A CONTROLLED STUDY OF COLD DUST CONTENT IN GALAXIES FROM Z=0-2

Allison Kirkpatrick[1,2], Alexandra Pope[1], Anna Sajina[3], Daniel A. Dale[4], Tanio Díaz-Santos[5], Christopher C. Hayward[6,7], Yong Shi[8], Rachel S. Somerville[9], Sabrina Stierwalt[10], Lee Armus[11], Jeyhan S. Kartaltepe[12], Dale D. Kocevski[13], Daniel H. McIntosh[14], David B. Sanders[15], Lin Yan[16]

Draft version June 5, 2017

ABSTRACT

At $z = 1 - 3$, the formation of new stars is dominated by dusty galaxies whose far-IR emission indicates they contain colder dust than local galaxies of a similar luminosity. We explore the reasons for the evolving IR emission of similar galaxies over cosmic time using: 1) Local galaxies from GOALS ($L_{IR} = 10^{11} - 10^{12} L_\odot$); 2) Galaxies at $z \sim 0.1 - 0.5$ from the 5MUSES ($L_{IR} = 10^{10} - 10^{12} L_\odot$); 3) IR luminous galaxies spanning $z = 0.5 - 3$ from GOODS and *Spitzer* xFLS ($L_{IR} > 10^{11} L_\odot$). All samples have *Spitzer* mid-IR spectra, and *Herschel* and ground-based submillimeter imaging covering the full IR spectral energy distribution, allowing us to robustly measure L_{IR}^{SF}, T_{dust}, and M_{dust} for every galaxy. Despite similar infrared luminosities, $z > 0.5$ dusty star forming galaxies have a factor of 5 higher dust masses and 5 K colder temperatures. The increase in dust mass is linked with an increase in the gas fractions with redshift, and we do not observe a similar increase in stellar mass or star formation efficiency. L_{IR}^{SF}/L_{70}^{SF}, a proxy for T_{dust}, is strongly correlated with L_{IR}^{SF}/M_{dust} independently of redshift. We measure merger classification and galaxy size for a subsample, and there is no obvious correlation between these parameters and L_{IR}^{SF}/M_{dust} or L_{160}^{SF}/L_{70}^{SF}. In dusty star forming galaxies, the change in L_{IR}^{SF}/M_{dust} can fully account for the observed colder dust temperatures, suggesting that any change in the spatial extent of the interstellar medium is a second order effect.

1. INTRODUCTION

As the Universe ages, a phenomenon known as cosmic downsizing shifts the bulk of star formation to smaller galaxies. Additionally, as the Universe expands, the merger rate of galaxies decreases, and mergers become a less important triggering mechanism for star formation

[1] Department of Astronomy, University of Massachusetts, Amherst, MA 01002, USA
[2] Yale Center for Astronomy & Astrophysics, Physics Department, P.O. Box 208120, New Haven, CT 06520, USA, allison.kirkpatrick@yale.edu
[3] Department of Physics & Astronomy, Tufts University, Medford, MA 02155, USA
[4] Department of Physics & Astronomy, University of Wyoming, Laramie, WY 82071, USA
[5] Núcleo de Astronomía de la Facultad de Ingeniería, Universidad Diego Portales, Av. Ejército Libertador 441, Santiago, Chile
[6] Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA
[7] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[8] School of Astronomy and Space Science, Nanjing University, Nanjing 210093, China
[9] Department of Physics and Astronomy, Rutgers, The State University of New Jersey, 136 Frelinghuysen Rd, Piscataway, NJ 08854 USA
[10] Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA
[11] Spitzer Science Center, California Institute of Technology, Pasadena, CA 91125, USA
[12] School of Physics and Astronomy, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, NY 14623, USA
[13] Department of Physics and Astronomy, Colby College, Waterville, ME 04901, USA
[14] Department of Physics and Astronomy, University of Missouri-Kansas City, 5110 Rockhill Road, Kansas City, MO 64110, USA
[15] Institute for Astronomy, 2680 Woodlawn Drive, University of Hawaii, Honolulu, HI 96822, USA
[16] Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA

(Conselice et al. 2008, 2009; Bluck et al. 2012). These changes mean that, in certain respects, the composition of galaxies in the local Universe is quite different from what we would see if we existed in a galaxy at cosmic noon, 8-10 billion years ago ($z = 1 - 2$), when the Universe was forming most of its stars (Madau & Dickinson 2014). Galaxies at cosmic noon appear to be more compact than their present day counterparts of similar mass (e.g., Trujillo et al. 2007; van Dokkum et al. 2008; Williams et al. 2014), suggesting that the inner parts of galaxies are in place before the outer parts form (e.g., Hopkins et al. 2009; Nelson et al. 2016).

On the other hand, the cosmic noon Universe shares some key similarities with our present day Universe. For example, the backbone of the Hubble sequence was already in place by $z \sim 2$, so that blue, disky galaxies were strongly star forming (e.g. Lee et al. 2013). Crucially, the main sequence of galaxy formation already existed by $z \sim 3$ (e.g. Behroozi et al. 2013; Suzuki et al. 2015). The main sequence is the tight empirical relationship between SFR and M_* that defines the so-called "normal mode" of secular star formation, as opposed to short-lived intense starbursts, typically triggered by mergers (e.g., Brinchmann et al. 2004; Noeske et al. 2007; Elbaz et al. 2011).

At cosmic noon, the galaxies dominating the buildup of stellar mass are massive, dusty, infrared luminous systems ($L_{IR} > 10^{11} L_\odot$), but today, less massive, relatively isolated galaxies such as our Milky Way are the dominant sites of star formation (Murphy et al. 2011; Madau & Dickinson 2014). IR luminous galaxies exist in the local Universe as well, but they contribute little to the current cosmic SFR density. Historically, galaxy IR spectral energy distributions (SEDs) were parameterized only by L_{IR} (Sanders & Mirabel 1996; Dale et al. 2001; Chary & Elbaz 2001; Rieke et al. 2009). However, *Her-*

schel brought new insight into the peak of the SED in distant galaxies, and we now know that at $z \sim 1-2$, Ultra Luminous Infrared Galaxies (ULIRGs) have colder far-IR emission than galaxies with the same luminosity at $z \sim 0$ (Chapman et al. 2002; Pope et al. 2006; Sajina et al. 2006; Clements et al. 2010; Nordon et al. 2010; Symeonidis et al. 2009; Magnelli et al. 2011; Elbaz et al. 2011; Sajina et al. 2012; Kirkpatrick et al. 2012).

The shape of the ULIRG SED evolves strongly from $z \sim 0-2$, but local ULIRGs represent the most extreme objects. They are merging, compact starbursts, which are heavily obscured and even optically thick at far-IR wavelengths (e.g. Sakamoto et al. 2008; Scoville et al. 2015). These objects also lie well above the main sequence, while ULIRGs at $z \sim 1-2$ lie on the main sequence, indicative of a slower, secular star formation history (Elbaz et al. 2011). However, local LIRGs, which are an order of magnitude less luminous than local ULIRGs, share certain dust emission characteristics with their high redshift luminosity-based counterparts indicating they may have similar ISMs despite the fact that local LIRGs also lie above the main sequence. For example, $z \sim 1-2$ ULIRGs and LIRGs have similar polycyclic aromatic hydrocarbon (PAH) emission strength, parameterized by $6.2\,\mu$m equivalent widths and $L_{6.2}/L_{\rm IR}$, as local LIRGs (Pope et al. 2013; Stierwalt et al. 2014; Kirkpatrick et al. 2014b). The similarity of these features indicates a similar photodissociation region structure in these galaxies, hinting at similarities between their interstellar media (Pope et al. 2013; Stierwalt et al. 2014). Therefore, despite evolution of the main sequence with lookback time, it is reasonable to examine how closely dusty galaxies resemble each other from $z = 0-2$.

Far-IR/submm emission is sensitive to dust temperature, emissivity, dust mass, the geometry of the star forming regions within the ISM, and the incident radiation field (Witt & Gordon 2000; Gordon et al. 2001; Misselt et al. 2001), although the far-IR SED is most commonly used to calculate three parameters: $L_{\rm IR}(8 - 1000\,\mu$m); $M_{\rm dust}$, measured from the Rayleigh-Jeans tail; and $T_{\rm dust}$, which characterizes the peak of the SED. The change of SED shape, along with the similarity and differences between the distant and local Universe, leads to a key question in galaxy evolution: What is the root cause of the colder dust emission at $z \sim 1-2$?

Recently, it has been suggested that $L_{\rm IR}$ *and* ISM geometry (including surface density and compactness) fully account for the shape of the SED; that is, high redshift IR luminous galaxies are colder because they are more extended (Elbaz et al. 2011; Rujopakarn et al. 2013). Additionally, the overall structure of the ISM in high redshift galaxies may be different. A larger fraction of gas is likely in H_2 rather than H_I, the mass function of giant molecular clouds evolves, and proportionally more of the dust may arise from the diffuse ISM in distant galaxies (Popping et al. 2014a,b; Scoville et al. 2016). However, theoretical simulations and radiative transfer calculations contend that $L_{\rm IR}$ and $M_{\rm dust}$ or total mass, M, are the primary parameters for determining a galaxy's global SED (Chakrabarti & McKee 2005; Safarzadeh et al. 2016). Then, compactness of the ISM, mass function of molecular clouds, and possibly surface density, have second order effects over the shape of the SED. Indeed, Lee et al. (2016) demonstrate at $z = 0$ that heavily dust obscured galaxies with $L_{\rm IR} = 10^{11}-10^{12}\,L_\odot$ have similar dust masses and temperatures as their less obscured counterparts, even though the extremely dust obscured galaxies are likely much more compact. This would suggest that the geometry of the ISM in IR luminous galaxies has little overall effect on the measured dust mass and temperature.

In this work, we test how the dust temperatures and dust masses change with $L_{\rm IR}$ and redshift for IR luminous galaxies from $z \sim 0-2$. We utilize three extensively studied samples from the literature, which are large ($\gtrsim 200$ galaxies in each sample) and have deep multiwavelength data sets that eliminate any need for photometric stacking. We analyze all the SEDs in a self-consistent manner. Crucially, all of our galaxies have mid-IR spectroscopy, which we use to remove AGN and compact starbursts, which excludes a potential source of uncertainty when comparing dust masses and temperatures (e.g. Lee et al. 2016). In Section 2, we describe our three data sets and summarize the relevant literature pertaining to these samples; in Section 3, we discuss how we remove AGN and measure dust masses self-consistently; in Section 4, we characterize the shape of the far-IR SED using broadband colors; in Section 5, we measure an increase in the dust mass of DSFGs with redshift and parameterize the relationship between $L_{\rm IR}/M_{\rm dust}$ and $T_{\rm dust}$; in Section 6, we discuss the origin of the increased dust masses with redshift and whether ISM size or merger classification can influence the SED, and we summarize in Section 7. Throughout this work, we assume a flat cosmology with $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\rm M} = 0.3$, and $\Omega_\Lambda = 0.7$.

2. SAMPLE DESCRIPTION

In this study, we compare the far-IR/submm properties of three samples at different redshifts extensively described in the literature (Figure 1): 1) $z < 0.088$: Great Observatories All Sky Survey (GOALS; PI L. Armus Armus et al. 2009); 2) $z = 0.05 - 0.75$: 5 mJy Unbiased *Spitzer* Extragalactic Survey (5MUSES; P.I. G. Helou Wu et al. 2010); 3) $z = 0.3 - 3.0$: *Spitzer* IRS Supersample (Kirkpatrick et al. 2015). These samples comprise a case study of dusty star forming galaxies where the far-IR SEDs are well sampled and stacking is not required. Indeed, these samples are ideal for comparison as they each contain massive ($M_* \gtrsim 10^9\,M_\odot$), infrared luminous ($L_{\rm IR} = 10^{10} - 10^{13}\,L_\odot$) galaxies with far-IR imaging from the *Spitzer Space Telescope* and the *Herschel Space Observatory* and mid-IR spectroscopy from the *Spitzer* IRS instrument (Houck et al. 2004), useful for quantifying and removing dust heating due to an obscured AGN. We briefly summarize the basic properties of each sample in the following subsections, and in Section 3, we discuss how we arrive at the final sample sizes of 46 galaxies from GOALS, 30 galaxies from 5MUSES, and 51 galaxies from the Supersample.

