A REDSHIFT SURVEY OF *HERSCHEL* FAR-INFRARED SELECTED STARBURSTS AND IMPLICATIONS FOR OBSCURED STAR FORMATION

C. M. Casey[1], S. Berta[2], M. Béthermin[3,4], J. Bock[5,6], C. Bridge[5], J. Budynkiewicz[1,7], D. Burgarella[8], E. Chapin[9,10], S. C. Chapman[11,12], D. L. Clements[13], A. Conley[14], C. J. Conselice[15], A. Cooray[16,5], D. Farrah[17], E. Hatziminaoglou[18], R. J. Ivison[19,20], E. le Floc'h[3], D. Lutz[2], G. Magdis[3,21], B. Magnelli[2], S. J. Oliver[22], M. J. Page[23], F. Pozzi[24], D. Rigopoulou[21,25], L. Riguccini[3,26], I. G. Roseboom[22,20], D. B. Sanders[1], Douglas Scott[9], N. Seymour[23,27], I. Valtchanov[10], J. D. Vieira[5], M. Viero[5], and J. Wardlow[16]

[1] Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822
[2] Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstrasse, 85748 Garching, Germany
[3] Laboratoire AIM-Paris-Saclay, CEA/DSM/Irfu - CNRS - Université Paris Diderot, CE-Saclay, pt courrier 131, F-91191 Gif-sur-Yvette, France
[4] Institut d'Astrophysique Spatiale (IAS), bâtiment 121, Université Paris-Sud 11 and CNRS (UMR 8617), 91405 Orsay, France
[5] California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125
[6] Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109
[7] Department of Astronomy, University of Massachusetts, 710 North Pleasant St, Amherst, MA 01003
[8] Laboratoire d'Astrophysique de Marseille - LAM, Université d'Aix-Marseille & CNRS, UMR7326, 38 rue F. Joliot-Curie, 13388 Marseille Cedex 13, France
[9] Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada
[10] Herschel Science Centre, European Space Astronomy Centre, Villanueva de la Cañada, 28691 Madrid, Spain
[11] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[12] Department of Physics and Atmospheric Science, Dalhousie University, 6310 Coburg Rd, Halifax, NS B3H 4R2, Canada
[13] Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK
[14] Center for Astrophysics and Space Astronomy 389-UCB, University of Colorado, Boulder, CO 80309
[15] School of Physics and Astronomy, University of Nottingham, NG7 2RD, UK
[16] Dept. of Physics & Astronomy, University of California, Irvine, CA 92697
[17] Department of Physics, Virginia Tech, Blacksburg, VA 24061
[18] ESO, Karl-Schwarzschild-Str. 2, 85748 Garching bei München, Germany
[19] UK Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[20] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[21] Department of Astrophysics, Denys Wilkinson Building, University of Oxford, Keble Road, Oxford OX1 3RH, UK
[22] Astronomy Centre, Dept. of Physics & Astronomy, University of Sussex, Brighton BN1 9QH, UK
[23] Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK
[24] Dipartimento di Fisica e Astronomia, Viale Berti Pichat, 6/2, 40127 Bologna, Italy
[25] RAL Space, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, UK
[26] NASA Ames, Moffett Field, CA 94035
[27] CSIRO Astronomy & Space Science, P.O. Box 76, Epping, NSW 1710, Australia

ABSTRACT

We present Keck spectroscopic observations and redshifts for a sample of 767 *Herschel*-SPIRE selected galaxies (HSGs) at 250, 350, and 500 μm, taken with the Keck I Low Resolution Imaging Spectrometer and the Keck II DEep Imaging Multi-Object Spectrograph. The redshift distribution of these SPIRE sources from the Herschel Multitiered Extragalactic Survey peaks at $z = 0.85$, with 731 sources at $z < 2$ and a tail of sources out to $z \sim 5$. We measure more significant disagreement between photometric and spectroscopic redshifts ($\langle \Delta z/(1 + z_{\rm spec}) \rangle = 0.29$) than is seen in non-infrared selected samples, likely due to enhanced star formation rates and dust obscuration in infrared-selected galaxies. The infrared data are used to directly measure integrated infrared luminosities and dust temperatures independent of radio or 24 μm flux densities. By probing the dust spectral energy distribution (SED) at its peak, we estimate that the vast majority (72%–83%) of $z < 2$ *Herschel*-selected galaxies would drop out of traditional submillimeter surveys at 0.85–1 mm. We find that dust temperature traces infrared luminosity, due in part to the SPIRE wavelength selection biases, and partially from physical effects. As a result, we measure no significant trend in SPIRE color with redshift; if dust temperature were independent of luminosity or redshift, a trend in SPIRE color would be expected. Composite infrared SEDs are constructed as a function of infrared luminosity, showing the increase in dust temperature with luminosity, and subtle change in near-infrared and mid-infrared spectral properties. Moderate evolution in the far-infrared (FIR)/radio correlation is measured for this partially radio-selected sample, with $q_{\rm IR} \propto (1 + z)^{-0.30 \pm 0.02}$ at $z < 2$. We estimate the luminosity function and implied star formation rate density contribution of HSGs at $z < 1.6$ and find overall agreement with work based on 24 μm extrapolations of the LIRG, ULIRG, and total infrared contributions. This work significantly increased the number of spectroscopically confirmed infrared-luminous galaxies at $z \gg 0$ and demonstrates the growing importance of dusty starbursts for galaxy evolution studies and the build-up of stellar mass throughout cosmic time.

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

Online-only material: color figures, machine-readable table

1. INTRODUCTION

Ultraluminous Infrared Galaxies (ULIRGs; $L_{IR} > 10^{12}$ $L_\odot$) exhibit the most extreme star formation rates (SFRs) in the universe (see an overview in Lonsdale et al. 2006). At early epochs ($z > 1$), ULIRG activity contributes significantly to the build-up of stellar mass presumably through intense star-forming bursts (with $\tau \lesssim 100$ Myr and SFR $\gtrsim 500\,M_\odot$ yr^{-1}, e.g., see Sanders & Mirabel 1996; Blain et al. 2002; Smail et al. 2002). Since the observed properties of these starbursts are short lived and intense, they are thought to be triggered by the collision of gas-rich disk galaxies (Engel et al. 2010) and serve as a fundamental transition phase to luminous active galactic nuclei (AGNs) or quasars (Sanders et al. 1988). Although the merger history of high-z ULIRGs has recently come into question, with some evidence pointing to a substantial (perhaps >50%) fraction of ULIRGs building stellar mass through minor mergers or passive gas accretion (Daddi et al. 2010; Elbaz et al. 2011; Rodighiero et al. 2011), there is little doubt that ULIRGs contribute non-negligibly to the star formation history of the universe and the formation of massive elliptical galaxies at the present day (e.g., Kartaltepe et al. 2010). Unfortunately, much about the infrared starburst population[28] is still unknown due to limitations in far-infrared (FIR) observations, strong selection biases, and sample inhomogeneity.

Galaxies that have been called "submillimeter galaxies" (SMGs) are selected at wavelengths around 1 mm, particularly in the atmospheric window at 850 μm. Such "classical SMGs" with $S_{850} \gtrsim 5$ mJy (Smail et al. 2002) have put powerful constraints on galaxy evolution theories and the environments of heavy star formation since their initial discovery a decade ago (Smail et al. 1997; Hughes et al. 1998; Barger et al. 1999). However, their selection at wavelengths 850 μm–1.4 mm is susceptible to strong temperature biasing (Blain et al. 2004; Chapman et al. 2004, 2010; Casey et al. 2009a; Magdis et al. 2010). This leaves the possibility that a significant fraction of high-z ULIRGs have yet to be discovered and characterized. Building a comprehensive sample of spectroscopically confirmed SMGs or extreme starbursts is paramount for determining the evolutionary histories of ULIRGs by breaking the $T_{dust}/(1 + z)$ degeneracy, for carrying out stellar population analysis, and for measuring the AGN stage and contribution to luminosity.

The *Herschel Space Observatory* (Pilbratt et al. 2010) has identified thousands of galaxies at 70–500 μm, wavelengths previously near-inaccessible from the ground, sampling galaxies' emission at the peak of their spectral energy distributions (SEDs) at $z \sim 2$, when the importance of dusty starbursts in the global context of the universe's star formation is most evident (Chapman et al. 2005). The SPIRE instrument (Griffin et al. 2010) will map $\sim$350 deg^2 in sky near the confusion limit at 250, 350, and 500 μm as part of the Herschel Multi-tiered Extragalactic Survey (HerMES; Oliver et al. 2012), covering areas much larger than SCUBA, MAMBO, AzTEC, or LABOCA. With much larger areas, the rarest sources can be uncovered and the dynamic range of sources thereby expands, from nearby luminous infrared galaxies (LIRGs; >10^{11} $L_\odot$) to distant hyperluminous infrared galaxies (HyLIRGs; >10^{13} $L_\odot$) and lensed sources. Working toward completeness in high-redshift starburst samples, and removing the impact of selection biases introduced by prior starburst selection techniques, is a key long-term goal.

Understanding ULIRG completeness and, in turn, constraining key astrophysical quantities of the luminous starburst population, is only possible with a spectroscopic census of a diverse population of FIR-luminous galaxies. Redshift identification is a crucial piece of information for a high-redshift dusty galaxy, since it allows the measurement of its luminosity and SFR, and is a prelude to subsequent interferometry (often dependent on a known redshift) in order to constrain the vast reservoirs of molecular gas that fuel extreme starbursts. Such subsequent studies cannot be completed using photometric data alone, and to date, Keck Low Resolution Imaging Spectrometer (LRIS) and DEep Imaging Multi-Object Spectrograph (DEIMOS) multi-slit spectroscopy is the most efficient method for uncovering large samples of galaxy redshifts, for both normal star-forming $z >$ 1 galaxies (Cowie & Hu 1998; Cowie et al. 1999, 2001; Steidel et al. 1996, 1999) and heavily dust-obscured ULIRGs (Barger et al. 1998, 1999, 2000; Cowie et al. 2002; Chapman et al. 2005).

This paper presents the first results from a large spectroscopic redshift survey of 1594 *Herschel*[29]-SPIRE selected galaxies (HSGs). We measure redshifts for 767 of 1594 targeted HSGs, describe their bulk infrared properties, address their relationship to the now well-studied SMGs, and assess their contribution to cosmic star formation. The results of this spectroscopic survey have been split into two papers of which this is the first, presenting the details of source selection, completeness, spectroscopic confirmations, and associated results for $z < 2$ sources. An accompanying paper presents the $2 < z < 5$ sub-sample in more detail. Throughout we use a flat ΛCDM cosmology (Hinshaw et al. 2009) with $H_0 = 71$ km s^{-1} Mpc^{-1} and $\Omega_M = 0.27$.

2. *HERSCHEL*-SELECTED GALAXY SAMPLE

The sources observed in this paper were detected by the *Herschel Space Observatory* SPIRE instrument as part of the HerMES (Oliver et al. 2012). SPIRE, the Spectral and Photometric Imaging Receiver (Griffin et al. 2010), is designed for wide-field mapping at 250, 350, and 500 μm. The beamsizes at these respective wavelengths are 18″, 25″, and 36″ with measured mean point-source confusion noise uncertainties of $\sigma_{250} = 3.8$ mJy, $\sigma_{350} = 4.6$ mJy, and $\sigma_{500} = 5.2$ mJy, which dominate over instrumental noise (these values are for $3\sigma_{conf}$ cuts used for deboosted photometric measurements, see the SPIRE Observers' Manual and Griffin et al. 2010; Nguyen et al. 2010). We make use of SPIRE maps as described by Levenson et al. (2010).

In this paper, deep ancillary data, particularly radio and 24 μm, are essential for optical spectroscopic surveying. We observe sources in the Lockman Hole North (LHN) whose *Spitzer* imaging comes from the Spitzer Wide-Area Infrared Extragalactic (SWIRE) survey (Lonsdale et al. 2003) and GO MIPS programs (PI: Owen) and very deep 1.4 GHz mapping from the Very Large Array (VLA; Owen & Morrison 2008). LHN has additional coverage with *Herschel*-PACS from HerMES (PI: G. Magdis). In the Great Observatories Origins Deep Survey North (GOODS-N) field, deep 1.4 GHz radio mapping comes from the VLA (Morrison et al. 2010) and

---

[28] In this paper, we use the term "starburst" to refer to high-SFR galaxies (SFR > 100 $M_\odot$ yr^{-1}). This differs from the recent definition of "starburst" as a combination of SFR and stellar mass (e.g., Noeske et al. 2007; Rodighiero et al. 2011).

Spitzer coverage of the GOODS-N center is from FIDEL (M. Dickinson et al. 2013, in preparation) and some of the extended area is from *Spitzer* program ID83 (PI: Rieke; D. L. Shupe 2012, private communication). We also observe sources in the ELAIS-N1 (EN1) and extended UKIDSS Ultra Deep Field (UDS)/*XMM* fields, both extragalactic areas in the SWIRE survey. The UDS has additional coverage from the *Spitzer* Legacy Program (SpUDS; PI: Dunlop). Radio coverage of the EN1 field is substantially more sparse than in GOODS-N or LHN, with the only mapping taken with the GMRT at 610 MHz (Garn & Alexander 2008) and 325 MHz (Sirothia et al. 2009), whose depths effectively translate to $100\,\mu$Jy rms at 1.4 GHz assuming a synchrotron slope of $\alpha = 0.75$, where $S \propto \nu^{-\alpha}$ (e.g., Dale et al. 2007). Our UDS observations sit on the edge of a new deep VLA radio map with $\sim 7\,\mu$Jy rms (V. Arumugam et al. 2013, in preparation). In the Cosmic Evolution Survey field (COSMOS; Scoville et al. 2007), radio coverage in the central $1\,\mathrm{deg}^2$ has a depth of $\sim 10.5\,\mu$Jy rms (Schinnerer et al. 2007) and *Spitzer* coverage is described in Sanders et al. (2007), Le Floc'h et al. (2009), and Frayer et al. (2009). COSMOS is also covered by PACS as part of the PEP program (PI: D. Lutz). We additionally observed sources in the Extended Chandra Deep Field South or ECDF-S region, which has deep radio coverage (Miller et al. 2008; Biggs et al. 2011), as well as *Spitzer*-MIPS $24\,\mu$m from FIDEL and IRAC (Damen et al. 2011).

We use the photometric redshift catalogs from SWIRE (EN1) described by Rowan-Robinson et al. (2008), the deep LHN catalog in Strazzullo et al. (2010), and the extensive COSMOS (Ilbert et al. 2010) and ECDF-S (Cardamone et al. 2010) catalogs. In GOODS-N, since we survey radio galaxies outside of the central deep region, the photometric catalog is limited, so we exclude it from photometric redshift analysis.

2.1. Source Extraction and Photometry

SPIRE point-source photometry is performed by flux extraction at positions of known $24\,\mu$m sources or radio 1.4 GHz sources. This cross-identification prior source extraction (XID) method is described in detail in Roseboom et al. (2010) with some follow-up discussion in Roseboom et al. (2012). The disadvantage of the XID technique is that it excludes any sources that are not $24\,\mu$m or radio identified. This is particularly problematic for potentially high-redshift sources that drop out of both $24\,\mu$m and radio surveys, and might have ambiguous near-IR counterparts; since $24\,\mu$m and radio source dropouts are excluded from this sample, and their influence is undoubtedly more significant at high redshifts than at $z < 2$ (e.g., Magdis et al. 2011; Béthermin et al. 2012), the high-z sample is treated in a separate paper (Casey et al. 2012; henceforth C12).

The advantage of cross-identification with $24\,\mu$m/IRAC and radio sources is that it can correct for confusion boosting in the extracted SPIRE flux densities by estimating the flux contributions from nearby sources within one beamsize. It also reduces the confusion noise by a factor of ~ 2 by pushing slightly below the nominal confusion limit using the Least Absolute Shrinkage and Selection Operator (LASSO) method to assign SPIRE flux densities to an overdense prior source list (see Roseboom et al. 2010 for method details). The LASSO algorithm combines strengths of model prediction and filter prediction source identification (i.e., balancing source priors to the SPIRE map flux distribution with choosing the brightest source correspondence between $24\,\mu$m/radio and SPIRE). The algorithm upweights "rare" sources, therefore radio sources are preferred SPIRE counterparts over $24\,\mu$m sources. This makes
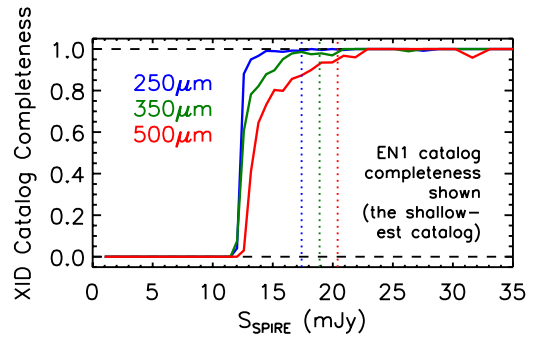

Figure 1. XID catalog completeness curves in the EN1 field as a function of flux density at SPIRE wavelengths $250\,\mu$m (blue), $350\,\mu$m (green), and $500\,\mu$m (red). The dotted vertical lines mark the 3σ noise limits at the respective wavelengths. This completeness measure is a reflection of the XID catalog's ability to assign and identify $24\,\mu$m or radio counterparts for all SPIRE sources at a given flux density.

(A color version of this figure is available in the online journal.)

sense given the expectation that radio sources are FIR-luminous (Helou et al. 1985; Condon 1992). The procedure assumes that the ancillary data are of adequate depth to identify the vast majority of FIR emitting sources, so it is only practical in deep legacy survey fields.

2.2. Completeness of Source Catalog

The most crucial aspect of a redshift survey is having a clear understanding of survey completeness and biases. This subsection addresses our survey's completeness in identifying strong SPIRE sources using the XID flux extraction method.

Since this is a test of the robustness of the XID technique, the results will vary by field (LHN, GOODS-N, EN1, COSMOS, CDFS, and UDS), based on survey depths, and is a function of SPIRE flux density. Roseboom et al. (2010) demonstrate that the robustness of the XID catalog depends greatly on the depth of ancillary data available in the field to act as source priors. Beyond a fairly standard depth at $24\,\mu$m of $S_{24} \sim 150\,\mu$Jy (or a sky density of $\gtrsim 3000\,\mathrm{deg}^{-2}$), XID source flux extraction will be >95% complete at the 3σ limit of SPIRE (where σ includes instrumental and confusion noise, and 3σ roughly corresponds to a $\gtrsim 15$ mJy cutoff). At higher source densities XID is more robust, even when the source density of priors exceeds the number of SPIRE-bright sources.

