± 6.6	...	2.0 ± 0.8	21.1 ± 2.0	21.7 ± 2.7	20.7 ± 4.0	<8.4	...	...	6.9	29.5 ± 3.8
GH500.33[a]	12:37:30.78	62:12:58.7	2.41	184 ± 7.7	4.9 ± 0.4	23.4 ± 1.1	52.3 ± 2.3	55.0 ± 2.1	44.8 ± 3.8	12.21 ± 0.77	4.1 ± 1.1	114 ± 7	9.3	28.3 ± 0.9
GH500.34	12:37:38.15	62:17:37.1	3.19[c]	39.3 ± 6.6	...	...	9.6 ± 2.1	12.4 ± 2.8	7.3 ± 3.2[a,b,c]	4.7 ± 2.3	3.4 ± 1.0	26.0 ± 5.9	4.0	28.6 ± 5.8
GH500.35	12:38:00.27	62:16:21.2	2.4	140.5 ± 8.3	2.1 ± 0.4	6.0 ± 1.0	11.2 ± 2.1	12.7 ± 2.8	11.6 ± 3.6	<10.5	...	...	2.9	32.4 ± 4.9

Notes.
[a] IR SED fittings are very poor, thus the inferred far-IR luminosity and dust temperature are highly uncertain. We did not consider the two sources in our analysis.
[b] Redshift from the far-IR SED fit. Although it is likely at high-z, we did not include this source in our analysis.
[c] Spectroscopic redshift from Barger et al. (2008), Daddi et al. (2009a, 2009b).

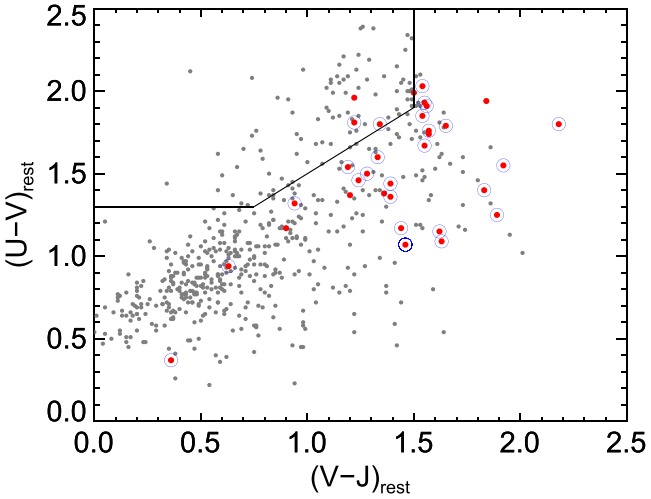

Figure 10. Rest-frame $(U - V)$ vs. $(V - J)$ color diagram for all IRAC sources within the search radii of 15″ (gray dots). Red solid circles represent secure *Spitzer* (radio)-identified 500 μm counterparts (Table 2 in the text), and those satisfying IRAC color selection for SMGs are highlighted in blue (Biggs et al. 2011).

UVJ selection is still poorly explored (e.g., Straatman et al. 2014), whether they can be considered quenched remains further investigations. Indeed, their specific SFRs ($8–40 \times 10^{-9}\,\mathrm{yr}^{-1}$) are already $>2\times$ higher than the sSFR $= 5 \times 10^{-9}\,\mathrm{yr}^{-1}$ of similarly massive star-forming galaxies at $z > 3$ (Schreiber et al. 2015).

The *Herschel*/SPIRE fluxes are from the GOODS-*Herschel* catalog (Elbaz et al. 2011), or derived from PSF fitting using GALFIT (Peng et al. 2002) for cases where the flux could be wrongly attributed to nearby 24 μm sources due to blending, or is not measured in the catalog. A description of the *Herschel*/SPIRE flux measurements is presented in Appendix B. Note that due to the poor resolution of the *Herschel* data, measuring reliable SPIRE flux densities for individual galaxies, especially those at 500 μm, is a challenging task. The de-blending approach of using 24 μm data of better resolution would be still affected by the degeneracy between the positions of the sources and flux uncertainties attributed by faint sources. Given the case of heavily blending, the SPIRE flux measurements should be treated with some caution. The thumbnail images for each identified source are presented in Figure 20 in Appendix C. Among 14 sources that have two or more candidate radio/24 μm counterparts, four (GH500.13, GH500.26, GH500.28, GH500.35) have counterparts separated by only ∼5″ or less. Three of these cases are treated as single systems because the counterparts are consistent with being at the same redshift and are assumed to be interacting galaxies. These three sources are so closely separated that deblending the *Herschel*/SPIRE flux is impossible, hence we report the combined fluxes of the two counterparts in Table 2. For GH500.35, we cannot assess whether the three counterpart galaxies are interacting systems or spatially clustered, and we assume that the source with the brightest 24 μm flux is the only counterpart, which should be treated with caution.

2.3.4. Contaminants

We have visually inspected all sources which pass our S/N cut in order to search for noise artifacts and blends of bright sources. To check for false detections caused by map noise artifacts and/or from sources well below the nominal "flux" detection limit, we run our search procedure on the $S_{500\,\mu\mathrm{m}}/S_{24\,\mu\mathrm{m}}$ map by removing all individually detected 500 μm sources in the catalog used in Elbaz et al. (2011). The P(D) plot of the $S_{500\,\mu\mathrm{m}}/S_{24\,\mu\mathrm{m}}$ ratio map omitting all 500 μm detections in the existing catalog is shown in Figure 3 (red curve). It can be seen that few pixels have values above our S/N cut of $S_{500\,\mu\mathrm{m}}/S_{24\,\mu\mathrm{m}} = 20.4$, suggesting that the contamination from noise artifacts is negligible. Figure 4 (lower panel) shows the number of spurious sources as a function of detection threshold. As one would expect, the number of false detections depends strongly on the chosen S/N threshold. At a 2σ cut, the result suggests that two detections will be spurious in our ratio map, corresponding to a spurious source rate of ∼6% (2/36). Note that while increasing the S/N requirement would decrease the number of false sources, it would also remove a fraction of real but slightly fainter sources from the map. Therefore, we adopt 2σ as the detection S/N cut for sources in the ratio map because it provides a good compromise between the catalog size and source reliability.

The above estimate, however, provides only an lower limit to the number of contaminants. As shown in Figure 1, the diversity in IR SED of galaxies means that lower redshift sources would have probabilities of being detected in the ratio map if they have similarly high $S_{500\,\mu\mathrm{m}}/S_{24\,\mu\mathrm{m}}$ ratios, contaminating the selection of $z \gtrsim 2$ sources. We account for this effect statistically by performing simulations to measure the recovery rate for injected low-z sources of varying flux density, using the similar procedure described in Section 2.3.1. For each injected source we assumed an IR SED randomly selected from the CE01 libraries to predict its 24 μm flux density at a given redshift. Note that we did not assume any peculiar IR SED for injected sources to avoid bias in detections. Although very "cold" galaxies have more chance of being detected, they are expected rare given the small area covered by GOODS-North. Since lower redshift sources are not expected to be detected in our data except those with abnormal high $S_{500\,\mu\mathrm{m}}/S_{24\,\mu\mathrm{m}}$ ratios, any source recovered in the map would be considered as a contaminant. The average recovery rate for low-z contaminants as a function of flux density is shown in Figure 4 (lower right panel). It can be seen that sources below $z \sim 1.2$ have a low probability to be detected, but the detection rate increases with redshift above $z \sim 1.4$ due to the increased fraction of sources with higher $S_{500\,\mu\mathrm{m}}/S_{24\,\mu\mathrm{m}}$ ratios. The low-z contaminants become more significant at higher flux density, making them non-negligible effect in selecting $z \gtrsim 2$ galaxies. In fact, such low-z sources have been identified in the ratio map. As we mentioned above, we found 6 out of 36 sources at $z < 2$ (∼17%) that could be contaminants in our catalog, which is consistent with the simulation result of ∼20% detection rate at ∼15 mJy, a median flux for catalog sources.

3. BASIC OBSERVED SOURCE PROPERTIES

3.1. Redshift Distribution

In this section, we present the redshift distribution for our $z \gtrsim 2$ *Herschel* 500 μm sources. To derive the redshift distribution, we take spectroscopic redshifts if available, based on optical and/or CO spectroscopic observations (Barger et al. 2008; Daddi et al. 2009a, 2009b), and otherwise photometric redshifts based on a K_s-selected multi-wavelength photometric catalog (Pannella et al. 2015). Spectroscopic redshifts are

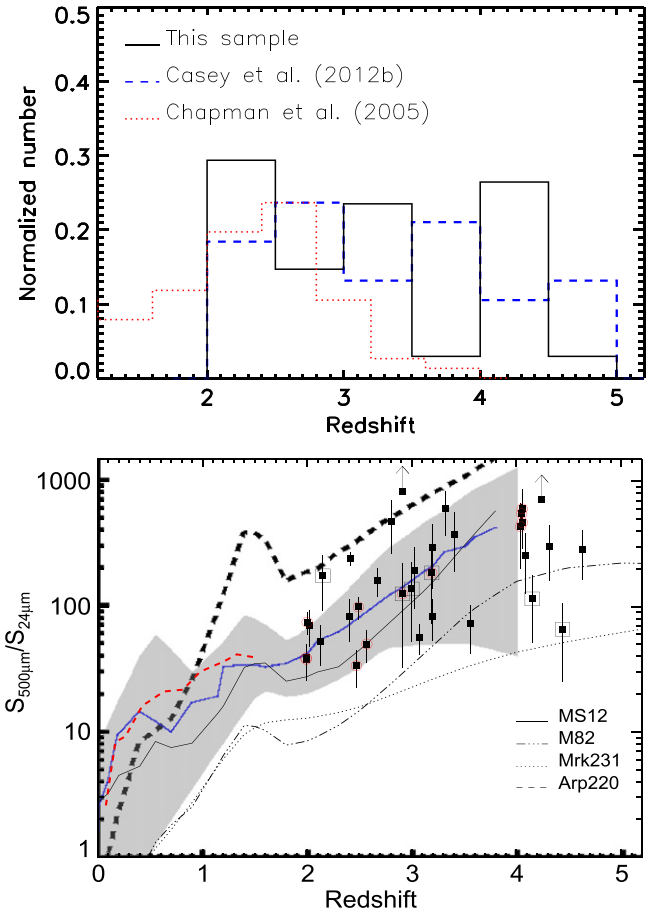

Figure 11. Upper: the redshift distribution of the $S_{500\,\mu m}/S_{24\,\mu m}$-selected sources in the GOODS-North field (black line, normalized by the total). The red dotted histogram is the redshift distribution of SCUBA 850 μm sources from Chapman et al. (2005). For comparison, the blue dashed histogram shows the distribution of spectroscopic redshifts for $2 < z < 5$ SPIRE sources from Casey et al. (2012b). The excess number of objects at $4.0 < z < 4.5$ may be due to cosmic variance and the presence of a high-redshift proto-cluster (Daddi et al. 2009b). Lower: $S_{500\,\mu m}/S_{24\,\mu m}$ flux ratio as a function of redshift. Black curves are model SEDs of Arp220 (dashed), a main-sequence galaxy template (solid) from Magdis et al. (2012), M82 (dotted–dashed) and Mrk231 (dotted). The shaded region represents the range of the CE01 models, and the blue solid is its mean trend. The red dashed line is the median trend of $z < 2$ *Herschel*/SPIRE sources from Casey et al. (2012a). Filled squares are $z \gtrsim 2$ sources identified from the ratio map, with spectroscopically confirmed ones highlighted by red circles. Large open squares denote galaxies whose 500 μm emission is less than 3σ.

available for 12 out of the 35 *Herschel* sources (33%) in the sample. Note that the photo-z for one source (GH500.15) is not available, as it is not detected in the K_s band. Although SED fittings in the far-IR suggests it likely being at $z \gtrsim 4$ (Section 4), we exclude the source in the following analysis due to the relatively poor constraint on the photo-z.

The redshift distribution for the 34 $z \gtrsim 2$ *Herschel* sources identified from the ratio map is shown in the upper panel of Figure 11. For comparison, we also plot the redshift distribution for SMGs (red) from Chapman et al. (2005). Each distribution is normalized to the total number of sources in the sample. The redshift distribution derived by Chapman et al. (2005) is based on a sample of SMGs drawn from various SCUBA/850 μm surveys with radio counterparts targeted for spectroscopic redshift follow-ups. It can be seen that the distribution of our $S_{500\,\mu m}/S_{24\,\mu m}$-selected sample is in sharp contrast to that SCUBA/850 μm SMGs. The median redshift for our sample is $z_{\rm med} = 3.07$, with a significant fraction of sources (10/34, 29%) at $z \gtrsim 4$. The radio-identified SCUBA SMGs have a lower median redshift of $z_{\rm med} = 2.6$ (considering only sources at $z \gtrsim 2$), for which only 18% of the population is at $z > 3$. Part of this difference could be due to the identification bias in the sample of Chapman et al. The redshift distribution of SCUBA SMGs includes only sources with robust radio detections; the highest-redshift sources may fall below the depth of radio data and are likely to be missed. On the other hand, the sample considered here is confined to 500 μm sources significantly detected in the ratio map which efficiently selects sources at the highest redshifts. Lower-redshift submm sources are down-weighted in the ratio map in an effort to obtain a homogenous sample of galaxies at $z \gtrsim 2$ as possible.

A recently published spectroscopic survey of *Herschel*/SPIRE sources with optical counterparts determined from high-resolution radio/24 μm imaging found a population of ULIRGs at $2 < z < 5$ (Casey et al. 2012b) and a median redshift of $z_{\rm med} = 3.38$ (blue dashed line in Figure 11, upper). It should be noted that this population is only a tail of spectroscopically confirmed *Herschel* sources in their survey, with the majority of them (~95%) being at $z < 2$ (Casey et al. 2012a). Though the sample of $z > 2$ *Herschel* sources in Casey et al. (2012b) is likely biased toward the optically bright population for which spectroscopic identifications are made, the redshift distribution appears similar to that derived from our sample. A Kolmogorov–Smirnov test shows that the probability that both samples are drawn from the same underlying parent distribution is 0.43. Note that while the Casey et al. (2012b) redshift distribution is relatively flat from $z = 2$ to 5, our sample shows a tentative excess at $z \gtrsim 4$. This could be due to cosmic variance, since our sample probes a small volume in a single field, while the Casey et al. data cover several deep fields over a much larger ($\gtrsim$5 times) sky area.[19] In fact, Daddi et al. (2009b) have reported a proto-cluster structure at $z = 4.05$ in the GOODS-North field (see their Figure 13) which includes three CO-detected SMGs within ~25″ at the same redshift.