2.1. *GOALS*

The GOALS sample is comprised of 180 luminous IR galaxies (LIRGs; $L_{\rm IR} = 10^{11}\,L_\odot$) and 22 ultra luminous IR galaxies (ULIRGs; $L_{\rm IR} > 10^{12}\,L_\odot$). Many of these systems are interacting, resulting in 244 individual galactic nuclei with *Spitzer* IRS spectroscopy. These galaxies
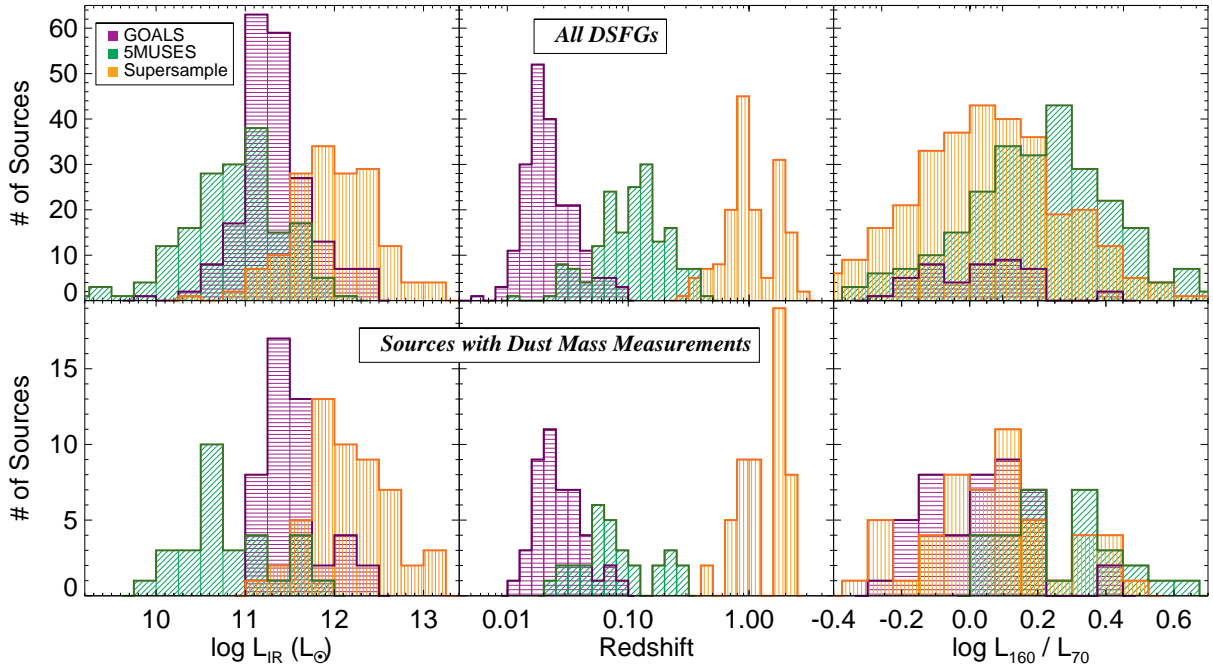
FIG. 1.— L_{IR} (*left*), redshift (*middle*), and $L_{160}/L_{70}(\propto T_{dust})$ (*right*) distributions of the dusty star forming galaxies (DSFGs) in the GOALS, 5MUSES, and Supersample. We include all sources with *Spitzer* IRS spectroscopy and $f(AGN)_{MIR} < 0.5$ (see Section 3.1) in the upper panels. The lower panels show the distributions for sources in which we were able to measure the dust mass (see Section 3.3).

are a representative subset of the IRAS Bright Galaxy Survey (Sanders et al. 2003) and were selected to have $S_{60} > 5.24$ Jy. GOALS sources cover the distance range 15 Mpc < D < 400 Mpc, which corresponds to $z < 0.088$.

The IRS spectra have been previously analyzed in depth (Díaz-Santos et al. 2010, 2011; U et al. 2012; Stierwalt et al. 2014). Here, we make use of the IRS SL staring observations, $\lambda = 5.5 - 14.5\,\mu m$, to identify and remove AGN from the sample (Section 3.1). The GOALS galaxies have global flux densities from all three *Spitzer* MIPS bandpasses (Mazzarella et al. 2017, in prep). $L_{IR}(8 - 1000\,\mu m)$ was calculated following the formula in Sanders & Mirabel (1996) using all four *IRAS* flux densities (Díaz-Santos et al. 2010).

The UV-IR SEDs were analyzed in detail for a subset of 64 galaxies (see U et al. 2012, for a discussion of the selection of these objects). These galaxies have stellar mass measurements, so they form our initial dust mass sample as well. From the 64 galaxies, 58 sources have $850\,\mu m$ from the *James Clerk Maxwell Telescope* or $250\,\mu m$ photometry from *Herschel* (Chu et al. 2017), which we use to calculate M_{dust}. Five of these are double nuclei sources that we remove, and seven are hosting AGN. This leaves a final sample size of 46 galaxies.

2.2. *5MUSES*

5MUSES is a *Spitzer* IRS mid-IR spectroscopic survey of 330 galaxies selected from the SWIRE and *Spitzer* Extragalactic First Look Survey fields (details in Wu et al. 2010). It is a flux limited sample selected at $24\,\mu m$, with $S_{24} > 5$ mJy. 280 sources have optical spectroscopic redshifts (Wu et al. 2010). 5MUSES is a representative sample at intermediate redshift (the median redshift of the sample is 0.14) of galaxies with $L_{IR} \sim 10^{10} - 10^{12}\,L_{\odot}$, bridging the gap between local (U)LIRGs and high redshift observations.

For the present study, we make use of the *Spitzer* IRS SL spectroscopy to quantify AGN emission. Complete details of the *Spitzer* data reduction are found in Wu et al. (2010). IRS spectra combined with *Spitzer* MIPS imaging is used to calculate $L_{IR}(5 - 1000\,\mu m)$ in Wu et al. (2010), with a slightly different cosmology ($\Omega_M = 0.27, \Lambda = 0.73$). We adjust the published L_{IR}s by multiplying by 0.99 to account for the difference in cosmology and 0.95 to scale to $L_{IR}(8 - 1000\,\mu m)$, which we have determined using a purely star forming template from Kirkpatrick et al. (2015). We use *Herschel* SPIRE observations of a subset of 188 sources in the *Herschel* Multi-Tiered Extragalactic Suvery (HerMES Oliver et al. 2010, 2012; Magdis et al. 2013) to calculate M_{dust}.

Of the 188 *Herschel* galaxies, 40 have $500\,\mu m$ detections, necessary to calculate M_{dust}. Six of these are removed because they host AGN. We remove an additional 4 sources where the $500\,\mu m$ photometry appears blended with a nearby source. This leaves a final sample of 30 galaxies.

2.3. *Supersample*

We have assembled a multi-wavelength data set for a sample of 343 high redshift ($z \sim 0.3 - 4.0$) (U)LIRGs in the Great Observatories Origins Deep Survey North (GOODS-N), Extended *Chandra* Deep Field Survey (ECDFS), and *Spitzer* xFLS fields. All sources are selected to have mid-IR spectroscopy from *Spitzer* IRS, which is used to measure the redshift (Kirkpatrick et al. 2012). Our sample contains a range of sources from individual observing programs, each with differing selection criteria. However, the overarching selection criterion is a $24\,\mu m$ flux limit of 0.9 mJy for the xFLS galaxies, which are taken from a shallower survey, and 0.1 mJy for the GOODS and ECDFS galaxies, compiled from a deep sur-

vey. In addition to IRS spectra, these sources all have *Spitzer* MIPS and *Herschel* PACS and SPIRE imaging (Sajina et al. 2012; Kirkpatrick et al. 2015). We calculate individual $L_{\rm IR}(8 - 1000\,\mu{\rm m})$ by fitting a full suite of *Spitzer* and *Herschel* photometry with the library of templates from Kirkpatrick et al. (2015). Redshifts are determined by fitting the main PAH emission features, or using optical spectroscopy in the case of a featureless MIR spectrum.

In this work, we also utilize $870\,\mu{\rm m}$ photometry of GOODS-S from LABOCA on APEX (Weiß et al. 2009), the combined AzTEC+MAMBO 1.1mm map of GOODS-N (Penner et al. 2011), and MAMBO 1.2 mm imaging of xFLS (Lutz et al. 2005; Sajina et al. 2008; Martínez-Sansigre et al. 2009), which exists for 169 galaxies. For 4 of these, the source redshift is high enough that there is no data beyond $\lambda = 250\,\mu{\rm m}$, which is our threshold for calculating dust mass. We also remove any sources from the sample where the submm emission is blended with a nearby source (10 sources). We remove 57 galaxies for hosting an AGN. Finally, we remove 47 galaxies for have submm photometry with a SNR < 2, leaving a final sample of 51 galaxies.

2.4. *Selection Effects*

We are comparing the far-IR/submm properties of three different samples with different selection criteria, so we must understand whether our results are biased due to the selection criteria. The GOALS sample is selected at $60\,\mu{\rm m}$ while 5MUSES and the Supersample are selected at observed frame $24\,\mu{\rm m}$; this selection criteria alone could produce samples of galaxies that are warmer or colder for the same $L_{\rm IR}$ range.

We consider the 5MUSES and GOALS sources together to make up our low z sample. All of the 5MUSES sources with $L_{\rm IR} > 10^{11}$ (same as GOALS) would be selected at rest frame $S_{60} > 5.24$ Jy, the GOALS selection criterion. Similarly, all GOALS galaxies meet the 5MUSES selection criterion of $S_{24} > 5$ mJy.

We use the mid-IR spectrum to determine if a 5MUSES source meets the Supersample selection criterion, since at $z = 1 - 2$, S_{24} covers the PAH features. Only 28% of the 5MUSES sample would be selected at $z = 1$, and 11% at $z = 2$. This is not surprising, since the 5MUSES galaxies are on average less luminous. Only 32% of the 5MUSES galaxies overlap in the $L_{\rm IR}$ range spanned by the Supersample. However, of the 60 5MUSES galaxies with $L_{\rm IR} > 2 \times 10^{11}\,L_{\odot}$ (the same range as the bulk of the Supersample), 50 would be selected as Supersample galaxies at $z = 1$ and 30 would be selected at $z = 2$, due to similar $L_{\rm PAH}/L_{\rm IR}$ ratios as $z \sim 1 - 2$ (U)LIRGs (Pope et al. 2013; Kirkpatrick et al. 2014b; Battisti et al. 2015). Estimating whether the GOALS sources would be included in our Supersample is more nuanced, since we lack global IRS spectra for the GOALS galaxies. Instead, we use the Chary & Elbaz (2001) library of templates which were derived from the IRAS BGS. At $z = 1$, templates with $\log L_{\rm IR} > 11.11 L_{\odot}$ meet the selection threshold, but at $z = 2$, this increases to $\log L_{\rm IR} > 12.21 L_{\odot}$.

Next, we estimate what fraction of the Supersample would be selected as a 5MUSES or GOALS galaxy, if the Supersample existed at a different redshift. This is largely an academic exercise, since the Supersample contains the most luminous galaxies, and Symeonidis et al.

(2011) demonstrated that the IRAS selection criteria are sensitive to ULIRGs ($L_{\rm IR} > 10^{12}\,L_{\odot}$) with $T = 17 - 87$ K. Nevertheless, such a check will ensure that the Supersample contains enough warm dust to be luminous at rest frame $60\mu{\rm m}$ and $24\,\mu{\rm m}$. We calculate synthetic $60\,\mu{\rm m}$ photometry following the prescription in Section 3.2, for a redshift of $z = 0.015$, which is the peak of the GOALS redshift distribution in Figure 1. 97% of the Supersample meets the GOALS selection criterion. The 5MUSES selection criterion is $S_{24} > 5$ mJy. We shift our synthetic rest frame photometry to $z = 0.15$ by scaling down by 1.62, a ratio determined with the Kirkpatrick et al. (2015) template library. Again, 97% would be selected as a 5MUSES galaxy.

From this, we conclude that if cold, massive galaxies, like most of the Supersample, exist at lower redshift, they are easily detectable in current surveys. Similarly, at $z \sim 1$ at least, we should be able to select a population of galaxies with the range of dust masses and temperatures exhibited by local LIRGs, if they exist. However, at $z \sim 2$, galaxies with spectral shapes similar to local LIRGs are undetectable in current $24\,\mu{\rm m}$ surveys.

2.5. *ULIRG nomenclature*

Throughout the remainder of this paper, we dispense with the (U)LIRG nomenclature for clarity. In the local Universe, the term ULIRG evokes a massive merging system, where the merger gives rise to a starburst followed by an AGN (Sanders & Mirabel 1996). At earlier times, the correlation between ULIRG and merger still exists, but is more tenuous, and the link between merger and starburst is even less clear in these systems (Hopkins et al. 2010; Kartaltepe et al. 2012; Hayward et al. 2013). LIRG and ULIRG are simply luminosity cuts, so to avoid any preconceived notions between these luminosity cuts and physical properties, we follow the convention in Casey et al. (2014) and refer to objects in all three samples as dusty star forming galaxies (DSFG).

3. DATA AND PHYSICAL PROPERTIES

In this paper, we compare three samples of IR luminous galaxies at three different epochs. For a fair comparison, we shift all IR photometry to the rest frame, we remove any AGN contribution using *Spitzer* spectroscopy, and we derive dust masses self consistently. By using the mid-IR spectrum to look for a hot dust continuum, which we attribute to an AGN, we are also removing galaxies with a hot central starburst.