The completeness of the XID process is tested by first producing residual maps at the SPIRE wavelengths using the best-fit XID solution, and then re-injecting sources into these maps using a number count and clustering model consistent with the real data. The XID process is repeated on these simulated images and the results are assessed to determine the number of sources returned at >3σ as a function of injected flux density. This process is repeated several times for different realizations of the SPIRE maps to build up suitable statistics across a wide range of SPIRE flux densities. The results of this process are shown in Figure 1 for the field with the shallowest ancillary data, EN1; deeper fields have completeness curves approaching boxcar functions. This completeness curve is generated using the same analysis from Roseboom et al. (2010) with improvements described in Roseboom et al. (2012). An $S_{\mathrm{SPIRE}} > 15$ mJy limit should be >80% complete at $500\,\mu$m (>95% for $250\,\mu$m and $350\,\mu$m), and a 20 mJy limit should be >95% complete in all fields. Additional sources are identified down to flux densities comparable to the confusion limits (5–6 mJy). This is below the nominal 3σ cutoff and is achieved by using positional priors
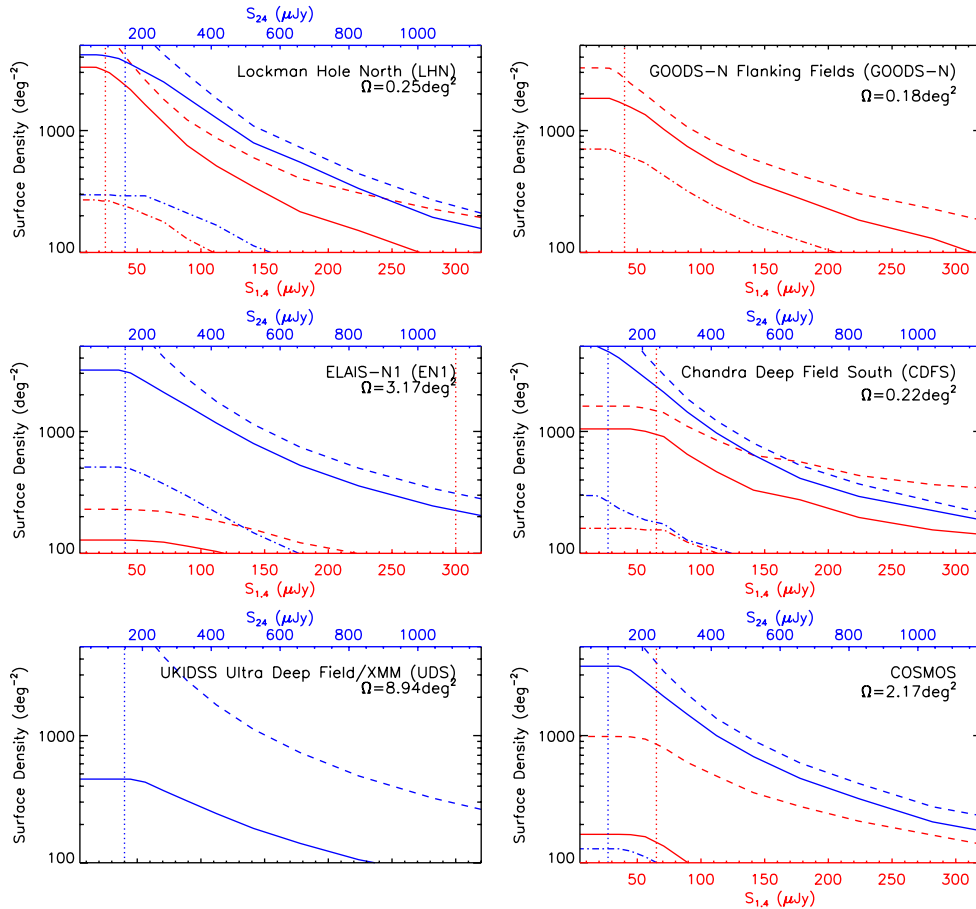
Figure 2. Cumulative surface density of 24 μm and radio sources with flux density. In other words, as a function of radio *or* 24 μm flux density, we plot the surface density of sources with flux densities above that given flux density. Dashed lines represent the parent population of all 24 μm sources (blue) and radio sources (red) in the whole field. Each panel is labeled at top with a 24 μm flux density scale and at bottom with a 1.4 GHz flux density scale (the scales are omitted if the data are insufficient/do not exist). Solid lines represent the source density for sources which are $>3\sigma$ significant in at least one of the three SPIRE bands. Dot-dashed lines represent sources $>3\sigma$ significant in all three SPIRE bands. Vertical dotted lines mark where 24 μm or radio catalogs become incomplete in flux density.
(A color version of this figure is available in the online journal.)

of many galaxies thought only to emit at SPIRE wavelengths at levels $\sim$0.1–1 mJy; the flux density of SPIRE-bright sources is then "deboosted" using the density of source priors that are thought to be SPIRE faint.

Note that the ability to match counterparts down to 15 mJy does not necessarily mean that counterpart matching is always correct (an issue that is so far largely unconstrained, but is starting to be addressed through systematic interferometric work, e.g., Wang et al. 2011). This question, particularly as it applies to the application of the XID extraction technique, is discussed in Roseboom et al. (2010).

2.3. Completeness in Radio, 24 μm Samples

The source density of $>3\sigma$ SPIRE-significant XID sources is constant within a factor of $\sim$2 (1500–3000 deg^{-2}) despite large variations in ancillary data across the different fields. For example, the LHN—a field with ultra-deep radio coverage and deep *Spitzer* mid-infrared coverage—has a source density of $\sim$3000 deg^{-2} for sources above $S_{1.4} \sim 25$ μJy or $S_{24} \sim 150$ μJy, which is the same density measured in ELAIS-N1—a field with only very shallow radio coverage—and in the GOODS-N flanking fields—an area not completely covered by *Spitzer*. Figure 2 shows the cumulative surface densities of sources selected at 24 μm and radio, then selected to be $>3\sigma$ in at least one of the SPIRE bands. The surface density of 24 μm SPIRE-

significant sources is about a factor of two times the surface density of radio SPIRE-significant sources, assuming a rough correlation between 24 μm and 1.4 GHz radio flux density of $S_{24} \approx 3.7 \, S_{1.4}$ (which we measure from Figure 3). Figure 2 also highlights that the surface density of any one field is dependent on the depth of the prior catalog, although across all fields this is consistent within a factor of about two. With comparable surface densities $\sim$1500–3000 deg^{-2}, one might then ask whether this means that 24 μm identified sources in EN1/COSMOS/CDFS/UDS and the radio-identified (>40 μJy) sources in LHN/GOODS-N/COSMOS/CDFS are drawn from the same population of IR galaxies.

Figure 3 investigates the relation between 24 μm and 1.4 GHz flux density for the LHN sub-sample that contains the deepest ancillary data. Our SPIRE targets are shown against normal field galaxies, both carving out a similar parameter space, indicating no clear SPIRE-bright/faint bias. Using our LHN sample as a guide, we find that 70% $\pm$ 20% of $S_{1.4} \geqslant 40$ μJy sources have $S_{24} \geqslant 150$ μJy (see Figure 4), suggesting that most sub-mJy radio sources are sub-mJy 24 μm sources. This highlights (1) that the incidence of SPIRE detection is higher in faint radio sources than in 24 μm sources and (2) that nearly all radio sources ($\sim$70%) are 24 μm detected and a large fraction of 24 μm sources are radio detected (from Figure 4). Since our sample is rest-frame FIR selected and is likely to obey the
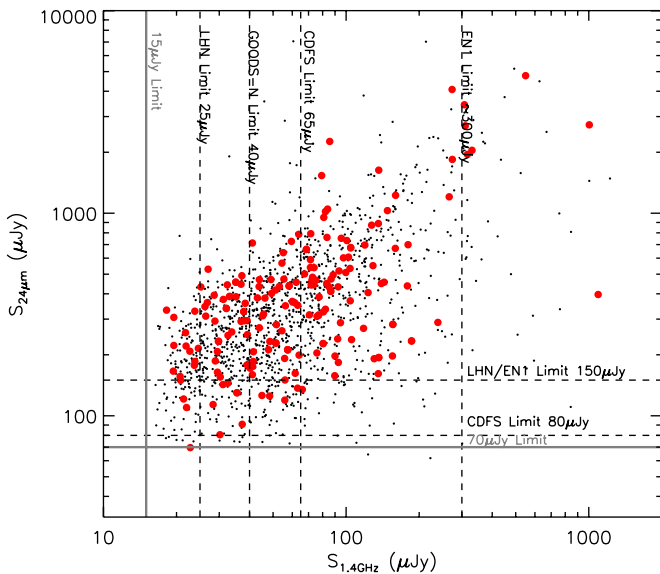
Figure 3. Distribution of LHN targets in 1.4 GHz radio flux density ($S_{1.4}$) vs. 24 μm flux density (S_{24}). The various flux limits in the three different fields are marked. Black points represent field galaxies while red points denote SPIRE-selected galaxies. Note that SPIRE detection is largely uncorrelated with 24 μm or radio flux density. However, faint 24 μm sources tend also to be faint in the radio (roughly corresponding to $S_{24} \sim 3.7 \times S_{1.4}$ as measured here), leading us to conclude that the majority of faint 24 μm EN1 sources are drawn from the same population as the GOODS-N radio sources (also see the work of Magdis et al. 2011).

(A color version of this figure is available in the online journal.)

Figure 4. Fraction of radio sources (red) or 24 μm sources (blue) that are SPIRE detected as a function of radio or 24 μm flux density. The "x"s represent the fraction of 24 μm SPIRE sources that are also radio-detected (blue) and radio SPIRE sources that are also 24 μm detected. Here we use $S_{24} \sim 3.7 \times S_{1.4}$, which is the rough relation we observe for field galaxies in Figure 3; this scaling allows us to judge the relative depths of the catalogs. "Depth A" corresponds roughly to $S_{1.4} \sim 40\,\mu$Jy and $S_{24} \sim 150\,\mu$Jy, "Depth B" corresponds roughly to $S_{1.4} \sim 25\,\mu$Jy and $S_{24} \sim 100\,\mu$Jy, and "Depth C" corresponds roughly to $S_{1.4} \sim 15\,\mu$Jy and $S_{24} \sim 70\,\mu$Jy. "Depth A" corresponds to the EN1 and UDS 24 μm limit, "Depth B" corresponds to LHN 24 μm and COSMOS/GOODS-N/CDFS radio limits, while "Depth C" corresponds to the LHN radio limit and GOODS-N(center)/CDFS 24 μm limit.

(A color version of this figure is available in the online journal.)

FIR/radio correlation (Helou et al. 1985), our sources are more likely to be radio detected than random field galaxies. Also, given the beamsize of *Herschel* to be 18″–36″, the probability of random coincidence with a radio galaxy is >8 times lower than with a 24 μm galaxy, thus the probability of correct counterpart identification is >8 times higher in radio galaxies than it is in 24 μm galaxies (although that probability itself cannot be constrained without interferometric infrared observations).

The typical range of surface densities of SPIRE galaxies, ~1500–3000 deg^{-2}, translates to an expected number of SPIRE-detected sources per slit-mask of $N_{\mathrm{DEIMOS}} = 33^{+15}_{-8}$ sources per mask and $N_{\mathrm{LRIS}} = 17^{+8}_{-4}$ sources per mask across all fields (derived from mean and standard deviations of source densities between the six survey fields). This agrees with our actual spectroscopic sampling per slit-mask; we average 33 ± 7 significant SPIRE sources per DEIMOS slit-mask and 20 ± 6 sources per LRIS slit-mask.

It is important to note that this analysis only tests the completeness and source density for galaxies already detected at mid-infrared or radio wavelengths. Some sources will be excluded from the XID catalog since they will not be detected in the ancillary data. Important examples are very high redshift sources that are radio faint and 24 μm faint and cannot be identified due to a lack of a multi-wavelength counterpart(s). These high-redshift sources will have a profound effect on the derived IR luminosity function (LF) at $z > 3$; therefore, this is a significant limitation of this HSG sample. However, at lower redshift we suspect that, at the least, the XID catalog is 80% complete out to $z = 2$ (again, see more details on estimated completeness as a function of redshift in Roseboom et al. 2010). The completeness of the XID catalog with respect to the low-redshift sources, $z \lesssim 2$, has also been shown to be ≈95% (Magdis et al. 2011).

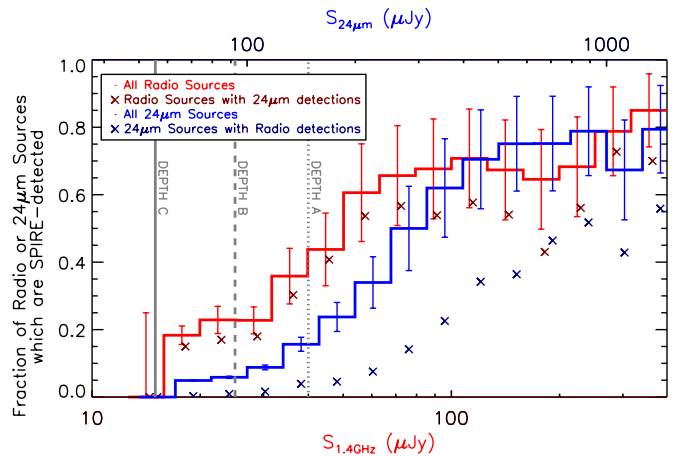### 2.4. Spectroscopic Target Characteristics

Slit-masks for the Keck LRIS (Oke et al. 1995) and DEIMOS (Faber et al. 2003) were populated by a prioritization scheme, whereby sources were graded in priority from 0 to 1000 by their SPIRE photometry, radio detection or non-detection, and 24 μm flux density. No optical magnitude selection or prioritization was made. Note that there was no filtering of the sample to remove low-redshift targets or quasars (as was often done for SMGs, e.g., Chapman et al. 2005). The i-band magnitudes of our spectroscopic targets are plotted against the brightest SPIRE flux in Figure 5. Since our spectroscopic sampling is exclusively driven by SPIRE detectability and not by optical magnitude, there is no clear relationship perceived in Figure 5. When there was no i-band or z-band counterpart to center our slit on, we used the IRAC 3.6 μm position or radio VLA position, which are both good to ~0″.7 given their relatively small beamsizes.

Sources detected at >3σ in at least one out of the three SPIRE bands were given a priority = 300, and sources detected in all three SPIRE bands at >3σ were given a priority = 500. Mask centers and orientations were chosen based on the positions of rare, "red" 500 μm peaking sources ($S_{250} < S_{350} < S_{500}$), thought to be the highest-redshift SPIRE-bright galaxies (Cox et al. 2011). These red sources were given very high priority, >800, and if their multi-wavelength properties were consistent with a high-z source, e.g., $i_{\mathrm{AB}} > 22$ and $S_{24} < 500\,\mu$Jy, priority was graded even higher, at 1000. The prioritization scheme is linear, such that a source with priority 1000 will be assigned a slit in favor of two $p = 300$ sources, or one $p = 300$ source plus one $p = 500$ source, however, a $p = 800$ source would be passed up in lieu of two $p = 500$ sources. While this scheme could accidentally remove very high priority sources from our masks, we adjusted mask position angles manually to ensure optimal spatial sampling, and were only minimally affected by source overlap (≲5% of slits were conflicted).
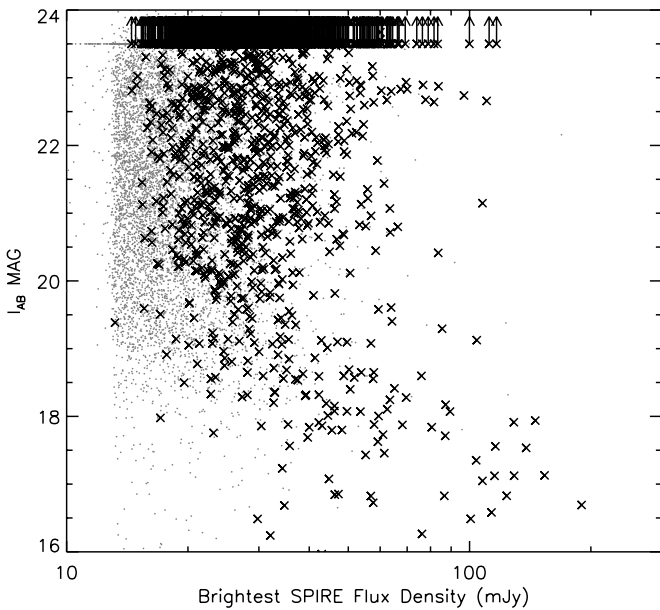
Figure 5. Optical *i*-band magnitude (AB) against SPIRE flux density (the brightest flux density is taken from SPIRE 250 μm, 350 μm or 500 μm) for sources in EN1 (similar results for other fields). Black points are all significant SPIRE sources used to fill our slit-masks. The selection of these SPIRE sources is not a function of optical magnitude.

Figure 6. Top: distribution of 24 μm and 1.4 GHz radio flux density of our targets. Of our sample, 1319 were primarily identified through 24 μm emission and 633 of were radio selected. The gray distributions reflect the sources with confirmed spectroscopic redshifts (this number exceeds 767 since some sources are identified as both 24 μm and radio sources). The spectroscopic success rate for both samples is ~55%–60% and does not correlate with flux density. Dashed vertical lines represent average flux density limits of the radio data (~30 μJy) and the 24 μm data (~150 μJy) in our fields. Bottom: each target in our survey must be detected at >3σ in at least one of the three SPIRE bands at 250, 350, or 500 μm. Here we show the 250 μm signal-to-noise ratio (S/N) against 350 μm S/N. The 500 μm S/N scales with the size of the gray circles; if a target is under 3σ at 500 μm, it has no gray circle. Spectroscopically confirmed sources are then marked with red crosses, showing no obvious correspondence between FIR flux density and spectroscopic success rates.

(A color version of this figure is available in the online journal.)

The density of our sources, both low and high priority, is high enough and comparatively uniform over the LHN, GOODS-N, EN1, UDS, CDFS, and COSMOS fields such that our mask coverages constitute a random sampling, and completeness estimates are performed with respect to the sky area probed by the masks alone (which total ~0.93 deg²). Note that the centering of our masks around high-priority "red" sources might give a high-redshift bias to our sample, since we set out to find some of the rarest, high-*z* HyLIRGs. What we find (to be discussed later in the paper and more in C12) is that the redshift distribution of "red" targets is *not* strongly biased toward high-*z*. Fewer high-*z* sources are red than not, leading us to believe that the sky sampling is essentially random, despite our efforts to detect more high-*z* galaxies. The density of >3σ SPIRE sources is also low enough so that our radio, 24 μm or SPIRE color prioritization does not introduce statistically significant selection biases into the sample; in other words, nearly all (>90%) the >3σ SPIRE sources within areas covered by slit-masks were spectroscopically observed.

We assigned slits for 100% of the very high priority "red" targets, 95% of the high-priority targets (>500), and 90% of the lower priority targets (300 < *p* < 500), filling any free space with additional objects not discussed in this paper (radio sources or 24 μm galaxies with insignificant SPIRE fluxes). Any high-priority SPIRE sources not observed were only excluded on the basis of the mask configuration, where observing them would bump another high-priority source off a slit. However, the exclusion of high-priority SPIRE sources is rare enough that having a few missing from our sample does not impact our completeness, particularly since we find no significant bias toward high-redshift spectroscopic identifications in the high-priority target sub-sample (see C12 for more details on the *z* > 2 sub-sample).

Our targets' distributions in radio flux density and in 24 μm flux density and SPIRE signal-to-noise ratio (S/N) are shown
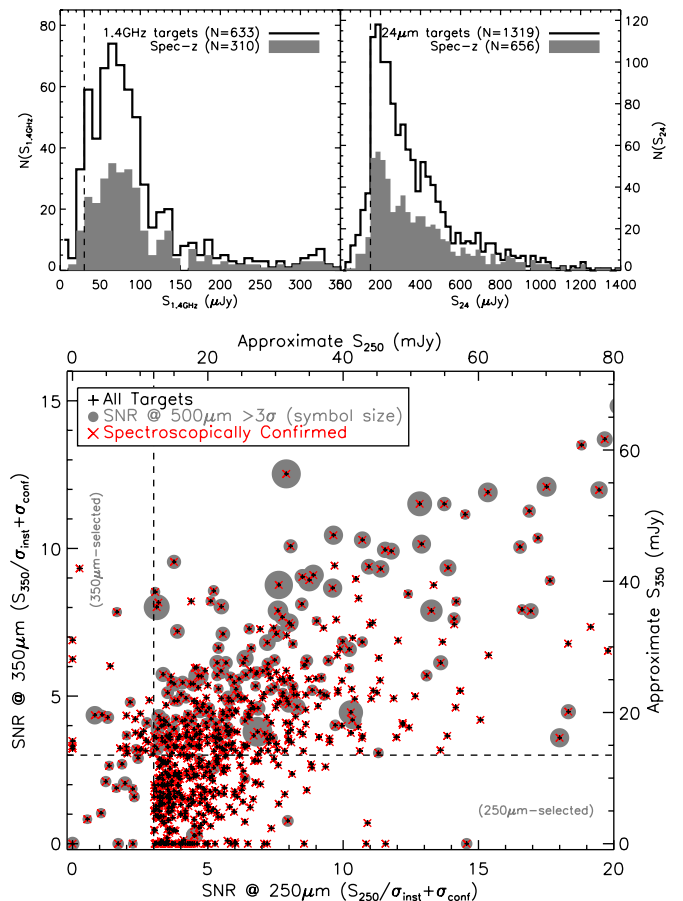
in Figure 6. Out of 1594 targets, 633 are 1.4 GHz identified, 1319 are 24 μm identified, and 588 are both radio and 24 μm identified. The sources which are selected in both radio and 24 μm maps come primarily from the LHN, COSMOS, or CDF-S fields, since EN1 and UDS lacks deep radio coverage and many sources in the flanking fields of GOODS-N lack deep *Spitzer* 24 μm coverage.

Throughout this paper, we refer to our sample as HSGs. We prefer not to call them SMGs for sake of confusion with the historical definition of SMG selected at 850 μm–1.4 mm. We also choose not to use the term ULIRG for our sources since it places the special qualification of a luminosity cut on the sample (10^{12} $L_\odot$ < L < 10^{13} $L_\odot$). For the rest of the paper, any reference to SMGs refers to a population selected at 850 μm with $S_{850} \geqslant 5$ mJy. We use this strict definition of the SMG population in order to draw comparisons between "classic SMGs" and HSGs.