Using a $S_{500\,\mu m}/S_{24\,\mu m}$ ratio method, we find 10 *Herschel* 500 μm sources at $z \gtrsim 4$ in the GOODS-North field, corresponding to ~29% of the whole sample. Eight of these $z \gtrsim 4$ candidates have $S_{500\,\mu m} > 10$ mJy. This suggests significantly more ULIRGs at the highest-redshift end compared to other surveys (e.g., Smolčić et al. 2012a and references therein). This will enrich our understanding of obscured star formation in the early universe. Excluding four spectroscopically confirmed $z > 4$ sources at the same redshift which are likely associated with a protocluster at $z \sim 4.05$ (Daddi et al. 2009a, 2009b), we can place a lower limit on the surface density of $z > 4$ ULIRGs at $\gtrsim 6/0.05 = 120$ deg^{-2}. This value is more than an order of magnitude higher than what is expected in cosmological models (Baugh et al. 2005; Swinbank et al. 2008; Coppin et al. 2009; Béthermin et al. 2012). However, given the very small area covered by GOODS-N and lack of the spectroscopic confirmation for most candidates, these values should be taken with caution. A complete view of the properties of the $z > 4$ ULIRG population will require a detailed submm follow-up and

[19] Casey et al. (2012b) sample comes from a spectroscopic survey of bright SPIRE sources (>10–12 mJy) in four HerMES legacy fields, including the Lockman Hole, COSMOS, ELAIS-N1, and ECDFS (see their Table 1).

robustly determined redshifts. Notably, recent ALMA follow-up of a bright sample of lensed SMGs discovered with the SPT yielded at least 10 spectroscopically confirmed sources at $z > 4$ (Vieira et al. 2013; Weiß et al. 2013), challenging the current models for galaxy formation and evolution.

In Figure 11 (lower panel), we show the measured $S_{500\,\mu m}/S_{24\,\mu m}$ ratio (Table 2) for each of our $z \gtrsim 2$ sources as a function of redshift, along with the various model SED predictions. The $S_{500\,\mu m}/S_{24\,\mu m}$ ratios of our galaxies are all over 30, and are in good agreement with the dusty SEDs allowed by current models, albeit with a large scatter. The large $S_{500\,\mu m}/S_{24\,\mu m}$ ratios therefore support the expectation that most sources, if not all, in our sample are indeed at $z \gtrsim 2$. In the next section, we will investigate the far-IR properties of individual galaxies.

3.2. Infrared Luminosities and Dust Properties

To derive the far-IR emission and SED characteristics of the sample, we fit the photometry with a library of galaxy templates including CE01 and the well-studied SEDs of the local starburst M82, the ULIRG Arp220, and the dusty active galactic nucleus (AGN) Mrk231. The SEDs for the latter three local galaxies are taken from the SWIRE template library (Polletta et al. 2007). Our library consists of a total of 108 templates with characteristic dust temperatures (T_d) in the range $T_d \sim$ 20–60 K. The SED fitting results are shown in Figure 21 in the appendix, and their infrared luminosities and dust temperatures are given in Table 1.

We compute the infrared luminosity of each source by integrating the best-fit SED between rest-frame 8 and 1000 μm. The luminosities for this sample range from 1.5×10^{12}–$1.8 \times 10^{13}\,L_\odot$, implying infrared SFRs of ~260–3100 $M_\odot\,\mathrm{yr}^{-1}$ with a Salpeter IMF. Note that the IR luminosities can be determined to less than 50% accuracy when excluding SPIRE 350 and 500 μm measurements for 60% of sources where the data at longer wavelengths (>500 μm) are available. This suggests that they are among the most extreme star-forming galaxies seen in the early universe (e.g., Casey et al. 2014 and references therein). Figure 12 shows the infrared luminosity against redshift (filled red circles). The distribution of the Casey et al. (2012b) spectroscopically confirmed sample of $2 < z < 5$ SPIRE sources is shown for comparison (open circles). Our $S_{500\,\mu m}/S_{24\,\mu m}$-selected population and the Casey et al. (2012b) sample each probe different regions of parameter space. Though the redshift distribution is similar (Section 3.1), our sources are systematically lower in luminosity by ~0.1–0.5 dex at a fixed redshift. While the Casey et al. (2012b) $z > 2$ sample may not be representative of all high-z 500 μm sources due to their selection of brighter objects for spectroscopy follow-ups, our selection offers a unique probe of intrinsically fainter objects, which is important for a complete view of the properties of the $z \gtrsim 2$ *Herschel* population. Note that, as we mentioned above, it is important to understand the different volumes probed for a more meaningful comparison of the typical luminosities in the two samples.

The dust temperature is derived using the code of Casey (2012), which fits the far-IR photometry with an SED consisting of a modified blackbody spectrum and a power-law component in the mid-IR portion ($\lesssim$40 μm). The combination of a graybody and a mid-IR power law takes into account both galaxy-wide cold dust emission and smaller-scale warm dust emission (e.g., Kovács et al. 2006; Casey 2012). To reduce the number of free parameters, we fix the slope of the mid-IR power law to $\alpha = 2$ for sources which are not detected in PACS. In addition, the emissivity index of the blackbody spectrum on the Rayleigh–Jeans portion is fixed to $\beta = 1.5$ (a commonly chosen value of β in the literature, e.g., Chapman et al. 2005; Pope et al. 2005; Younger et al. 2009), since we have not enough data points at longer wavelengths ($\gtrsim$ 850 μm) to constrain β in a meaningful way. However, we note that there is very little change in the derived T_d by fixing β within the range of 1–2.

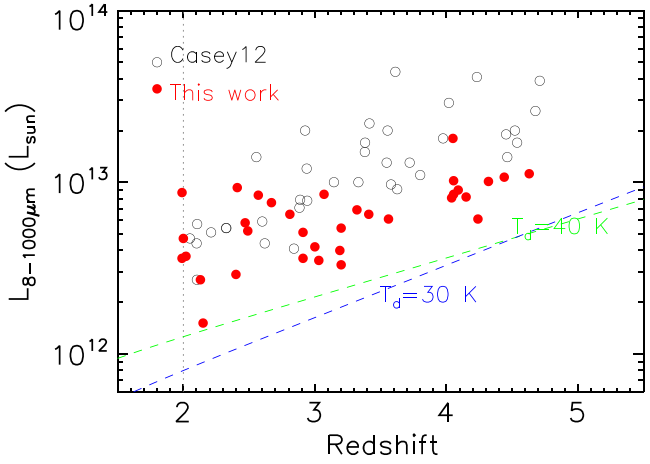

Figure 12. Infrared luminosity (8–1000 μm) against redshift for the $S_{500\,\mu m}/S_{24\,\mu m}$-selected $z \gtrsim 2$ ULIRGs (filled red circles) in the GOODS-North field. The *Herschel*/SPIRE-selected, spectroscopically confirmed sources at $2 < z < 5$ (open circles; Casey et al. 2012b) are overplotted for comparison. The detection boundaries as a function of redshift for $T_d = 30$ K (blue dashed line) and $T_d = 40$ K (green dashed line) are illustrated, assuming a flux limit of 9 mJy at 500 μm.

We present in Figure 13 the dust temperature against infrared luminosity for the $z \gtrsim 2$ sources in our sample which have a meaningful constraint on the dust temperature from the submm and/or mm measurements. We include in the comparison the measurements for SCUBA SMGs from Chapman et al. (2005) as well as the $z = 0$–1 SPIRE galaxies from Symeonidis et al. (2013), which appear to closely follow the relation that the dust temperature increases with infrared luminosity, as observed in the local IRAS 60 μm selected sample (e.g., Chapman et al. 2003). To examine whether our high-z sample has temperatures similar to those of local samples and at $z < 1$, we divided our sample into two bins of $L_{\rm IR}$ with roughly equal numbers of sources: $12.2 \leqslant \log10(L_{\rm IR}/L_\odot) \leqslant 12.8$ and $12.8 \leqslant \log10(L_{\rm IR}/L_\odot) \leqslant 13.4$. The derived median temperatures and luminosities are $T_d = 31.0 \pm 1.6$ and 36.5 ± 6.3 K for $L_{\rm IR} = 4.0$ and $8.7 \times 10^{12} L_\odot$, respectively. As Figure 13 shows, at a fixed luminosity the high-z galaxies in our sample tend to have cooler dust temperatures ($\delta T_d \sim$ 5K) compared to that expected from the $L_{\rm IR}$–T_d relation of $z = 0 - 1$ SPIRE galaxies (Symeonidis et al. 2013). This could be partially due to the selection in the 500 μm band, which is biased against the warmer sources, even at these high redshifts.

By comparing the dust properties between local galaxies ($z < 0.1$) observed with *AKARI* and galaxies in the redshift range $0.1 \lesssim z \lesssim 2.8$ selected by *Herschel*, Hwang et al. (2010) found modest changes in the $L_{\rm IR}$–T_d relation as a function of redshift. *Herschel*-selected galaxies appear to be 2–5 K colder than that of *AKARI*-selected local galaxies with similar

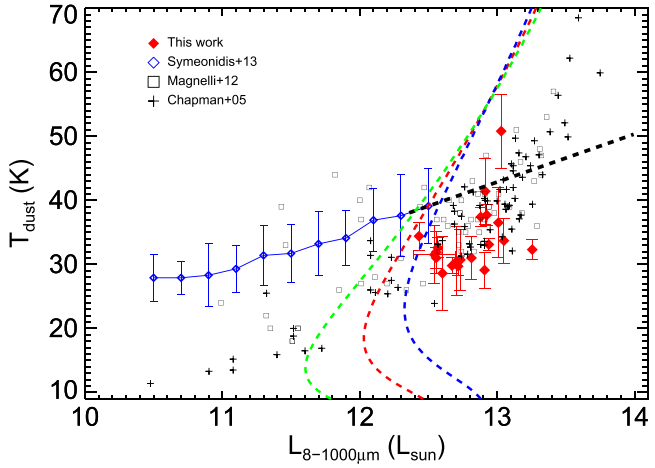

Figure 13. Effective dust temperature (T_d) vs. total far-IR luminosity ($L_{8-1000\,\mu m}$), for $z < 1$ galaxies (blue line and diamonds: SPIRE sample of Symeonidis et al. 2013; black crosses: SMG sample of Chapman et al. 2005), $z = 1$–3 SMGs+*Herschel* galaxies (open squares, Magnelli et al. 2012), and our $z \gtrsim 2$ $S_{500\,\mu m}/S_{24\,\mu m}$-selected sample (red filled diamonds). Colored dashed lines: transposition of our detection limit of $S_{500\,\mu m} = 9$ mJy for three redshifts, $z = 2$ (green), $z = 3$ (red), and $z = 4$ (blue). Black dashed line: extrapolation of the T_d–L_{IR} relation for $z < 1$ SPIRE sample of Symeonidis et al. (2013).

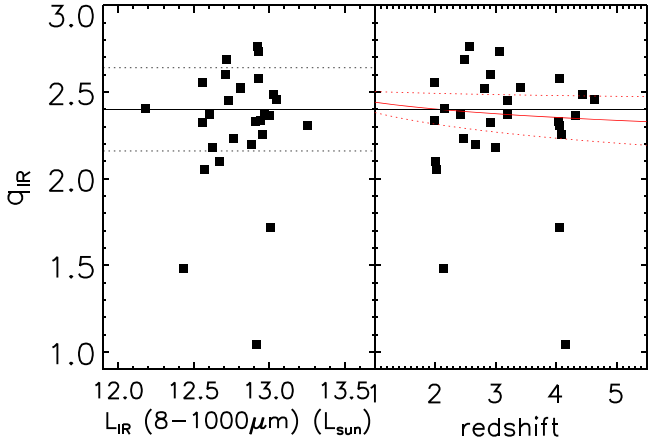

Figure 14. Radio-to-far-IR correlation coefficient (q_{IR}) as a function of far-IR luminosity (left) and redshift (right). The solid and dashed lines represent the median and 2σ values for SPIRE sources from Ivison et al. (2010b). Good agreement is seen with the Ivison et al. (2010b) results, and no significant evolution with redshift is observed (right). The red line shows the moderate redshift evolution of $q \propto (1+z)^{-0.04\pm0.03}$ obtained by Ivison et al. (2010b) for comparison.

luminosities. Such evolution of the L_{IR}–T_d relation with redshift has also been reported by the studies of Rex et al. (2010) and Chapin et al. (2011), and recently by Symeonidis et al. (2013) and Swinbank et al. (2014) with a much larger *Herschel* data set. The offset in the L_{IR}–T_d relation may imply that high-z ULIRGs have a weaker average interstellar radiation field than local analogs of a similar luminosity, which could be due to a more extended distribution of dust and gas (e.g., Hwang et al. 2010). However, we note that dust temperatures derived in this work and others in the literature suppose that the emission-weighted sum of all the dust components could be well fitted by a single modified blackbody model. In reality, the far-IR emission of a galaxy results from the combination of a series of blackbodies associated with a temperature distribution depending on a number of factors, including dust emission spectral index, dust grain distribution, and geometry (e.g., Casey 2012; Magnelli et al. 2012), that is often summarized by an effective dust temperature associated with the peak emission of the far-IR SED.

3.3. High-z Radio/far-IR Relation

For the 28 radio-detected sources, we measured the far-IR/radio ratio, q_{IR}, to investigate the far-IR/radio correlation for ULIRGs at $z > 2$. The tightness of the far-IR/radio relation is very useful in many aspects in the study of galaxy evolution, e.g., in estimates of the SFRs of dusty starbursts in the absence of far-IR data (Condon 1992; Barger et al. 2012), and in the search for radio-excess AGNs (Del Moro et al. 2013). However, controversial results have been obtained in literature studies on the evolution of such a relation. Recent studies based on *Herschel* data claimed a redshift evolution of the correlation (e.g., Ivison et al. 2010a; Magnelli et al. 2010, 2012, 2014; Casey et al. 2012a, 2012b; Thomson et al. 2014). On the contrary, Sargent et al. (2010) showed that by taking into account the sample selection effect, the far-IR/radio relation is roughly invariant up to redshift $z = 1$, and possibly up to higher redshift (in agreement with Bourne et al. 2011; Barger et al. 2012; Pannella et al. 2015). Both results are inconsistent with the theoretical prediction that this tight relationship should break down at high redshift due to rapid Compton cooling of the relativistic electrons (e.g., Condon 1992).