3.1. *Identifying and Removing AGN*

The mid-IR spectrum ($\lambda = 3 - 20\,\mu{\rm m}$) is rich for identifying star formation features and AGN. The most prominent dust emission complexes are produced by polycyclic aromatic hydrocarbons (PAHs, $\lambda = 6.2, 7.7, 11.3, 12.7\,\mu{\rm m}$) which are abundant in galaxies with metallicity close to solar, such as high redshift DSFGs (Magdis et al. 2012). These molecules are preferentially located in the photodissociation regions surrounding star forming regions (Helou et al. 2004). As such, PAHs are good tracers of ongoing star formation in a galaxy (Peeters et al. 2004). However, the mid-IR spectrum can also exhibit continuum emission due to the torus enveloping the AGN.

We perform spectral decomposition of the mid-IR spectrum ($\sim 5-15\,\mu m$ rest frame) for each source in all three samples in order to disentangle the AGN and star forming components. Pope et al. (2008) explain the technique in detail, and we summarize it here. We fit the individual spectra with a model comprised of four components: (1) the star formation component is represented by the mid-IR spectrum of the prototypical starburst M 82; (2) the AGN component is determined by fitting a pure power-law (λ^α) with the slope and normalization (N) as free parameters; (3,4) extinction curves from the Weingartner & Draine (2001) dust models for Milky Way (MW) type dust is applied to the AGN component and star forming component. The full model is then

$$S_\nu = N_{\rm AGN}\lambda^\alpha\, e^{-\tau_{\rm AGN}} + N_{\rm SF}S_\nu({\rm M82})e^{-\tau_{\rm SF}} \qquad (1)$$

We fit for the $N_{\rm AGN}$, $N_{\rm SF}$, α, $\tau_{\rm AGN}$, and $\tau_{\rm SF}$ simultaneously.

For each source, we quantify the strength of the AGN, $f({\rm AGN})_{\rm MIR}$, as the fraction of the total mid-IR luminosity ($\lambda = 5 - 15\,\mu m$) coming from the power-law continuum component. For the GOALS sample, our AGN identification technique selects the same sources as the $S_{15}/S_{5.5}$ v. $S_{6.2}/S_{5.5}$ diagnostic applied in Petric et al. (2011). For the 5MUSES sources, we also identify the same AGN as Wu et al. (2010), where the authors classify AGN on the basis of $6.2\,\mu m$ PAH feature equivalent width alone.

In Kirkpatrick et al. (2015), we found that an AGN had a significant contribution to the far-IR emission, quantified through dust temperatures and luminosities, when $f({\rm AGN})_{\rm MIR} \geq 0.5$. Recently, Lee et al. (2016) showed that local dust obscured galaxies hosting an AGN had lower dust masses than their star forming counterparts. Consequently, we remove all 5MUSES, GOALS, and Supersample galaxies with $f({\rm AGN})_{\rm MIR} \geq 0.5$, so that we are only comparing the far-IR/submillimeter properties of strongly star forming systems. We note that because the GOALS sample is nearby, in some cases IRS observations only cover the central region of the galaxy, in contrast to the IRS observations of 5MUSES and the high z Supersample. This could introduce a slight bias as we may remove sources that have a nucleus dominated by AGN emission, whereas the galaxy integrated emission is dominated by star formation. We only remove 13% of the GOALS sample, so the effects of this potential bias are small.

Our AGN identification technique only relies on detecting hot continuum emission in the mid-IR. While we attribute this to an AGN, in principle, a very compact, dust enshrouded starburst would display the same mid-IR signature. By removing AGN, we are also then removing these compact starbursts, which will have different ISM properties than galaxies where the star forming regions are located throughout the galaxy, due to higher gas densities and a harsher radiation field. However, we note that although such compact starbursts exist locally (e.g. Díaz-Santos et al. 2011), so far such galaxies are rare at high redshift and still have prominent PAH features (Nelson et al. 2014). From the Supersample, we remove 47% for having $f({\rm AGN})_{\rm MIR} \geq 0.5$, and 62% of these removed sources have X-ray detections consistent with being an AGN (Kirkpatrick et al. 2015), while the remaining sources may be Compton thick based on predicted X-ray luminosities (Kirkpatrick et al. 2012). As such, we do not think that we are missing a large population of compact starbursts by removing sources with $f({\rm AGN})_{\rm MIR} \geq 0.5$.

3.2. *Calculating Rest Frame Photometry*

In Section 4, we compare the far-IR SED shape of the GOALS, 5MUSES, and Supersample utilizing far-IR colors. To make a fair comparison between three samples at different redshift ranges, we require synthetic rest frame photometry in the MIPS and SPIRE photometric bandpasses for the 5MUSES and Supersample. GOALS is at low enough redshift that the offset in rest frame wavelengths is negligible. We fit templates to the Supersample and 5MUSES from the MIR-based Library in Kirkpatrick et al. (2015). The MIR-based Library is a suite of 11 empirical templates that characterize the full shape of the IR SED based on the relative amounts of PAH and continuum emission in a source's mid-IR spectroscopy. We select the appropriate template for each source from the spectral decomposition described above, fit to the far-IR photometry ($\lambda > 20\,\mu m$), and then convolve each template with the MIPS and SPIRE transmission filters to create synthetic rest frame photometry at 70, 160, and $250\,\mu m$. We plot the Supersample intrinsic SEDs and estimated rest frame photometry in Appendix A. The biggest concern when interpolating between data points is whether we have chosen a template with the correct far-IR shape. However, the majority of galaxies have a sufficiently well-sampled far-IR SED to mitigate this uncertainty, as the estimated photometry lies very close to observed data points.

It is unlikely that an AGN that is not significantly affecting the mid-IR emission would have a strong effect on the far-IR emission. Even so, we correct the rest-frame broadband colors and $L_{\rm IR}$s of all SFGs with $0.0 < f({\rm AGN})_{\rm MIR} < 0.5$ in order to account for any scatter introduced by a buried AGN that may contribute to, but not dominate, the IR luminosity. We have calculated these corrections by decomposing empirical templates into their relative star formation and AGN components as described in detail in Kirkpatrick et al. (2015). The AGN corrections are listed below, for the rest frame luminosities. In Kirkpatrick et al. (2015), we discuss ways to estimate $f({\rm AGN})_{\rm MIR}$ in the absence of mid-IR spectroscopy, so these corrections can be applied to sources with only broadband photometry.

$$L_{\rm IR}^{\rm SF} = L_{\rm IR} \times (1 + 0.035f({\rm AGN})_{\rm MIR} - 0.66f({\rm AGN})_{\rm MIR}^2)$$
$$L_{24}^{\rm SF} = L_{24} \times (1 - 1.81f({\rm AGN})_{\rm MIR} + 0.94f({\rm AGN})_{\rm MIR}^2)$$
$$L_{70}^{\rm SF} = L_{70} \times (1 - 0.08f({\rm AGN})_{\rm MIR} - 0.16f({\rm AGN})_{\rm MIR}^2)$$
$$L_{160}^{\rm SF} = L_{160} \times (1 - 0.036f({\rm AGN})_{\rm MIR})$$
$$L_{250}^{\rm SF} = L_{250} \times (1 - 0.018f({\rm AGN})_{\rm MIR}) \qquad (2)$$

3.3. *Dust Mass*

With our three samples, we can compare how $M_{\rm dust}$ evolves in dusty galaxy populations with redshift. Dust mass is calculated as

$$M_{\rm dust} = \frac{S_\nu D_L^2}{\kappa_\nu B_\nu(T)} \qquad (3)$$

where $B_\nu(T)$ is the Planck equation, κ_ν is the dust opacity, D_L is the luminosity distance, and S_ν is the flux density. For S_ν, we use the longest far-IR/submm data available, which varies with each sample. We only use sources whose rest frame measurement is $\lambda \geq 250\,\mu$m, otherwise the dust emission may not be tracing the coldest dust component (Scoville et al. 2016). When we combine this criterion with the $f(\mathrm{AGN})_{\mathrm{MIR}} < 0.5$ criterion, we have the final sample sizes of 46 galaxies from GOALS, 30 galaxies from 5MUSES, and 52 galaxies from the Supersample.

We take κ_ν, the dust opacity, from the Weingartner & Draine (2001)[17], models at the rest wavelength of S_ν for each individual source. We assume MW-like dust and $R_V = 3.1$, although in these models, the far-IR/submm opacities have negligible changes for different R_V and are similar for the MW-, LMC-, and SMC-type models (at $850\,\mu$m, the opacity differs by $<10\%$).

Scoville et al. (2016) argue for using fixed T_{cold}, since presumably the temperature in the diffuse ISM should be relatively similar for all massive, dusty galaxies, assuming similar radiation fields and dust grain mixtures. Indeed, in Kirkpatrick et al. (2015) we found that the average cold dust temperature in SFGs in the Supersample was remarkably constant, $T \sim 26\,$K, and did not scale with L_{IR}. Accordingly, we set $T_{\mathrm{cold}} = 25\,$K for all three samples and derive M_{dust} using Equation 3. As a test of this assumption, we fit each galaxy with a two temperature modified blackbody with $T_{\mathrm{cold}} = 25\,$K and achieve excellent fits for all sources. We explore how the dust masses change if T_{dust} is allowed to vary in Appendix B and find that all conclusions in this paper hold regardless of the specific method of measuring dust mass as long as these measurements are consistent across all three samples. In this study, we are primarily concerned with comparing the dust masses at high redshift relative to low redshift. By fixing T_{cold} and using the same κ model for all samples, we are able to probe how dust masses change with redshift if the other dust properties of these galaxies remain the same.

We use L_{IR} for each sample compiled from the literature (see Section 2). In all cases, the L_{IR} was *not* determined using the submm data that goes into the M_{dust} calculation, so in this way, L_{IR} and dust mass are independent. L_{IR} for the 5MUSES galaxies and Supersample were measured by fitting template libraries, so the main source of concern is how the choice of β used in constructing the templates might effect L_{IR}, as M_{dust} also depends on β. We calculate that varying $\beta \in [1.5, 2.0]$ changes L_{IR} by 5%, so using different values of β to calculate L_{IR} and M_{dust} is not a significant source of scatter in the resulting relationships. In the analysis below, we correct L_{IR} for any contribution from nuclear activity as indicated by the presence of a hot dust continuum in the mid-IR spectrum, so that we are actually comparing M_{dust} with $L_{\mathrm{IR}}^{\mathrm{SF}}$. The correction is at most 0.07 dex. As all of the GOALS and 5MUSES photometry and redshifts used in this paper are published elsewhere in the litera-

ture (Wu et al. 2010; Magdis et al. 2013; U et al. 2012), we only list the M_{dust} and L_{IR} values that we have calculated for the Supersample in Table 1.

3.4. *Stellar Mass*

Stellar masses are compiled from the literature. For GOALS, U et al. (2012) calculates M_* by fitting Bruzual & Charlot (2003) stellar population models to 10 different broadband UV-NIR photometry points (from *GALEX*, ground-based observatories, and *Spitzer*). The star formation history is parameterized as SFH$= e^{-t/\tau}$ with $\tau \in [1, 30]\,$Gyr. Metallicity is a free parameter. The authors choose a Chabrier initial mass function (IMF) (Chabrier 2003) with a mass range of $0.1\,M_\odot$ to $100\,M_\odot$.

Shi et al. (2011) calculate stellar masses for the 5MUSES sample by fitting Bruzual & Charlot (2003) stellar population models to UV-IR data assuming a Chabrier IMF. The SFH has $\tau \in [0.03, 22.4]\,$Gyr, and metallicity is a free parameter.

For the GOODS and ECDFS sources (we do not have M_* for the six xFLS sources included in the final high z sample), Kirkpatrick et al. (2012) describes the procedure for calculating M_* by fitting Bruzual & Charlot (2003) models to suite of photometry from the U band to $4.5\,\mu$m (see also Pannella et al. 2015). For the SFH, $\tau \in [0.1, 20]\,$Gyr, and metallicity is fixed to solar. For the Supersample, a Salpeter IMF is used (Salpeter 1955), which results in higher stellar masses than a Chabrier IMF. We convert the Supersample masses to a Chabrier framework using $M_*^{\mathrm{Cha}} = 0.62\,M_*^{\mathrm{Sal}}$ (Zahid et al. 2012; Speagle et al. 2014).

We are drawing from three different samples fit with three different methods, so we verify the stellar masses by comparing with rest frame K-band estimated M_* (Howell et al. 2010). In all cases, the stellar masses are consistent to within 0.6 dex, which is the largest deviation. We further check the consistency of the Supersample masses, since these were initially calculated with a Salpeter IMF and require a conversion. We fit the UV-submm SEDs of 10 randomly selected galaxies with the MAGPHYS code, which employs an energy-balance technique to simultaneously fit the UV/optical/IR emission (da Cunha et al. 2008). The MAGPHYS derived stellar masses are on average 0.2 dex larger than the masses from Kirkpatrick et al. (2012), converted to a Chabrier IMF.

All three samples cover a similar range of $M_* = 1 \times 10^{10} - 3 \times 10^{11}\,M_\odot$, as can be seen in Figure 8.