3. SPECTRAL OBSERVATIONS AND REDSHIFT IDENTIFICATION

Optical spectroscopic observations were carried out at the W. M. Keck Observatory using LRIS on Keck I and DEIMOS on Keck II in 2011 and 2012.

LRIS observations were carried out in adequate conditions on 2011 February 6 with $\sim 1''$ seeing and cloud cover, on 2012 January 26 and 27 with $\sim 0''.8$–$1''.2$ seeing with minor to no cloud cover, and on 2012 February 27 with $\sim 0''.5$–$1''.0$ seeing and no cloud cover. We used the 400/3400 grism for maximum wavelength coverage in the blue. The 2011 observations used the 600/7500 grating in the red with a central "multi-slit" wavelength of 6500 Å; the 2012 January observations used the 400/8500 grating in the red with a central wavelength of 8400 Å, and the 2012 February observations used a central wavelength of 8000 Å. All observations used clear filters for both red and blue arms and the 560 nm dichroic. These setups give a 1.09 Å dispersion in the blue (e.g., shortward of 5600 Å) and a 0.80 Å dispersion in the red. Wavelength coverage for each source varied by its position on the $5'.5 \times 7'.8$ slit-mask, the mean wavelength coverage ranged from 2500–8200 Å in 2011 and from 2500 Å–1 μm in 2012 but varied up to ~ 1300 Å for sources on each mask. Due to the dichroic, some sources near the edge of the slit-mask have gaps in wavelength coverage ~ 800 Å wide in the vicinity of 5600 Å.

We observed a total of 25 LRIS multi-slit masks, 13 of which were observed in near-photometric conditions. LRIS data reduction, including bias subtraction, flat fielding, wavelength calibration, and sky subtraction were all performed using custom-built IDL routines. Of 664 LRIS targets with $>3\sigma$ SPIRE detections, 268 were spectroscopically identified. One hundred ninety-eight of the identified sources (74%) were confirmed in near-ideal conditions; we add asterisks in the NAME column of Table 1 to distinguish these sources from sources confirmed in poorer weather.

DEIMOS observations were carried out in good conditions on 2011 May 28 and 29 with $\sim 0''.6$–$0''.7$ seeing, on 2011 November 28 in average to cloudy conditions with $\sim 1''.0$–$1''.3$ seeing, and on 2012 February 16 and 17 in very cloudy conditions with $1''.0$–$3''.0$ seeing; we used the 600 lines mm^{-1} grating with a 7200 Å blaze angle (resulting in dispersion of 0.65 Å) and the GG455 filter to block out higher-order light. Wavelength coverage varied with source position on the $5' \times 16'.7$ slit-mask from 4400–9200 Å to 5200–9900 Å and averaged to 4850–9550 Å.

Sixteen of 29 DEIMOS multi-slit masks were observed in near-photometric conditions, with integration times ~ 2700–4800 s. We used the DEEP2 DEIMOS data reduction pipeline to reduce these data.[30] Of 930 DEIMOS HSG targets, 499 were spectroscopically confirmed. Three hundred twenty-four of those 499 (65%) were confirmed in near-ideal conditions and are also marked with asterisks in Table 1.

Redshift identification was carried out through the identification of multiple spectral signatures, primarily with the [O II] doublet, [O III], Hβ, Hα, [N II], Ca H & K absorption and the Balmer break, Hγ, Lyα, and the Lyman break (given in order of decreasing occurrence in the sample). A minority of sources were identified by C III], C IV, and He II emission (all sources identified via features in the rest-frame ultraviolet are discussed in C12). Full spectroscopic details of all confirmed $z < 2$ HSGs are given in Table 1 (available online in full). The reliability

of redshift identifications is also given in Table 1, rated on a scale of $q_z = 1$–5, where 5 is best. Sources with multiple feature identifications have $q_z \geqslant 3$, and sources with single line identifications have $q_z = 1$–2. When visually inspecting the spectra for our targets, redshifts were graded on a wider scale, with some sources having $q_z = 0$ (poor) or –1 (non-existent). These sources' potential redshifts are not reported in this paper due to unreliability.

There are more HSGs with unconfirmed redshifts (826) than there are with redshifts (767). The vast majority of the unconfirmed sample is unconfirmed due to poor weather including cloud cover and poor seeing (54% of sources were observed in sub-optimal conditions). The average spectroscopic yield during good weather was 60% ± 20% while the poor weather yield was only 15% ± 10%. The remaining unconfirmed sources either very red faint continuum without identifiable spectral features (20%) or no continuum whatsoever (80%). These sources are potentially misaligned on the slit (which is based on an IRAC and i- or z-band image when available) or are too optically obscured to be detected in bright emission lines with a ~ 1–2 hr integration. They could also be at $z \gtrsim 3$, thus intrinsically much more difficult to detect optically.

Since we have very limited constraints on the spectroscopic failures, we emphasize that our sample is spectroscopically incomplete. We use photometric redshifts and inferences on the redshift distribution itself in Section 3.2 to estimate the spectroscopic completeness as a function of redshift.

3.1. Biases in Spectroscopic Confirmations

Here we quantify the biases of our spectroscopic observations by analyzing sources that failed to yield spectroscopic identifications. This is arguably the most difficult completeness to quantify, since it requires some knowledge of the redshift distribution of sources that are (1) the most optically obscured, (2) have featureless continua, or (3) have observed-frame emission lines outside of the wavelength range of our observations (this redshift range is $1.6 \lesssim z \lesssim 3.2$ for DEIMOS observations and $1.6 \lesssim z \lesssim 1.7$ for LRIS observations). Figure 7 shows the optical i-band magnitude distributions for sources spectroscopically confirmed and unconfirmed, and i-band magnitude against redshift. This makes it clear that optical i-band magnitude need not be very bright for a spectroscopic identification based on emission lines, and that optical magnitude itself does not constrain redshift, given the number of sources at low-z that are undetected in the i band. However, it is clear that there is an overall trend with redshift at brighter magnitudes $i < 22$, with very few sources at these magnitudes at $z > 0.5$. In that sense, it is also clear from Figure 7 that optically bright sources are more likely to be spectroscopically confirmed than those that are faint; the mean i-band magnitude of spectroscopically confirmed sources is $i_{AB} = 22.1$, while for unconfirmed source it is $i_{AB} = 23.8$.

Prior studies of optical photometric properties of infrared starburst galaxies find poor correspondence between optical magnitude, IR luminosity, and redshift (e.g., Chapman et al. 2004, 2005). This makes it quite difficult to estimate the redshift distribution and IR luminosities of the optically faint targets for which we fail to measure redshifts. Without redshift information, we must rely on color information to infer whether the optically fainter targets have a significantly distinct redshift distribution (this also assumes SPIRE color varies with redshift and dust temperature is roughly fixed). Figure 8 shows how SPIRE color relates to i-band magnitude for the

8

Table 1
Spectroscopically Identified 0 < z < 2 SPIRE-selected Galaxies

NAME	z_{spec}	CONF	z_{phot}	S_{24} (μJy)	S_{100} (mJy)	S_{160} (mJy)	S_{250} (mJy)	S_{350} (mJy)	S_{500} (mJy)	$S_{1.4\,GHz}$ (μJy)	L_{IR} ($L_{\odot}$)	T_{dust} (K)
1HERMES X1.4 J021856.18−043546.5*...	1.102^D	1	...^{U}	275.2 ± 18.1	−	−	16.4 ± 12.4	5.7 ± 8.2	...	−	$(2.1^{+2.4}_{-1.1}) \times 10^{12}$	57.8 ± 7.9
1HERMES X1.4 J021856.37−043538.0*...	0.470^D	3	0.47^U	336.3 ± 18.6	−	−	10.2 ± 12.4	35.9 ± 8.2	40.6 ± 5.5	−	$(1.1^{+1.2}_{-0.5}) \times 10^{11}$	8.90 ± 1.7
1HERMES X1.4 J021901.73−043449.0*...	1.010^D	3	...^{U}	676.3 ± 17.3	−	−	31.0 ± 12.4	25.9 ± 7.0	8.5 ± 5.3	−	$(1.3^{+5.9}_{-1.0}) \times 10^{12}$	28.4 ± 30.6
1HERMES X1.4 J021921.93−043335.2*...	0.922^D	2	...^{U}	337.2 ± 16.1	−	−	32.6 ± 12.4	27.1 ± 7.1	12.1 ± 5.5	−	$(8.7^{+25.}_{-6.5}) \times 10^{11}$	25.2 ± 19.7
1HERMES X24 J021926.00−043435.4...	1.302^D	4	...^{U}	342.6 ± 19.4	−	−	24.1 ± 12.4	14.6 ± 7.1	27.9 ± 5.2	−	$(1.1^{+1.6}_{-0.6}) \times 10^{12}$	16.3 ± 5.0
1HERMES X24 J021926.85−043044.0...	1.029^D	3	...^{U}	368.5 ± 21.0	−	−	34.2 ± 12.4	29.4 ± 7.1	20.1 ± 5.2	−	$(9.1^{+13.}_{-5.5}) \times 10^{11}$	21.2 ± 9.0
1HERMES X24 J021933.91−043335.3...	0.582^D	1	...^{U}	1164. ± 21.1	−	−	48.2 ± 12.4	50.8 ± 7.0	29.9 ± 5.4	−	$(4.0^{+3.4}_{-1.8}) \times 10^{11}$	16.3 ± 3.9
1HERMES X1.4 J021937.20−042931.2*...	0.295^D	2	...^{U}	191.0 ± 17.0	−	−	38.8 ± 12.4	56.8 ± 7.0	58.5 ± 5.3	−	$(5.1^{+2.5}_{-1.6}) \times 10^{10}$	9.10 ± 0.9
1HERMES X24 J032628.73−293757.4...	0.326^D	4	...^{S}	1073. ± 17.9	−	−	27.6 ± 3.7	4.0 ± 7.7	...	...	$(3.3^{+0.6}_{-0.5}) \times 10^{11}$	38.9 ± 1.1
1HERMES X24 J032713.82−271332.8...	1.042^L	4	...^{S}	411.9 ± 16.9	−	−	13.9 ± 3.8	4.1 ± 3.7	...	...	$(2.2^{+0.8}_{-0.6}) \times 10^{12}$	58.2 ± 3.5

Notes. Source names indicate both detection band in the XID catalog (24 μm or 1.4 GHz) and source position based on near-infrared/optical counterparts. Asterisks after the name indicate the source was observed in near-photometric conditions. Superscripts in the z_{spec} column indicate the instrument with which observations were obtained, LRIS (L) or DEIMOS (D). The CONF column indicates the confidence in the spectroscopic redshift identification, and ranges here from 1 to 5, 5 being the most confident (further identifications at CONF < 1 have been excluded from this paper). Photometric redshift superscript indicates which field the data are in, UDS (U), CDFS (S), COSMOS (C), LHN (L), GOODS-N (G), or Elais-N1 (E). PACS flux densities come from PEP (Lutz et al. 2011) in COSMOS and from HerMES (Oliver et al. 2012) in LHN. Ellipsis (...) denote that the given source is undetected at the corresponding wavelength or has no photometric redshift despite having the necessary imaging or coverage whereas dashes (−) denote that no data exist.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)
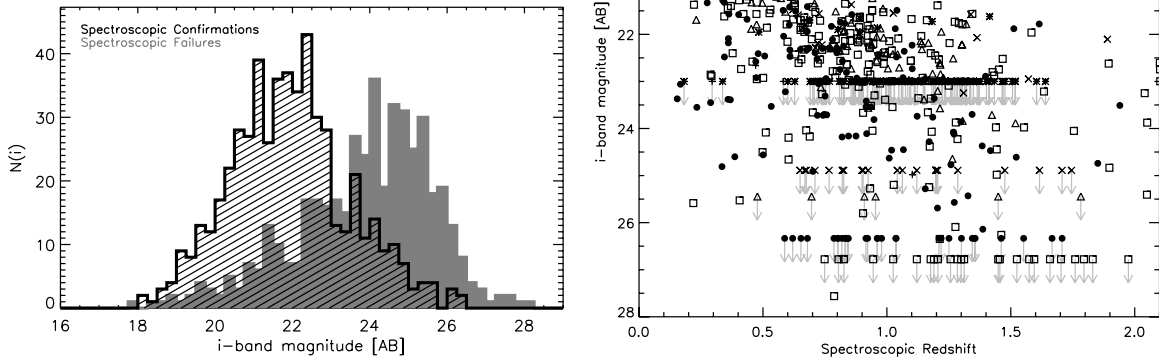
Figure 7. Left: distribution in *i*-band magnitudes of spectroscopically confirmed (hashed area) and unconfirmed sources (filled gray area). There is clearly a bias in sources with spectroscopic confirmations being brighter at *i*-band magnitude, also much more likely to sit at lower redshifts. However, note that for sources not included on this plot—those undetected in *i* band—the spectroscopic success rate is 39%, lower than the same fraction for sources with *i*-band detections, 54%, although not as low as one might expect for very optically faint sources. This is likely caused by many sources being confirmed through bright emission lines that can contribute minimally to broad-band photometry. Right: the *i*-band magnitude against redshift for sources with confirmed redshifts. At the bright end, $i < 22$, sources show a clear trend with redshift, while many sources (marked as upper limits, according to the imaging depths in each field) have no available *i*-band photometry. Clearly *i*-band magnitude strongly impacts the likelihood of measuring a spectroscopic redshift, a concept that is explored when measuring the spectroscopic completeness, shown in Figure 12.
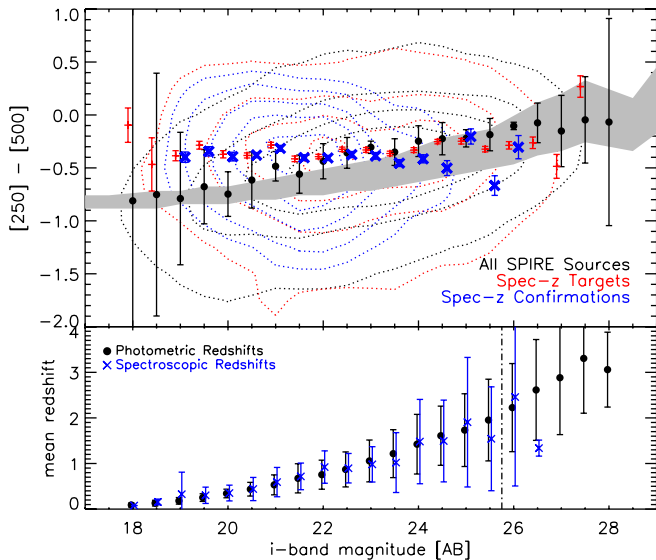
Figure 8. Relationship between *i*-band magnitude and SPIRE color, $[250] - [500] = -\log(S_{250}/S_{500})$. SPIRE color becomes redder with fainter magnitudes, which is expected given the redshift–*i*-band magnitude relationship seen in Figure 7 and an assumption of constant dust temperature from low to high redshifts. The locus of all $>3\sigma$ SPIRE sources in the upper panel consists of $\sim$5000 sources in the COSMOS field (black dotted contours and black points, the mean SPIRE color for *i* magnitude bins). Red dotted contours and points represent the distribution of spectroscopic targets ($\sim$1600 sources across all fields) and blue represent sources with spectroscopic confirmations, clearly skewed toward brighter *i*-band magnitudes. In the bottom panel we compute the mean photometric and spectroscopic redshift for the same colored samples from the plot above, and see agreement down to $i_{AB} \approx 25.5$ (vertical dot-dashed line). In the top panel, the gray band represents the expected SPIRE colors at the mean photometric redshift per *i*-band magnitude bin for dust temperatures ranging from 30 K to 50 K. Within uncertainty, the full sample follows this expectation, while the spectroscopic sample deviates toward redder colors at brighter *i*-band magnitudes.

(A color version of this figure is available in the online journal.)

entire parent galaxy sample of $>3\sigma$ SPIRE sources, spectroscopic targets, and also spectroscopic confirmations. This indicates that SPIRE colors should become redder with fainter *i*-band magnitude, assuming a constant temperature and redshift range. The mean redshifts—photometric and spectroscopic—per *i*-band bin are self-consistent, indicating no strong bias in the spectroscopic redshift distribution with *i* magnitude. Note however that the analysis as presented in Figure 8 excludes sources (a) without photometric redshifts and (b) without *i*-band counterparts. The variation of SPIRE color with redshift, along with the perceived bias of our spectroscopic sample toward redder colors at bright magnitudes, is addressed further in Section 3.5.

Of the spectroscopically confirmed sub-sample, $65\% \pm 15\%$ are $250\,\mu$m peaking (meaning $S_{250} > S_{350} > S_{500}$), $23\% \pm 14\%$ are $350\,\mu$m peaking ($S_{350} > S_{250}$ and $S_{350} > S_{500}$), and $13\% \pm 9\%$ are $500\,\mu$m peaking ($S_{250} < S_{350} < S_{500}$). These values are comparable to the spectroscopic targets: $72\% \pm 12\%$ were $250\,\mu$m peaking, $17\% \pm 9\%$ were $350\,\mu$m peaking, and $11\% \pm 6\%$ were $500\,\mu$m peaking. This leads us to conclude that successful spectroscopic identification does not have a bias with respect to FIR SED shape.

3.2. Redshift Distribution

The redshift distribution of the 767 spectroscopically confirmed HSGs in this survey is shown in Figure 9. The vast majority, 731, are at $z < 2$. The distribution peaks at $z = 0.85$ with a tail of sources extending out to higher redshifts, discussed fully in a separate accompanying paper, C12. Also plotted is the distribution in photometric redshifts for all spectroscopic targets. The photometric redshift distribution peaks at the same epoch, but with a much higher fraction of sources at $z \gtrsim 2$. The deficit in spectroscopic redshifts at $z > 2$ is caused by (1) the DEIMOS "redshift desert" (DEIMOS observations comprise 66% of our sample and no strong emission lines are visible within DEIMOS wavelength coverage from $1.6 < z < 3.2$); and (2) by a decreasing spectroscopic completeness due to enhanced obscuration in
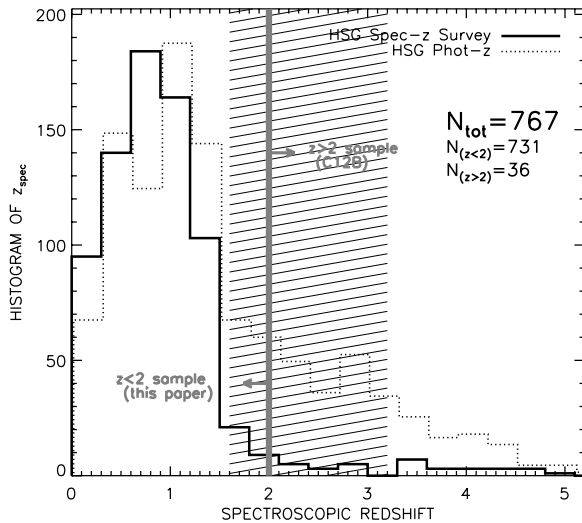
Figure 9. Spectroscopic redshift distribution of HSGs. Since the majority of our 767 spectroscopically confirmed sources were observed with DEIMOS (66%), the redshift range $z = 1.6$–3.2 (hashed area) is sparsely populated (only two sources within this range are DEIMOS identified, both bright quasars). Distribution of photometric redshifts for the sample of spectroscopically targeted HSGs is shown as the dotted line, consistent at $z \lesssim 1.5$ but suggesting a significantly decreased spectroscopic yield at $z > 1.5$.

the rest-frame ultraviolet relative to rest-frame optical. The latter point relates to the very obscured nature of infrared-luminous galaxies: The increased presence of dust implies more significant extinction in the UV and optical, with more substantial effects at bluer wavelengths.

Note that predicted redshift distributions for *Herschel* sources have been studied in detail by Amblard et al. (2010) and Béthermin et al. (2011). However, these are based on different SPIRE selection methods and are thus not directly comparable to the HSG sample discussed herein. For instance, the Amblard et al. distribution of $350 \, \mu m$ selected sources peaks at $z \sim 2.2$, as seen in the upper panel of Figure 10. We extract HSGs from our sample that would satisfy their detection criteria ($S_{350} > 35$ mJy, with $250 \, \mu m$ and $500 \, \mu m$ S/N > 3) for comparison and find a statistically distinct distribution from Amblard et al.