We use the ratio (q_{IR}) of rest-frame 8–1000 μm flux to 1.4 GHz radio flux as defined in Ivison et al. (2010a). The rest-frame radio power was computed assuming $S_\nu \propto \nu^{-\alpha}$ and a radio spectral index of $\alpha = 0.8$ (e.g., Thomson et al. 2014). For these radio detected sources, the measured q_{IR} ranges from 1.04 to 2.76 with a median value of 2.37. In Figure 14, we plot the q_{IR} versus (left) their far-IR luminosities and (right) their redshifts. Using the *Herschel* data in the GOODS North field, Ivison et al. (2010b) performed a stacking analysis of 24 μm selected galaxies at $z = 0$–2, and found a median value of $q_{IR} = 2.4$ (solid line in Figure 14) and a scatter of 0.24 (dashed lines). Our $z \gtrsim 2$ galaxies have q_{IR} values in good agreement with this range. Only three galaxies (GH500.21, GH500.27c, and GH500.30) in Figure 14 lie off the far-IR/radio correlation. This is likely due to the existence of a strong AGN, which contributes significantly to the radio emission hence lowers the q_{IR} (e.g., Del Moro et al. 2013). This is consistent with the observed excess of 24 μm emission in GH500.30 (see Figure 21 in the appendix). We note that our selection is not sensitive to high-z AGNs, which are expected to be bright at 24 μm and thus have depressed $S_{500\,\mu m}/S_{24\,\mu m}$ ratios unless they are extremely luminous in the far-IR.

Since the radio continuum of galaxies without strong AGNs is predominantly produced by supernova remnants from young stellar populations, our results suggest that the far-IR luminosity for high-z *Herschel* sources is dominated by heated dust emission due to star-forming process in massive galaxies, with little contribution from old stars. Parameterizing the q_{IR} evolution in the form $q_{IR} \propto (1+z)^\gamma$, and excluding three galaxies with a strong AGN contribution to the radio emission, we find $\gamma = 0.05 \pm 0.05$, which is consistent with no evolution. For comparison, Ivison et al. (2010b) found a slight evolution with $\gamma = -0.04 \pm 0.03$ (red solid curve with dotted $\pm 1\sigma$ uncertainties). By extending to higher redshifts (out to $z = 4.6$) than most other studies, our data suggest no evolution of q_{IR}. This agrees with the recent finding by Pannella et al.

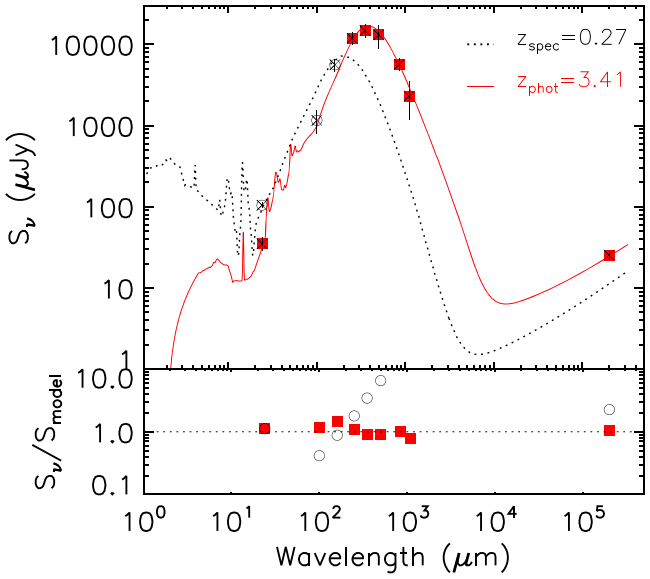

Figure 15. Far-IR SED of a mis-identified $z = 3.4$ *Herschel* source GH500.19 (red squares). The best-fit CE01 template at $z = 3.4$ is shown by the red line, while the black dotted line is the best fit at $z_{spec} = 0.27$ to *all* data points from 24 μm to radio. It is clear that the best-fit SED at $z_{spec} = 0.27$ ($\chi^2 = 51.2$) is much poorer than that at $z = 3.4$ ($\chi^2 = 1.38$). The $z_{spec} = 0.27$ source can only have a significant contribution to the *Herschel* fluxes at shorter wavelengths (black open circles).

(2015), who studied the far-IR/radio correlation up to $z \simeq 4$ in a stacking analysis of *Herschel* data for a mass-selected sample of star-forming galaxies.

4. DISCUSSION

4.1. Misidentified High-redshift Sources in the Herschel Catalog

The large beam sizes (FWHM $\sim$ 20″–40″) of the *Herschel*/SPIRE images mean that obtaining robust multi-wavelength identifications and photometry for SPIRE sources is challenging. *Herschel* point source photometry is frequently performed by flux extraction at positions of known 24 μm sources or 1.4 GHz radio sources (e.g., Roseboom et al. 2010, 2012; Elbaz et al. 2011), as these wavelengths have much better resolution ($\sim$1″–5″) and, like *Herschel*, are good tracers of star formation activity (e.g., Ivison et al. 2010a). However, this is particularly problematic for high-redshift sources that are faint or undetected in 24 μm and radio surveys, resulting in ambiguous and/or incorrect counterpart identifications (Roseboom et al. 2010; Yan et al. 2014).

Our map-based approach has the advantage of identifying the correct counterparts to high-z *Herschel* sources which are bright at 500 μm but relatively faint at 24 μm. When using 24 μm priors for a *Herschel* catalog, if the SPIRE beam contains several blended 24 μm sources that cannot be deblended, the brightest 24 μm source is usually chosen as the counterpart. This can lead to misidentifications for high-z sources as we show in Figure 15 for the case of GH500.19. Within the 500 μm beam, there are two closely separated (<10″) 24 μm sources, one spectroscopically confirmed at $z_{spec} = 0.27$ and another at $z_{phot} = 3.41$, both possibly contributing to the *Herschel* flux. While the far-IR emission is taken to be associated with the galaxy at $z_{spec} = 0.27$, the far-IR SED is inconsistent with the low-redshift solution. At $z_{spec} = 0.27$, the fit to the IR photometry has $\chi^2 > 50$ for all the templates in the library, with a minimized $\chi^2 = 51.2$ (dotted line). In contrast, the best-fitting template at $z_{phot} = 3.41$ yielded a much lower $\chi^2 = 1.38$ (red solid line). We therefore conclude that the $z_{phot} = 3.41$ galaxy is likely to be the true counterpart to the *Herschel* source, and the low-redshift candidate at $z_{spec} = 0.27$ is ruled out. Our method correctly identified the $z_{phot} = 3.41$ object as the counterpart, while in this case the usual radio identification would favor the low-redshift counterpart, which is the brightest radio source within the *Herschel* beam.

By effectively reducing the confusion from low-z sources using the map-based $S_{500\,\mu m}/S_{24\,\mu m}$ ratio, we have identified three potential high-z 500 μm sources that are not detected at 24 μm, of which two are likely to be at $z > 4$. In Figure 16, we present the images in the ground-based K_s band, IRAC 8 μm, MIPS 24 μm, *Herschel* 500 μm, the $S_{500\,\mu m}/S_{24\,\mu m}$ map, and the VLA radio image. There is no 24 μm source within 5″ of the radio counterpart, while a galaxy is clearly detected by *Spitzer* in all four IRAC bands. All three sources are relatively faint in the K_s band, and the extremely red colors of $K_s - [3.6\,\mu m] > 2$, if due to the strong, redshifted Balmer/4000 Å break of an evolved stellar population, imply that they are likely at $z > 3$. Indeed, the photometric redshift analysis suggests that two of them are at $z = 2.9$ and $z = 4.2$ respectively. The remaining source, which is not seen at K_s and shorter wavelengths, is a bright submm source (Pope et al. 2006; Perera et al. 2008; Barger et al. 2014). Since there is no photometric redshift available for this galaxy, we fitted its far-IR to radio photometry with a suite of CE01 templates, and found a best-fit redshift of $z = 4.4$. The high redshifts of 24 μm dropout *Herschel* sources which have properties very similar to the well-studied SMG GN10, if confirmed, would have important implications for galaxy formation and evolution.

To summarize, among 36 unique $z \gtrsim 2$ 500 μm sources identified in this paper, we find five cases ($\sim$14%) where the 500 μm flux is incorrectly associated with a brighter 24 μm prior in the official *Herschel* catalog. Three high-z 500 μm sources that are not detected at 24 μm are revealed in the $S_{500\,\mu m}/S_{24\,\mu m}$ map ($\sim$8%, see Section 4.2). In three further cases, the *Herschel* 500 μm flux in the catalog is probably given to nearby sources due to strong source blending ($\sim$8%). Therefore, using the deepest 24 and 500 μm imaging in the GOODS-N field, we estimate that at least 20% of $z > 2$ dust obscured galaxies in our sample were mis-identified or missed in the current catalog. This missing fraction of high-z ULIRGs population should be taken into account in studies of starburst galaxies in the early universe. Due to the small volume of the GOODS-N field, it is important to apply our method to larger survey fields with comparable depths (e.g., HerMES) to build a statistically significant and more complete sample of $z \gtrsim 2$ *Herschel* sources. However, detailed, high-resolution submm interferometric follow-ups, e.g., with NOEMA or ALMA, are required to *unambiguously* identify the correct SPIRE source counterparts, and resolve their multiplicities.

While *Herschel* is highly sensitive to $z \sim$ 1–2 galaxies whose SED peaks in the far-IR, the long-wavelength submm bands (such as 850 μm and 1 mm) are optimized for higher-redshift galaxies. It is thus interesting to compare the *Herschel* galaxies with submm sources detected by SCUBA (for example). Among our 34 *Herschel* sources, 20 have been previously reported as SMGs selected at 850 μm (Pope et al.

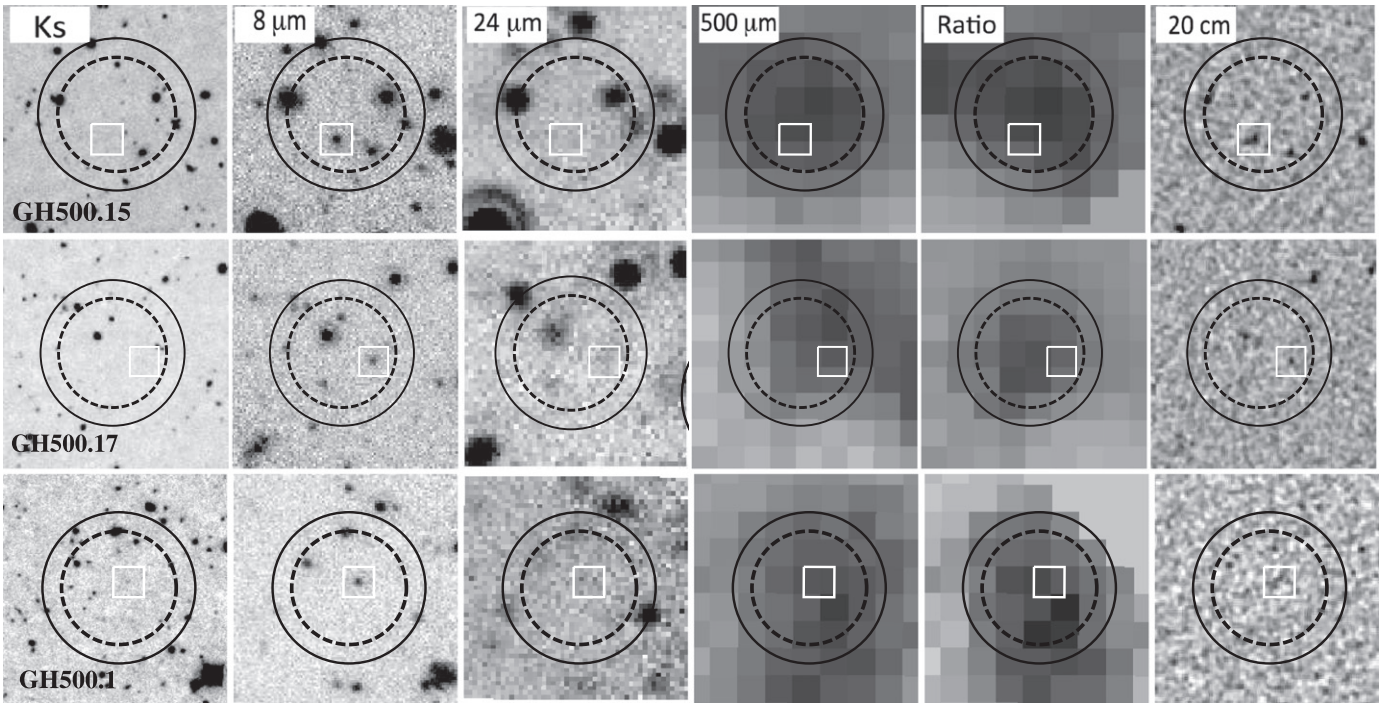

Figure 16. Cutout images for three 500 μm sources which are not detected at 24 μm, from left to right: ground-based K_s-band, IRAC/8 μm, MIPS/24 μm, SPIRE/500 μm, $S_{500\,\mu m}/S_{24\,\mu m}$, and VLA 1.4 GHz. The smaller white square indicates the counterpart identified from the radio or IRAC. The dashed circle represents the search radius of 15″, centered on the source position determined by SExtractor, while the solid circle indicates the beam FWHM of 500 μm with a radius of 18″. The non-detections at 24 μm and the red colors (K_s–[3.6 μm] > 2) suggest that they are likely to be extremely dust-obscured galaxies at $z \gtrsim 3$. The images are 1 arcmin on a side.

2006; Wall et al. 2008), although three are outside of the SCUBA coverage. Therefore, twelve of our sources are not listed in the 850 μm catalog. In Figure 17, we show the locations of sources with and without 850 μm detections on the $S_{850\,\mu m}$–z and $S_{850\,\mu m}$–L_{IR} planes. We find that there is a correlation between $S_{850\,\mu m}$ and L_{IR}, although it is subject to possible selection effects, as submm observations could be biased against galaxies with warm dust temperatures and therefore relatively faint 850 μm flux densities forthe same total luminosity (e.g., Chapman et al. 2010).