4. SED COMPARISON

We look for evolution in the IR SED of our samples of DSFGs by comparing rest frame colors, as this does not require any model assumptions. We show a sample SED at $z \sim 1$ and the MIPS and SPIRE $250\,\mu$m broadband photometry filters in Figure 2 as a visual guide when comparing galaxy colors below. We also plot a local template from the (Chary & Elbaz 2001) library to better illustrate how the far-IR colors can indicate SED shape. We plot a local template that peaks at the same wavelength as our $z \sim 1$ template, despite being nearly an order of magnitude lower in luminosity, so that the color L_{160}/L_{70} should be approximately equal for both the local and $z \sim 1$ template. However, the $z \sim 1$ template has proportionally more cold dust, as evidenced by an increase in the emission longwards of $\sim 200\,\mu$m.

---

[17] Recent results indicate a change in these model opacities (Dalcanton et al. 2015; Planck et al. 2016), but we are consistently using the same model for every source, so we do not expect any change in opacity to affect our dust mass comparison

TABLE 1
SUPERSAMPLE PROPERTIES

Name	RA	Dec	z	$\log L_{\rm IR}(L_\odot)$[a]	$\log M_{\rm dust}(M_\odot)$	$\log M_*(M_\odot)$[b]
GN_IRS2	12:37:02.74	+62:14:01.0	1.24	12.13	8.74 ± 0.17	10.87
GN_IRS5	12:36:20.94	+62:07:14.0	1.15	12.23	8.64 ± 0.24	10.79
GN_IRS10	12:37:07.21	+62:14:08.1	2.49	12.72	8.85 ± 0.12	11.07
GN_IRS11	12:36:21.27	+62:17:08.0	2.00	12.56 (12.55)	8.84 ± 0.14	11.20
GN_IRS15	12:37:11.37	+62:13:31.0	2.00	12.86 (12.85)	9.20 ± 0.06	11.28
GN_IRS18	12:37:16.59	+62:16:43.0	1.80	12.46 (12.43)	8.67 ± 0.21	11.52
GN_IRS21	12:36:18.33	+62:15:50.0	2.00	12.77 (12.71)	8.88 ± 0.13	10.75
GN_IRS23	12:36:19.13	+62:10:04.0	2.21	12.30	8.44 ± 0.33	11.02
GN_IRS25	12:37:01.59	+62:11:46.0	1.72	12.62	8.79 ± 0.16	11.54
GN_IRS26	12:36:34.51	+62:12:40.0	1.22	12.62	8.45 ± 0.36	10.94
GN_IRS27	12:36:55.94	+62:08:08.0	0.79	12.06 (12.03)	8.37 ± 0.37	11.10
GN_IRS31	12:36:22.66	+62:16:29.0	1.79	12.44	8.54 ± 0.29	11.15
GN_IRS38	12:36:29.10	+62:10:46.3	1.01	12.00	8.75 ± 0.16	11.14
GN_IRS42	12:36:46.72	+62:08:33.9	0.97	12.13 (12.12)	8.75 ± 0.16	11.05
GN_IRS45	12:38:21.76	+62:17:06.0	1.62	12.45 (12.41)	9.02 ± 0.19	11.09
GN_IRS47	12:35:53.81	+62:13:38.0	0.88	11.87 (12.86)	9.01 ± 0.09	10.88
GN_IRS49	12:37:05.49	+62:21:24.0	0.95	11.85	8.62 ± 0.24	11.11
GN_IRS60	12:36:49.72	+62:13:12.9	0.47	11.24	8.44 ± 0.20	9.99
GN_IRS62	12:36:29.54	+62:06:46.5	0.80	11.58	8.48 ± 0.31	10.56
GN_IRS66	12:37:05.85	+62:21:29.8	0.95	11.82	8.65 ± 0.23	10.47
GN_IRS67	12:36:19.14	+62:13:01.8	1.23	11.76	8.55 ± 0.28	10.47
GN_IRS68	12:36:19.50	+62:12:52.6	0.47	11.47	8.25 ± 0.31	10.56
GS_IRS20	03:32:47.58	−27:44:52.0	1.91	12.60 (12.59)	8.55 ± 0.24	10.77
GS_IRS23	03:32:17.23	−27:50:37.0	1.96	12.35	8.78 ± 0.14	10.99
GS_IRS25	03:32:18.70	−27:49:19.0	1.04	11.83	8.44 ± 0.28	10.60
GS_IRS26	03:32:20.70	−27:44:53.0	0.97	11.52	8.38 ± 0.31	10.69
GS_IRS28	03:31:35.26	−27:49:58.0	0.96	11.61	8.52 ± 0.23	10.62
GS_IRS29	03:32:22.53	−27:45:38.0	2.08	12.12 (12.10)	8.47 ± 0.27	10.62
GS_IRS30	03:32:43.78	−27:52:31.0	1.62	11.85	8.33 ± 0.38	11.07
GS_IRS31	03:32:06.00	−27:45:07.0	1.07	11.78	8.36 ± 0.34	10.57
GS_IRS34	03:32:39.00	−27:44:20.0	1.93	11.95	8.39 ± 0.33	10.11
GS_IRS39	03:31:48.18	−27:45:35.0	1.72	12.34 (12.31)	8.79 ± 0.13	⋯
GS_IRS43	03:31:48.93	−27:39:45.0	0.82	11.62 (11.58)	8.57 ± 0.19	10.34
GS_IRS45	03:32:17.45	−27:50:03.0	1.62	12.48	8.64 ± 0.19	10.39
GS_IRS46	03:32:42.71	−27:39:27.0	1.85	12.32	8.70 ± 0.16	⋯
GS_IRS52	03:32:12.52	−27:43:06.0	1.79	12.11	8.50 ± 0.25	10.43
GS_IRS53	03:32:12.13	−27:42:49.0	2.45	12.37 (12.31)	8.74 ± 0.14	10.90
GS_IRS56	03:32:13.58	−27:47:54.0	1.72	11.91	8.41 ± 0.31	11.22
GS_IRS62	03:32:22.48	−27:49:35.0	0.73	11.76	8.56 ± 0.17	10.76
GS_IRS63	03:32:22.59	−27:44:25.9	0.74	11.64	8.42 ± 0.24	10.63
GS_IRS64	03:32:17.28	−27:49:08.0	1.62	11.88	8.50 ± 0.26	10.45
GS_IRS70	03:32:27.71	−27:50:40.6	1.10	11.99 (11.98)	8.41 ± 0.31	10.78
GS_IRS73	03:32:43.24	−27:47:56.2	0.67	11.26	8.17 ± 0.41	10.47
GS_IRS74	03:32:44.32	−27:49:11.9	2.00	12.12	8.56 ± 0.22	⋯
GS_IRS76	03:32:48.83	−27:42:35.0	1.96	12.58 (12.57)	9.16 ± 0.06	⋯
GS_IRS80	03:32:36.52	−27:46:30.7	1.05	11.91 (11.89)	8.69 ± 0.16	⋯
MIPS289	17:13:50.02	+58:56:57.1	1.89	13.10 (13.07)	9.00 ± 0.07	⋯
MIPS8377	17:17:33.53	+59:46:40.4	0.84	12.05	8.44 ± 0.34	⋯
MIPS8493	17:18:05.06	+60:08:32.6	1.80	12.73 (12.70)	8.62 ± 0.24	⋯
MIPS8543	17:18:12.54	+59:39:23.0	0.65	12.03	9.06 ± 0.14	⋯
MIPS22530	17:23:03.33	+59:16:00.1	1.96	13.05 (13.04)	8.97 ± 0.12	⋯
19456000	17:14:29.66	+59:32:33.7	1.97	13.27 (13.26)	8.93 ± 0.07	⋯

[a] $L_{\rm IR}^{\rm SF}$ is listed in parenthesis for those sources requiring a correction to account for nuclear activity as indicated by the presence of a hot dust continuum in the mid-IR spectrum.
[b] M_* was initially derived using a Salpeter IMF (see Kirkpatrick et al. 2012, for details), and we have corrected it to a Chabrier IMF, which we list here.

In Figure 3, we compare $L_{160}^{\rm SF}/L_{70}^{\rm SF}$ and $L_{250}^{\rm SF}/L_{70}^{\rm SF}$. The peak of the SED is traced by $L_{160}^{\rm SF}/L_{70}^{\rm SF}$, which is also a proxy for $T_{\rm dust}$. $L_{250}^{\rm SF}/L_{70}^{\rm SF}$ is difficult to interpret decoupled from $L_{160}^{\rm SF}/L_{70}^{\rm SF}$. If $L_{160}^{\rm SF}/L_{70}^{\rm SF}$ indicates temperature, then for a given $T_{\rm dust}$, $L_{250}^{\rm SF}/L_{70}^{\rm SF}$ can be thought of as the amount of cold dust (traced by $L_{250}^{\rm SF}$) relative to warm dust (traced by $L_{70}^{\rm SF}$). SEDs with proportionally more warm dust will have smaller ratios of $L_{250}^{\rm SF}/L_{70}^{\rm SF}$. All three samples have a tight relationship between $L_{160}^{\rm SF}/L_{70}^{\rm SF}$ and $L_{250}^{\rm SF}/L_{70}^{\rm SF}$.

Physically, this relationship can be described by how the incident radiation field in the ISM drives the far-IR emission (Witt & Gordon 2000; Misselt et al. 2001;

Dale & Helou 2002; Dale et al. 2014), which we illustrate using the suite of IR star forming templates from Dale et al. (2014). These model SEDs are created by combining the emission from different dust mass elements, heated by an incident radiation field of strength U, in a power law distribution:

$$dM_{\rm dust} \propto U^{-\alpha} dU \qquad (4)$$

We chose the Dale et al. (2014) templates due to the simplicity of interpretation. In this formulation, α characterizes the relative contributions of each dust mass element to the total SED, so that a smaller α means more of the dust in a galaxy is heated by a harsher radia-
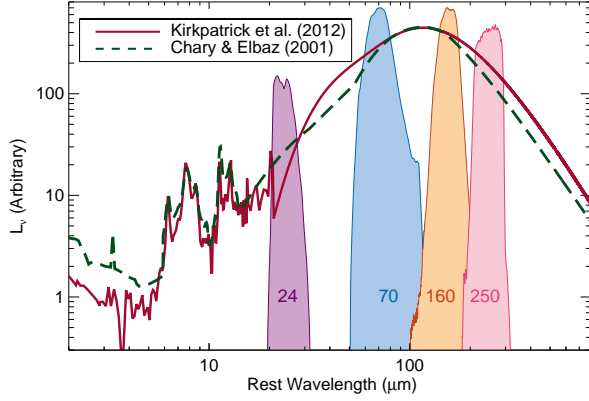
FIG. 2.— We illustrate where the broadband MIPS 24, 70, 160 μm and SPIRE 250 μm transmission filters sample the SED using a representative SFG template (red line) at $z \sim 1$ (Kirkpatrick et al. 2012) with $\log L_{IR} = 11.62 L_\odot$. To illustrate how the far-IR SED changes, we also plot a template (green dashed line) from the local Chary & Elbaz (2001) library, with $\log L_{IR} = 10.79 L_\odot$. The templates have been arbitrarily normalized so that the peaks coincide. The $z \sim 1$ template has proportionally more cold dust, notably at 250 μm, than the local template.

tion field. We show the location of the Dale et al. (2014) templates in Figure 3 using the thick grey-scale line. Our sources cluster around this line and have colors suggestive of $\alpha \in [1.125, 3.375]$, with the low z (GOALS + 5MUSES) and high z samples spanning a similar range of α.

The two low redshift samples show a distinct offset along the dark line, which is not unexpected given the different luminosities probed by 5MUSES and GOALS. We measure this offset to be 0.2 dex by fitting a line to the high z and low z (GOALS + 5MUSES) samples separately. Supersample sources in general have a higher L_{250}^{SF}/L_{70}^{SF} than GOALS or 5MUSES sources at fixed L_{160}^{SF}/L_{70}^{SF} (or T_{dust}), although both the high z and low z samples span the same range in each individual color. This suggests an evolution with redshift of the longer wavelength dust emission, due either to increased dust mass or possibly a colder dust component boosting the submm emission (e.g., Galametz et al. 2014).

4.1. Trends with L_{IR}^{SF}

Our higher z sample is also at a higher L_{IR}, which can confuse any redshift evolution. We compare L_{160}^{SF}/L_{70}^{SF} with L_{IR}^{SF} in Figure 4. Figure 4 shows a clear trend between L_{160}^{SF}/L_{70}^{SF} and L_{IR}^{SF}, otherwise known as the $L_{IR} - T_{dust}$ relation (Chapman et al. 2003; Casey et al. 2012; Magnelli et al. 2014; Lee et al. 2016). We overplot the mean of each sample in $\log(L_{IR}^{SF})$ bins of 0.25 as the larger symbols.