We also compare data with the model predictions of the expected SPIRE distributions from Lagache et al. (2004), Negrello et al. (2007), and Béthermin et al. (2011). While our data (both of the limited $S_{350} > 35$ mJy sample and the distribution of all photometric redshifts) are inconsistent with most models (particularly at the $z > 2$ end), our results are most consistent (tested via a Kolmogorov–Smirnov statistic) with the predicted distribution of Lagache et al. (2004); the number of sources we observe at $z < 1$ is notably different than the predictions from Negrello et al. (2007) or from the long-wavelength photometric redshift based work of Amblard et al. (2010).

In the middle panel of Figure 10 we compare our results to those of Chapin et al. (2011) for BLAST-detected galaxies in ECDF-S, a combination of photometric redshifts (Dunlop et al. 2010; Ivison et al. 2010a) and spectroscopic redshifts (Casey et al. 2011a). Their selection is based on three different flux cuts at 250, 350, and $500 \, \mu m$, and when applying the same selection to our sample, we find an agreement between the median redshifts, both of which have $\langle z \rangle \approx 1$ but with differences at both low and high redshift. There is an additional peak in the Chapin et al. sample at $z < 0.5$ that is not observed in our sample. This could be due to cosmic variance, since the
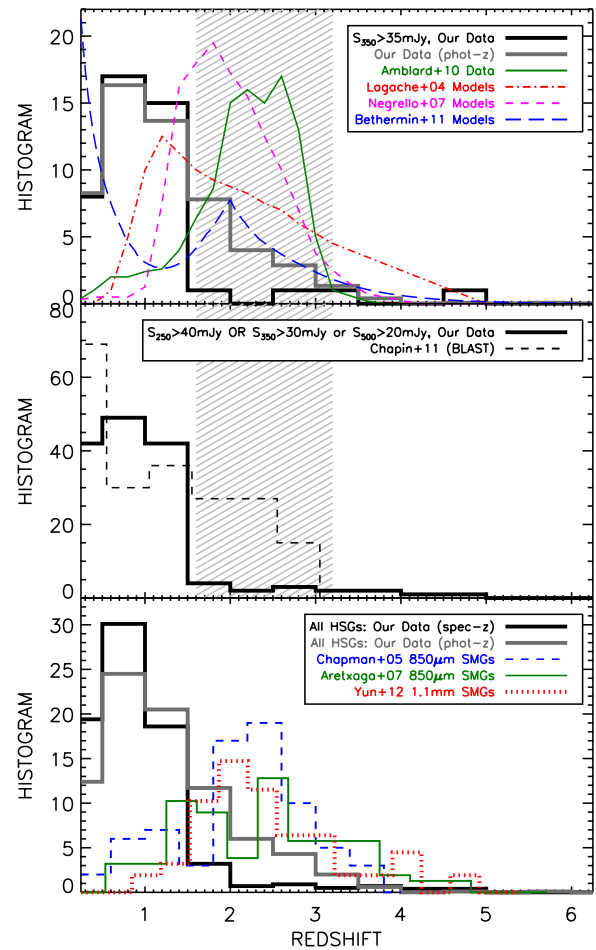

Figure 10. Comparison of our redshift distributions to those in the literature; In the top panel, we compare to the Amblard et al. (2010) distribution (green) for sources selected at $350 \, \mu m$ above 35 mJy (with $>3\sigma$ significant detections at both $250 \, \mu m$ and $500 \, \mu m$). We also compare the full distribution of photometric redshifts of our targets (thick, solid gray) to the model distributions from Lagache et al. (2004), Negrello et al. (2007), and Béthermin et al. (2011; dot-dashed red, dashed magenta, and long-dashed blue, respectively). Neither observed distribution (spectroscopic and photometric) shows good agreement with the models, however the Lagache et al. model (which peaks at $z \sim 1$) agrees best with our data. The middle panel compares to the redshift distribution of Chapin et al. (2011) for BLAST-bright galaxies in ECDF-S. Given the limited number of sources in the BLAST analysis (and that they are all identified in ECDF-S, a small volume), we conclude that these two distributions are consistent within uncertainty. In the bottom panel, we compare the HSG redshift distribution to those of $850 \, \mu m$–1.1 mm selected SMGs, which peak at $z \approx$ 2.2–2.6 due to the longer wavelength selection. The HSG distribution is taken from Figure 9 and scaled down 10 times in total number for direct comparison.
(A color version of this figure is available in the online journal.)

BLAST ECDF-S sample probes a small volume in a single field, while our data sample multiple deep fields over a larger sampling volume. Given the DEIMOS redshift desert, comparison cannot be drawn fairly beyond $z \sim 1.6$.

Worth noting is the contrast of these redshift distributions with the $850 \, \mu m$ selected SMG redshift distribution (Chapman et al. 2005; Yun et al. 2012), which peaks at $z \sim 2.2$ (see the bottom panel of Figure 10). The peak in $850 \, \mu m$ selected galaxies occurs at earlier epochs due to the selection wavelength: $850 \, \mu m$ is more sensitive to lower luminosity/colder sources at higher redshifts and less sensitive to warmer sources at low redshift. SPIRE selection is not constant in luminosity out to very high redshifts like $\sim$1 mm selection is, but it has the benefit of probing the peak of the infrared SED at $z \sim 1$–2, thus is not
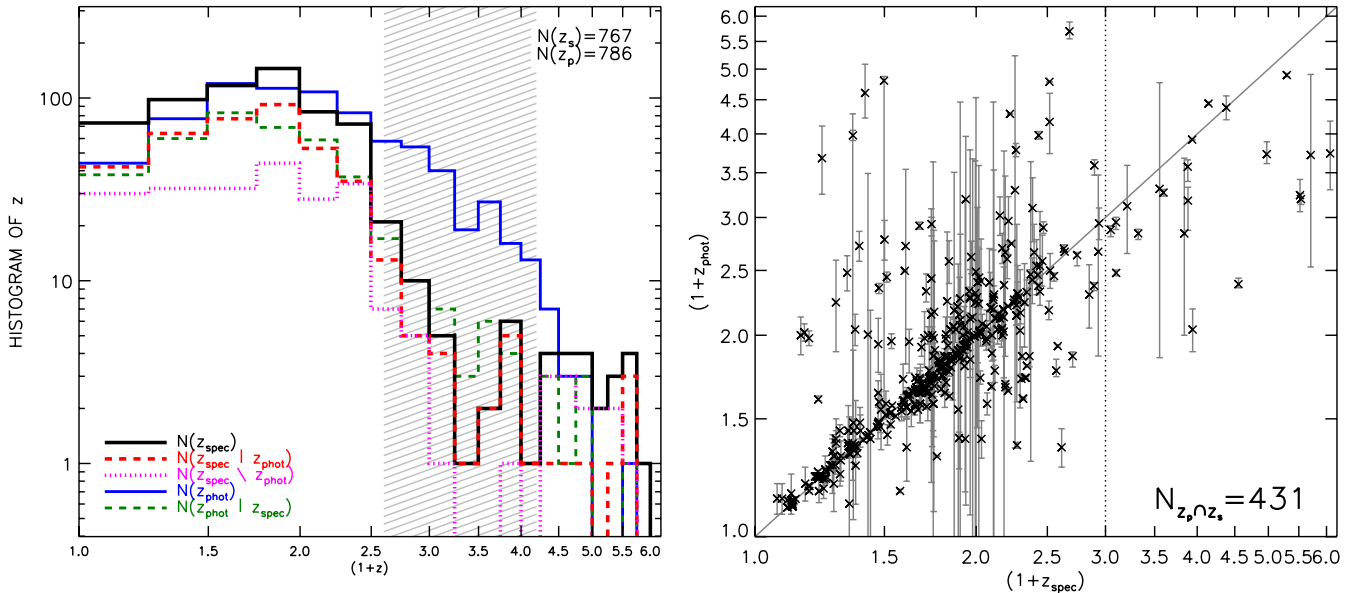
Figure 11. Left: redshift distributions for HSGs shown on a logarithmic scale: the full spectroscopic sample (solid black); spectroscopically confirmed sources with photometric redshifts ($z_{\rm spec} \mid z_{\rm phot}$, in other words "$z_{\rm spec}$ *given* $z_{\rm phot}$," dashed red); spectroscopically confirmed sources without photometric redshifts ($z_{\rm spec} \setminus z_{\rm phot}$, in other words "$z_{\rm spec}$ *take − away* $z_{\rm phot}$," dotted magenta); photometric redshifts (solid blue); and photometric redshifts for spectroscopically confirmed sources ($z_{\rm phot} \mid z_{\rm spec}$, in other words "$z_{\rm phot}$ *given* $z_{\rm spec}$," dashed green). The DEIMOS redshift desert is shown as a hashed area. Right: spectroscopic vs. photometric redshift for successfully identified targets *which have* photometric redshifts (431 sources out of 767). The photometric redshifts were taken from Strazzullo et al. (2010; Lockman Hole North), Rowan-Robinson et al. (2008; ELAIS-N1, UDS), Ilbert et al. (2010; COSMOS), and Cardamone et al. (2010; ECDF-S) with associated uncertainties on $z_{\rm phot}$ shown. The mean rms scatter at is $\Delta z/(1 + z_{\rm spec}) = 0.29$, notably worse than the expected photometric redshift accuracy of photometric redshifts of "normal" starburst galaxies.

(A color version of this figure is available in the online journal.)

likely to miss or select against sources at a given luminosity due to their warmer SED shape.

3.3. Comparison to Photometric Redshifts

Photometric redshifts are far easier to obtain on larger samples of galaxies than spectroscopic redshifts, and whenever the latter become available, it is important to test the reliability of the former. Many large statistical studies are now motivated exclusively by use of large catalogs of photometric redshifts; here we explore possible underlying biases that might persist in HSG photometric redshifts and what might cause them. Note that all photometric redshifts used in this paper are calculated from ultraviolet through near-infrared photometry and exclude any long-wavelength data $> 8\,\mu$m (the catalogs used in this analysis are described at the end of Section 2).

Spectroscopic and photometric redshifts are plotted against one another in Figure 11. From this we measure that HSG photometric redshifts have a mean rms scatter of $\sigma = \Delta z/(1 + z_{\rm spec}) = 0.29$, characteristically a factor 3–4 times worse than the photometric redshifts for regular field galaxies (Ilbert et al. 2010 measure $\Delta z/(1 + z_{\rm spec}) = 0.07$ for galaxies with spectral confirmations in COSMOS). This might be surprising given that the majority of sources in this comparison come from fields with deep, multi-band photometry (e.g., COSMOS, LHN). Why are the photometric redshifts of infrared-bright galaxies substantially worse than those of most field galaxies?

One might think that the disagreement originates from source blending or mismatching due to the large beamsize of infrared observations. However, we do not expect blending or mismatching to occur since the photometric redshifts taken from the source catalogs are matched to our spectroscopic targets' position within $\lesssim 1''$ (i.e., even if a SPIRE source's counterpart is mistakenly identified, we would still expect the false counterpart's photometric redshift to agree with its spectroscopic

redshift). Since mismatched counterparts are not likely to be the source of the photometric redshift scatter, the scatter is likely due to some other property of the infrared-selected sample.

If the $z_{\rm phot} - z_{\rm spec}$ disagreement is not caused by mismatched counterparts, then what intrinsic physical processes could explain less reliable photometric redshifts in infrared-selected galaxies? Direct detection at far-infrared wavelengths implies two things about a galaxy: (1) it has a significant dust reservoir, which has absorbed more energetic light and re-radiated it in the far-infrared, and (2) it has a high SFR (since FIR luminosity scales directly to SFR; Kennicutt 1998a). The first point highlights that the galaxy's emission at rest-frame ultraviolet and optical wavelengths—emitted by young, hot stars—is being obscured and scattered by dust; as a result, infrared-selected galaxies are optically fainter than their less-dusty counterparts of similar redshifts and stellar masses. The second point that infrared galaxies have intrinsically higher SFRs than most "normal" field galaxies follows from the implied infrared luminosities of the infrared-selected samples. The vast majority of field galaxies have SFRs $< 10\,M_\odot$ yr^{-1} (e.g., the vast array of 200 K galaxies in COSMOS which are fit with stellar population templates from Bruzual & Charlot 2003). In contrast, most infrared-selected galaxies satisfy $L_{\rm IR} > 10^{11}\,L_\odot$, which implies SFRs of $> 17\,M_\odot$ yr^{-1}, or $L_{\rm IR} > 10^{12}\,L_\odot$, which implies $> 170\,M_\odot$ yr^{-1}. Higher SFRs translate to brighter rest-optical emission lines and higher line-to-continuum ratios.

How might dust obscuration and enhanced line-to-continuum ratios impact the reliability of photometric redshifts for HSGs? The first effect is straightforward: The fainter a galaxy is in the optical, the more difficult it is to put a reliable constraint on its photometric redshift. Since HSGs are optically fainter than "normal" galaxies that harbor less dust, this is one reason HSG photometric redshifts are less reliable. The second effect comes from the "contamination" of bright emission lines of

the optical broadband filters used to compute the sources' photometric redshifts. Although emission lines can contaminate the broadband magnitudes for both normal and dusty galaxies, dusty galaxies have comparably fainter continuum (or rather, a more significant contribution from emission lines). Also, despite some efforts to account for emission lines in photometric redshift code algorithms (by including them in the templates used for the fitting), none of the existing optical templates in the literature (e.g., Calzetti et al. 1994; Bruzual & Charlot 2003) include models with extremely high SFRs and dust content as exists in HSGs, hence the poorer $z_{phot} - z_{spec}$ agreement.

The influence of dust obscuration on photometric redshift estimates is not completely straightforward in that dust is not necessarily expected to extinct optical flux uniformly across all wavelength regimes. Sources with high phot-z's and low spec-z's can be explained by differential blue-to-red obscuration whereby the galaxies dropout in several blue bands due to stronger absorption of higher energy photons. The handful of sources with low photometric redshifts and high spectroscopic redshifts ($z > 2$) are described in detail in C12; these are thought to disagree due to differential obscuration of resonant emission-line photons and continuum (e.g., Neufeld 1991), whereby emission lines are not extincted as significantly as stellar emission continuum. This is also highly dependent on viewing geometry, but the observation holds that photometric redshifts of dusty galaxies are significantly worse than they are for normal field galaxies, which contain less dust. Caution should be exercised when using large photometric data sets to quantify the aggregate properties of infrared-selected samples.

3.4. Spectroscopic Completeness

With a lack of any better constraint on the redshift distribution of unconfirmed HSGs, we use our targets' photometric redshifts (those with and without spectroscopic redshifts) to constrain spectroscopic completeness as a function of redshift. Figure 12 (top panel) contrasts the spectroscopic and photometric redshift distributions by instrument, since LRIS and DEIMOS observations have different wavelength coverage, thus different redshift completeness levels. Measuring spectroscopic completeness requires an estimate of the underlying redshift distribution for the whole population, including those without photometric redshifts. The only constraint we have on sources without photometric redshifts is from those spectroscopically observed, shown in Figure 11 (left panel), consistent with the distribution in photometric redshifts. Therefore, a scaling correction factor, C, is applied to the distribution of photometric redshifts to account for the sources without photometric redshifts. Although this assumption might not be entirely correct (i.e., there might be a much higher fraction of $z > 2$ sources that do not have photometric redshifts), the assumption that the redshift distribution is weighted at lower redshifts effectively places the lower limit of spectroscopic completeness at $z < 2$. If the fraction of sources without photometric redshifts is higher at high-z, then the spectroscopic completeness measurement at $z < 2$ will increase. The spectroscopic completeness is measured as a function of redshift by dividing the distribution in spectroscopic redshifts by that of the scaled up photometric redshifts (i.e., $N(z_{spec})/(C \times N(z_{phot}))$). This spectroscopic completeness estimate is shown in the bottom panel of Figure 12. As expected, it declines over $0 < z < 2$ and then is unconstrained at $z > 2$ due to limited samples. The DEIMOS completeness estimate has two "outlier" points at $z \sim 0.8$ and $z \sim 1.3$; this is caused by the enhanced sensitivity of DEIMOS from 6500–7400 Å (where
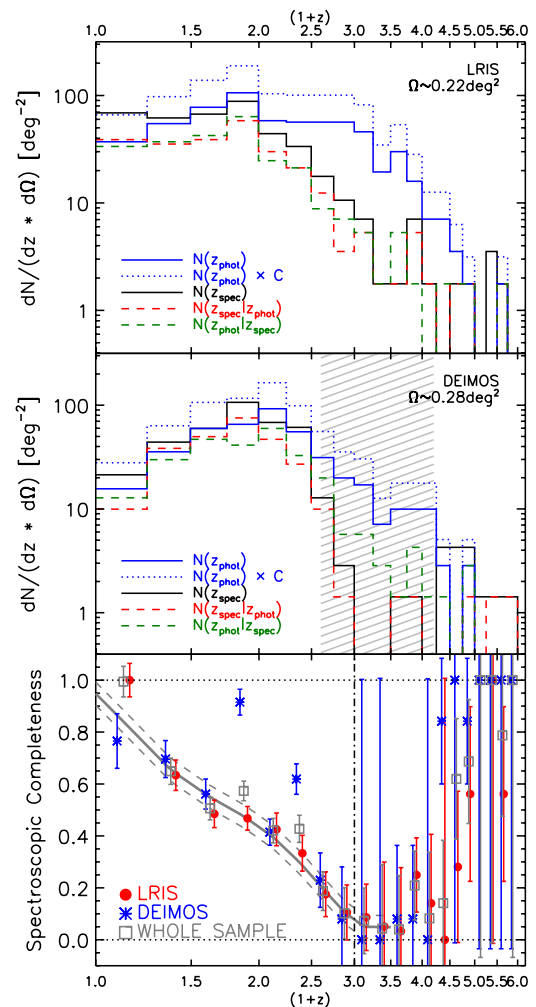

Figure 12. Top and middle panels show the redshift distributions of HSGs split into LRIS and DEIMOS samples. DEIMOS redshift desert is marked in the middle panel. Assuming the redshift distribution for sources without photometric redshifts resembles the distribution in photometric redshifts (see Figure 11, left panel), then we can estimate the spectroscopic completeness of our survey, bottom panel, by dividing the spectroscopic redshift distributions (by instrument) by the photometric redshift distribution scaled up to account for sources without photometric redshifts. From this, we determine that survey completeness cannot be constrained at $z > 2$ (due to small number statistics), which is why that sample is discussed in a separate paper. The best-fit curve at $z < 2$ (plotted in gray with dashed lines showing uncertainty) is used later to estimate the total contribution of HSGs to the SFRD. The two "outlier" high-completeness DEIMOS points are discussed in the text.

(A color version of this figure is available in the online journal.)

LRIS often has a gap in coverage), and 7600–9000 Å (where our LRIS observations were slightly more prone to skyline contamination). Since the completeness becomes unconstrained at $z > 2$, we choose $z = 2$ as a natural boundary at which to divide the sample, addressing the well-constrained $z < 2$ population in this paper, while discussing the full sample of spectroscopically incomplete $z > 2$ sources in C12.

3.5. SPIRE Color with Redshift

Recalling from the discussion in Section 3, our spectroscopic HSGs sample the underlying distribution of all SPIRE sources well in SPIRE color, despite a bias in i-band magnitude (where brighter optical sources are more likely to be spectroscopically confirmed). Figure 8 showed that the mean redshift and SPIRE color per i-band magnitude of our spectroscopic HSGs is
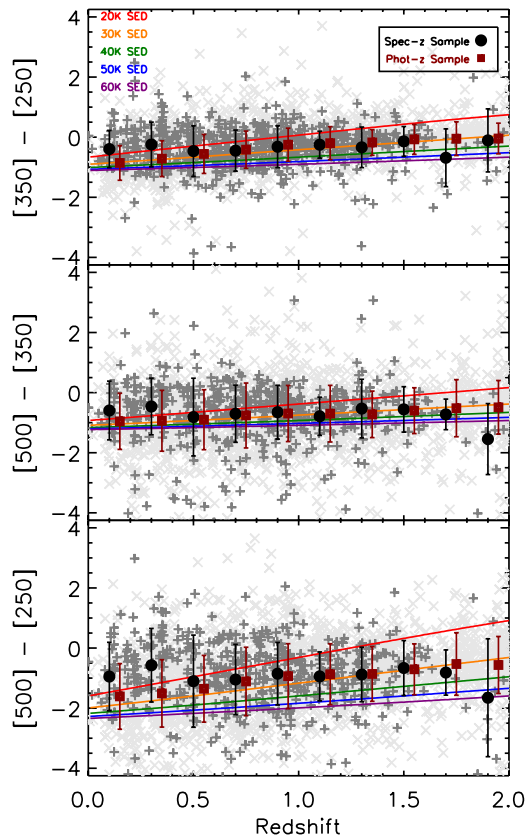
Figure 13. *Herschel*-SPIRE color with redshift, where color is given as a magnitude difference between SPIRE bands, such that high values correspond to red colors. In light gray are SPIRE colors of the entire photometric HSG sample from the COSMOS field; dark gray points are our spectroscopic sample. The mean of the spectroscopic sample and photometric sample is shown as black circles and dark red squares, respectively, with associated uncertainties. Overplotted are tracks of SEDs with fixed dust temperature, 20 K (red), 30 K (orange), 40 K (green), 50 K (blue), and 60 K (purple). We see no significant color evolution in the HSG sample with redshift; this is primarily due to selection properties of the sample.