Using the best-fitting dust SED, the predicted 850 μm flux densities (open circles in the lower panel of Figure 17) for our *Herschel* sample appear to be consistent with the $S_{850\,\mu m}$–L_{IR} relation. These sources have fitted SED dust temperatures in the range of 30–40 K (Table 1), hence their non-detections in 850 μm cannot be simply attributed to the selection effect. In fact, most galaxies that have not been detected in 850 μm lie at the edge of the GOODS-North field where the local noise level of the SCUBA map is relatively high (e.g., Pope et al. 2006). The median noise value for the non-detections is 3.1 mJy, compared to 0.9 mJy for those detected by SCUBA in our sample. As illustrated in Figure 17 (lower panel), most of the non-detections could be due to the relatively low S/N given their expected 850 μm flux densities. More sensitive submm imaging of the GOODS-North field, e.g., from SCUBA-2 (Geach et al. 2013), may detect the 850 μm emission in these high-z *Herschel* sources and enable a more detailed study of their dust properties.

4.2. Implications for Star Formation-rate Density (SFRD)

The SFRD represents the total star formation occurring per unit comoving volume. The importance of ULIRGs to the buildup of stellar mass can be determined by comparing their SFRD contribution to other galaxy populations (i.e., UV-selected). This is particularly important for understanding the role of dust obscuration at high redshifts. At lower redshifts ($z \lesssim 1$), ULIRGs are rare and their contribution to the SFRD is negligible. However, the importance of ULIRGs grows toward $z \sim 2$, where they could contribute half of the total SFRD (e.g., Le Floc'h et al. 2005; Magnelli et al. 2013). At $z \gtrsim 2$, the contribution of ULIRGs to the SFRD is highly uncertain due to their limited numbers (Chapman et al. 2005; Michałowski et al. 2010a; Wardlow et al. 2011). Our $z \gtrsim 2$ *Herschel* sample therefore provides an important opportunity to investigate the SFRD contribution from the high-z ULIRG population.

To estimate the contribution of our $z \gtrsim 2$ *Herschel* sources to the cosmic SFRD, we used the Kennicutt (1998) relation to convert the far-IR (8–1000 μm) luminosity into a dust-obscured SFR, assuming a non-evolving Salpeter IMF (Section 3.2). We split the individual SFR measurements into two redshift bins: $1.9 < z < 3.1$ and $3.1 < z < 4.7$ with 18 sources in the former and 16 in the latter. The SFRD in each redshift bin was measured by dividing the SFR of each galaxy in that bin by the comoving volume and summing them. We correct this number by the ratio V/V_{max} for each individual source, where V_{max} is the maximum accessible volume within which a source could reside and still be detectable by our survey. V_{max} is calculated using the survey area of GOODS-North and the luminosity detection limit which is determined by setting a flux limit of 9 mJy at 500 μm (3σ detection threshold, Elbaz et al. 2011).

We note that the SFRs of individual sources could be affected by gravitational lensing, which would make the observed IR luminosities artificially large, thus overestimate the SFRs. The probability of lensing as a function of 500 μm

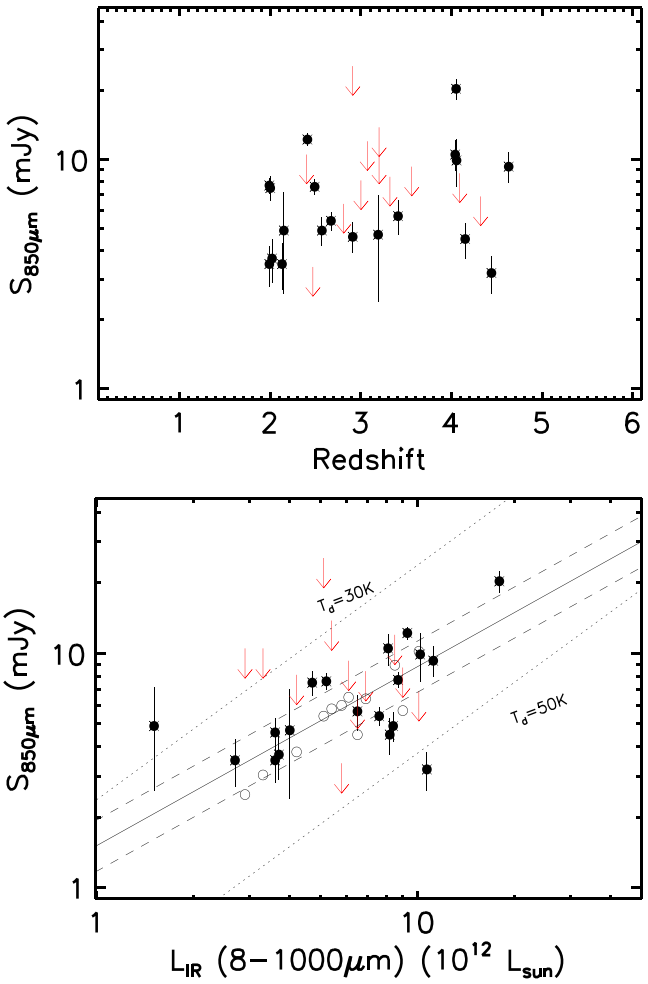

Figure 17. Submm flux densities as a function of redshift (upper panel) and far-IR luminosity (lower panel). For sources not detected in SCUBA 850 μm, we plot the 3σ upper limits on flux (downward arrows) and IR SED-predicted 850 μm flux densities (open circles). Solid and dashed lines show the linear fit (in log space) to the $S_{850\,\mu m}$–L_{IR} relation for the galaxies detected in the submm. Dotted lines represent the $S_{850\,\mu m}$–L_{IR} relation for a galaxy at $z = 2$, with a single modified blackbody at 30 and 50 K, respectively.

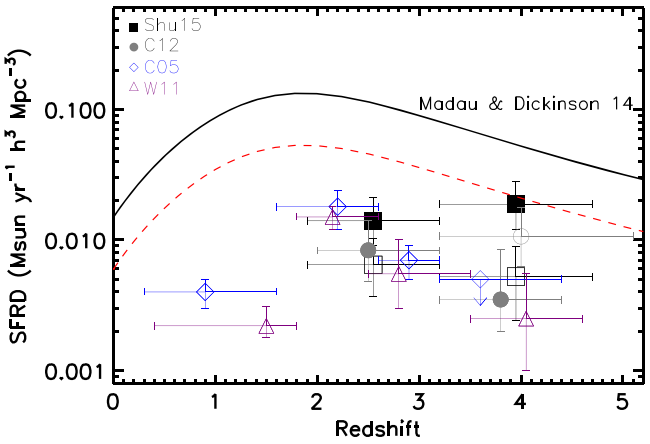

Figure 18. Estimated contribution of $S_{500\,\mu m}/S_{24\,\mu m}$-selected galaxies at $z \gtrsim 2$ (black filled squares) to the cosmic SFRD. The open squares are for the spectroscopically confirmed subsample. The data points are shown at the median redshift in each redshift bin, and error bars on SFRD are a combination of 1σ Poissonian error on the number of sources and the typical IR luminosity error of ~0.15 dex. The black curve shows the UV-based SFRD from a compilation by Madau & Dickinson (2014), while the red curve is its rescaling by a factor of 0.4. The literature SFRD values for comparison are from Chapman et al. (2005) 850 μm selected SMGs (blue diamonds), Wardlow et al. (2011) 870 μm selected SMGs (purple triangles), and *Herschel*-SPIRE spectroscopically confirmed $2 < z < 5$ sources (filled gray circles) from Casey et al. (2012b). We observe an increase in the ULIRGs' contribution to the SFRD from $z \sim 2.5$ to $z \sim 4$, consistent with the trend seen in Casey et al. (2012b), when taking into account the sources with less confident spectroscopic redshifts in their sample (open circle). Note that Barger et al. (2012, 2014) suggested that the SMG contribution to the SFRD at higher redshifts could be higher when using a more unbiased and sensitive selection of the sample.

flux density (see Wardlow et al. 2013, Figure 9) suggests that galaxies with $S_{500} < 80$ mJy have a low probability of $\lesssim$1% being lensed by a factor of $\geqslant$2. Wardlow et al. (2013) also show that there is little dependence of the conditional magnification probability on redshift at $z > 2$. Note that the brightest source in our sample (GN500.19, also known as GN20) has $S_{500} = 46.6$ mJy, and only three sources are above ~30 mJy (see Table 1). For the median value of $S_{500} = 13.9$ mJy for $z \gtrsim 2$ SPIRE sources in this sample, the predicted lensed factor would be less than 2. Therefore, the gravitational lensing effect (if there is any) would not change significantly the *Herschel* source contribution to the SFRD. Source clustering can boost the flux at the 500 μm by ~20% on average (Schreiber et al. 2015), but since such clustering only affects sources at the same redshift, it does not change our SFRD estimation for the given redshift bin.

The contribution of our sample to the SFRD is shown in Figure 18 as black filled squares. We compare our results with previous works based on the SMG samples of Chapman et al. (2005, blue diamonds) and Wardlow et al. (2011, purple triangles). Using our new technique to explore the dusty, high-redshift star-forming galaxies seen by *Herschel*, we recover more of the obscured cosmic star formation as shown in Figure 18. In particular, we observe an increase of the ULIRGs' contribution to the SFRD from $z \sim 2.5$ to $z \sim 4$ (filled squares). Note that Casey et al. (2012b) found a similar increase in the SFRD contribution from their $z > 2$ SPIRE-selected galaxies (open circle in their Figure 9) when taking into account the sources with less confident spectroscopic redshifts in the bin $3.2 < z < 5.0$. The results imply that extremely dusty starbursts may contribute significantly to the build-up of stellar mass in the early universe (about 40% of extinction-corrected UV samples, e.g., Madau & Dickinson 2014).[20] This result may be affected by cosmic variance in the GOODS-North field where the number of high-z ULIRGs is systematically high compared to other surveys (Section 3.1). Conversely, if we consider only the spectroscopically confirmed subsample, the SFRD contribution of our *Herschel*-selected ULIRGs (open squares) is similar to that of SMGs, which peaks at $z \sim 2.2$. Future spectroscopic observations of the other candidates in our sample, in particular those at $z \gtrsim 4$, are therefore required to more robustly constrain their SFRD contribution, and to understand the importance of dust-obscured star-forming activity in ULIRGs in the first few Gyr of the universe.

5. CONCLUSIONS

In this paper we present a new, map-based method to select candidate $z \gtrsim 2$ ULIRGs using the $S_{500\,\mu m}/S_{24\,\mu m}$ color. Our approach offers a relatively unbiased search for ULIRGs at

[20] Our SFRD is strictly a lower limit to the contribution from the overall ULIRG population, since we have not corrected for incompleteness in the sample.

$z \gtrsim 2$, and recovers a fraction of 500 μm sources which would otherwise be misidentified and/or missed due to incorrect 24 μm counterpart associations or non-detections at 24 μm. We provide a catalog of 34 $z \gtrsim 2$ dusty galaxies (all ULIRGs) uniformly selected in the GOODS-North field, and the main results are as follows.

1. From a combination of spectroscopic (comprising 34% of the sample) and photometric redshifts, we found that our *Herschel*-selected galaxies have a median redshift of $z = 3.1 \pm 0.8$ and 10 of them are at $z \gtrsim 4$. The redshift distribution is different from that of the SCUBA 850 μm-selected SMGs ($S_{850\,\mu m} \gtrsim 5$ mJy), but is consistent with SPIRE-selected galaxies between $2 < z < 5$ from the spectroscopic survey of Casey et al. (2012b), though the latter sample has IR luminosities systematically higher by ~0.6 dex on average. The observed $S_{500\,\mu m}/S_{24\,\mu m}$ colors are in agreement with the dusty SEDs allowed for normal low-z star-forming galaxies, albeit with a large scatter.
2. We observe a correlation between dust temperature and infrared luminosity. The median far-IR luminosity is $L_{\text{far-IR}} = 6.5 \times 10^{12} L_{\odot}$, which corresponds to a median SFR $= 1100\,M_{\odot}\,\text{yr}^{-1}$, with a range of 260–3100 $M_{\odot}\,\text{yr}^{-1}$ (assuming a Salpeter IMF). The median dust temperature is $T_d = 32.3 \pm 5.2$ K. In comparison with $z \lesssim 1$ *Herschel*-selected galaxies from Symeonidis et al. (2013), we find that at a fixed luminosity, our sample has cooler dust temperatures than that extrapolated from the L_{IR}–T_d relation at $z \lesssim 1$, which may be attributed to the more extended gas and dust distribution in the high-z ULIRGs. The result could also be partially explained by the selection at 500 μm, even at these high redshifts, which biases toward galaxies with colder dust temperatures for a fixed luminosity.
3. The radio-detected subset of our sample (28/34, excluding three likely strong AGNs) follows the far-IR/radio correlation from previous work (Ivison et al. 2010a, 2010b) with a median value of $q_{\text{IR}} = 2.37$. The result is consistent with no evolution of the correlation with redshift out to $z \sim 5$.
4. By splitting the sample into two redshift bins ($1.9 < z < 3.1$ and $3.1 < z < 4.7$), we find that the contribution from our sample to the overall SFRD at $z \sim 2.5$ is comparable to 850 μm selected SMGs and *Herschel*/SPIRE-selected ULIRGs. However, in the higher-redshift bin ($z \sim 4$), the SFRD contribution of our sample appears to be a factor of 5–10 higher than that of SMGs, and is about 40% of that of UV-selected star-forming galaxies. We caution that this result may be affected by cosmic variance and apparent overdensity of $z \sim 4$ ULIRGs in the GOODS-North field. Future exploration of the $z \gtrsim 4$ ULIRGs population in larger survey fields and spectroscopic follow-ups will be crucial to set more robust constraints on their contribution to the cosmic SFRD.

We thank the anonymous referee for helpful comments that improved the content of the paper significantly. This work is based on observations taken by the GOODS-*Herschel* Program, and is supported by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 312725. This paper used data from the SCUBA-2 Cosmology Legacy Survey (MJLSC01) on the James Clerk Maxwell Telescope (JCMT). The JCMT has historically been operated by the Joint Astronomy Centre on behalf of the Science and Technology Facilities Council of the United Kingdom, the National Research Council of Canada, and the Netherlands Organisation for Scientific Research. X.W.S. thanks the support from Chinese NSF through grant 11573001, 11103017, 11233002, 11421303, and Research Fund for the Doctoral Program of Higher Education of China (No. 20123402110030). Partial support for this work was provided by National Basic Research Program 2015CB857004. R.L. acknowledges support from the FONDECYT grant 3130558.