The 5MUSES and GOALS galaxies follow the same general trend, but there is a clear offset between the low z samples and Supersample, due to evolution of the $L_{IR} - T_{dust}$ relationship with redshift (Chapin et al. 2009; Casey et al. 2012; Magnelli et al. 2014). At a given L_{IR}^{SF}, the difference between low z and high z is ~ 0.2 dex, which corresponds to approximately a 5 K temperature difference. The Supersample galaxies have a much larger scatter than the GOALS sample, illustrating the increasing diversity of DSFGs with redshift (Symeonidis et al. 2011). Interestingly, the 5MUSES galaxies with $\log L_{IR}^{SF} > 11.4$ overlap with the Supersam-
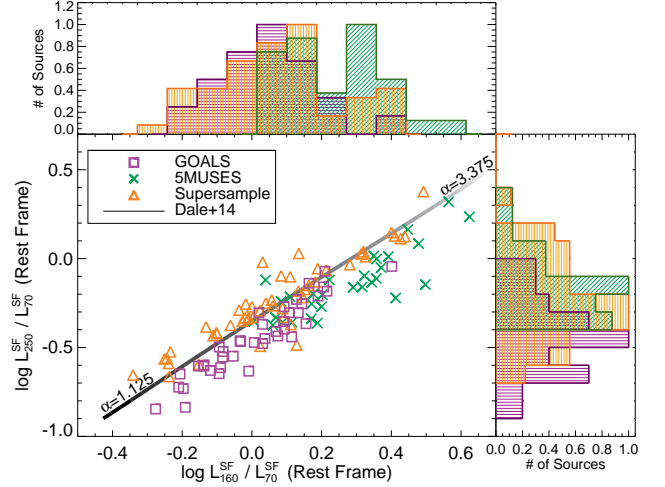

FIG. 3.— L_{160}^{SF}/L_{70}^{SF}, a proxy for T_{dust}, versus L_{250}^{SF}/L_{70}^{SF}, which measures the width of the SED. These colors are tightly correlated and increase with decreasing strength of the incident radiation field, as parameterized by the Dale et al. (2014) library (thick black/grey line). There is an offset between the low z galaxies (purple squares and green crosses) and the high z galaxies (orange triangles), with the high z sample having higher L_{250}^{SF}/L_{70}^{SF} for a given L_{160}^{SF}/L_{70}^{SF}, indicating enhanced submm emission and possibly higher dust masses.
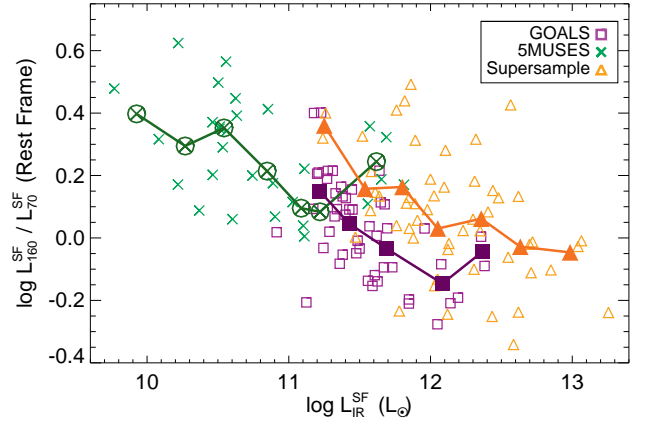

FIG. 4.— L_{160}^{SF}/L_{70}^{SF} vs. L_{IR}^{SF} for the GOALS and Supersample. We overplot the means as the filled symbols. This correlation is stronger for the low z samples (which have a Kendall's τ of -0.51), illustrating the increasing diversity of of DSFGs at high z (a Kendall's τ test shows $\tau = -0.33$ for the Supersample. There is a clear offset between the samples, with the Supersample having redder colors (colder T_{dust}), due to the evolution of the $L_{IR} - T$ relationship with redshift.

ple rather than the GOALS sample. These are also the highest redshift 5MUSES galaxies, with $z > 0.15$, and they also overlap with the Supersample below in Figure 5. Overall, we see a similar slope in the mean L_{160}^{SF}/L_{70}^{SF} vs. L_{IR}^{SF} points, with a normalization that depends on redshift, in agreement with the findings in Magnelli et al. (2014) and Casey et al. (2012).

5. EVOLVING TRENDS WITH DUST MASS

We now consider the physical parameter, M_{dust}. We have calculated M_{dust} in a self-consistent manner for all galaxies, while making as few assumptions as possible. Figure 5 shows a clear, strong trend between M_{dust} and L_{IR}^{SF} (Kendall's $\tau = 0.68$). The dashed line is fit to the 5MUSES and GOALS galaxies, and almost all the Su-
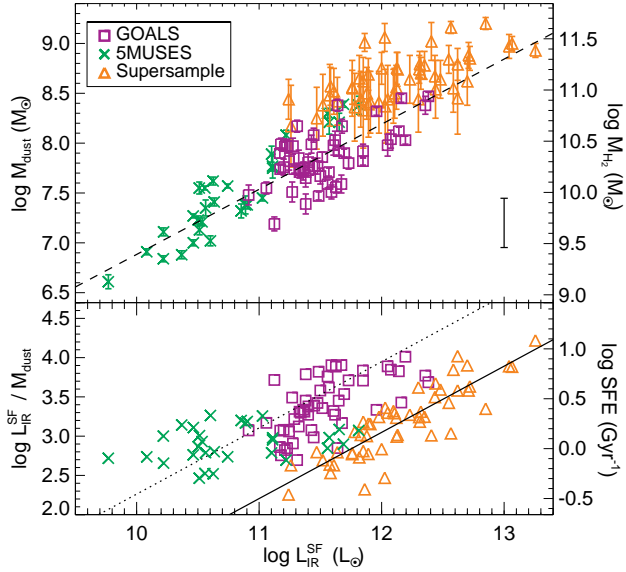
FIG. 5.— *Top*–$M_{\rm dust}$ v. $L_{\rm IR}^{\rm SF}$ for $T_{\rm cold} = 25\,$K. There is a tight correlation, which we parameterize with the dashed line, fit to the GOALS (purple squares) and 5MUSES (green crosses) sources. The Supersample (orange triangles) lies above this relation. In this and subsequent Figures 6 and 2, the uncertainties on $M_{\rm dust}$ simply reflect the photometric uncertainties on S_ν. The effect of changing $T_{\rm cold}$ by $\pm 5\,$K is illustrated by the solid bar in the bottom right corner. *Bottom*– $\log L_{\rm IR}^{\rm SF}/M_{\rm dust}$ as a function of $L_{\rm IR}^{\rm SF}$. The evolution with redshift is clearly seen, as the Supersample lies distinctly offset from the 5MUSES and GOALS samples.

persample sources lie above this relation. The fit is sub linear, with

$$M_{\rm dust} = 2.26 \times (L_{\rm IR}^{\rm SF})^{0.653} \qquad (5)$$

For clarity, the error bars on each point only reflect the error of the submm flux used to calculate $M_{\rm dust}$. The effect of changing $T_{\rm cold}$ by $\pm 5\,$K is illustrated by the solid bar in the bottom right corner of Figure 5 (Han et al. in prep). At first approximation, one might expect that $L_{\rm IR}/M_{\rm dust} \propto T_{\rm dust}^{4+\beta}$ in the optically thin regime, if the dust emission can be represented with a single temperature modified blackbody (Lanz et al. 2014; Magnelli et al. 2014). This simplification would imply a linear relationship between $L_{\rm IR}$ and $M_{\rm dust}$ for a fixed $T_{\rm dust}$. Since we measure a sub linear trend, this means that this model is unphysical since dust has a range of temperatures and other effects, such as optical depth or geometry, are affecting the far-IR/submm emission.

Although in the top panel of Figure 5, it appears that all samples may actually lie along the same relationship between $L_{\rm IR}^{\rm SF}$ and $M_{\rm dust}$, two separate relationships become evident in the bottom panel of Figure 5, where we show $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ v. $L_{\rm IR}^{\rm SF}$. There is a strong trend between $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ and $L_{\rm IR}^{\rm SF}$, with more luminous galaxies having proportionally less dust for a given radiation field, which results in warmer temperatures (Kendall's $\tau = 0.49$ (0.67) for the low (high) z samples). However, this trend (plotted as the solid line) shifts with redshift, so that Supersample galaxies with $L_{\rm IR}^{\rm SF} = 10^{12}\,L_\odot$ have $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ ratios similar to GOALS galaxies with $L_{\rm IR}^{\rm SF} = 2 \times 10^{11}\,L_\odot$ (shown by the dotted line); a complementary result was found for submillimeter galaxies in Chakrabarti & McKee (2008). The 5MUSES galax-

ies have a flatter distribution, although the galaxies that clearly overlap the Supersample are the galaxies at higher redshifts ($z > 0.15$). Similarly, Magdis et al. (2014) found that intermediate redshift DSFGs had larger gas reservoirs and lower dust temperatures than their local counterparts by a redshift of $z \sim 0.33$.

Submm fluxes can also be used to calculate gas mass (Scoville et al. 2016), although this requires extra assumptions, such as the expected dust-to-gas ratio. We illustrate on the right axis in Figure 5 what the H_2 gas mass would be, following the formalism of Scoville et al. (2016):

$$M_{\rm H_2} = \alpha_{850}^{-1} L_{850} \qquad (6)$$

where $\alpha_{850} = 6.7 \times 10^{19}\,{\rm ergs\,s^{-1}\,Hz^{-1}\,M_\odot^{-1}}$. We have CO luminosities for a handful of 5MUSES and Supersample galaxies (Kirkpatrick et al. 2014b; Yan et al. 2010), and $M_{\rm H_2}$ calculated from $L_{\rm CO}$ for these galaxies agrees with $M_{\rm H_2}$ calculated from L_{850}. Furthermore, the gas masses we derive here for our high z sample are completely consistent with the gas masses in Scoville et al. (2016) for galaxies spanning the same $L_{\rm IR}$ and redshift ranges. We then convert $L_{\rm IR}^{\rm SF}$ to SFR with SFR $[M_\odot\,{\rm yr}^{-1}] = 1.59 \times 10^{-10} \times L_{\rm IR}^{\rm SF}\,[L_\odot]$ from Murphy et al. (2011). In the bottom panel, we show how $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ translates to SFR/$M_{\rm H_2}$, which is the star formation efficiency (SFE). The relationship between SFE and $L_{\rm IR}^{\rm SF}$ evolves with redshift. That is, at a given $L_{\rm IR}^{\rm SF}$, local galaxies convert gas to stars *more* efficiently than high z galaxies. Assuming a dust-to-gas ratio of ~ 100, the $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ relation for the Supersample is in excellent agreement with that predicted by hyrdodynamical simulations combined with radiative transfer (Fig. 9 of Hayward et al. 2012). We remind the reader that we are comparing galaxies at a fixed luminosity, rather than a fixed location on the main sequence, which we examine in Section 6.3. At $L_{\rm IR}^{\rm SF} > 10^{11}$, local DSFGs lie predominantly above the main sequence, while $z \sim 1-2$ DSFGs lie on it, which would account for the difference in SFEs.

We now return to the question of what drives the cooler dust temperatures observed in higher redshift galaxies. Based on simulations, Safarzadeh et al. (2016) predict that $L_{\rm IR}$ and $M_{\rm dust}$ alone determine the shape of the SED, independent of the size of the galaxy. That is, $T_{\rm dust}$ depends solely on the total luminosity absorbed by dust ($L_{\rm IR}$) and the amount of absorbing material ($M_{\rm dust}$), in which case it is not necessary to measure all three parameters for a galaxy, since the third parameter can be inferred from the other two. We test this prediction with observations in Figure 6, where we plot $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ vs. $L_{160}^{\rm SF}/L_{70}^{\rm SF}$, a proxy for $T_{\rm dust}$.

There is a clear correlation between the parameters (Kendall's $\tau = -0.51$), and crucially, there is no offset with redshift. Magdis et al. (2012) found that $L_{\rm IR}/M_{\rm dust}$ is proportional to $\langle U \rangle$, the average interstellar radiation field. In turn, $\langle U \rangle$ determines $T_{\rm dust}$, giving rise to the strong correlation (Draine & Li 2007). The dot-dashed line shows the relationship:

$$\log\left(\frac{L_{\rm IR}^{\rm SF}}{M_{\rm dust}}\right) = -1.24 \times \log\left(\frac{L_{160}^{\rm SF}}{L_{70}^{\rm SF}}\right) + 3.68 \qquad (7)$$

which we have derived using all three samples, so this parameterization holds over three orders of magnitude

($L_{\rm IR}^{\rm SF} = 10^{10} - 10^{13} L_\odot$) and from $z \sim 0 - 2$. This parameterization can be used to estimate $M_{\rm dust}$ from far-IR data without requiring a longer wavelength ($\lambda > 500\,\mu{\rm m}$) observation. This relationship is also observed on much smaller scales ($\leq 1\,{\rm kpc}^2$) in local resolved galaxies from the KINGFISH survey (Kirkpatrick et al. 2014a). We calculate the residuals around the best fit line and show these in the bottom panel of Figure 6. Although the Supersample have a mean residual of -0.14 dex, indicating they lie slightly below the best fit line, the scatter of the residuals is quite large (standard deviation of 0.28 dex). Combined with the large uncertainties on the dust masses, this argues against any strong redshift evolution in Equation 7.