(A color version of this figure is available in the online journal.)

drawn from the parent sample of SPIRE sources without significant bias. While this hinted at the evolution of SPIRE color with redshift, we can test this evolution directly using our spectroscopic HSGs and larger photometric redshift samples from COSMOS (Ilbert et al. 2010).

Assuming that most infrared starbursts can be described with a 35 K dust SED (Chapman et al. 2005; Rieke et al. 2009), one would expect low-redshift sources to have blue SPIRE colors and high-redshift sources to have red SPIRE colors (e.g., Cox et al. 2011; Roseboom et al. 2012). Figure 13 plots the [350]–[250], [500]–[350], and [500]–[250] colors against redshift, where [250] is defined as a 250 μm magnitude, i.e., $-2.5\log(S_{250})$. We see no strong trend in SED shape with spectroscopic redshift. While it might be expected that high-z sources are significantly more "red" than "blue" low-z sources, this interpretation is simplistic since it assumes no evolution in dust temperature, no relationship between dust temperature and luminosity, and ignores the impact of population selection effects. Overplotted in Figure 13 are the color–redshift tracks for SEDs of fixed dust temperature systems from 20–60 K; they evolve more strongly with redshift than our data, particularly notable in [500]–[250] color.

The dispersion on the SPIRE colors in both the spectroscopic and photometric samples implies a wide range of dust tempera-

tures. This highlights that far-infrared flux densities and colors cannot be used exclusively as a proxy for redshift. Notwithstanding the large dispersion in SPIRE colors, two main observations can be drawn from Figure 13: (1) that the mean SPIRE colors do not redden significantly with redshift as might be expected; and (2) at low-z ($z < 0.5$) our spectroscopic sample is redder, i.e., cooler than the photometric redshift sample.

Both of these observations can be understood by investigating how SPIRE selection works in the L_{IR}–T_{dust} plane, discussed fully in Section 4.4. Dust temperature increases with luminosity (which is partially a selection effect against high-z cold systems and partially thought to be a real correlation). Since the highest-redshift sources in our sample are significantly more luminous than our lower redshift sources, the correlation between L_{IR} and T_{dust} translates to a correlation between z and T_{dust}. This results in roughly constant SPIRE colors with redshift. At $z < 0.5$, we note that our spectroscopic survey is limited in the number of very luminous ($\gtrsim 10^{12}\ L_\odot$) galaxies it can detect simply due to sky area probed ($\approx 1\ \mathrm{deg}^2$); since (in Section 4.4) we will observe that more luminous HSGs are hotter, it is natural that a sample of HSGs with photometric redshifts (covering a much larger area than our spectroscopic survey) will probe naturally warmer HSGs. This gives rise to the "redder" $z < 0.5$ spectroscopic sample. Although the two redshift bins at $1.6 < z < 2.0$ seem bluer than at any other redshift, we note that the number of sources in these redshift bins is significantly fewer (and the observation is not the same for the much larger sample of sources with photometric redshifts in that range).

4. RESULTS

4.1. Summary of Spectra and Redshifts

Table 1 lists the positions, redshifts, SPIRE flux densities, 24 μm and radio flux densities, luminosities, and dust temperatures for the full sample of $z < 2$ spectroscopically confirmed HSGs. Sample spectra of a variety of spectroscopically confirmed targets of the $z < 2$ confirmed HSG sample are shown in Figure 14. Of 1594 targets with $>3\sigma$ SPIRE detections, 767 have spectroscopically identified redshifts from these data, 731 of which are at $z < 2$. By field, there are 288 in COSMOS, 163 in LHN, 119 in GOODS-N, 139 in EN1, 51 in CDFS, and 8 in the UDS. The total sky area probed by this redshift survey is $\sim 0.93\ \mathrm{deg}^2$.

Besides the selection and spectroscopic biases mentioned in the previous section, the spectroscopic yield for sources varied strongly as a function of observing conditions and integration time, anywhere from 15% to 80%. However, unlike the selection and spectroscopic biases, this variation in yield is much easier to correct. While the full 0.93 deg^2, 767 source sample is used in most analysis, only the sources surveyed under photometric conditions to full survey depth, i.e., 13 LRIS masks and 16 DEIMOS masks, covering an area of 0.43 deg^2, are used to measure quantities that rely on accurate measurements of source density (e.g., the integrated infrared LF and star formation rate density, SFRD).

4.2. Infrared SED Fitting

We note that best-fit SED template fits from Chary & Elbaz (2001), Dale et al. (2001), Dale & Helou (2002), Siebenmorgen & Krügel (2007), or Draine & Li (2007) recover L_{IR} well, however we choose not to fit our data to these complex models given the limitations of our data and the variation in SED shapes. As discussed in Casey (2012) at length, the templates should be
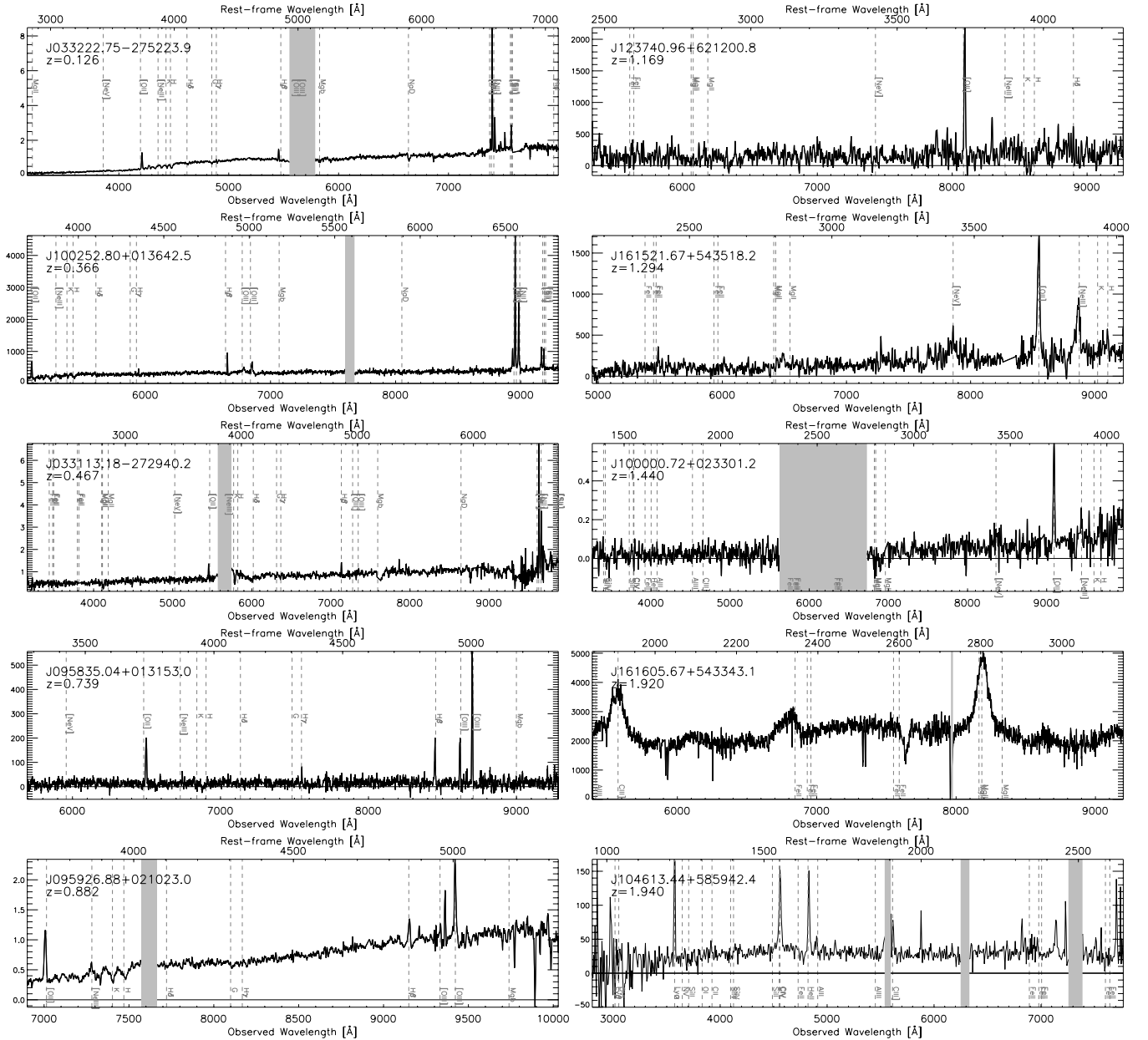
Figure 14. Ten sample spectra of *Herschel*-SPIRE selected galaxies ranging in redshift from $z = 0.126$ to $z = 1.940$. These sample spectra span a range of spectral types and include both LRIS and DEIMOS observations. The flux scales are in arbitrary, uncalibrated flux units. Spectra of the $z > 2$ HSGs are given in full in C12. Individual spectra for sources given in Table 1 are available on request.

used with caution when the number of free parameters in the templates exceeds the number of data points; this is certainly the case for those sources that have three to five photometric FIR data points. The dust temperature range of these SED templates is restricted to the 20 K $< T_{\rm dust} <$ 60 K range and quantized. For these reasons, we opt to use the functional fits in lieu of FIR SED templates.

Measuring infrared luminosities and dust temperature for SPIRE galaxies is performed by fitting a modified blackbody SED extrapolated over the rest-frame FIR to existing photometric data. In addition to the SPIRE flux densities, we use photometry from *Herschel*-PACS at 100 μm, and 160 μm in LHN and COSMOS, and *Spitzer*-MIPS at 24 μm, and 70 μm where available.

We fit the data for each galaxy to a modified blackbody fit of general opacity coadded to a mid-infrared power law, where the flux density, S_ν at rest-frame frequency ν is represented by

$$S_\nu \propto (1 - e^{-\tau(\nu)})B_\nu(T)$$
$$= N_{\rm bb}\frac{(1 - e^{-\tau(\nu)})\nu^3}{e^{h\nu/kT} - 1} + N_{\rm pl}\nu^{-\alpha}\, e^{-(\nu/\nu_c)^2}, \quad (1)$$

where $\tau(\nu) = (\nu/\nu_0)^\beta$ is the optical depth (Draine 2006), β is the emissivity, the physical dust temperature is T, and $\nu = c/\lambda_0$ ($\lambda_0 = 200\,\mu$m) is the frequency where optical depth is unity (see Casey 2012 for full details on the SED fitting method). Note that the physical dust temperature, T, used in this equation is not the same as the dust temperature used for the remainder of the paper; we choose to measure a "peak-SED" dust temperature, called $T_{\rm dust}$, which is the dust temperature measured from the peak of the SED via Wien's law. $T_{\rm dust}$ is less dependent on model parameters than T, so is more easily
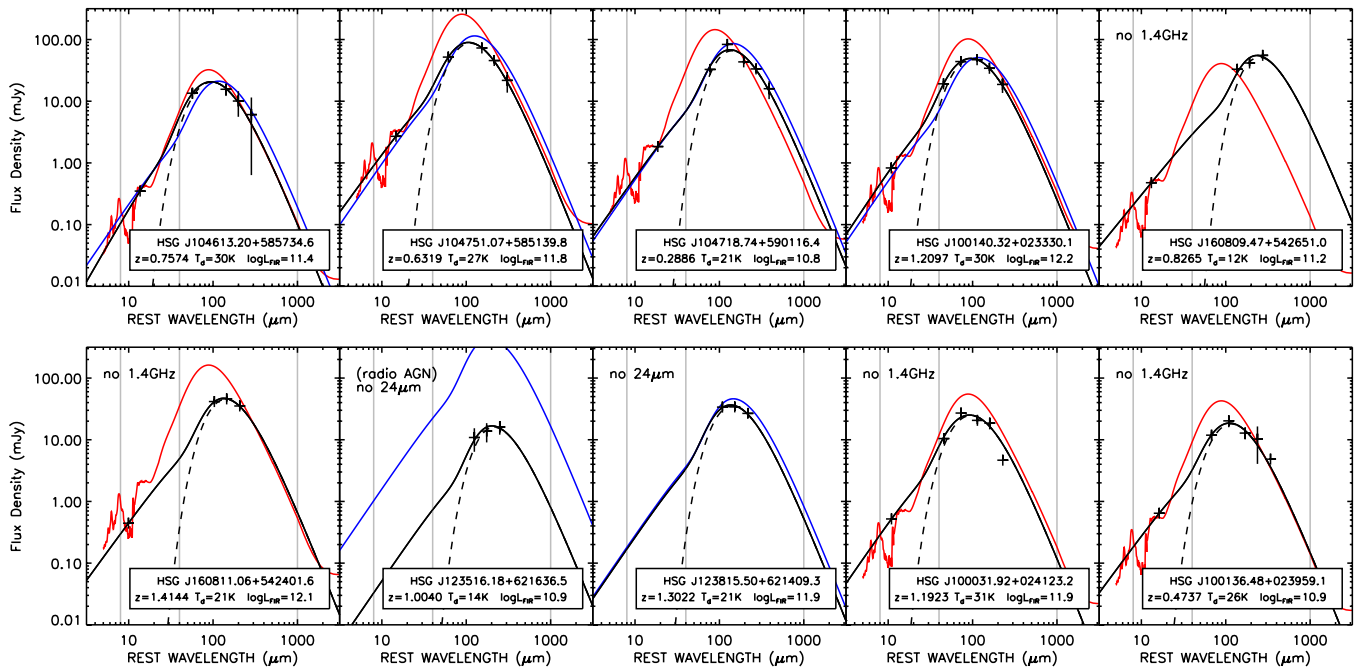
Figure 15. Ten example FIR SED fits to SPIRE-selected starbursts: three from LHN, three from COSMOS, two from EN1, and two from GOODS-N. The FIR data points (SPIRE and PACS, where available) are shown with 1σ uncertainties, and our best-fit modified blackbody SED is shown in black (see Section 4.2 for details on SED fitting). The underlying cold-dust blackbody dominating the FIR portion $>40\,\mu$m of our SED is shown as a dashed line. For sources which have *Spitzer*-MIPS coverage, we overplot the mean composite SED of SMGs from Pope et al. (2008) normalized to $24\,\mu$m (red). Sources with 1.4 GHz VLA coverage have another SED overplotted (blue): a modified blackbody inferred from assuming the radio/FIR correlation holds (with q_{IR} as given in Ivison et al. 2010b) with T_{dust} best fit to the data. Gray vertical lines mark rest-frame $8\,\mu$m, $40\,\mu$m, and $1000\,\mu$m. We note that the cold-dust blackbody only dominates the SED in the 40–$1000\,\mu$m wavelength range, although IR luminosity is computed across the range 8–$1000\,\mu$m, using the combination of IR modified blackbody and mid-infrared power law.

(A color version of this figure is available in the online journal.)

compared to dust temperatures measured via other means (see Casey 2012, Figure 2 for more detail). N_{bb} and N_{pl} are the coefficients of the modified blackbody and power-law terms, where N_{pl} is set such that the two functions are equal at the frequency, ν_c, where the gradient of the blackbody is equal to the slope of the power law, α. We allow α to vary with the range $0.5 < \alpha < 5.5$ for sources with $24\,\mu$m measurements (all sources except those in GOODS-N). If we have no mid-infrared measurements for the source, we fix $\alpha = 2.0$, which is the mean value for the rest of the SPIRE galaxies in our sample and a common value used throughout the literature (see Younger et al. 2009; Magnelli et al. 2010; Hilton et al. 2012). Since our observations cover a relatively narrow wavelength range without measurements on the Rayleigh–Jeans tail in the millimeter (from $\approx 850\,\mu$m–2 mm), we cannot constrain β in a meaningful way (except for very low redshift sources where the SPIRE points sit on the Rayleigh–Jeans tail). For consistency in our fitting technique, we decided to fix the emissivity to $\beta = 1.5$ (this is a commonly chosen value of β in the literature, e.g., Chapman et al. 2005; Pope et al. 2005; Younger et al. 2009). There is very little change in derived L_{IR} or T_{dust} by fixing β within the 1–2 range, so we determine this procedure to be reasonable.

Figure 15 illustrates 10 example SED fits using our method, contrasting them with both radio-implied FIR SEDs (from the radio/FIR correlation) and extrapolations from $24\,\mu$m assuming a nominal template for high-z ULIRGs, like the Pope et al. (2008) composite SED.

Table 1 lists the 8–$1000\,\mu$m IR luminosities and dust temperatures for the whole sample, and Figure 16 shows how the sources fall in luminosity versus redshift with respect to the various selection criteria. Note that the SPIRE selection criteria are not fixed in infrared luminosity since luminosity depends both
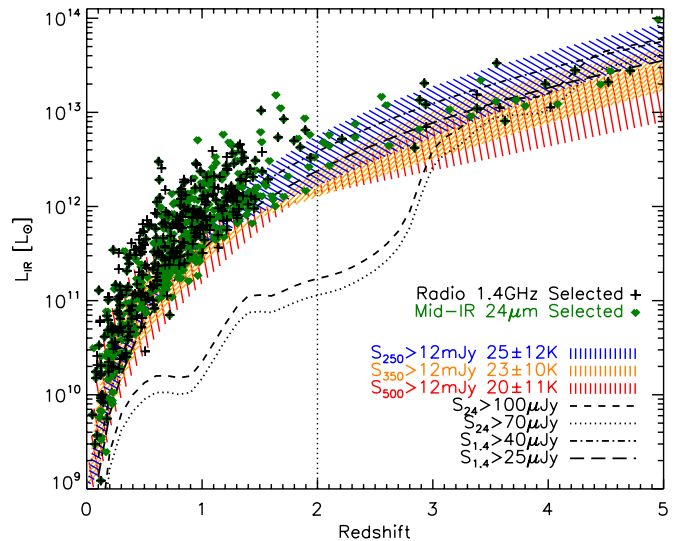

Figure 16. 8–$1000\,\mu$m integrated infrared luminosity against redshift for *Herschel*-SPIRE spectroscopically identified sources. The lower luminosity limits at the selection wavelengths are marked, both for the prior source catalogs at $24\,\mu$m ($S_{24} > 100\,\mu$Jy, dashed line) and 1.4 GHz ($S_{1.4} > 25$ or $40\,\mu$Jy, long dashed and dot-dashed lines), and SPIRE wavelengths (colored hashed regions). The detection limits at the SPIRE bands change with SED shape; we measure the dust temperatures of galaxies (via Wien's law) selected in each filter to determine the range of peak-SED dust temperatures to integrate between to calculate the IR luminosity limits at $250\,\mu$m (blue), $350\,\mu$m (yellow), and $500\,\mu$m (red). Sources selected as radio sources are marked as crosses and $24\,\mu$m sources are marked as green diamonds.

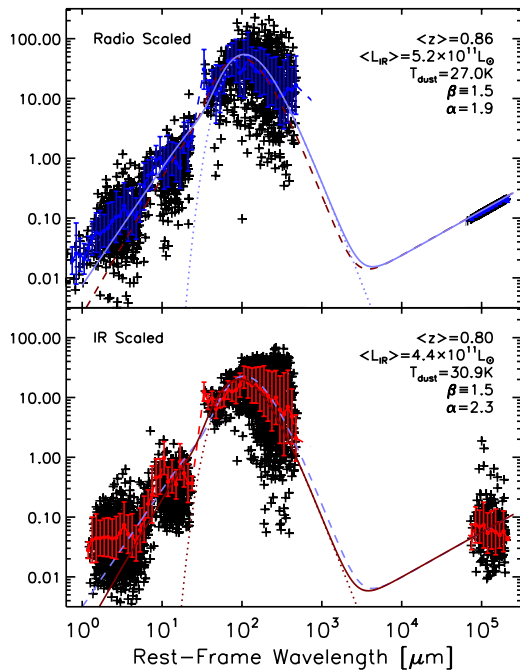(A color version of this figure is available in the online journal.)