APPENDIX A
NOTES ON INDIVIDUAL SOURCES

Here we present detailed notes on individual *Herschel* 500 μm sources found in the $S_{500\,\mu m}/S_{24\,\mu m}$ map. The candidate counterparts in the higher-resolution bands are identified with (modified) *P*-statistics (Section 2.3.2 and Table 1).

GH500.1: This is a single 500 μm source ($S_{500\,\mu m} \sim 15$ mJy) revealed with the $S_{500\,\mu m}/S_{24\,\mu m}$ method. There are two 24 μm sources within the search radius of 15″, but niether has a modified *P*-value less than 0.1. We therefore searched the *Spitzer*/IRAC catalog for candidate counterparts and identified an IRAC source with photometric redshift $z_{\text{phot}} = 4.24$ as the counterpart based on the *P*-statistics. The source is weakly detected in the K_s band, but appears relatively bright in the *Spitzer*/IRAC bands, strongly suggesting that it is a candidate high-z dusty star-forming galaxy.

GH500.2: There is a single unambiguous 24 μm counterpart with photometric redshift of $z_{\text{phot}} = 3.03$. There is a faint 1.2 mm source (GN 1200.38, Greve et al. 2008) 3 arcsec away from the 24 μm position. The source is not detected in the radio or other submm bands.

GH500.3: There are seven 24 μm sources within the search radius, but *P*-statistics suggest that a $z_{\text{spec}} = 1.6$ source is the likely counterpart to the 500 μm source that is detected in the $S_{500\,\mu m}/S_{24\,\mu m}$ map. The source is also detected in the radio.

GH500.4: There are five 24 μm sources within the search radius, but only two of them, one at $z_{\text{phot}} = 3.2$ and another at $z_{\text{phot}} = 2.81$, have a modified *P*-value less than 0.1. While both sources are detected in the radio, the one at $z_{\text{phot}} = 3.2$ is also identified as the counterpart to a 1.1 mm source (AzGN13, Chapin et al. 2009). The *Herschel* flux in the official GOODS-*Herschel* catalog (Elbaz et al. 2011) was given to a nearby $z_{\text{spec}} = 0.98$ source. By fitting its far-IR photometry with a suite of CE01 templates, we found that there is a significant excess of SPIRE flux compared to that predicted from the best template, suggesting that it is likely a wrong association.

GH500.5: There are four 24 μm sources within the search radius, but only the one at $z_{\text{phot}} = 3.2$ has a high probability to be associated with the 500 μm emission. The source is not detected in the radio 1.4 GHz and other submm bands.

GH500.6: There is no source in the 24 μm and IRAC 3.6 μm catalog with a modified *P*-value less than 0.1. Therefore, we searched for a candidate counterpart in the radio 1.4 GHz catalog. There are two radio sources within the search radius, one at $z_{\text{phot}} = 1.54$ and another at $z_{\text{phot}} = 1.33$. Since the latter has a much lower probability ($P_{1.4\,\text{GHz}} = 0.056$) of being a chance association, we consider it as the probable counterpart. This source is also identified as the counterpart to the SCUBA 850 μm emission (GN02, Pope et al. 2006).

GH500.7: Like GH500.6, there is no source in the 24 μm and IRAC 3.6 μm catalog with a modified P-value less than 0.1. The candidate counterpart is therefore identified from the radio 1.4 GHz catalog with a photometric redshift $z_{\rm phot} = 2.15$. The source is also identified as the counterpart to the SCUBA source GN 23 by Pope et al. (2006). Nevertheless, we found that the fit with $z_{\rm phot} = 2.15$ to the far-IR photometric data is poor ($\chi^2 = 12.5$) and there is a significant excess of emission above 250 μm compared to the best-fit template. Allowing the redshift to vary in the fit, we found a much better fit at $z = 3.2$ ($\chi^2 = 0.49$). Either its photometric redshift was wrong, or it has extremely cold SED that was not included in the CE01 libraries. The source is also listed in the SCUBA-2 catalog with source name CDFN43, and the 1.1 mm catalog (ID AzGN24), for which the same identification was proposed. Note that the *Herschel* flux in the official GOODS-*Herschel* catalog was assigned to a nearby $z = 1.46$ source, which may suggest severe blending of *Herschel* flux from multiple counterparts.

GH500.8: There are three 24 μm sources within the search radius, and two of them have low P-values, with one at $z_{\rm phot} = 3.56$ and another at $z_{\rm phot} = 4.54$. Although the latter is significantly detected at 1.4 GHz, it is further away from the 500 μm source centroid (~10″). Our careful flux decomposition suggests that the source has little contribution to the 500 μm emission (~15%). The strong radio emission is likely associated with an AGN. We therefore considered the $z_{\rm phot} = 3.56$ as the most probable counterpart to the 500 μm source ($S_{500\,\mu{\rm m}} = 11$ mJy).

GH500.9: There are two heavily blended 24 μm sources within the search radius, one has a spectroscopic redshift of $z_{\rm spec} = 1.99$ and the other has $z_{\rm spec} = 2.578$ (CDFN29, Barger et al. 2014). The latter has a radio detection and a modified P-value less than 0.1, thus was chosen as counterpart. Barger et al. (2014) also identified it as the counterpart to the 850 μm emission (CDFN29, also known as GN04). Since the blending is an issue, there may be a large uncertainty in the 500 μm flux.

GH500.10: There are five 24 μm sources in the search radius, but only two of them have a modified P-value less than 0.1. While both sources are also detected in the radio, neither is detected in the other submm bands.

GH500.11: Like GH500.9, there are two heavily blended 24 μm sources within the search radius, one has a redshift of $z_{\rm phot} = 1.99$ and the other has $z_{\rm phot} = 1.61$. The latter has a modified P-value less than 0.1, thus was chosen as the counterpart.

GH500.12: There are five 24 μm sources within the search radius, but only one of them has a modified P-value less than 0.1, with a spectroscopic redshift of $z_{\rm spec} = 2.0$. The source is also identified as the counterpart to the SCUBA-2 850 μm emission based on its strong radio emission (CDFN14, Barger et al. 2014). The source is also known as a submm source 850–9 (Barger et al. 2012) or GN 19 (Pope et al. 2006).

GH500.13: There are five 24 μm sources within the search radius, but only two of them are identified as candidate counterparts based on the P-statistics, with one at $z_{\rm spec} = 1.99$ and another at $z_{\rm phot} = 2.32$. Both sources are also detected in the radio, but separated by only 2.″7, with the latter (CDFN34, Barger et al. 2014) being four times brighter than the former. While the two sources cannot be resolved by SCUBA (GN7, Pope et al. 2006), Barger et al. presented higher-resolution SMA measurements, separating it into two distinct sources, each with an 860 μm flux of ~3.5 mJy. As it is impossible to deconvolve the *Herschel* flux, we report the total flux of the two.

GH500.14: Like GH500.13, there are five 24 μm sources within the search radius, but only two of them are considered as candidate counterparts. Both sources are detected in the radio, with one being four times brighter than the other. The brighter radio source has a spectroscopic redshift of 2.47, while the fainter one has a photometric redshift of $z = 4.63$. Barger et al. (2014) measured the SCUBA-2 850 μm flux of 5.2 ± 0.7 mJy for the fainter radio source (CDFN21, also known as GN9). Because of their close separation (<10″), the SPIRE flux density from the *Herschel* catalog is given to the $z_{\rm spec} = 2.46$, brighter 24 μm counterpart.

GH500.15: This source is a 24 μm dropout galaxy, but detected in the AzTEC 1.1 mm Survey of the GOODS-N field (AzGN06, Perera et al. 2008). It is also listed in the SCUBA-2 sample of 850 μm sources (Name: CDFN8, $S_{850\,\mu{\rm m}} = 9.09$ mJy, Barger et al. 2014). This galaxy is not detected in the K_s or bluer bands, but is detected in all four *Spitzer*/IRAC bands, suggesting it is a possible candidate at $z > 4$. Using the 1.4 GHz to 850 μm flux ratio, Barger et al. estimated that the source has a millimetric photometric redshift of 3.7. By fitting its far-IR to radio photometry with a suite of CE01 templates, we found a best-fit redshift of $z = 4.4$ for the source.

GH500.16: This object has a single and clear radio counterpart, and is also detected with SCUBA-2 at 850 μm (Source ID: CDFN39).

GH500.17: There is single 24 μm source at $z_{\rm spec} = 1.22$ within the search radius, but the probability of this source being a chance alignment with the 500 μm emission is high. We therefore searched for candidate counterparts in the IRAC 3.6 μm catalog, and found a counterpart at $z_{\rm phot} = 4.31$ based on the P-statistics. By fitting the far-IR photometry with a range of CE01 templates, we found that the fit with $z_{\rm spec} = 1.22$ is very poor ($\chi^2 = 192$), andthe predicted flux based on the best-fitted SED suggests that the source contributes no more than 10% to the measured 500 μm emission. Therefore, the 24 μm dropout, at a redshift of $z_{\rm zphot} = 4.31$, is likely to be the dominant 500 μm emitter.

GH500.18: This object is part of the Pope et al. (2006) catalog (GN10) and Perera et al. (2008) catalog (AzGN3). Similar to GN20, Daddi et al. (2009a) have identified that it has a spectroscopic redshift of 4.042 from a CO emission line. The source is also listed in the Greve et al. (2008) 1.2 mm catalog and SCUBA-2 catalog. We note that the SPIRE flux densities in the catalog are assigned to a nearby, brighter 24 μm source at $z_{\rm spec} = 1.45$. By fitting the far-IR photometry with a range of CE01 templates, we find that the predicted 500 μm emission for the $z = 1.45$ source is less than 20% of that reported in the catalog, suggesting an incorrect association, and that most of the 500 μm flux should be associated with GN10.

GH500.19: There are three 24 μm sources within the search radius, but only one at $z_{\rm phot} = 3.41$ was identified as a candidate counterpart based on the P-statistics. It is detected in the AzTEC 1.1 mm imaging by Perera et al. 2008 (source ID: AzGN28), but not in the SCUBA map. New SCUBA-2 imaging reveals that the source has an 850 μm flux of 5.66 ± 0.97 mJy (ID: CDFN19, Barger et al. 2014). The SPIRE fluxes in the catalog of Elbaz et al. (2011) were mis-associated with a $z_{\rm spec} = 0.27$ galaxy, which is ~10″ away from the $z_{\rm phot} = 3.41$ source.

GH500.20: Like GH500.9, there are four heavily blended 24 μm sources within the search radius. Only one source at $z_{\rm phot} = 2.67$ was identified as the candidate counterpart based on the *P*-statistics. This sources is also known as GN 12 in the Pope et al. (2006) catalog and CDFN20 in the Barger et al. (2014) catalog. Chapin et al. (2009) identified this source as the counterpart to the 1.1 mm emission (AzTEC 8), though it is ~7 arcsec away from the 1.1 mm source centroid.

GH500.21: There are four 24 μm sources within the search radius, two of which are detected in the radio. However, only one object at $z_{\rm phot} = 2.13$ was chosen as a candidate counterpart based on the modified *P*-statistics computed from the 24 μm catalog. This source is also identified as the counterpart to the 850 μm emission (CDFN44 and GN16, Barger et al. 2014).

GH500.22: Like GH500.21, there are five 24 μm sources within the search radius, but only one ($z_{\rm phot} = 2.01$) is a probable counterpart to the 500 μm emission. This source is also detected in the radio and at 850 μm (CDFN22 and GN17, Barger et al. 2014).

GH500.23: Among three 24 μm sources, the one at $z_{\rm phot} = 3.0$ is identified as a candidate counterpart. There is another IRAC 3.6 μm source ($z_{\rm phot} = 5.9$) separated by only ~4″, which possibly contributes to the 500 μm emission, although it is not detected in the 24 μm and radio maps.

GH500.24: There are two heavily blended 24 μm sources within the search radius. Though both are detected in the radio map, neither has a modified *P*-value less than 0.1 based on the 24 μm and IRAC 3.6 μm catalog. We chose the brighter radio source ($z_{\rm spec} = 0.8$) as the candidate counterpart.

GH500.25: This source has a single and unambiguous 24 μm/radio identification, which is also associated with the SCUBA 850 μm emission (GN18, Pope et al. 2006).

GH500.26: This 500 μm source appears to have two 24 μm/radio counterparts to the 500 μm flux. Nevertheless, both objects are confirmed to have a spectroscopic redshift of 2.484 (Pope et al. 2006), thus the 500 μm emission comes from two interacting galaxies at the same redshift. While Barger et al. (2014) identified a single counterpart to the 850 μm emission (CDFN14), Pope et al. (2006) reported two radio identifications to the SCUBA 850 μm emission (GN19).

GH500.27: There are four 24 μm identifications to the 500 μm emission, three of them are also known as GN20, GN20a and GN20b. They are associated with a $z = 4.05$ protocluster in the GOODS-North field (Daddi et al. 2009b), and have spectroscopic redshifts of 4.055 based on CO lines. They are among the brightest submm sources detected with SCUBA. There is another IRAC 3.6 μm source at a lower redshift ($z_{\rm spec} = 1.63$) likely contributing to the 500 μm emission, though the probability of random association is relatively high for the 24 μm catalog ($P^C_{24\,\mu{\rm m}} = 0.187$).

GH500.28: This object is a significant SCUBA source, GN 39, described in Wall et al. (2008), and has two radio counterparts, both confirmed to lie at the same redshift ($z_1 = 1.996$ and $z_2 = 1.992$, Barger et al. 2008). As the two sources are separated by only ~6″, it is impossible to deconvolve the SPIRE fluxes, and we report the total flux of the two from the *Herschel* catalog.