With the data available to us, the source of scatter in this Figure is open to interpretation. One likely source is the geometry of the ISM, which can explain how galaxies with the same $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ can have different $L_{160}^{\rm SF}/L_{70}^{\rm SF}$ ratios. A galaxy whose ISM approximates a shell geometry, where the stars are enclosed by a spherical shell of dust, will always peak at shorter wavelengths than an ISM geometry where the stars and dust are well mixed, all other parameters being equal (Misselt et al. 2001). This is because the shell geometry absorbs all of the incident radiation from stars, heating the dust to higher temperatures. Testing whether different geometries can account for all of the scatter between $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ and $L_{160}^{\rm SF}/L_{70}^{\rm SF}$ will require resolved observations of the ISMs with ALMA.

The relationship between $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ and $L_{160}^{\rm SF}/L_{70}^{\rm SF}$ can also explain the offset between the low and high z samples in Figure 4. It is possible that the normalization of the $L_{\rm IR} - T_{\rm dust}$ relation evolves with redshift because galaxies at $z = 1 - 2$ have higher dust masses at a given $L_{\rm IR}$ (Magdis et al. 2012). If at fixed $L_{\rm IR}$, the typical dust mass is higher in high redshift DSFGs than their $z \sim 0$ counterparts, then the dust temperature will be lower, shifting the normalization of the $L_{\rm IR} - T_{\rm dust}$ relation (Safarzadeh et al. 2016). This explanation for the shift does not require any change in the mode of star formation or ISM density (compact starburst vs. main sequence) or possibly in the dust composition. To produce colder temperatures for a given radiation field, an increase in the amount of large dust grains, relative to small ones, is required if the far-IR emission is optically thin. However, in the ULIRG regime, the high dust masses can cause the dust to be optically thick even at submm wavelengths (e.g., González-Alfonso et al. 2004), in which case simply adding more dust (and not changing the proportion of large to small grains) will decrease the temperature.

6. DISCUSSION

6.1. *Potential Evolution in Dust Properties?*

The DSFGs at $z \sim 1 - 2$ have higher dust masses for a given $L_{\rm IR}^{\rm SF}$ than local DSFGs. Here we want to make an important distinction. The observed quantity in our sample is the luminosity density in the Rayleigh Jeans tail, and it is the luminosity density that is actually higher in high redshift galaxies (as we have used the same assumptions for all other parameters when calculating $M_{\rm dust}$). The luminosity density in the Rayleigh Jeans tail is directly proportional to $M_{\rm dust} \times \kappa_\nu$, meaning that in principal, $M_{\rm dust}$ and κ_ν are degenerate. Thus far, we
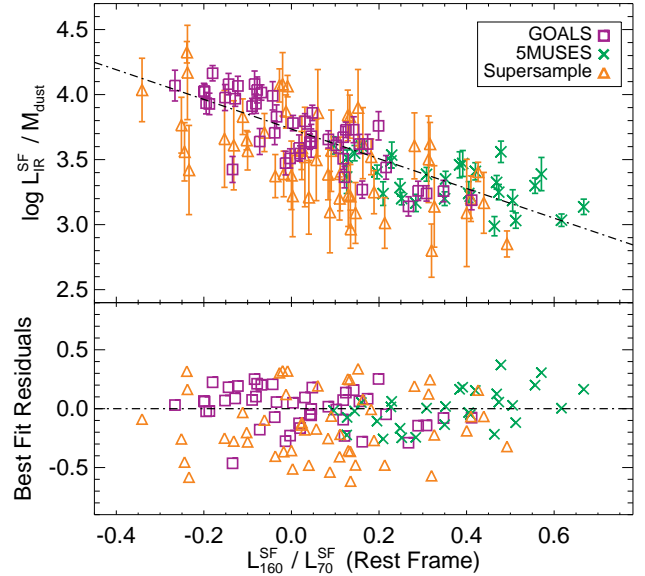

FIG. 6.— *Top–* $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ vs. $\log L_{160}^{\rm SF}/L_{70}^{\rm SF}$. There is a strong correlation between the two parameters, as exhibited by the dot dashed line, which is the best fit to all the data (Equation 7). The relationship between $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ and $L_{160}^{\rm SF}/L_{70}^{\rm SF}$ does not evolve with redshift, and so $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ alone can account for the colder temperatures in high z galaxies. *Bottom–* Residuals around the best fit line shown in the top panel. On average the Supersample has a mean residual of -0.14 dex, indicating they lie slightly below the low z samples. However, the scatter is quite large, with a standard deviation of 0.28 dex, arguing against any clear redshift evolution.

have been assuming the same opacity relationship for all galaxies, but in reality, the opacity might be evolving with redshift rather than the $M_{\rm dust}$. Dust opacity has a wavelength dependence

$$\kappa_\nu = \kappa_0 \left(\frac{\lambda}{\lambda_0}\right)^{-\beta} \quad (8)$$

so that either the normalization, κ_0, or the slope, β, could be evolving.

Although the Weingartner & Draine (2001) models all have similar κ_ν values in the submillimeter, grain compositions and distributions that differ considerably from these models can give different submillimeter opacities. The apparent increase in dust mass could possibly be due to a change in the dust grain properties with redshift (Dwek et al. 2014). For example, ice mantles can affect β, causing opacity to increase, and ice mantles should be prevalent in cold star forming regions (Tielens et al. 1984; Bergin et al. 2000; Spoon et al. 2002). The presence of ice mantles can increase κ_ν by a factor of 4 (Preibisch et al. 1993; Pollack et al. 1994) A factor of 4 increase in κ for the high redshift galaxies would reconcile the difference in $M_{\rm dust}$ for the GOALS and Supersample. The strength of water ice, silicate absorption, and $3.4\,\mu{\rm m}$ PAH emission features has been measured in a handful of Supersample sources and closely resembles what is measured in local GOALS galaxies with similar obscuration, providing some indication that the amount of ice in dusty galaxies does not strongly evolve with redshift (Sajina et al. 2009).

If β in our galaxies differs significantly from β in the Weingartner & Draine (2001) models, for example, varying with redshift, then our assumed κ_ν will be incorrect.

The emissivity measure from the Rayleigh Jeans tail is observed to be shallower in lower metallicity galaxies (Galametz et al. 2011; Kirkpatrick et al. 2013b), though there is little evidence that metallicity evolves strongly in massive, dusty galaxies out to $z \sim 2$ with similar M_*, SFR, and $M_{\rm mol}$ (Mannucci et al. 2010; Magdis et al. 2012; Bothwell et al. 2016). In fact, when we measure β directly for individual galaxies by fitting a modified blackbody with $T = 25\,{\rm K}$, we find $\beta = 1.8$ for the Supersample and $\beta = 2.02$ for the GOALS sample. We are not fully sampling the Rayleigh Jeans tail in the Supersample, and longer wavelength observations are required to more accurately measure the effective β, but with current observations, we see little evidence supporting a drastic increase in opacity required to reconcile dust mass measurements of our samples. Longer wavelength observations, such as with ALMA or the Large Millimeter Telescope, are required to better measure the effective β at $z = 1 - 2$.

6.2. Potential Role of Mergers or Compactness?

Additional insight into the ISM of galaxies can be gained by examining morphologies and merger signatures, which we have for the GOALS and Supersample sources. Galaxies hosting a compact central starburst will have a higher temperature than galaxies at the same $L_{\rm IR}$ which are disks (Chanial et al. 2007). In local LIRGs and ULIRGs, a compact starburst is triggered by a major merger, as the transfer of angular momentum funnels gas to the central kpc of the merging system.

Stierwalt et al. (2013) classified the GOALS sample into merger classifications using HST $B-$, $I-$, and $H-$band imaging (see also Haan et al. 2011) and IRAC $3.6\,\mu$m imaging. All 46 of the galaxies in the present work also have a merger classification. The merger classifications we use here are **0:** no merger or massive neighbor; **1:** galaxy pair; **2:** early-stage (disk still in tact); **3:** mid-stage (amorphous disk, tidal tails); **4:** late-stage (two nuclei in common envelope). Morphology analysis of the Supersample is less straightforward due to our combining different surveys. GOODS-S and GOODS-N galaxies have been visually classified with HST WFC3 images as part of the Cosmic Assembly Near-ir Deep Extragalactic Legacy Survey (CANDELS; P.I. S. Faber & H. Ferguson). The CANDELS tool gives classifiers a choice of 5 different interaction stages, similar to those listed above. The morphology catalogs then contain the fraction of classifiers that selected each interaction stage (Kartaltepe et al. 2015). We use these fractions to determine a weighted average interaction stage, from $0-4$ (see also Rosario et al. 2015, which create a weighted average on a 0-1 scale). Zamojski et al. (2011) classifies 134 xFLS sources into merger classifications using $Hubble$ NICMOS imaging, from 0-5. Here, we combine old mergers (5) with advanced mergers (4), to match the categories described above. In total, 17 Supersample galaxies have dust masses and a merger classification. 5MUSES galaxies lack a morphology classification.

The top panel of Figure 7 shows interaction stage as a function of $L_{\rm IR}^{\rm SF}/M_{\rm dust}$, and the points are shaded by $L_{160}^{\rm SF}/L_{70}^{\rm SF}$. There is no correlation between the parameters, for either the GOALS or Supersample. Galaxies at a given $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ occupy all merger classifications, as do galaxies at a given $L_{160}^{\rm SF}/L_{70}^{\rm SF}$. DSFGs exhibit a va-

riety of $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ ratios and a variety of major merger classifications, indicating that the major merger scenario cannot directly explain all of the far-IR observed properties of these galaxies. In the same vein, using hydrodynamical simulations post-processed with dust radiative transfer, Lanz et al. (2014) found that the effective dust temperature of mergers does not reflect the merger stage except for a very short period ($\lesssim 100\,{\rm Myr}$) during the coalescence-induced starburst.

The most direct test of how ISM geometry affects $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ and $T_{\rm dust}$ is to compare the spatial extent of the dusty ISM. Measuring submm sizes on a resolved scale at low and high redshift requires the sensitivity of ALMA. To date, only a handful of DSFGs have resolved ALMA imaging, but we can compare other tracers of the ISM until we have a larger ALMA sample. Rujopakarn et al. (2011) compiles from the literature ISM diameters for the GOALS sample measured from Paα, 8.4 Ghz, and CO (3-2). The authors also compile CO (3-2) and 1.4 GHz sizes for galaxies in the GOODS-N field. The diversity of measurements underscores the need for a homogeneous ALMA comparison. The bottom panel of Figure 7 compares ISM diameter with $L_{\rm IR}^{\rm SF}/M_{\rm dust}$. The GOALS DSFGs are all much smaller than the high z DSFGs. However, this does not translate to different $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ or $L_{160}^{\rm SF}/L_{70}^{\rm SF}$ ratios. This echoes the result in Lee et al. (2016), where the authors find that local dust obscured galaxies, likely very compact, with $L_{\rm IR} = 10^{11} - 10^{12}$ have similar dust masses and temperatures as their less obscured, more extended counterparts. Similarly, using SEDs output from hydrodynamical simulations, Martínez-Galarza et al. (2016) recently found that mergers and compactness alone are not able to reproduce the SEDs of local LIRGs, but that increasing the gas fraction was required.

However, without a large sample of measurements of galaxy size in the dust continuum (i.e., with ALMA), it is very difficult to distinguish between size and $M_{\rm dust}$ as the driver of the colder dust temperatures, since these effects are linked. Figure 7 clearly demonstrates that the high z galaxies are larger. If a galaxy can be represented as a central heating source surrounded by a disk of dust, then naturally a more extended disk will produce lower $T_{\rm dust}$, as the outskirts of the disk will see a diminished radiation field (e.g. Misselt et al. 2001). But, this is not a realistic description of most galaxies. Instead, to first order, both the SFR and dust densities are correlated with gas density, and so a mixed geometry of SFR regions within the ISM is more likely. In this geometry, the extent of the system doesn't affect the global effective dust temperature (Misselt et al. 2001; Safarzadeh et al. 2016). How size and increased dust/gas mass are linked is also difficult to unravel. Locally, compaction can lead to a more efficient transformation of atomic gas into molecular gas (Larson et al. 2016), so a decrease in size could be actually the trigger of increased molecular gas masses in many of the GOALS galaxies (Díaz-Santos et al. 2010, 2011). Alternately, smaller size translates to a higher SFR surface density, which would boost $L_{\rm IR}^{\rm SF}$ without requiring a boost in $M_{\rm gas}$, leading to higher SFE at smaller size (Hayward et al. 2011, 2012). These effects are observed locally and in simulations, but our small sample of high z galaxies do not appear to show any of the same trends between small size and higher
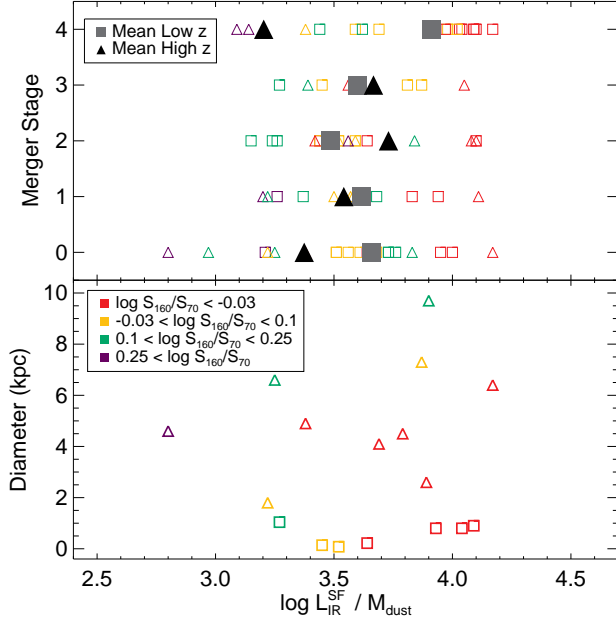
FIG. 7.— *Top panel* – merger classification vs. L_{IR}^{SF}/M_{dust}, where sources are shaded by L_{160}/L_{70}^{SF}. There is no correlation between merger classification and L_{IR}^{SF}/M_{dust} or L_{160}/L_{70}^{SF}. To help guide the eye, we plot the mean L_{IR}^{SF}/M_{dust} in each merger stage for the low z galaxies (large grey squares) and high z galaxies (large black triangles). *Bottom panel* – ISM diameters from Rujopakarn et al. (2011), measured using CO, radio, or Paα, as a function of L_{IR}^{SF}/M_{dust}. While the high z galaxies are all clearly larger than the GOALS galaxies, there is no correlation between size and L_{IR}^{SF}/M_{dust}. The far-IR/submm data seems to tell observers very little about the compactness or merger stage of galaxies.

temperatures or L_{IR}^{SF}/M_{dust}. These galaxies also have optical radii measurements from the CANDELS collaboration, and we find no correlation between optical radius measured in the observed frame H-band and L_{IR}^{SF}/M_{dust}. On the other hand, Scoville et al. (2016) argue that high z galaxies have more turbulent ISMs, leading to compression in the ISM and enhancing the star formation efficiency per unit mass. This efficient mode of star formation can occur throughout the galaxy, so no obvious correlation between galaxy size and dust temperature may be expected.