Figure 17. Composite infrared to radio spectral energy distributions of HSGs scaled to their radio flux densities (top; 331 radio-detected HSGs) and integrated infrared luminosities (bottom; all 767 HSGs). The best-fit SED shapes (light blue, top, dark red, bottom) are remarkably similar between the two samples, with best-fit dust temperatures of 27 K and 30 K and mid-infrared slopes of 1.9 and 2.1, respectively.

(A color version of this figure is available in the online journal.)

on dust temperature and flux density. Similarly, the radio and $24\,\mu$m detection limits are not absolute, since survey depths at both wavelengths vary between fields.

Figure 16 shows all spectroscopic HSGs in L_{IR}–z space. Despite the fact that our sources are identified initially by $24\,\mu$m or radio source observations, the detection threshold in L_{IR}–z space corresponds well with the SPIRE $> 3\sigma$ luminosity limit alone (at $z < 3$, where *Spitzer* is effectively more sensitive than SPIRE).

4.3. Aggregate Infrared SEDs

While individual galaxies in our sample have at most five flux density measurements in the FIR, some with very low S/N, we can combine the measurements from many sources to infer the aggregate infrared properties of our sample in greater detail.

Figure 17 shows two composite near-infrared through radio SEDs for the entire $0 < z < 2$ HSG sample. The first comprises the 331 sources that are radio detected at $z < 2$: Their flux densities are scaled so that the K-corrected radio flux density equals the mean of the sample, $133\,\mu$Jy, at the mean redshift of the sample, $\langle z \rangle = 0.89$. The second composite SED is constructed of all 731 $z < 2$ sources, with flux densities re-normalized to the mean infrared luminosity, $4.1 \times 10^{11}\,L_\odot$, and redshift, $\langle z \rangle = 0.80$, of the sample. The two SEDs are remarkably similar, with dust temperatures of 27 K and 30 K, and mid-infrared slopes of $\alpha = 1.9$ and 2.1, respectively. The most noticeable difference lies in the near-infrared, where radio normalization seems to artificially wash out the stellar mass bump. The other observation to make is the spread in the radio flux densities in the IR scaled SED; although this is large ($\sim$2 dex, most falling within $\sim$1 dex scatter), we note that this could very well be due to individual source variation in
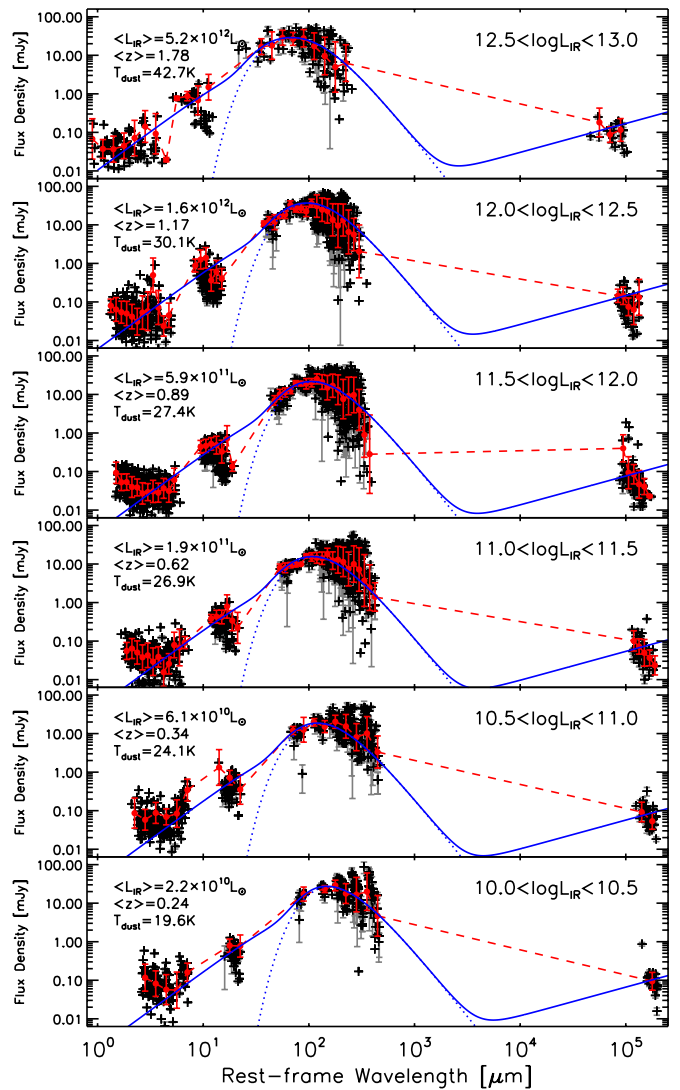

Figure 18. Composite infrared to radio spectral energy distributions of HSGs in six luminosity bins, from $L_{IR} = 10^{10}\,L_\odot$ to $10^{13}\,L_\odot$. Flux densities are re-normalized first to the mean luminosity and redshift of each bin. Infrared dust SEDs are fit to binned data (red points) from rest-frame $8\,\mu$m to $1000\,\mu$m, with fixed $\alpha = 2.0$ and $\beta = 1.5$. The radio portion of the SEDs is generated by assuming the FIR/radio correlation holds; overall, these radio SEDs agree with radio data. From the SED fits, we measure a steady increase in dust temperature with luminosity, from $T = 19$ K at $2 \times 10^{10}\,L_\odot$ at $\langle z \rangle = 0.23$ to $T = 46$ K at $5.2 \times 10^{12}\,L_\odot$ at $\langle z \rangle = 1.81$. We also see evidence for the evolution of near- to mid-infrared properties, most prominent in the highest luminosity bin, with detection of the $\approx 10\,\mu$m Si absorption feature (from a $24\,\mu$m flux density deficit of sources around $z \sim 1.4$).

(A color version of this figure is available in the online journal.)

the synchrotron slope or bright radio AGN. Since the mean value of the radio flux densities falls within ± 0.2 dex of the expected radio flux density (from the FIR radio correlation), we deduce agreement. The overall agreement between the two radio scaled and IR scaled SEDs re-enforces that (1) the FIR radio correlation seems to hold in this sample within uncertainty; and that (2) radio selection is not clearly biased in integrated infrared properties in comparison to $24\,\mu$m selected HSGs.

To test for luminosity and redshift evolution of infrared SED type, Figure 18 illustrates the near-infrared through radio SEDs of all $0 < z < 2$ sources split into several infrared luminosity bins. We measure a steady increase in dust temperature with luminosity (and also consequently redshift) from $\approx$20 K at

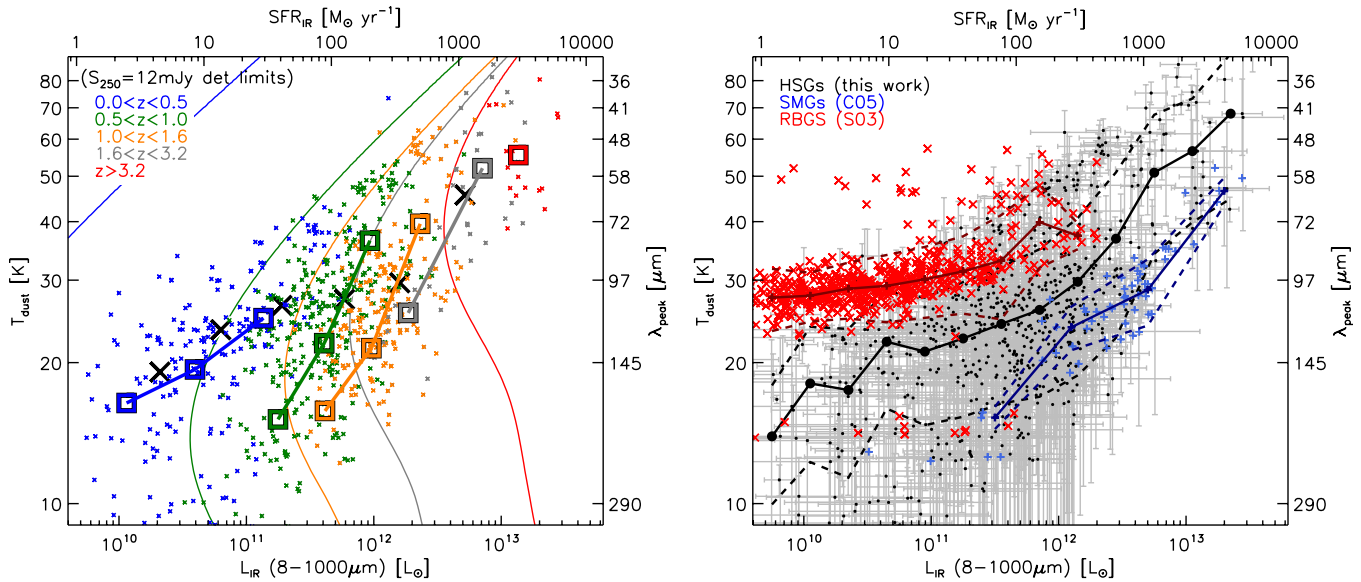**Figure 19.** L_{IR} against T_{dust} as an indicator of the variation of SED type in our sample. Dust temperature is estimated via Wien's Law (inversely proportional to peak of the SED). Colors on the left denote redshift bins, and the mean L_{IR}–T_{dust} relation for each redshift slice is shown as large squares. This relation changes with increasing redshift, however this can be attributable to the selection bias of SPIRE selection: the 250 μm detection limits at fixed redshifts are shown as solid lines, whose color corresponds to the lower redshift limit of the bin. At a fixed temperature, only sources to the right of the line (i.e., at higher luminosities) are detectable. The mean luminosities and temperatures for the composite SEDs in Figure 18 are large crosses. At right, we illustrate the L_{IR}–T_{dust} relation for SPIRE galaxies (black) in comparison to local *IRAS* galaxies, from the RBGS and GOALS samples (red; Sanders et al. 2003; Armus et al. 2009), and SMGs (blue; Chapman et al. 2005; Pope et al. 2006). We refit SEDs for all samples consistently so that both dust temperatures and luminosities are directly comparable. This demonstrates that our SPIRE galaxy sample is statistically colder than local *IRAS* galaxies of similar luminosities but warmer than SMGs of similar luminosities. Although we observe a different slope to the L_{IR}–T_{dust} relation than is seen locally, our sample is not large enough to measure evolution, since luminosity is largely a function of redshift (e.g., Seymour et al. 2010).

(A color version of this figure is available in the online journal.)

10^{10} $L_\odot$, to ≈ 30 K at 10^{12} $L_\odot$, and a jump to 45 K at $10^{12.5}$–$10^{13.0}$ $L_\odot$. Understanding whether or not this increase in dust temperature is due to redshift evolution, luminosity, or selection effects is difficult, due to the redshift dependence of luminosity; this issue is discussed more in context of the L_{IR}–T_{dust} plane in the next section.

An interesting conclusion can be drawn from the composite SEDs about the evolution of near- and mid-infrared properties. At lower luminosities (up to 10^{12} $L_\odot$), the near-infrared data $< 10\,\mu$m increases toward shorter wavelengths, indicative of emission from old stars (i.e., the stellar mass bump) with a local minimum at ~ 5–6 μm. In the highest luminosity bin, the prominence of the stellar bump disappears, as the emission seems to be dominated by a power law. Furthermore, the rest-frame mid-IR is unremarkable at low luminosities, largely consistent with the expected continuum flux density from our SED fit, while the highest luminosity bin shows a strong absorption feature at $\approx 10\,\mu$m, thought to be 9.7 μm Si absorption, that is seen in some of the most luminous local infrared galaxies, including Arp 220 (Charmandaris et al. 1997). This feature is detected in this composite SED only due to sources at the corresponding redshift ($z \sim 1.4$) having a 24 μm flux density deficit relative to sources at redshifts above and below $z \sim 1.4$.

4.4. The Temperature–Luminosity Plane

Figure 19 shows our sample in L_{IR}–T_{dust} space. Overall, dust temperature increases with luminosity, which is partially attributable to our selection and partially to a real physical effect. The selection effect (that more luminous sources are hotter) stems from the nature of selection at wavelengths on the Rayleigh–Jeans tail, where hot-dust SEDs are selected against (at comparable IR luminosities). This dust-temperature

selection effect was much more pronounced in the SMG population selected at 850 μm (see Blain et al. 2004; Chapman et al. 2004; Casey et al. 2009b; Chapin et al. 2009) than it is in SPIRE samples, which probe the dust SED closer to its peak and even beyond, at $z \gtrsim 3$. The solid lines in the left panel of Figure 19 represent SPIRE detection limits, or contours of fixed 250 μm flux density and redshift. When split into different redshift bins, the L_{IR}–T_{dust} relation appears to evolve, although this is primarily attributable to luminosity-limit selection effects. In order to probe intrinsic redshift evolution of sources in L_{IR}–T_{dust}, a larger dynamic range of luminosities is needed over narrow redshift ranges (e.g., expanding to $> 10^{12}$–10^{13} $L_\odot$ sources at $0.0 < z < 0.5$ would help greatly). In contrast to 850 μm selection, 250 μm selection is relatively unbiased with dust temperature (Casey et al. 2011a).

There has been work that argues that the selection biases in L_{IR}–T_{dust} space are minor compared with physical effects (Sajina et al. 2007). The argument is that there is more intense dust heating in more extreme star-forming environments. Our data support this, since the dust temperature bias effects are less prominent using SPIRE selection. Although selection effects play a role in our perceived relation, we fit dust temperature as a function of L_{IR} and find a significant correlation for SPIRE selected samples of $\langle T_{dust} \rangle = 0.47^{+0.09}_{-0.06} (L_{FIR})^{0.144 \pm 0.006}$ [K $L_\odot^{-1}$]. The slope of this correlation, 0.14, is steeper than is seen in the local *IRAS* sample. We highlight that the difference between the two samples can be attributed to two factors. First, the dust temperature bias of the *IRAS* sample, selected at 60 μm with $S_{60} > 5.24$ Jy, which selects against low dust temperature sources (see Figure 19, right). Second, the local sample is a single-redshift snapshot at $z = 0$, while our sample is spread over a wide epoch range, $0 < z < 2$ (with 36 sources at
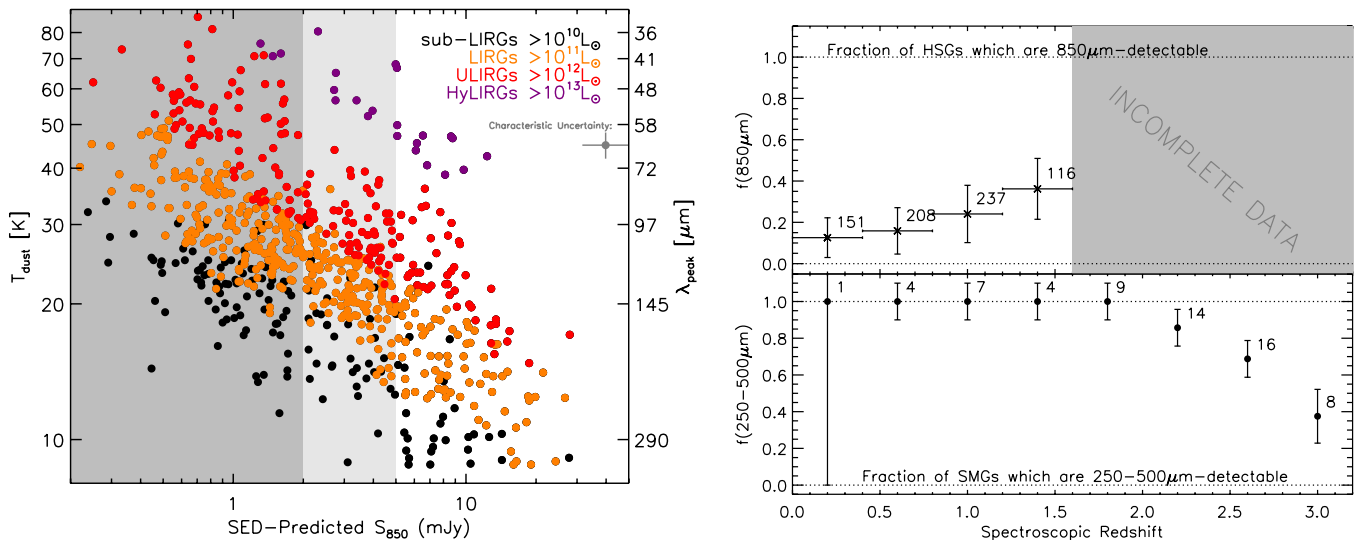
Figure 20. Predicted 850 μm flux densities for our SPIRE sample against the fitted SED dust temperatures. The characteristic uncertainty in either measurement is shown in the upper right. Note that the classic SMG selection criteria, $S_{850} > 5$ mJy would exclude 79^{+4}_{-8}% of SPIRE sources from detection in traditional 850 μm based SMG surveys: 83^{+4}_{-7}% of HSGs at $z < 1$ and 72^{+5}_{-11}% of HSGs at $1 < z < 2$. This is significantly higher than previous estimates of the "missing" fraction of warm-dust ULIRGs at $z > 1$ missed by SMG surveys. In the right panel, we show the fraction of SPIRE galaxies which are 850 μm detectable and 850 μm selected SMGs which are SPIRE detectable, as a function of redshift. Here SPIRE detectability is defined as S_{250}, S_{350}, or $S_{500} > 12$ mJy, and qualification as an SMG is $S_{850} > 5$ mJy. This illustrates that 60%–80% of SPIRE galaxies are undetectable at 850 μm out to $z \sim 1.6$, and that SMGs are only expected to drop out of the SPIRE bands at $z > 2$ for 20%–60% of sources. The numbers next to each point represent the number of galaxies used in each bin. There are too few SPIRE galaxies to reliably determine this statistic at $z > 1.6$ (gray shaded region).

(A color version of this figure is available in the online journal.)

$2 < z < 5$). Fitting the whole sample together washes out any evolution in the $L_{IR}-T_{dust}$ relation due to the strong redshift dependence of luminosity (e.g., Seymour et al. 2010). Eventually, much larger samples will make it possible to probe the $L_{IR}-T_{dust}$ relation as a function of redshift.

4.5. Comparison to SMGs and "Missing" Warm-dust ULIRGs

SMGs selected at 0.85–1.3 mm by SCUBA, LABOCA, AzTEC, or MAMBO have dominated the studies of high-z starbursts for the past decade. While *Herschel*-selected galaxies overlap with the SMG population (Roseboom et al. 2012), the correspondence between source detectability at 250–500 μm and $\sim$850 μm is not obvious. Although *Herschel* detects galaxies at their SED peak at $z \sim 1$–2 (and is more sensitive to lower luminosities at $z < 2$ than SCUBA), the submillimeter has the advantage of preferentially detecting higher-redshift galaxies and detecting them to lower luminosity limits than *Herschel*.

Unfortunately, very few of our SPIRE galaxies were directly observed with SCUBA, LABOCA, AzTEC, or MAMBO on account of small survey areas for those instruments, so it is difficult to assess population overlap directly. However, using the SED fits from Section 4.2, we can estimate 850 μm flux densities for every HSG. These SED-predicted 850 μm flux densities span a wide range, from $\sim$0.2–30 mJy. Figure 20 shows the relation between fitted dust temperature and extrapolated 850 μm flux density for SPIRE galaxies; there is a clear relation between dust temperature and 850 μm detectability.

We infer that 79^{+4}_{-8}% of all SPIRE galaxies are *undetectable* at 850 μm at a flux cutoff of $S_{850} < 5$ mJy. Even considering a submillimeter detection threshold as low as 2 mJy (the lowest 3σ detection limit for SCUBA, which had a confusion noise of $\sim$0.7 mJy at 850 μm; Blain et al. 1999), 47^{+27}_{-14}% of SPIRE galaxies would be undetectable at 850 μm. While this is a large fraction, it could be that most of these are at low redshift, and therefore low luminosities and low SFRs. The "missing" fraction

as a function of redshift is plotted on the right panel of Figure 20. Indeed the fraction of HSGs not detectable at 850 μm is high at $z < 1$, 83^{+4}_{-7}%, yet even at higher redshifts, $1 < z < 2$, the fraction is substantial: 72^{+5}_{-11}% of HSGs have $S_{850} < 5$ mJy.