GH500.29: There are two spectroscopically confirmed 24 μm sources identified as candidate counterparts (one at $z = 3.2$ and the other at $z = 2.91$). The $z = 2.91$ source has a radio detection (hence the counterpart identification is secure),

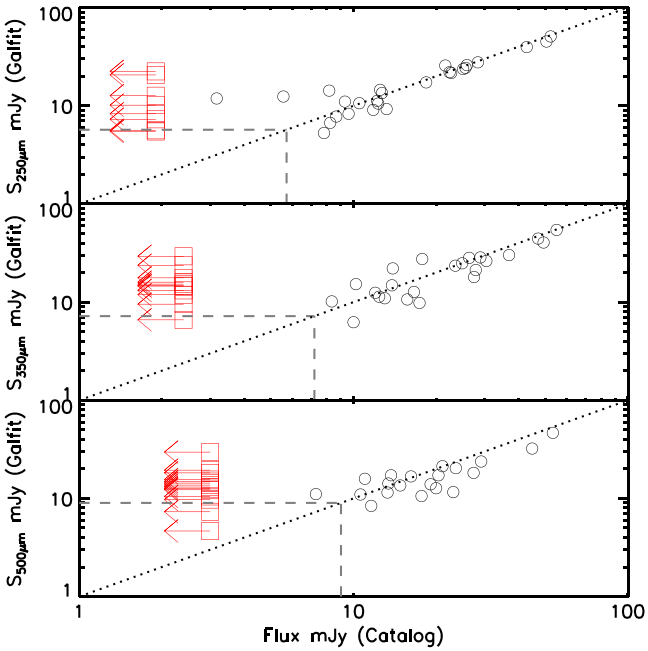

Figure 19. Comparison of flux densities derived from GALFIT with those from the catalog used in Elbaz et al. (2011). The dotted line in each panel is the one-to-one relation, while the dashed lines represent the 3σ flux limits of 5.7 mJy, 7.2 mJy, and 9 mJy at 250 μm, 350 μm, and 500 μm, respectively. When a source is not detected in the catalog, a 2σ upper limit is shown for clarification (red open squares).

and is associated with a submm source at 850 μm (CDFN27, Barger et al. 2014) and 1.2 mm (GN1200.29, Greve et al. 2008). The $z = 3.2$ source is, however, not detected in these (sub)mm bands, but shows a tentative detection in the radio. Deeper radio imaging will help to confirm the counterpart.

GH500.30: The source has a robust identification in the 24 μm and radio catalogs, with a photometric redshift of $z_{\rm phot} = 4.15$. Note that the source is very bright in the radio ($S_{1.4\,\rm GHz} = 626\,\mu$Jy) with a strong radio excess ($q < 1.5$), suggesting a significant contribution from an AGN. The source was reported as an additional 850 μm source from the Pope et al. (2006) sample with $S_{850\,\mu\rm m} = 8.8$ mJy (Wall et al. 2008), and was also detected at 1.1 mm (ID AzGN26, Perera et al. 2008), and 1.2 mm (GN 1200.6, Greve et al. 2008). However, it was later reported in the SCUBA-2 catalog with an 850 μm flux of 4.5 mJy (Name CDFN28, Barger et al. 2014).

GH500.31: A robust and clean $z_{\rm phot} = 4.3$ galaxy identified with a 24 μm/radio counterpart, which is, however, not detected in any submm bands. Although the *P*-statistics for the IRAC 3.6 μm catalog revealed another candidate counterpart at $z_{\rm phot} = 4.6$, the galaxy is not detected in the radio and is far from the 500 μm source centroid, thus its contribution to the 500 μm emission is marginal.

GH500.32: There are seven heavily blended 24 μm sources within the search radius, but only one at $z_{\rm phot} = 3.3$ was identified as a candidate counterpart based on the *P*-statistics. The source is not detected in the radio and other submm bands.

GH500.33: The source has a robust and unambiguous counterpart identification the 24 μm/radio. The galaxy does not appear in the Pope et al. (2006) catalog (it is at the edge of SCUBA coverage where the sensitivity is low). Barger et al. (2014) reported the significant SCUBA-2 850 μm detection of

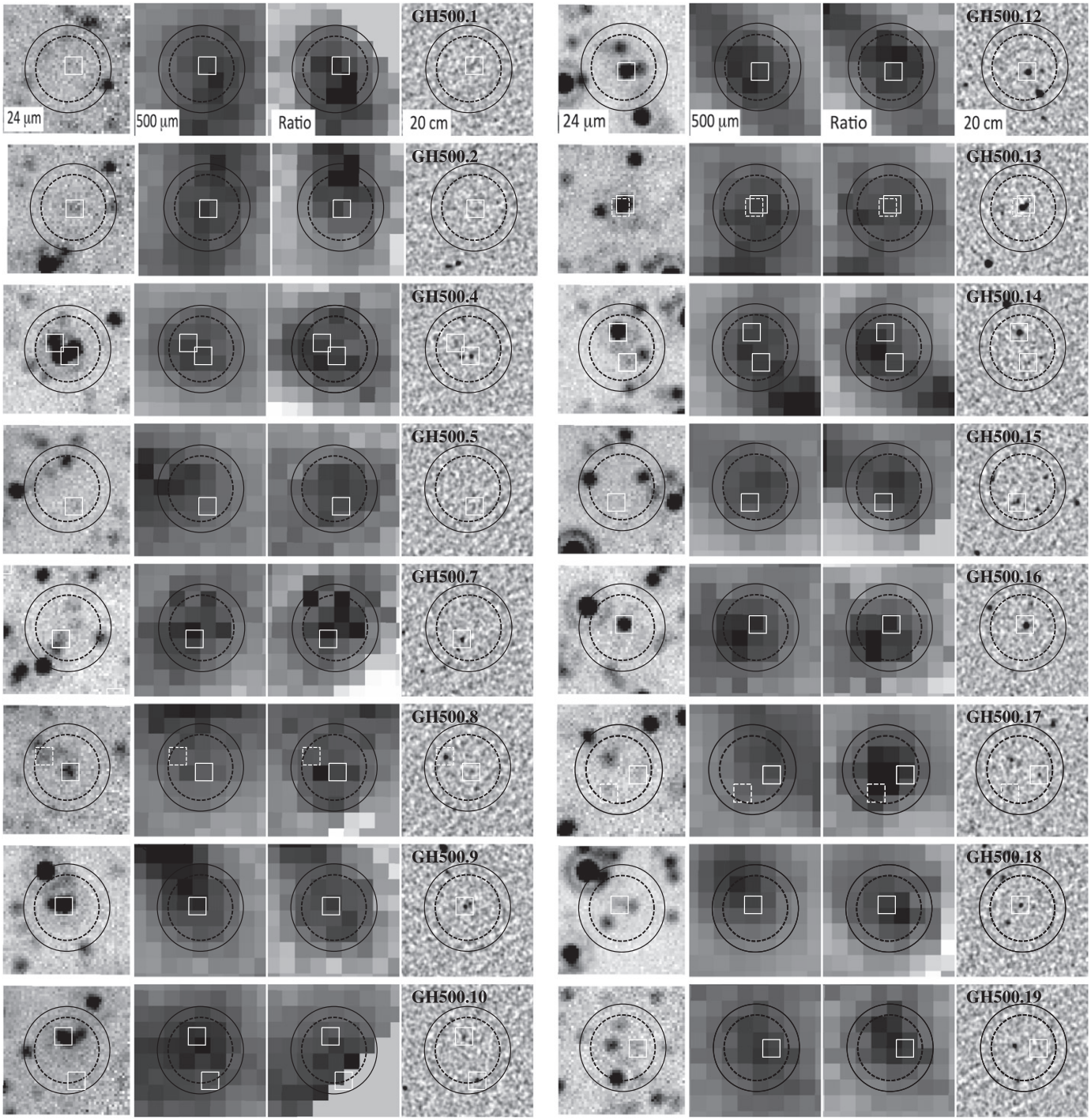

Figure 20. Cutout images of *Herschel* 500 μm counterparts, from left to right: MIPS 24 μm, *Herschel* 500 μm, the $S_{500\ \mu m}/S_{24\ \mu m}$ map, and the VLA 1.4 GHz. All images are 60 arcsec on a side and the black dashed circle indicates the search radius of 15″. The smaller white squares are secure counterparts, while the dashed squares represent tentative identifications (Table 1).

this source (ID: CDFN4) at $z_{\rm phot} = 2.41$, with flux of ∼12 mJy, making it one of the brightest submm sources (>10 mJy) in their catalog. This source also appears in the 1.1 mm catalog of Perera et al. (2008), known as AzGN05.

GH500.34: The galaxy has a robust identification in the radio and 24 μm with a spectroscopic redshift $z = 3.19$. It is also detected at 850 μm, as GN37 in the Pope et al. (2006) catalog, and at 1.1 mm (ID AzGN9, Perera et al. 2008). Like GH500.31, the *P*-statistics for the IRAC 3.6 μm catalog revealed another candidate at $z_{\rm phot} = 4.3$, but it is not detected in the radio. We chose the $z = 3.19$ source detected in both 24 μm and radio as the candidate counterpart.

GH500.35: There are three 24 μm sources identified as candidate counterparts based on the *P*-statistics, but none of them has a radio or submm detection. They are estimated to have photometric redshifts of 2.4, 2.2, and 1.98, and may contribute equally to the SPIRE flux. Therefore, we are unable to obtain a unique and secure identification. However, by inspecting the radio image (Morrison et al. 2010), we find a tentative radio detection at the position of source at $z = 2.4$. As

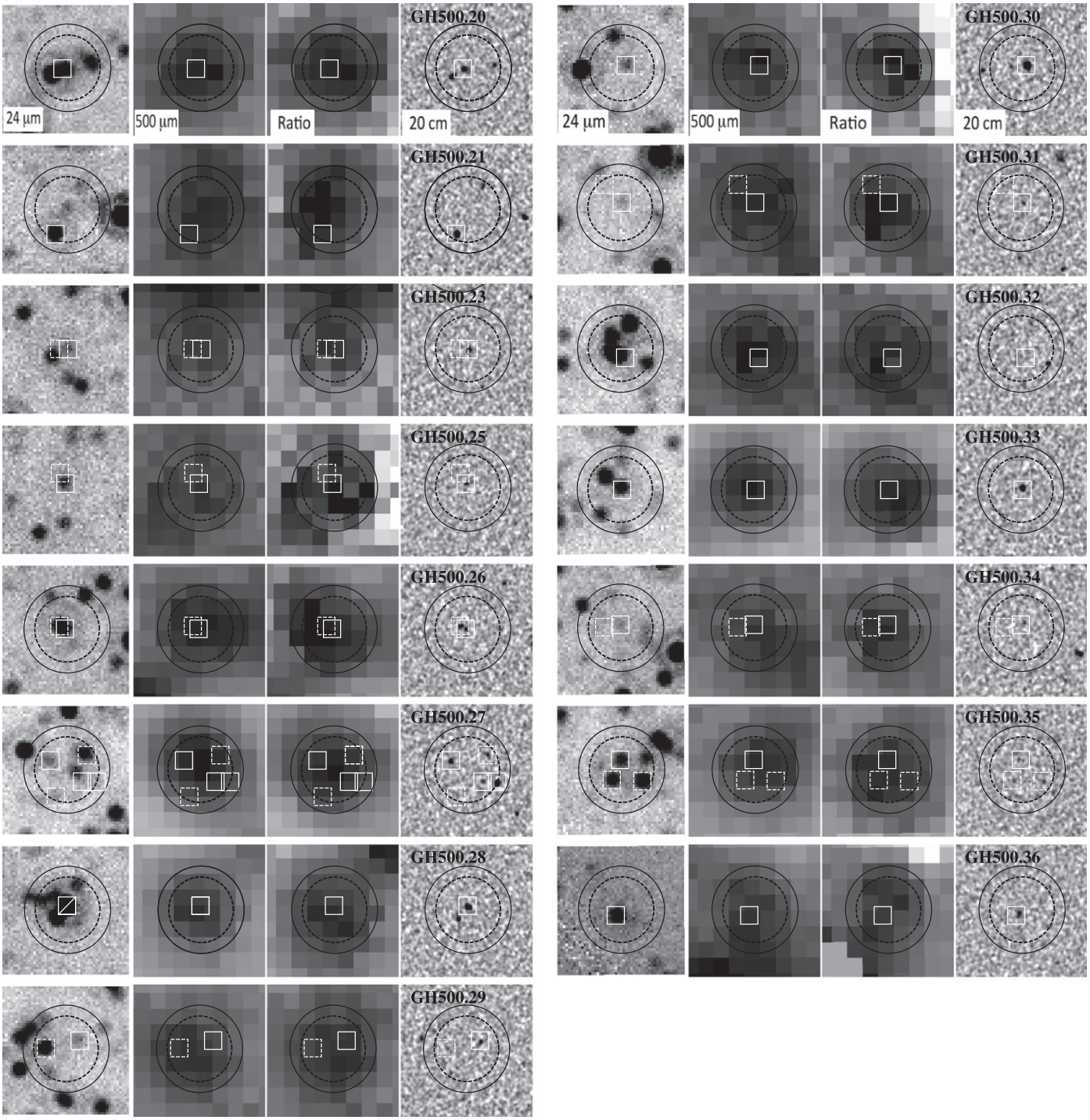

Figure 20. (Continued.)

this source is closest to the peak of the $S_{500\,\mu m}/S_{24\,\mu m}$ map, we consider it as the tentative counterpart of the 500 μm source. Using deeper VLA interferometry, or mm interferometric follow-up, it will be possible to further localize the 500 μm emission.

GH500.36: The source is near the edge of the GOODS-N region and has a photometric redshift of $z = 0.6$. However, similar to GH500.7, the best fit with a suite of CE01 library templates underestimates the SPIRE 350 and 500 μm flux significantly. Although allowing the redshift to vary yields a better fit ($\Delta\chi^2 > 10$) at a redshift of 2.6, it is possible that the source has an unusual IR SED, which is much colder than any CE01 templates.