Local galaxies can be resolved, allowing their ISM geometry to be measured in more detail than for high redshift galaxies (e.g., Barcos-Muñoz et al. 2015). Given the quality and abundance of observations, it is more straightforward to link mergers with compact starbursts with dust heating in the GOALS samples (Díaz-Santos et al. 2010). However, without resolved observations of the ISM in high redshift galaxies, we see no obvious link between mergers, size, L_{IR}^{SF}/M_{dust}, and T_{dust} in our sample. Therefore, we must conclude that a galaxy's global far-IR/submm emission, parameterized through L_{IR}, T_{dust}, M_{dust}, or f_{gas} may not tell observers anything about the ISM geometry, extent, or merger stage of that galaxy.

6.3. *The Effect of Increased Gas Fraction*

There are two likely explanations for the observed increase in M_{dust} with redshift. The first is that DSFGs at $z \sim 1 - 2$ are simply more massive overall, that is, they
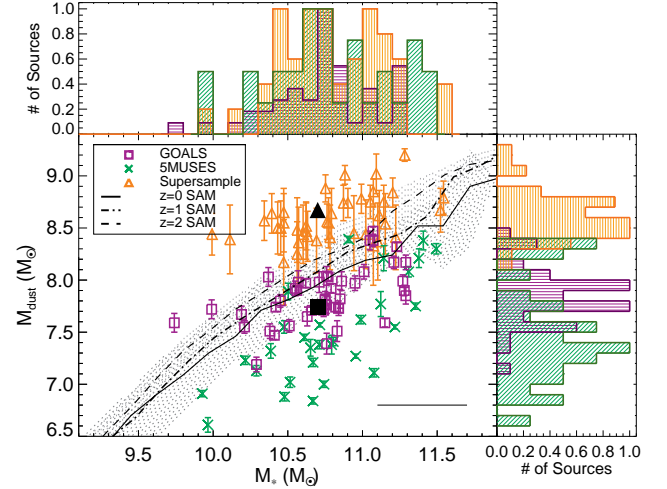

FIG. 8.— We explore how much of the increase in dust mass is related to the evolution of the main sequence (MS) with redshift. The Supersample (orange triangles) span the same range of stellar mass as the GOALS galaxies (purple squares) and 5MUSES galaxies (green crosses). We overplot the predicted relationship between M_{dust} and M_* for 3 different redshifts from the models of Popping et al. (2016). Although generally consistent with the lower redshift galaxies, the Supersample galaxies have significantly more dust than predicted by the models. We also predict what M_{dust} a galaxy with $\log M_* = 10.7\,M_\odot$ would have at $z = 0$ and $z = 1.2$ based on the evolution of f_{gas} with z and assuming a constant dust to gas ratio. The dark triangle ($z = 1.2$) and square ($z = 0$) demonstrate the predictions are consistent with the measurements. The bar in the lower right indicates the estimated errors on the 5MUSES and GOALS galaxies based on comparison with $K-$band derived stellar masses. If the 5MUSES galaxies are overestimated, this would explain the lack of consistency with the (Popping et al. 2016) models at $z \sim 0$.

have a higher stellar mass. M_* is expected to broadly scale with M_{dust}, since a higher M_* in a DSFG implies a higher metallicity, supplying more metals to form dust in the ISM.

In Figure 8, we compare M_{dust} with M_*. The three samples span the same range of $\log M_* \sim 10 - 11.5\,M_\odot$. However, the Supersample have roughly an order of magnitude higher dust masses. The solid, dotted, and dashed lines indicate how the relationship M_{dust} v. M_* is predicted to evolve with redshift using the semi-analytic models of Popping et al. (2016). We find an increase in the M_{dust} with redshift larger than predicted by those models. The Supersample is offset from the GOALS galaxies by ~ 0.7 dex, which is the same amount of offset seen in the bottom panel of Figure 5. The Popping et al. (2016) models shown in our comparison are based on a semi-analytic model of galaxy formation in a cosmological context, which includes a standard suite of physical processes (gas accretion and cooling, star formation, stellar feedback, chemical enrichment, etc). In addition, the Popping et al. (2016) model includes self-consistent tracking of the main processes thought to produce and destroy dust in galaxies, including dust condensation in stellar ejecta, dust growth through accretion in the ISM, dust destruction by supernovae, and ejection of dust by stellar driven winds. The much milder evolution of M_{dust}/M_* with redshift predicted by these models relative to our findings is interesting, as it indicates that one or more of the model ingredients needs to be revised.

The main sequence, which is the relationship between SFR and M_*, evolves with redshift (e.g. Whitaker et al.

2012). That is, for a given M_*, galaxies had a higher SFR at $z \sim 1 - 2$ than today. A higher SFR can be tied to an increase in dust mass indirectly, as the majority of grain growth is predicted to occur in the ISM, and a more gas rich ISM will lead to higher dust masses and higher SFRs (e.g., Dwek 1998; Draine 2003; Santini et al. 2014; McKinnon et al. 2016a). The increased dust mass in the Supersample might then be a natural consequence of the increase in gas fractions with lookback time. As mentioned above, submm data can also be converted to M_{H_2}, and we use this parameterization to calculate gas fractions $f_{gas} = M_{H_2}/(M_* + M_{H_2})$. These gas fractions are consistent with what we derive using CO observations for several 5MUSES and Supersample galaxies (Yan et al. 2010; Kirkpatrick et al. 2014b).

We also calculate the distance of these galaxies from the main sequence. The main sequence evolves with redshift, and it flattens at higher M_*. Therefore, we follow the method in Scoville et al. (2017) to calculate the main sequence. We use the relationship between M_* and SFR parameterized in Lee et al. (2015) at $z = 1.2$:

$$\log \mathrm{SFR_{MS}} = 1.72 - \log\left[1 + \left(\frac{M_*}{2 \times 10^{10} M_\odot}\right)^{-1.07}\right] \quad (9)$$

and then we renormalize depending on redshift (Speagle et al. 2014):

$$\mathrm{SFR_{MS}}(z) = \left(\frac{1+z}{1+1.2}\right)^{2.9} \times \mathrm{SFR_{MS}}(z=1.2) \quad (10)$$

Then, we calculate sSFR = SFR/M_* for every source and for the main sequence (MS) at each redshift.

We plot f_{gas} vs. ΔsSFR = sSFR/sSFR$_{MS}$ in the top panel of Figure 9. The Supersample is offset relative to the low z samples, so that at a given ΔsSFR, the Supersample sources have higher gas fractions, consistent with previous results in the literature (Magdis et al. 2012; Genzel et al. 2015). This means that massive galaxies at $z \sim 1.2$ have $4 - 5\times$ higher gas fractions, and hence gas masses, since the stellar masses are roughly consistent. Scoville et al. (2017) parameterize the evolution of the molecular gas mass as a function of redshift, M_*, and ΔsSFR. In particular, the authors find that M_{H_2} evolves as $(1 + z)^{1.84}$. For $\log M_* = 10.7 M_\odot$, and $z = 0$, $z = 1.2$, we calculate the gas masses based on Equation 6 in Scoville et al. (2017) and overplot these predictions as the dashed and dot-dashed lines. The predicted redshift evolution of $(1 + z)^{1.84}$ is consistent with our measurements, given the uncertainties in our stellar masses.

This redshift evolution is best interpreted by also considering star formation efficiency, SFE = SFR/M_{H_2}. In the case of a constant dust to gas ratio, SFE $\propto L_{IR}^{SF}/M_{dust}$; massive dusty galaxies show little variation in metallicity, even at $z \sim 1 - 2$, so a constant dust to gas ratio is a fair assumption (Magdis et al. 2012). We plot SFE vs. ΔsSFR in the bottom panel of Figure 9. The obvious correlation (Kendall's $\tau = 0.45$) is similar to the well established dependence of the gas depletion timescale ($t_{dep} = \mathrm{SFE}^{-1}$) on sSFR (Saintonge et al. 2011; Magdis et al. 2012; Tacconi et al. 2013; Sargent et al. 2014; Huang & Kauffmann 2014). We overplot the relation derived in Genzel et al. (2015), where the authors use CO observations to calculate

molecular gas mass for 500 SFGs from $z = 0 - 3$. There is no offset in any of our samples, consistent with the very mild redshift evolution between t_{dep} and ΔsSFR found by Genzel et al. (2015). However, this is in contrast to Scoville et al. (2017), which measure an increase in SFE $\propto (1 + z)^{1.05}$ for a given stellar mass. We plot what this evolution looks like as the arrow in the bottom right corner for $z = 0$, $z = 1.2$ and $\log M_* = 10.7 M_\odot$.

The clear increase in gas fractions with redshift, without corresponding increases in SFE, suggests that the higher SFRs (corresponding to higher L_{IR}) and higher M_{dust} of the Superample at $z \sim 1 - 2$ are driven by an increase in the gas supply rather than an increase in SFE, M_*, or changing mode of star formation with cosmic time (Bouché et al. 2010; Tacconi et al. 2013; Davé et al. 2012; Genzel et al. 2015; Schinnerer et al. 2016). In fact, the bottom panel of Figure 5 demonstrates that at a given L_{IR}^{SF} (SFR), SFEs are *lower* at $z \sim 1 - 2$. This increase in gas supply is also predicted to by the underlying cause of the main sequence evolution, SFR/M_*, with redshift (Bouché et al. 2010; Tacconi et al. 2013; Davé et al. 2012; Genzel et al. 2015). We can directly test whether an increase in f_{gas} fully explains the increase in M_{dust} using the relationship between f_{gas} and z, parameterized as $f_{gas} = 0.1(1 + z)^2$ (Geach et al. 2011). If we assume a dust-to-gas ratio of 100 and $\log M_* = 10.7 M_\odot$, then at $z = 0$, $\log M_{dust} = 7.75 M_\odot$, and at $z = 1.2$, $\log M_{dust} = 8.67 M_\odot$. These predictions are plotted as the dark filled symbols in Figure 8, and they agree remarkably well with the GOALS and Supersample.

7. CONCLUSIONS

We have analyzed the far-IR/submm properties of three samples of infrared luminous galaxies. We combine GOALS and 5MUSES galaxies to form a low z sample ($z < 0.3$), and the Supersample comprises our high z sample ($z \sim 0.5 - 2$). Mid-IR spectra are available for every source, and we utilize these spectra to identify sources with a hot dust continuum, likely due to an AGN, and remove them from the sample. We measure L_{250}^{SF}/L_{70}^{SF}, L_{160}^{SF}/L_{70}^{SF}, M_{dust}, f_{gas} self-consistently for all galaxies. We find

1. L_{250}^{SF}/L_{70}^{SF} is tightly correlated with L_{160}^{SF}/L_{70}^{SF} for all galaxies and both ratios are sensitive to the strength of the ISRF, as parameterized by SED models. The low z and high z samples span the same range of colors, although for a given L_{160}^{SF}/L_{70}^{SF}, the high z galaxies have a higher L_{250}^{SF}/L_{70}^{SF}, indicating that they have proportionally more cold dust.

2. L_{160}^{SF}/L_{70}^{SF} is correlated with L_{IR}^{SF}, but this relationship evolves with redshift so that at $z \sim 1 - 2$, DSFGs are $\sim 5\,\mathrm{K}$ colder.