The fact that 850 μm selection misses a large fraction of infrared-bright starbursts is not new, but has been difficult to measure directly or estimate in the past. This dust-temperature selection effect was studied in detail in the pre-*Herschel* era (Blain et al. 2004; Chapman et al. 2004; Casey et al. 2009a, 2011a, 2011b) where conservative estimates of the "missing" warm-dust ULIRG population (often referred to as "OFRGs" or submillimeter-faint radio galaxies, "SFRGs") were on the order of 10%–20%, and later works went on to estimate the missing fraction at $\gtrsim$30% using photometric data and limited spectroscopic data from *Herschel* (Chapman et al. 2010; Magdis et al. 2010).

While we note that roughly half of SPIRE galaxies are below the detection threshold at 850 μm, we can also investigate the converse statistic. How many SMGs are not detectable by SPIRE? We use the original sample of SMGs described in Chapman et al. (2005), along with assumed IR luminosities estimated via the FIR/radio correlation to fit SEDs and extrapolate to SPIRE fluxes. We then use these hypothetical fluxes to calculate a SPIRE detectability, in other words, we determine what fraction of the SMGs would pass our $>3\sigma$ selection criteria. The percentage of SMGs that are 250–500 μm detectable is shown at bottom on the right panel of Figure 20. At $z < 2$, all SMGs are detected with SPIRE, but above $z > 2$, anywhere from 20% to 60% of SMGs are SPIRE dropouts.

We also investigate this fraction using real SPIRE coverage of SMG fields together with LHN data presented in Chapman et al. (2010) and SPIRE coverage around SMGs from Chapman et al. (2005). The statistic is consistent, roughly 1/3 of SMGs are not detected in SPIRE, and nearly all at $z > 2.5$. Since the sample size of SMGs with overlapping SPIRE data is about half
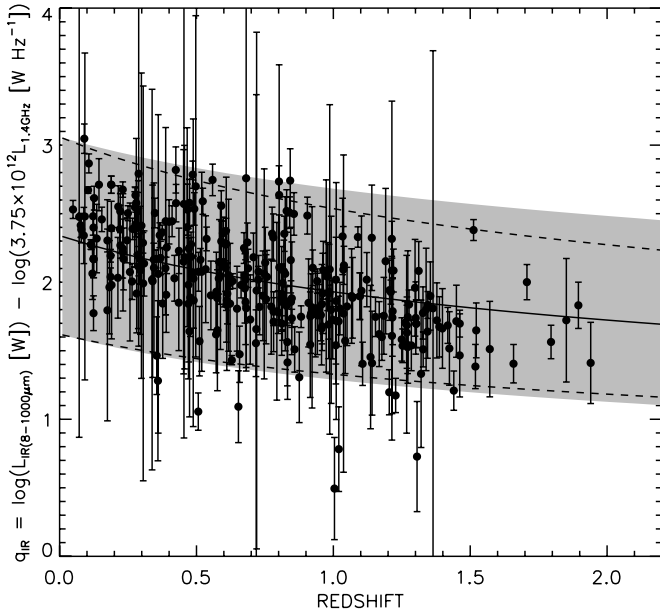
Figure 21. Ratio between FIR luminosity and radio luminosity in our sample, $q_{\rm IR}$, against redshift. We compare against the moderate evolution measured by Ivison et al. (2010a) $\propto (1 + z)^{-0.26 \pm 0.07}$ (gray band). We measure slightly stronger evolution, $q_{\rm IR} \propto (1 + z)^{-0.30}$.

the size of the full SMG sample, we use the extrapolated flux densities in Figure 20.

4.6. The FIR/Radio Correlation

Since IR luminosities were derived independently of radio luminosity, we measure the FIR/radio ratio, $q_{\rm IR}$, for radio-detected HSGs in order to assess the FIR/radio correlation for starbursts (see Helou et al. 1985; Condon 1992; Ivison et al. 2010a, 2010b) for HSGs. We use the bolometric $q_{\rm IR}$ ratio between 8–1000 μm flux and 1.4 GHz radio flux as defined in Ivison et al. (2010a), with IR luminosities measured from rest-frame 8 μm to 1000 μm. Figure 21 shows the $q_{\rm IR}$ ratio evolving with redshift for synchrotron slope $\alpha = 0.75$, where $S_\nu \propto \nu^{-\alpha}$. Our sample is consistent with the moderate evolution measured by Ivison et al. (2010a, 2010b) and Magnelli et al. (2010). We measure $q_{\rm IR} \propto (1 + z)^\gamma$ where $\gamma = -0.30 \pm 0.02$ using only the $0 < z < 2$ HSG sample (an independent assessment of the $z > 2$ sample is given in C12). Ivison et al. (2010b) measured $\gamma = -0.04 \pm 0.03$ at $z = 0$–2, but found that their $z < 0.5$ samples contaminated this measurement since they were few in number and not well matched in luminosity to the higher redshift sources; they measure $\gamma = -0.26 \pm 0.07$ for the $0.5 < z < 2.0$ sample, which is consistent with our finding, $\gamma = -0.30$. Either with or without our $z < 0.5$ sample, we measure $\gamma = -0.30$, likely because our sample is dominated (68%) by sources at $0.5 < z < 1.5$. Note that the evolution we measure is much more pronounced if the original luminosity limits defining $S_{\rm FIR}$ are used, 40–120 μm: $q_{\rm FIR} \propto (1 + z)^{-0.59 \pm 0.04}$. Note also that the significantly limiting factor of this measurement is that our sample is partially radio selected and is not a luminosity-matched sample.

5. DISCUSSION

The characterization of galaxies by *Herschel* is the largest advance in understanding star formation in the infrared since the discovery of SMGs by SCUBA. This paper has presented

Table 2
IR Luminosity Function from the *Herschel*-SPIRE Spectroscopic Sample

$0.0 < z < 0.4$, $N = 116$		$0.4 < z < 0.8$, $N = 168$	
$\log L_{\rm IR}$ ($L_\odot$)	$\log \Phi(L_{\rm IR})$ (h^3 Mpc^{-3} log L^{-1})	$\log L_{\rm IR}$ ($L_\odot$)	$\log \Phi(L_{\rm IR})$ (h^3 Mpc^{-3} log L^{-1})
9.55	(-2.49 ± 0.76)	10.66	(-3.84 ± 0.18)
9.92	(-2.57 ± 0.19)	10.97	(-3.52 ± 0.06)
10.21	(-2.88 ± 0.36)	11.26	-3.39 ± 0.06
10.41	(-2.49 ± 0.11)	11.54	-3.52 ± 0.05
10.69	-2.99 ± 0.32	11.82	-4.05 ± 0.09
10.94	-3.51 ± 0.28	12.07	-4.59 ± 0.12
11.11	-3.79 ± 0.19	12.40	-5.32 ± 0.13
11.42	-4.29 ± 0.27		
$\log \Phi^\star =$	-2.5	$\log \Phi^\star =$	-3.1
$\log L^\star =$	10.7	$\log L^\star =$	11.5
$\alpha \equiv$	-0.35	$\alpha \equiv$	-0.35

$0.8 < z < 1.2$, $N = 202$		$1.2 < z < 1.6$, $N = 92$	
$\log L_{\rm IR}$ ($L_\odot$)	$\log \Phi(L_{\rm IR})$ (h^3 Mpc^{-3} log L^{-1})	$\log L_{\rm IR}$ ($L_\odot$)	$\log \Phi(L_{\rm IR})$ (h^3 Mpc^{-3} log L^{-1})
11.31	(-4.05 ± 0.11)	12.18	-3.72 ± 0.05
11.51	(-3.55 ± 0.06)	12.54	-4.11 ± 0.10
11.70	(-3.38 ± 0.03)	12.91	-5.20 ± 0.54
11.89	-3.50 ± 0.06		
12.09	-3.83 ± 0.06		
12.28	-3.95 ± 0.05		
12.46	-4.70 ± 0.10		
12.69	-4.75 ± 0.11		
$\log \Phi^\star =$	-3.2	$\log \Phi^\star =$	-3.5
$\log L^\star =$	12.0	$\log L^\star =$	12.3
$\alpha \equiv$	-0.35	$\alpha \equiv$	-0.35

Notes. The luminosities in this table are integrated within 8–1000 μm. They represent the central luminosity of sources in each bin, in other words, the bin for luminosity L covers sources with luminosities between $L - \Delta L/2$ and $L + \Delta L/2$. The luminosity functions in this table are based exclusively on our SPIRE galaxy spectroscopic sample. We compute the luminosity function in units of h^3 Mpc^{-3} log L^{-1}.

a spectroscopic survey of 1594 SPIRE sources, 767 have identified spectroscopic redshifts from $0 < z < 5$, and 731 at $z < 2$. Securing spectroscopic redshifts is itself very valuable for follow-up studies of this population, from their metallicities, stellar populations, morphologies, and evolutionary histories to their molecular gas properties and dust content. However, arguably the impact of this population on the cosmic SFRD is the most relevant computation for galaxy formation and evolution studies in general. The LF has only ever been measured in integrated IR luminosity ($L_{\rm IR}(8$–1000 μm)) for limited populations of SMGs with known incompleteness (Chapman et al. 2005; Wardlow et al. 2011) or from extrapolations to the IR from *Spitzer* (Le Floc'h et al. 2005; Caputi et al. 2007; Magnelli et al. 2011) making use of model templates or explicitly assuming SEDs from libraries (Le Borgne et al. 2009; Béthermin et al. 2011; Marsden et al. 2011). This section presents the integrated IR LF of *Herschel* galaxies, a discussion of the completeness of our samples, the implications for the cosmic SFRD, and comparisons to these previous studies.

5.1. Luminosity Function

We compute the integrated IR (8–1000 μm) LFs of SPIRE-detected galaxies in four redshift bins where the spectroscopic completeness is constrained and the number of identifications is
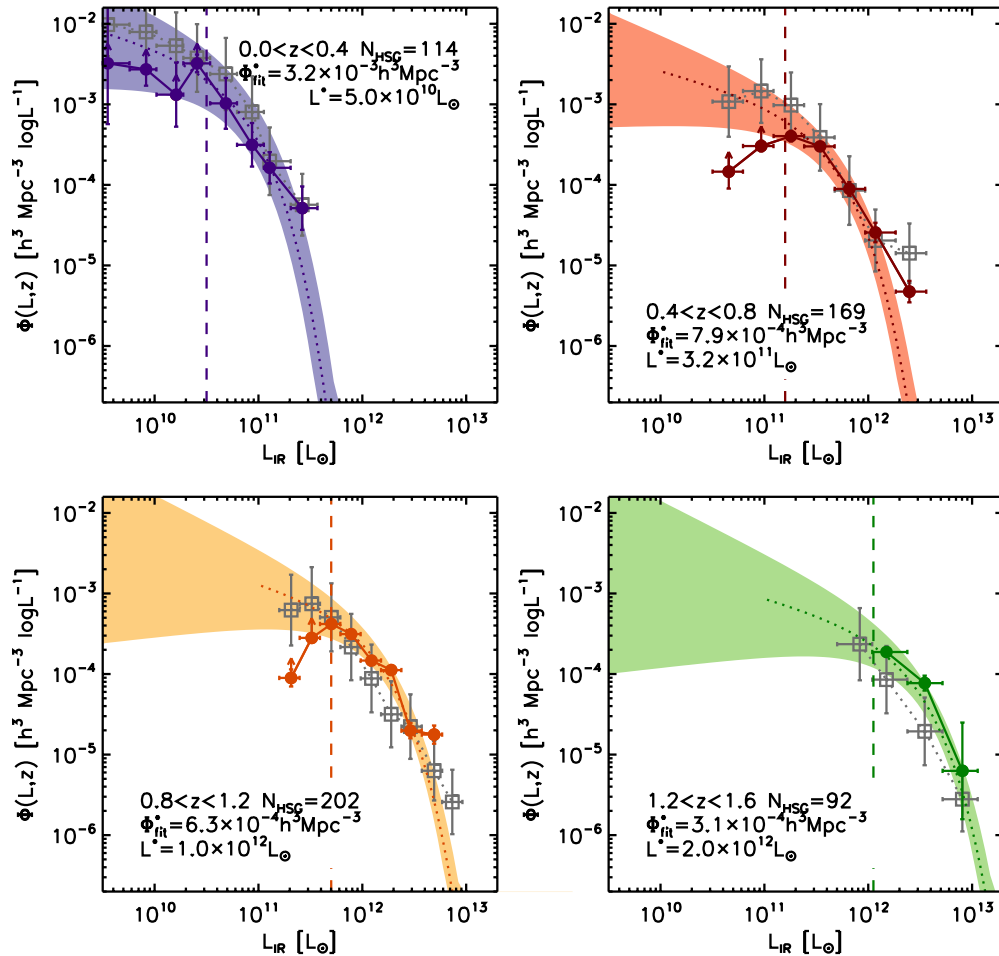
Figure 22. Evolving luminosity functions for our HSG spectroscopic sample. The redshift bins are $0.0 < z < 0.4$ (purple), $0.4 < z < 0.8$ (red), $0.8 < z < 1.2$ (yellow), and $1.2 < z < 1.6$ (green). We also include the estimated luminosity functions from photometric redshifts in COSMOS (Ilbert et al. 2010) as gray boxes and associated uncertainties. The best-fit Schechter functions are shown as dotted colored lines, with associated uncertainties (shaded regions) based on uncertainty in α, Φ^*, and L^*. The lower luminosity completeness limits of each bin are shown as dashed vertical lines. Below these luminosities, our SPIRE samples are incomplete (in terms of SPIRE detectability). Note that these luminosity functions are corrected for spectroscopic incompleteness (as detailed in Figure 12).

(A color version of this figure is available in the online journal.)

$\sim$100 per bin: $0.0 < z < 0.4$, $0.4 < z < 0.8$, $0.8 < z < 1.2$, and $1.2 < z < 1.6$. We use the $1/V_{\rm max}$ accessible volume technique:

$$\Phi_z(L)\Delta L = \sum_{i=0}^{n} \frac{1}{V_i(L)}, \qquad (2)$$

where $\Phi_z(L)$ is given in units of h^3 Mpc^{-3} log L^{-1} and $\Phi_z(L)\Delta L$ is the number density of sources with luminosities between L and $L + \Delta L$. Here $V_i(L)$ is the comoving volume within which the ith source is detectable in our survey. This accessible volume is determined by constructing a detection limit in $L_{\rm IR}$–z space for each source. As was evident in Figure 16, the SPIRE detection limit is highly dependent on dust temperature (i.e., 40 K galaxies will have a different $L_{\rm IR}$–z detection threshold than 20 K galaxies). The luminosity limit per source is determined by assuming an SED (of type given by Equation (1)) with fitted $T_{\rm dust}$, and the 3σ detection bounds at 250 μm, 350 μm, and 500 μm, where the lowest luminosity limit with redshift is adopted for the accessible volume calculation. The measured LF, after correction for spectroscopic incompleteness (as discussed earlier in the text and shown in Figure 12), is given in Table 2 and shown in Figure 22 as filled, colored circles.

As a check, we also compute the LFs for the same XID selection using the photometric redshifts in COSMOS (Ilbert

et al. 2010) and the same luminosity limit technique (for $\sim$6000 sources over $\sim$2.2 deg^2). The photometric redshift results are shown in Figure 22 as open gray squares. Within uncertainties, the spectroscopic and photometric LFs agree at each epoch.

Figure 22 clearly demonstrates evolution in the LF with redshift, with ultraluminous ($>10^{12}$ $L_\odot$) and hyperluminous ($>10^{13}$ $L_\odot$) galaxies becoming far more abundant with increasing z than in the local universe. At each of the four epochs, we fit Schechter functions of the form $\Phi = \Phi^* x^\alpha e^{-x}$, where $x = L/L^*$. The faint end of the LF, dominated by the x^α power law, is largely unprobed by our data set since L^*, the turnover luminosity, is approximately equal to the lower luminosity limit of SPIRE in each bin. Without constraints, we decided to fix $\alpha = -0.35$, our measurement from the local RBGS sample (Sanders et al. 2003) and the $0 < z < 0.4$ photometric redshift sample, which is also consistent with prior local and high-z measurements (Le Floc'h et al. 2005; Caputi et al. 2007). We add a large margin of uncertainty on the slope α, since its impact on integrated star formation density increases with redshift ($\sigma_\alpha = 0.1$ at $z \approx 0$, increasing to 0.5 at $z \approx 1.4$). At the top end of the LF, we note slight excesses at $0.0 < z < 0.4$, $0.4 < z < 0.8$, and $0.8 < z < 1.2$, which, in other samples, has been attributed to AGN contributions to infrared luminosity; while a double power-law LF might provide an alternate fitting method,
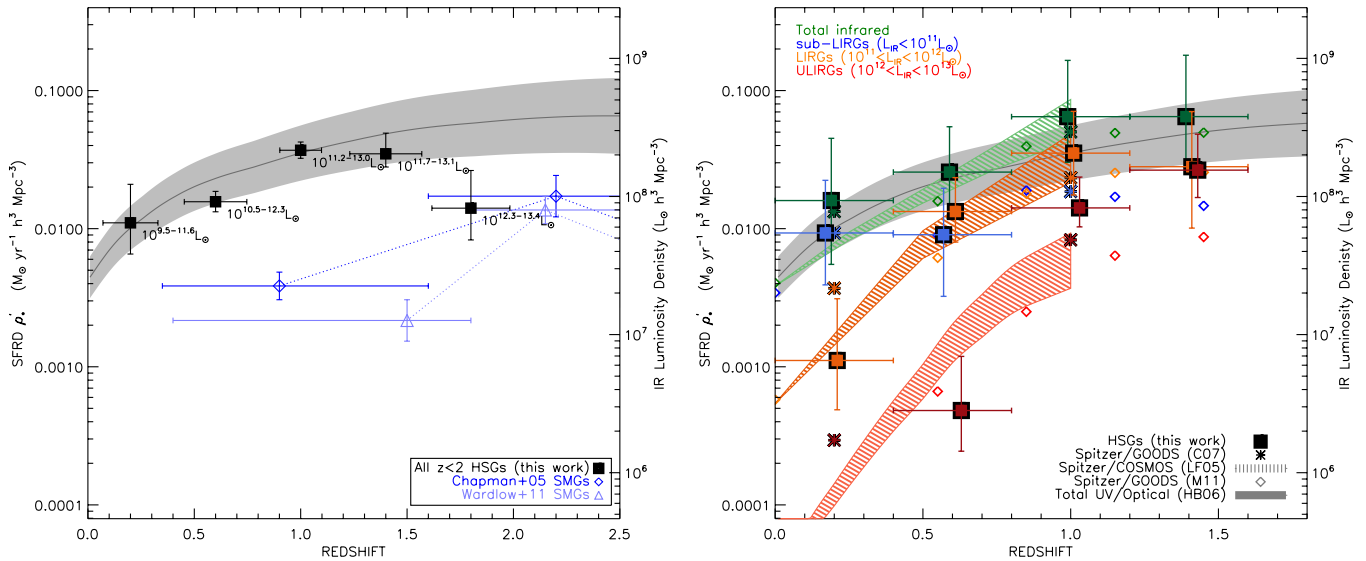
Figure 23. Contribution of SPIRE galaxies to the cosmic star formation rate density from $z = 0$ to $z = 2$. We compare against the compilation in Hopkins & Beacom (2006) of optical/UV SFRD estimates that are dust corrected (large gray area, representing the associated uncertainty of UV/optical estimates). We also overplot the SFRD points estimated by Chapman et al. (2005) and Wardlow et al. (2011) for SMGs which are substantially lower by $\sim 10\times$ than the optical/UV dust-corrected SFRD and HSGs at $z < 1.5$. At left, we measure the SFRD from this spectroscopic sample by converting the measured luminosity functions to star formation rate functions, then integrating. At right, we use the best-fit Schechter functions in Figure 22 to extrapolate over luminosities not directly probed by our survey, to measure the total infrared (green), sub-LIRG (blue), LIRG (orange), and ULIRG (red) contributions to the SFRD (colors are consistent across different data sets). The right y-axes give the value of the IR luminosity density that translates directly to SFRD via the Kennicutt (1998b) scaling. We gather similar computations from the literature for comparison: Le Floc'h et al. (2005) *Spitzer* samples in COSMOS (hashed regions), Caputi et al. (2007) *Spitzer* samples in GOODS (asterisks), and Magnelli et al. (2011) *Spitzer* samples in GOODS (diamonds).

(A color version of this figure is available in the online journal.)

the difference in this work is inconsequential considering the uncertainty on the LF points themselves.