APPENDIX B
NOTES ON HERSCHEL FLUX MEASUREMENTS

The *Herschel* flux measurements for the $z \gtrsim 2$ $S_{500\,\mu m}/S_{24\,\mu m}$-selected ULIRGs are mainly from the photometry catalog of Elbaz et al. (2011). However, as described in Section 4.1, around 30% of our sources (including three 24 μm

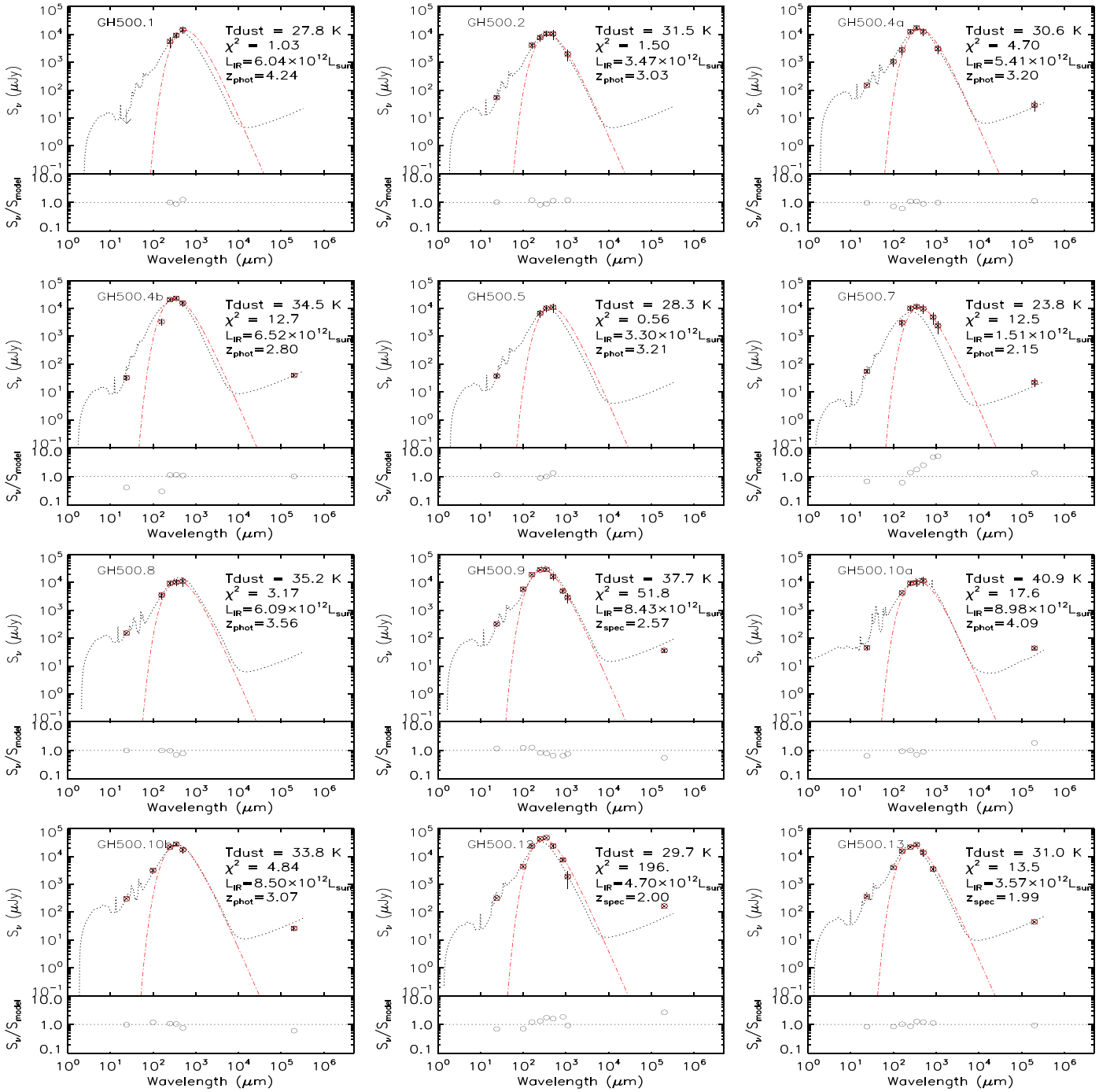

Figure 21. Spectral energy distribution of our $S_{500\ \mu m}/S_{24\ \mu m}$-identified $z \gtrsim 2$ *Herschel* sources (open circles). Dotted lines represent the best-fit SED template to the photometry. The modified blackbody emission which is used to fit the far-IR-to-submm data is shown by the red dashed line.

dropouts) have no measured 500 μm fluxes in the catalog. For these objects, *Herschel*/SPIRE flux densities are derived from PSF-fitting with GALFIT (Peng et al. 2002), using 24 μm prior positions. We used a very similar source extraction and photometry method to the one described in Elbaz et al. (2011), but fitted a stamp image of size 1.5 $\times$ 1.5 arcmin^2 centered on the source of interest for each of SPIRE bands. In brief, our fittings were performed using an "iterative" approach. We fit simultaneously all 24 μm priors above the 5σ limit within each of the 250, 350, and 500 μm stamps, but those with derived *Herschel* fluxes much smaller than the associated flux errors were removed from the prior list. The simultaneous fitting was then repeated using the refined list of 24 μm priors until no such low-signal-to-noise objects were left. Note that our restriction of the 24 μm priors to the 5σ limit would bias against some fainter 24 μm sources, especially those at high redshift contributing to the SPIRE emission. We therefore visually checked the residual image to check for residuals at the positions of 24 μm sources which were excluded from the original input list. If these were found, those objects

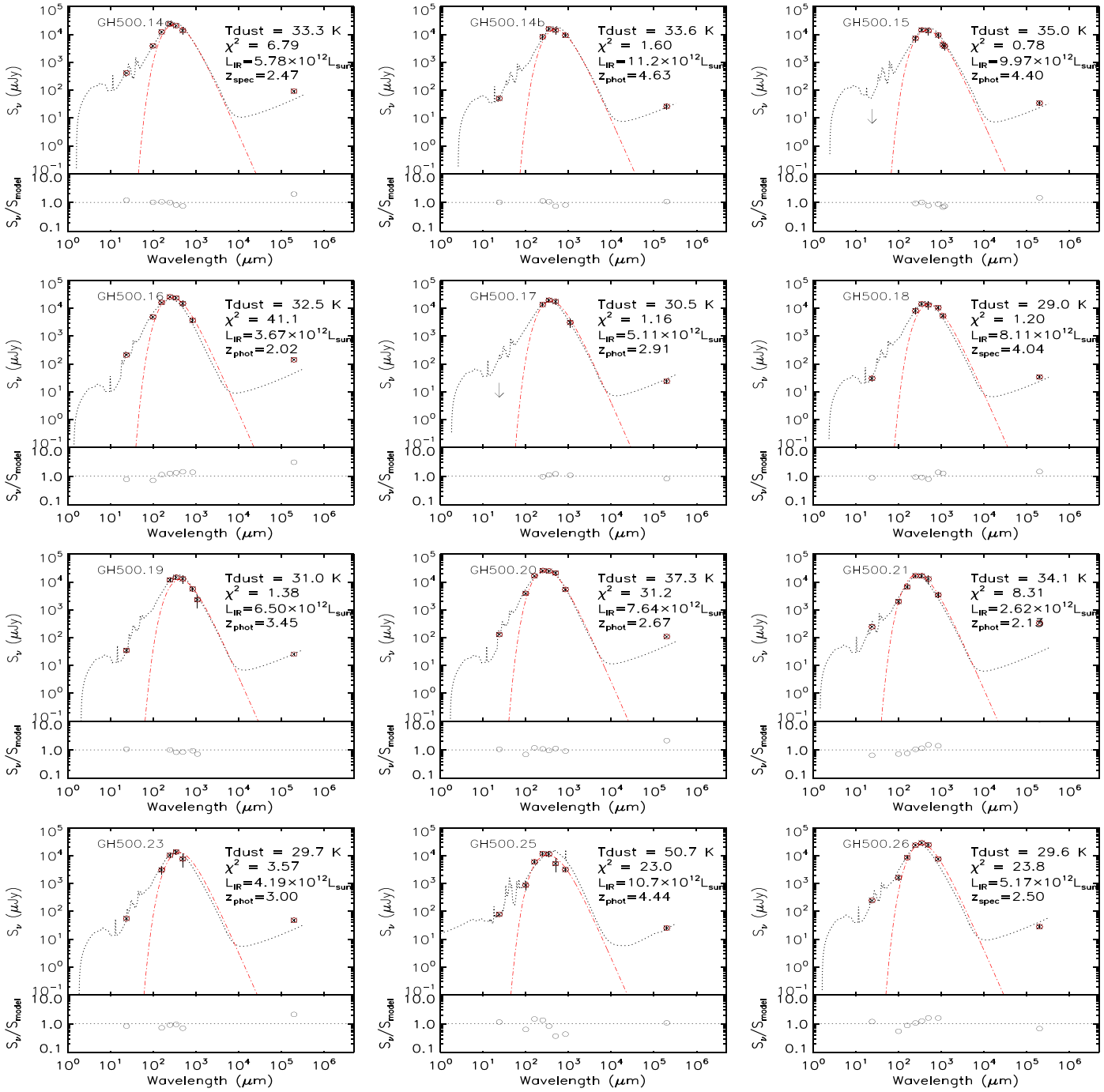

Figure 21. (Continued.)

were added back into the prior list and the fit was re-iterated. The flux density for each object was then found from the best fit (χ^2/dof < 2). To assess the reliability of our SPIRE photometry, we performed similar GALFIT fitting to all $z \gtrsim 2$ *Herschel* 500 μm sources identified in this paper and compared the results to the catalog of Elbaz et al. (2011), as shown Figure 19. Our measurements are in good agreement with the catalog used in Elbaz et al. (2011), except for high-redshift sources whose SPIRE photometry may be significantly under-estimated in the catalog due to their fluxes being assigned to neighbors with brighter 24 μm emission.

APPENDIX C
MULTIWAVELENGTH IMAGES OF HIGH-REDSHIFT *HERSCHEL* SOURCES

In Figure 20, we show the cutout images for 35 $z \gtrsim 2$ ULIRGs considered in our analysis, from left to right, MIPS 24 μm, *herschel* 500 μm, the $S_{500\ \mu m}/S_{24\ \mu m}$ map, and VLA 1.4 GHz. We mark the positions of secure counterparts as solid

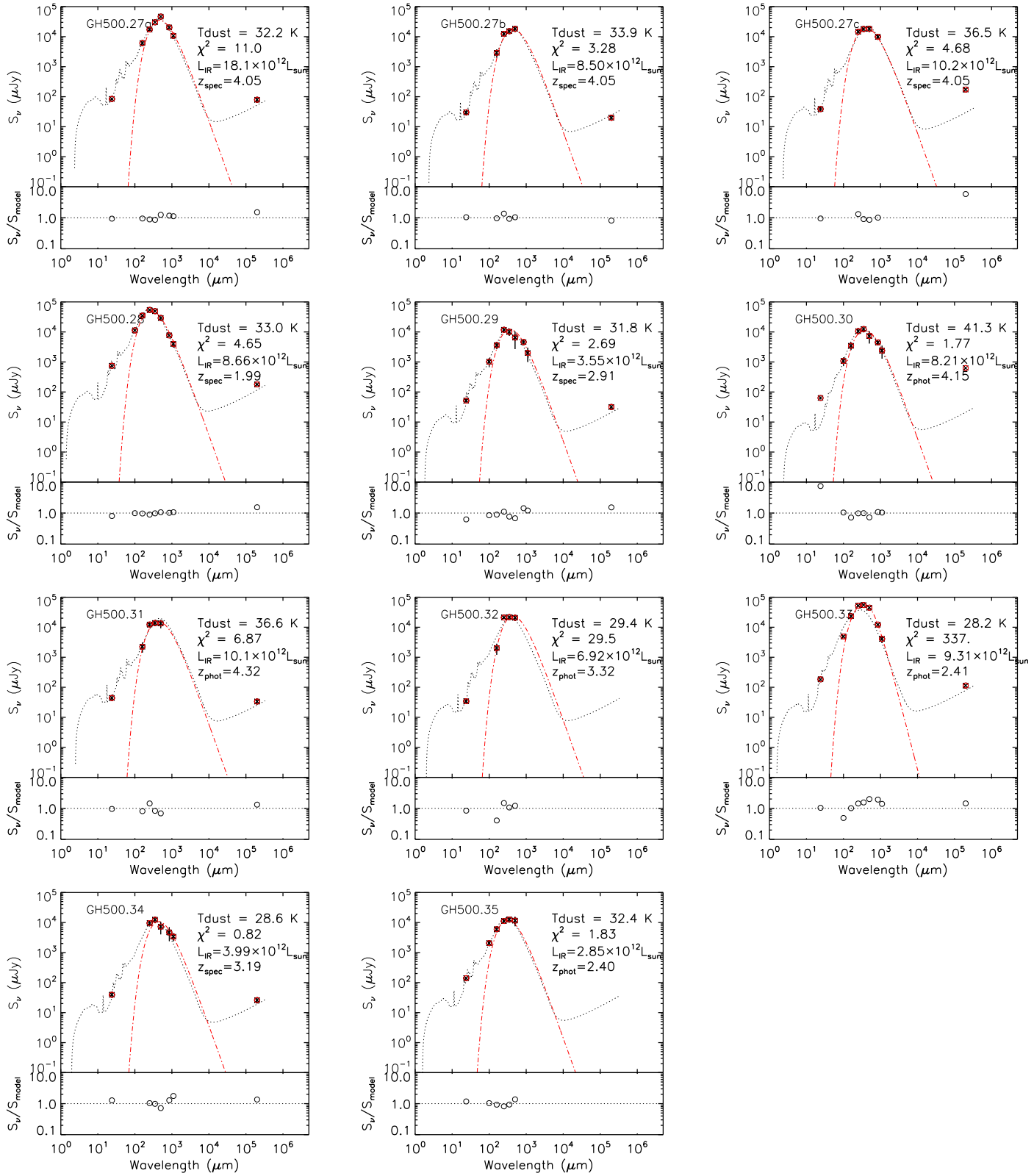

Figure 21. (Continued.)

white squares, and of tentative identifications as dashed squares. Our procedure for counterpart identification is outlined in Section 2.3. Note that four 500 μm sources have more than one secure 24 μm counterpart.

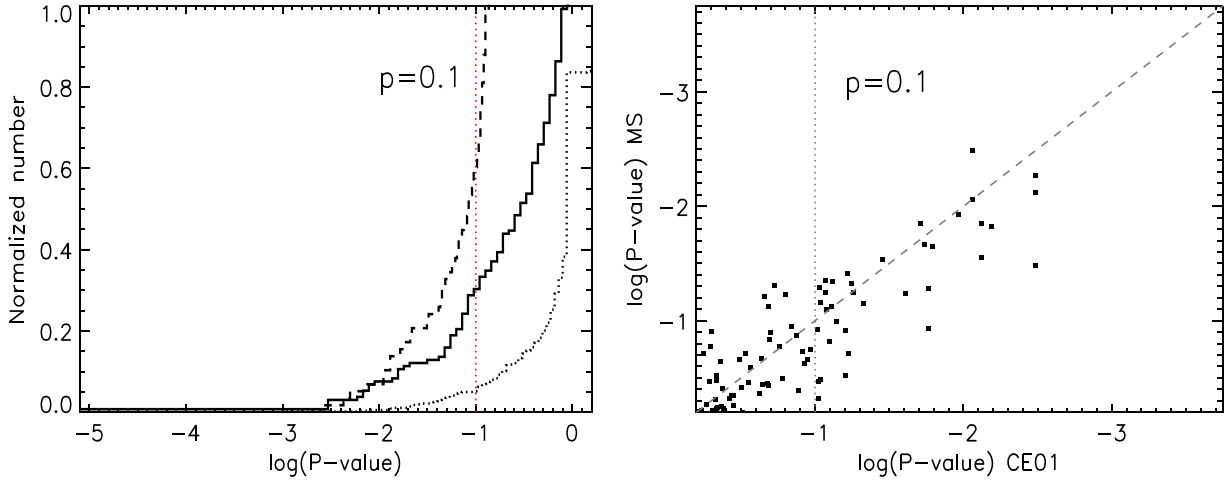

Figure 22. Left: cumulative distribution of *P*-values for all counterparts within the search area for our cataloged sources. The solid line shows the distribution for 24 μm counterparts, while the dotted and dashed line is for the counterparts at the IRAC 3.6 μm and radio, respectively. Right: comparison between the *P*-values calculated using the predicted $S_{500\,\mu m}$ starting from the 24 μm flux and that starting from the stellar masses (Schreiber et al. 2015). Gray dashed line is the one to one relation.