The parameters Φ^* and L^* are then varied to provide an optimal fit to data at each epoch. We measure evolution in both parameters, as given in Table 2. However, it should be stressed that the measurement of these parameters' values is highly degenerate with variation in the faint-end slope, α and with each other. Much larger data sets, particularly those which constrain the faint-end slope via stacking techniques, are needed for detailed analysis on the evolution of Φ^* and L^*. This detailed work is being done with larger samples from COSMOS (M. Vaccari & E. Le Floc'h 2012, private communication).

5.2. Star Formation Rate Density

To estimate the contribution of SPIRE-detected sources to the cosmic SFRD, our measured luminosity functions are converted to SFR functions via the Kennicutt (1998b) prescription assuming a non-evolving Salpeter initial mass function (IMF). Note that there is some recent discussion, particularly in relation to very high luminosity sources at $z > 2$, that the Kennicutt conversion from infrared luminosity to SFR might not hold due to IMF variation (Swinbank et al. 2008); however, in this paper we use it for consistency with previous work on the infrared SFRD (Le Floc'h et al. 2005; Hopkins & Beacom 2006; Caputi et al. 2007; Magnelli et al. 2011).

The SFRD contribution from HSGs is shown in Figure 23. The left panel shows the raw conversion of the LF data points to SFRD contribution, including all sources above and below the completeness luminosity limit (dashed lines in Figure 22), and without extrapolation to lower or higher luminosities. Although we excluded $1.6 < z < 2.0$ sources from the LF calculation due to small numbers, we bring them back for this computation. The measured HSG SFRD appears to increase from $z = 0$

to $z = 1$, and then fall sharply at $z = 1.5$; however, this is due entirely to the luminosity limits of the sample at the given epochs. For example, the highest redshift bin, $z \approx 1.8$, has a lower luminosity limit of log $L_{IR} = 12.3$, significantly higher than the lower redshift bins. Similar luminosity limit restrictions exist for the literature values of the SFRD found from SMGs, in Chapman et al. (2005), Wall et al. (2008), and Wardlow et al. (2011).

Although SMGs are thought of as a very rare class of galaxy in this context, it is important to point out that this depends entirely on the adopted definition of "SMG" and associated luminosity limits. The ~ 1 dex difference between the SFRD contributions of $850\,\mu$m selected SMGs and 250–$500\,\mu$m selected HSGs at $z \sim 1$ is due to the effective luminosity limits of SPIRE or $850\,\mu$m selection (where HSG selection probes luminosities ~ 5 times fainter) and also due to the dust temperature bias of $850\,\mu$m selection (another factor of ~ 2).

Drawing a fair comparison with other samples requires interpolation in the LF across equal luminosity bins via our best-fit Schechter functions. Several literature sources have estimated the total integrated SFRD from infrared sources using a variety of data sets: *Spitzer*-24 μm identified sources in COSMOS (Le Floc'h et al. 2005); *Spitzer*-24 μm sources in both GOODS fields (Caputi et al. 2007; Magnelli et al. 2011); and *Akari*-selected sources (Goto et al. 2010). We also draw on the full optical and ultraviolet data compilation of Hopkins & Beacom (2006), which is extinction corrected to account for infrared emission in optically bright galaxies. On the right panel of Figure 23, we show the breakdown of the SFRD in 1 dex luminosity bins, $<10^{11}\,L_\odot$ (sub-LIRGs), $10^{11}\,L_\odot < z <10^{12}\,L_\odot$ (LIRGs), $10^{12}\,L_\odot < z <10^{13}\,L_\odot$ (ULIRGs), and total integrated infrared. On a whole, we observe the same trends as other literature works: The total integrated infrared SFRD contribution is comparable to the total optical and ultraviolet contribution

(i.e., just as much star formation is obscured as is unobscured) LIRGs dominate at $z \approx 1$, and ULIRGs become increasingly important at $z > 1$. However, a few subtle differences stand out. For example, both Le Floc'h et al. (2005) and Caputi et al. (2007) find higher contributions from ULIRGs at $z < 0.5$ than is measured in the Magnelli et al. (2011) sample or in our sample. Since both prior samples were $24\,\mu m$ selected, this could be indicative of the disagreement between $24\,\mu m$ and total integrated luminosity. At $z > 1$, the Magnelli et al. (2011) sample shows a deficit in ULIRGs with respect to our sample (also seen in the models of Béthermin et al. 2011). While this deficit could be due to SED assumptions, it could also be due to cosmic variance and the size of the GOODS fields used in Magnelli et al. (e.g., as is known to be a problem in CDFS, which has a submillimeter deficit, see Weiß et al. 2009).

The uncertainties on the SFRD measurements are dominated by the uncertainty in the faint-end slope of the LF, which is not well constrained. At higher redshifts $L^\star$ increases; this impact of the faint-end slope uncertainty then grows, since sub-LIRGs and LIRGs are not constrained with our data beyond $z > 1.2$. This is not only true for our data set, but any data set (selected via $24\,\mu m$ or radio with a PACS or SPIRE detection) that does not employ stacking analysis to constrain faint sources. Further work is needed on stacking $24\,\mu m$ and radio-selected samples of galaxies to probe fainter luminosities out to high redshifts over large areas of sky less prone to cosmic variance.

It is clear from this work that the importance of infrared-luminous star formation in the context of cosmic star formation is high, and equally high as rest-frame UV/optical estimates, even though prior surveys of SMGs seem to imply that infrared starbursts are rare even at $z \sim 1$–2. *Herschel*-SPIRE has enabled us to directly constrain the far-infrared unlike previous work in the area, extrapolating from $24\,\mu m$. Our observations confirm the importance of dust obscured star formation across the first 10 billion years of the universe's history.

6. CONCLUSIONS

This paper has presented a large spectroscopic survey of galaxies selected in the *Herschel*-SPIRE 250–500 μm bands and followed up spectroscopically on Keck LRIS and DEIMOS. Out of 1594 spectroscopic targets spanning 0.93 deg^2 in multiple deep legacy fields, 767 sources had identifiable redshifts from their rest-frame UV or optical spectra, 731 of which are at $z < 2$. We present the following conclusions.

1. The redshift distribution of SPIRE-selected spectroscopically confirmed galaxies peaks at $z = 0.85$ and has a tail of sources extending out to $z \approx 5$. The vast majority of this spectroscopic sample ($731/767 \approx 96\%$) is at $z < 2$, mostly spectrally identified by [O II], [O III], Hβ, or Hα emission.

2. Our spectroscopic sample excludes sources without $24\,\mu m$ or radio 1.4 GHz counterparts. A negligible fraction of SPIRE sources in deep legacy fields are thought to drop out in the mid-IR and radio at $z < 2$, meaning our SPIRE targeting is close to complete at $z \lesssim 2$.

3. Only $\sim 50\%$ of spectroscopic targets yield redshifts (767 with identifications, 826 without). Under photometric observing conditions, $\sim 60\%$ yield identifications. We constrain the spectroscopic completeness at $z < 2$ to 20%–80%, depending on redshift, using large catalogs of photometric redshifts. The sources without spectroscopic identifications either were observed in poor weather condi-

tions, have no optical counterpart (i.e., no source picked up on the slit after 1–2 hr integration), or have spectra not easily identified with a single redshift (e.g., continuum without emission or absorption features).

4. We measure the accuracy of HSG optical/near-infrared photometric redshifts at $\Delta z/(1 + z_{\rm spec}) = 0.29$, a factor of $\sim$3–4 times worse than photometric redshifts for normal field galaxies. We determine that the lack of reliability in photometric redshifts for HSGs is due to more significant optical obscuration and higher overall SFRs (thus higher line-to-continuum ratios). We caution that future analysis of aggregate properties of infrared-selected samples like HSGs might be biased by only making use of large photometric redshift catalogs.

5. We observe a correlation between dust temperature and infrared luminosity, which is partially a selection effect and partially thought to be physically real. The selection effect is caused by the dependence of SPIRE flux density on dust temperature and the physical motivation comes from the local observation that more luminous galaxies have more compact, clumpy, and therefore hot dust.

6. SPIRE color does not evolve with redshift in our spectroscopic sample, due to the redshift dependence of luminosity and the luminosity dependence on dust temperature (a combination of selection effects and physical mechanisms).

7. We infer the aggregate properties of HSGs by combining the near-infrared to radio measurements by selection method and luminosity. We find little difference between the bulk infrared properties of radio-selected galaxies and $24\,\mu m$ selected galaxies, and good overall agreement with the FIR/radio correlation. Consistent with our results from individual sources, we see an increase in dust temperature with luminosity in the composite IR SEDs. At luminosities $> 10^{12.5}\ L_\odot$, we detect the $\approx 10\,\mu m$ Si absorption feature in the composite, resembling the SED of the local ULIRG Arp 220.

8. By extrapolating our infrared SED fits to $850\,\mu m$, we determine that only $79^{+4}_{-8}\%$ of HSGs would be detectable as "classic" SMGs ($S_{850} > 5$ mJy). This fraction is highest at low redshift, $83^{+4}_{-7}\%$ at $z < 1$, and still significant, $72^{+5}_{-11}\%$, from $1 < z < 2$.

9. For the 331 radio-selected sources with spectroscopic confirmations, we measure the evolution in the FIR/radio correlation and find $q_{\rm IR} = (1 + z)^{-0.30 \pm 0.02}$. This evolution is stronger than some previous measurements, however the sample is biased by being radio detected, thus the intrinsic evolution of $q_{\rm IR}$ with redshift is likely to be shallower.

10. We compute the LFs for HSGs in four redshift bins across $0 < z < 1.6$ and find agreement with predictions from photometric redshifts. The LFs are well fit by Schechter functions, with evolving parameters $\Phi^\star$ ($3.2 \times 10^3\ h^3$ Mpc^{-3} at $z = 0.2$ to $3.1 \times 10^{-4}\ h^3$ Mpc^{-3} at $z = 1.4$) and $L^\star$ ($5 \times 10^{10}\ L_\odot$ at $z = 0.2$ to $2 \times 10^{12}\ L_\odot$ at $z = 1.4$). The faint end of the LF is not well probed by our data and would only be accessible from *Herschel* data through stacking analyses.

11. We find the HSG contribution to the cosmic SFRD to be substantial at $z < 2$, essentially comparable to the SFRD from optical/near-IR surveys $\gtrsim 1 \times 10^{-2}\ M_\odot\ {\rm yr}^{-1}\ h^3$ Mpc^{-3}. Without relying on extrapolations from the mid-infrared or model template libraries, we measure the contribution from LIRGs ($10^{11}\ L_\odot < L < 10^{12}\ L_\odot$) to the SFRD peaks at $z \sim 1$, and the ULIRG ($> 10^{12}\ L_\odot$)

contribution increases with z, nearly surpassing the LIRG contribution at $z \sim 1.4$.

This work has demonstrated that infrared-luminous galaxies form an integral part of galaxy formation across a wide range of epochs, particularly at $0 < z < 2$. With spectroscopic redshifts in hand, follow-up studies can be more efficiently carried out to determine the *physical* and *evolutionary* origins of HSGs as a function of infrared luminosity, dust temperature, dust mass, and AGN content. This work will then begin to shed light on the dominant mechanisms producing obscured star formation in the universe.

REFERENCES

Amblard, A., Cooray, A., Serra, P., et al. 2010, A&A, 518, L9
Armus, L., Mazzarella, J. M., Evans, A. S., et al. 2009, PASP, 121, 559
Barger, A. J., Cowie, L. L., & Richards, E. A. 2000, AJ, 119, 2092
Barger, A. J., Cowie, L. L., Sanders, D. B., et al. 1998, Nature, 394, 248
Barger, A. J., Cowie, L. L., Smail, I., et al. 1999, AJ, 117, 2656
Béthermin, M., Dole, H., Lagache, G., Le Borgne, D., & Penin, A. 2011, A&A, 529, A4
Béthermin, M., Le Floc'h, E., Ilbert, O., et al. 2012, A&A, 542, 58
Biggs, A. D., Ivison, R. J., Ibar, E., et al. 2011, MNRAS, 413, 2314
Blain, A. W., Chapman, S. C., Smail, I., & Ivison, R. 2004, ApJ, 611, 725
Blain, A. W., Smail, I., Ivison, R. J., & Kneib, J. 1999, MNRAS, 302, 632
Blain, A. W., Smail, I., Ivison, R. J., Kneib, J., & Frayer, D. T. 2002, Phys. Rep., 369, 111
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582
Caputi, K. I., Lagache, G., Yan, L., et al. 2007, ApJ, 660, 97
Cardamone, C. N., van Dokkum, P. G., Urry, C. M., et al. 2010, ApJS, 189, 270
Casey, C. M. 2012, MNRAS, 425, 3094
Casey, C. M., Béthermin, M., Bock, J., et al. 2012, ApJ, 761, 139
Casey, C. M., Chapman, S. C., Beswick, R. J., et al. 2009a, MNRAS, 399, 121
Casey, C. M., Chapman, S. C., Muxlow, T. W. B., et al. 2009b, MNRAS, 395, 1249
Casey, C. M., Chapman, S. C., Neri, R., et al. 2011a, MNRAS, 415, 2723
Casey, C. M., Chapman, S. C., Smail, I., et al. 2011b, MNRAS, 411, 2739
Chapin, E. L., Chapman, S. C., Coppin, K. E., et al. 2011, MNRAS, 411, 505
Chapin, E. L., Hughes, D. H., & Aretxaga, I. 2009, MNRAS, 393, 653
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
Chapman, S. C., Smail, I., Blain, A. W., & Ivison, R. J. 2004, ApJ, 614, 671
Charmandaris, V., Mirabel, I. F., Tran, D., et al. 1997, in Extragalactic Astronomy in the Infrared, ed. G. A. Mamon, T. X. Thuan, & J. Tran Thanh Van (Paris: Editions Frontièrs), 283
Chary, R., & Elbaz, D. 2001, ApJ, 556, 562
Condon, J. J. 1992, ARA&A, 30, 575
Cowie, L. L., Barger, A. J., Bautz, M. W., et al. 2001, ApJ, 551, L9
Cowie, L. L., Barger, A. J., & Kneib, J.-P. 2002, AJ, 123, 2197
Cowie, L. L., & Hu, E. M. 1998, AJ, 115, 1319
Cowie, L. L., Songaila, A., & Barger, A. J. 1999, AJ, 118, 603
Cox, P., Krips, M., Neri, R., et al. 2011, ApJ, 740, 63
Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686
Dale, D. A., Gil de Paz, A., Gordon, K. D., et al. 2007, ApJ, 655, 863
Dale, D. A., & Helou, G. 2002, ApJ, 576, 159
Dale, D. A., Helou, G., Contursi, A., Silbermann, N. A., & Kolhatkar, S. 2001, ApJ, 549, 215
Damen, M., Labbé, I., van Dokkum, P. G., et al. 2011, ApJ, 727, 1
Draine, B. T. 2006, ApJ, 636, 1114
Draine, B. T., & Li, A. 2007, ApJ, 657, 810
Dunlop, J. S., Ade, P. A. R., Bock, J. J., et al. 2010, MNRAS, 408, 2022
Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119
Engel, H., Tacconi, L. J., Davies, R. I., et al. 2010, ApJ, 724, 233
Faber, S. M., Phillips, A. C., Kibrick, R. I., et al. 2003, Proc. SPIE, 4841, 1657
Frayer, D. T., Sanders, D. B., Surace, J. A., et al. 2009, AJ, 138, 1261
Garn, T., & Alexander, P. 2008, MNRAS, 391, 1000
Goto, T., Takagi, T., Matsuhara, H., et al. 2010, A&A, 514, A6
Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, A&A, 518, L3
Helou, G., Soifer, B. T., & Rowan-Robinson, M. 1985, ApJ, 298, L7
Hilton, M., Conselice, C. J., Roseboom, I. G., et al. 2012, MNRAS, 425, 540
Hinshaw, G., Weiland, J. L., Hill, R. S., et al. 2009, ApJS, 180, 225
Hopkins, A. M., & Beacom, J. F. 2006, ApJ, 651, 142
Hughes, D. H., Serjeant, S., Dunlop, J., et al. 1998, Nature, 394, 241
Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, ApJ, 709, 644
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., Sanders, D. B., Le Floc'h, E., et al. 2010, ApJ, 721, 98
Kennicutt, R. C., Jr. 1998a, ARA&A, 36, 189
Kennicutt, R. C., Jr. 1998b, ApJ, 498, 541
Lagache, G., Dole, H., Puget, J.-L., et al. 2004, ApJS, 154, 112
Le Borgne, D., Elbaz, D., Ocvirk, P., & Pichon, C. 2009, A&A, 504, 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
Levenson, L., Marsden, G., Zemcov, M., et al. 2010, MNRAS, 409, 83
Lonsdale, C. J., Farrah, D., & Smith, H. E. 2006, Ultraluminous Infrared Galaxies (Berlin: Springer), 285
Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, PASP, 115, 897
Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, A&A, 532, 90
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. 2011, A&A, 528, A35
Magnelli, B., Lutz, D., Berta, S., et al. 2010, A&A, 518, L28
Marsden, G., Chapin, E. L., Halpern, M., et al. 2011, MNRAS, 417, 1192
Miller, N. A., Fomalont, E. B., Kellermann, K. I., et al. 2008, ApJS, 179, 114
Morrison, G. E., Owen, F. N., Dickinson, M., Ivison, R. J., & Ibar, E. 2010, ApJS, 188, 178
Negrello, M., Perrotta, F., González-Nuevo, J., et al. 2007, MNRAS, 377, 1557
Neufeld, D. A. 1991, ApJ, 370, L85
Nguyen, H. T., Schulz, B., Levenson, L., et al. 2010, A&A, 518, L5
Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ, 660, L43
Oke, J. B., Cohen, J. G., Carr, M., et al. 1995, PASP, 107, 375

Oliver, S. J., Bock, J., Altieri, B., et al. 2012, MNRAS, 424, 1614
Owen, F. N., & Morrison, G. E. 2008, AJ, 136, 1889
Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1
Pope, A., Borys, C., Scott, D., et al. 2005, MNRAS, 358, 149
Pope, A., Chary, R., Alexander, D. M., et al. 2008, ApJ, 675, 1171
Pope, A., Scott, D., Dickinson, M., et al. 2006, MNRAS, 370, 1185
Rieke, G. H., Alonso-Herrero, A., Weiner, B. J., et al. 2009, ApJ, 692, 556
Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, ApJ, 739, L40
Roseboom, I. G., Ivison, R. J., Greve, T. R., et al. 2012, MNRAS, 419, 2758
Roseboom, I. G., Oliver, S. J., Kunz, M., et al. 2010, MNRAS, 409, 48
Rowan-Robinson, M., Babbedge, T., Oliver, S., et al. 2008, MNRAS, 386, 697
Sajina, A., Yan, L., Armus, L., et al. 2007, ApJ, 664, 713
Sanders, D. B., Mazzarella, J. M., Kim, D.-C., Surace, J. A., & Soifer, B. T.
    2003, AJ, 126, 1607
Sanders, D. B., & Mirabel, I. F. 1996, ARA&A, 34, 749
Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86
Sanders, D. B., Soifer, B. T., Elias, J. H., Neugebauer, G., & Matthews, K.
    1988, ApJ, 328, L35
Schinnerer, E., Smolčić, V., Carilli, C. L., et al. 2007, ApJS, 172, 46

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Seymour, N., Symeonidis, M., Page, M. J., et al. 2010, MNRAS, 402, 2666
Siebenmorgen, R., & Krügel, E. 2007, A&A, 461, 445
Sirothia, S. K., Dennefeld, M., Saikia, D. J., et al. 2009, MNRAS, 395, 269
Smail, I., Ivison, R. J., & Blain, A. W. 1997, ApJ, 490, L5
Smail, I., Ivison, R. J., Blain, A. W., & Kneib, J. 2002, MNRAS, 331, 495
Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M.
    1999, ApJ, 519, 1
Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L.
    1996, ApJ, 462, L17
Strazzullo, V., Pannella, M., Owen, F. N., et al. 2010, ApJ, 714, 1305
Swinbank, A. M., Lacey, C. G., Smail, I., et al. 2008, MNRAS, 391, 420
Wall, J. V., Pope, A., & Scott, D. 2008, MNRAS, 383, 435
Wang, W.-H., Cowie, L. L., Barger, A. J., & Williams, J. P. 2011, ApJ,
    726, L18
Wardlow, J. L., Smail, I., Coppin, K. E. K., et al. 2011, MNRAS, 415, 1479
Weiß, A., Kovács, A., Coppin, K., et al. 2009, ApJ, 707, 1201
Younger, J. D., Omont, A., Fiolet, N., et al. 2009, MNRAS, 394, 1685
Yun, M. S., Scott, K. S., Guo, Y., et al. 2012, MNRAS, 420, 957