APPENDIX D
INFRARED-TO-RADIO SEDS OF HIGH-REDSHIFT *HERSCHEL* SOURCES

In Figure 21, we present the measured IR-to-radio photometry and best-fitting SED for each of the 500 μm sources presented in Table 2. We find that the majority of our 500 μm sources are well represented by the best-fit SED templates, and only seven ULIRGs display evidence of a radio excess above the best-fit SED.

APPENDIX E
COMPARISON BETWEEN MIPS AND IRAC IDENTIFICATIONS

In Figure 22 (left), we plot the cumulative distribution of modified *P*-values for all 24 μm sources within the search radius (solid line), and that for the IRAC 3.6 μm counterparts (dotted line). The later were calculated using the main-sequence relation and a mass dependent dust extinction to predict the 500 μm flux from their stellar masses (see the text for details), in order to taking into account alternative counterparts for a few sources with no 24 μm counterpart satisfying $P \lesssim 0.1$. Using only the 24 μm image for counterpart identifications, we find 27 out of 36 sources detected in the ratio map having counterparts with $P \lesssim 0.1$. A *P*-value cut at 0.1 corresponds to only 30% of the cumulative distribution of *P*-values for the 24 μm catalog, and 7% for the IRAC catalog. In comparison to the distribution of the traditional *P*-values for the radio catalog (dashed line), it demonstrates that our approach using 24 μm data can be useful for identifying secure 500 μm counterparts in the absence of the radio data, by effectively reducing the number of foreground sources falling into the search area by chance. Note that *P*-statistics for the IRAC catalog that has a high source surface density (79/arcmin^2) is in agreement with the calculation starting with the 24 μm counterparts (Figure 22, right), albeit with a large scatter.

Several studies have shown that high-redshift submm-selected galaxies can be identified through their IR colors as measured by IRAC (e.g., Pope et al. 2006; Yun et al. 2008). By comparing with theoretical color tracks of dusty starbursts, Yun et al. (2008) proposed a dust-obscured young stellar population as the origin of the red IRAC color. In Figure 23, we show the 3.6- and 5.8-μm color-flux diagram for all IRAC sources within the search radius (gray dots), with secure counterparts identified with *P*-statistics highlighted in red circles (Table 2). It is apparent that candidate high-z 500 μm sources in the GOODS-North are typically redder than the field population. Within the region that was defined to select SMG counterparts (dotted line, Biggs et al. 2011), we recover 91% of the cataloged sources in Table 2, further supporting the reliability of our counterpart identifications.

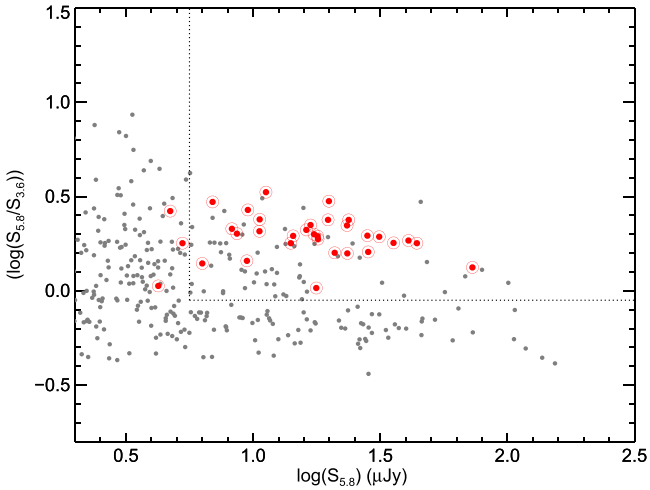

Figure 23. 3.6 and 5.8 μm color-flux diagram for IRAC sources falling into the search area. The secure radio and MIPS-identified counterparts are shown with red circles. The dotted line shows the region to select SMGs defined by Biggs et al. (2011).

REFERENCES

Barger, A. J., Cowie, L. L., Chen, C.-C., et al. 2014, ApJ, 784, 9
Barger, A. J., Cowie, L. L., & Wang, W.-H. 2008, ApJ, 689, 687
Barger, A. J., Wang, W.-H., Cowie, L. L., et al. 2012, ApJ, 761, 89
Baugh, C. M., Lacey, C. G., Frenk, C. S., et al. 2005, MNRAS, 356, 1191
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Béthermin, M., Daddi, E., Magdis, G., et al. 2012, ApJL, 757, L23
Biggs, A. D., Ivison, R. J., Ibar, E., et al. 2011, MNRAS, 413, 2314

Borys, C., Chapman, S., Halpern, M., & Scott, D. 2003, MNRAS, 344, 385
Bourne, N., Dunne, L., Ivison, R. J., et al. 2011, MNRAS, 410, 1155
Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, Natur, 470, 233
Casey, C. M. 2012, MNRAS, 425, 3094
Casey, C. M., Berta, S., Béthermin, M., et al. 2012a, ApJ, 761, 139
Casey, C. M., Berta, S., Béthermin, M., et al. 2012b, ApJ, 761, 140
Casey, C. M., Chapman, S. C., Beswick, R. J., et al. 2009, MNRAS, 399, 121
Casey, C. M., Chen, C.-C., Cowie, L. L., et al. 2013, MNRAS, 436, 1919
Casey, C. M., Narayanan, D., & Cooray, A. 2014, PhR, 541, 45
Chapin, E. L., Chapman, S. C., Coppin, K. E., et al. 2011, MNRAS, 411, 505
Chapin, E. L., Pope, A., Scott, D., et al. 2009, MNRAS, 398, 1793
Chapman, S. C., Helou, G., Lewis, G. F., & Dale, D. A. 2003, ApJ, 588, 186
Chapman, S. C., Blain, A. W., Smail, I., & Ivison, R. J. 2005, ApJ, 622, 772
Chapman, S. C., Ivison, R. J., Roseboom, I. G., et al. 2010, MNRAS, 409, L13
Chary, R., & Elbaz, D. 2001, ApJ, 556, 562
Condon, J. J. 1992, AR&AA, 30, 575
Coppin, K. E. K., Smail, I., Alexander, D. M., et al. 2009, MNRAS, 395, 1905
Daddi, E., Dannerbauer, H., Krips, M., et al. 2009a, ApJL, 695, L176
Daddi, E., Dannerbauer, H., Stern, D., et al. 2009b, ApJ, 694, 1517
Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Natur, 457, 451
Del Moro, A., Alexander, D. M., Mullaney, J. R., et al. 2013, A&A, 549, AA59
Dowell, C. D., Conley, A., Glenn, J., et al. 2014, ApJ, 780, 75
Downes, A. J. B., Peacock, J. A., Savage, A., & Carrie, D. R. 1986, MNRAS, 218, 31
Elbaz, D., Hwang, H. S., Magnelli, B., et al. 2010, A&A, 518, L29
Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119
Geach, J. E., Chapin, E. L., Coppin, K. E. K., et al. 2013, MNRAS, 432, 53
Greve, T. R., Pope, A., Scott, D., et al. 2008, MNRAS, 389, 1489
Heinis, S., Buat, V., Béthermin, M., et al. 2014, MNRAS, 437, 1268
Huang, J.-S., Rigopoulou, D., Magdis, G., et al. 2014, ApJ, 784, 52
Hwang, H. S., Elbaz, D., Magdis, G., et al. 2010, MNRAS, 409, 75
Ivison, R. J., Chapman, S. C., Faber, S. M., et al. 2007, ApJ, 660, L77
Ivison, R. J., Alexander, D. M., Biggs, A. D., et al. 2010a, MNRAS, 402, 245
Ivison, R. J., Magnelli, B., Ibar, E., et al. 2010b, A&A, 518, L31
Kartaltepe, J. S., Dickinson, M., Alexander, D. M., et al. 2012, ApJ, 757, 23
Kennicutt, R. C., Jr. 1998, ApJ, 498, 541
Kovács, A., Chapman, S. C., Dowell, C. D., et al. 2006, ApJ, 650, 592
Lagache, G., Puget, J.-L., & Dole, H. 2005, AR&AA, 43, 727
Le Floc'h, E., Aussel, H., Ilbert, O., et al. 2009, ApJ, 703, 222
Le Floc'h, E., Papovich, C., Dole, H., et al. 2005, ApJ, 632, 169
Leiton, R., Elbaz, D., Okumura, K., et al. 2015, A&A, 579, A93
Lutz, D. 2014, AR&AA, 52, 373
Madau, P., & Dickinson, M. 2014, AR&AA, 52, 415
Magdis, G. E., Daddi, E., Béthermin, M., et al. 2012, ApJ, 760, 6
Magdis, G. E., Elbaz, D., Dickinson, M., et al. 2011, A&A, 534, A15
Magdis, G. E., Elbaz, D., Hwang, H. S., et al. 2010, MNRAS, 409, 22
Magnelli, B., Elbaz, D., Chary, R. R., et al. 2009, A&A, 496, 57
Magnelli, B., Ivison, R. J., Lutz, D., et al. 2014, arXiv:1410.7412
Magnelli, B., Ivison, R. J., Lutz, D., et al. 2015, A&A, 573, A45
Magnelli, B., Lutz, D., Berta, S., et al. 2010, A&A, 518, L28
Magnelli, B., Lutz, D., Santini, P., et al. 2012, A&A, 539, A155
Magnelli, B., Popesso, P., Berta, S., et al. 2013, A&A, 553, A132
Michałowski, M., Hjorth, J., & Watson, D. 2010a, A&A, 514, A67
Michałowski, M. J., Dunlop, J. S., Ivison, R. J., et al. 2012, MNRAS, 426, 1845
Michałowski, M. J., Watson, D., & Hjorth, J. 2010b, ApJ, 712, 942
Miettinen, O., Smolčić, V., Novak, M., et al. 2015, A&A, 577, A29
Morrison, G. E., Owen, F. N., Dickinson, M., Ivison, R. J., & Ibar, E. 2010, ApJS, 188, 178
Nguyen, H. T., Schulz, B., Levenson, L., et al. 2010, A&A, 518, L5
Oliver, S. J., Bock, J., Altieri, B., et al. 2012, MNRAS, 424, 1614
Pannella, M., Elbaz, D., Daddi, E., et al. 2015, ApJ, 807, 141
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Perera, T. A., Chapin, E. L., Austermann, J. E., et al. 2008, MNRAS, 391, 1227
Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81
Pope, A., Borys, C., Scott, D., et al. 2005, MNRAS, 358, 149
Pope, A., & Chary, R.-R. 2010, ApJL, 715, L171
Pope, A., Scott, D., Dickinson, M., et al. 2006, MNRAS, 370, 1185
Rex, M., Rawle, T. D., Egami, E., et al. 2010, A&A, 518, L13
Riechers, D. A., Bradford, C. M., Clements, D. L., et al. 2013, Natur, 496, 329
Roseboom, I. G., Dunlop, J. S., Cirasuolo, M., et al. 2013, MNRAS, 436, 430
Roseboom, I. G., Oliver, S. J., Kunz, M., et al. 2010, MNRAS, 409, 48
Roseboom, I. G., Ivison, R. J., Greve, T. R., et al. 2012, MNRAS, 419, 2758
Rowan-Robinson, M., Roseboom, I. G., Vaccari, M., et al. 2010, MNRAS, 409, 2
Sanders, D. B., & Mirabel, I. F 1996, AR&AA, 34, 749
Sargent, M. T., Schinnerer, E., Murphy, E., et al. 2010, ApJL, 714, L190
Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74
Smolčić, V., Aravena, M., Navarrete, F., et al. 2012a, A&A, 548, A4
Smolčić, V., Karim, A., Miettinen, O., et al. 2015, A&A, 576, A127
Smolčić, V., Navarrete, F., Aravena, M., et al. 2012b, ApJS, 200, 10
Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, ApJ, 783, L14
Swinbank, A. M., Lacey, C. G., Smail, I., et al. 2008, MNRAS, 391, 420
Swinbank, A. M., Simpson, J. M., Smail, I., et al. 2014, MNRAS, 438, 1267
Symeonidis, M., Vaccari, M., Berta, S., et al. 2013, MNRAS, 431, 2317
Targett, T. A., Dunlop, J. S., Cirasuolo, M., et al. 2013, MNRAS, 432, 2012
Thomson, A. P., Ivison, R. J., Simpson, J. M., et al. 2014, MNRAS, 442, 577
Vieira, J. D., Marrone, D. P., Chapman, S. C., et al. 2013, Natur, 495, 344
Wall, J. V., Pope, A., & Scott, D. 2008, MNRAS, 383, 435
Walter, F., Decarli, R., Carilli, C., et al. 2012, Natur, 486, 233
Wardlow, J. L., Smail, I., Coppin, K. E. K., et al. 2011, MNRAS, 415, 1479
Wardlow, J. L., Cooray, A., De Bernardis, F., et al. 2013, ApJ, 762, 59
Weiß, A., De Breuck, C., Marrone, D. P., et al. 2013, ApJ, 767, 88
Whitaker, K. E., Kriek, M., van Dokkum, P. G., et al. 2012, ApJ, 745, 179
Yan, H., Stefanon, M., Ma, Z., et al. 2014, ApJS, 213, 2
Younger, J. D., Fazio, G. G., Huang, J.-S., et al. 2007, ApJ, 671, 1531
Younger, J. D., Fazio, G. G., Huang, J.-S., et al. 2009, ApJ, 704, 803
Yun, M. S., Aretxaga, I., Ashby, M. L. N., et al. 2008, MNRAS, 389, 333
Yun, M. S., Scott, K. S., Guo, Y., et al. 2012, MNRAS, 420, 957