3. There is a strong relationship between L_{IR}^{SF}/M_{dust} and L_{IR}^{SF}, and this relationship evolves with redshift. DSFGs at $z \sim 1 - 2$ have L_{IR}^{SF}/M_{dust} ratios similar to low z galaxies a factor of 5 less luminous. DSFGs also have higher dust masses than their local counterparts at the same luminosity.

4. The relationship between M_{dust} and M_* also evolves strongly with redshift, so that DSFGs at $z \sim 1 - 2$ have a > factor of 5 higher M_{dust} than
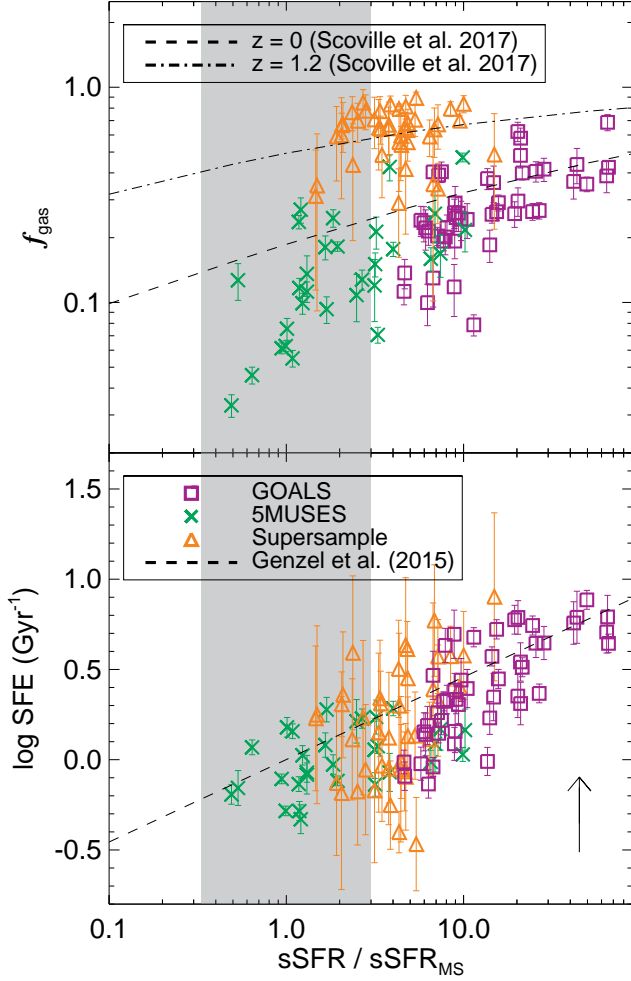
FIG. 9.— *Top panel–* $f_{\rm gas} = M_{\rm H_2}/(M_* + M_{\rm H_2})$ as a function of distance from the main sequence (grey shaded region). Starburst galaxies have higher gas fractions, and this increases with redshift. We overplot the predicted redshift evolution of $(1+z)^{1.84}$ from Scoville et al. (2017), and this evolution agrees with our samples. *Bottom panel–* Star formation efficiency as a function of distance from the main sequence. Now, there is no offset with redshift, in agreement with Genzel et al. (2015). In contract, the predicted evolution of $(1+z)^{1.05}$ measured by Scoville et al. (2017) is shown as the arrow in the lower right corner for $z = 0$ to $z = 1.2$.

local DSFGs in the same M_* range. This mirrors the measured increase in gas fraction with redshift from other studies.

5. We observe a redshift-independent correlation between distance from the main sequence (ΔsSFR) and star formation efficiency (SFE). The higher star formation rates (corresponding to higher $L_{\rm IR}$) in our $z \sim 1-2$ sample are then fueled by an increase in their gas content, rather than an increase in SFE.

6. We see a tight redshift-independent correlation between $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ and $L_{160}^{\rm SF}/L_{70}^{\rm SF}$, which is a proxy for $T_{\rm dust}$. We see no correlation between merger classification and either $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ or $L_{160}^{\rm SF}/L_{70}^{\rm SF}$. This is also true of galaxy size. Then, the change to colder $T_{\rm dust}$ in high z DSFGs can be explained simply by an increase in dust mass at fixed $L_{\rm IR}^{\rm SF}$, without requiring any change in the merger frac-

tion or extent of the ISM, although this may be a related effect.

7. In the subsample of galaxies for which we have size measurements, the large galaxy sizes at high z and lack of a correlation between size and L_{160}/L_{70} favor an ISM geometry in which the stars and dust are well mixed, so the expectation based on a central source surrounded by dust (that $T_{\rm dust}$ is inversely proportional to size) does not hold.

8. The far-IR/submm SED can be fully parameterized in terms of the observables $L_{\rm IR}^{\rm SF}$ and $L_{160}^{\rm SF}/L_{70}^{\rm SF}$ or, alternatively, the dust mass and the luminosity absorbed by dust (which is equal to $L_{\rm IR}^{\rm SF}$ unless dust self-absorption is non-negligible). As such, a galaxy's global far-IR/submm emission alone cannot be used to distinguish between ISM geometries.

APPENDIX

A. REST FRAME PHOTOMETRY

We plot the IR/submm SEDs of the Supersample galaxies listed in Table 1. We also plot the estimated rest frame 70, 160, 250 μm photometry. For two galaxies, GN_IRS49 and MIPS8543, we do not have enough coverage at $\lambda > 200\,\mu$m to reliably estimate a rest frame L_{250}. The errors on the estimated rest frame photometry are largely attributable to the uncertainties on the templates (Kirkpatrick et al. 2015) used to derive the photometry.
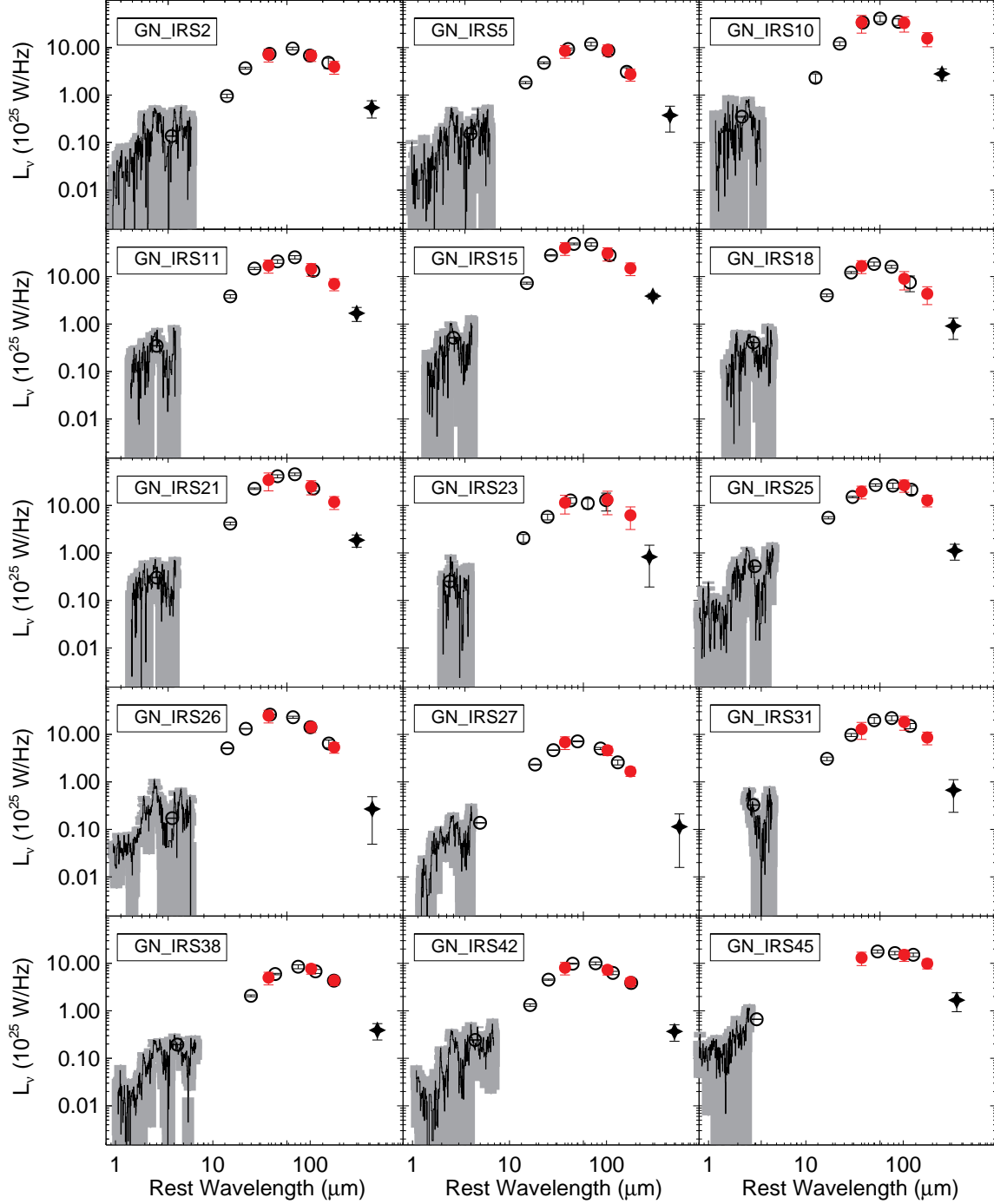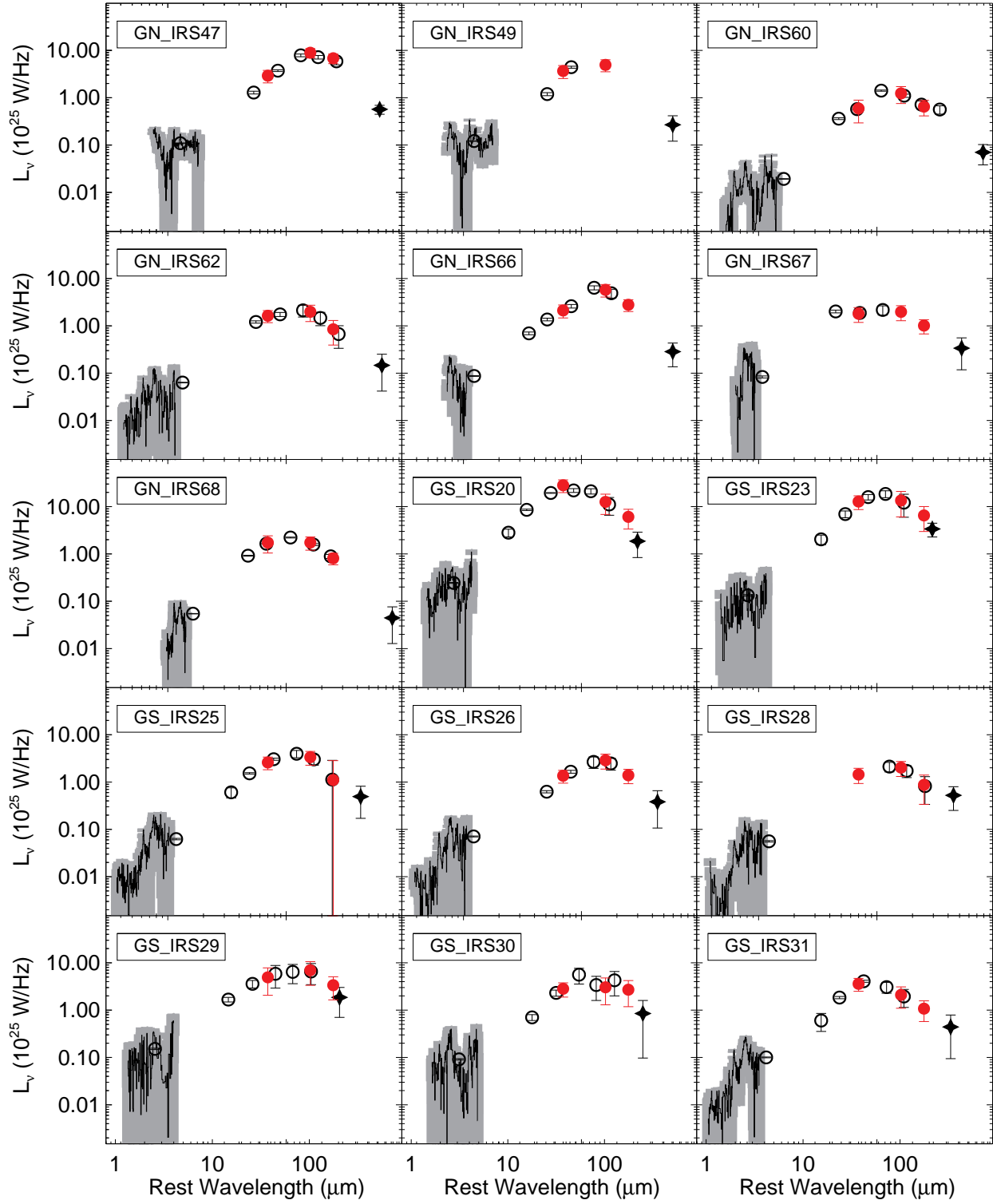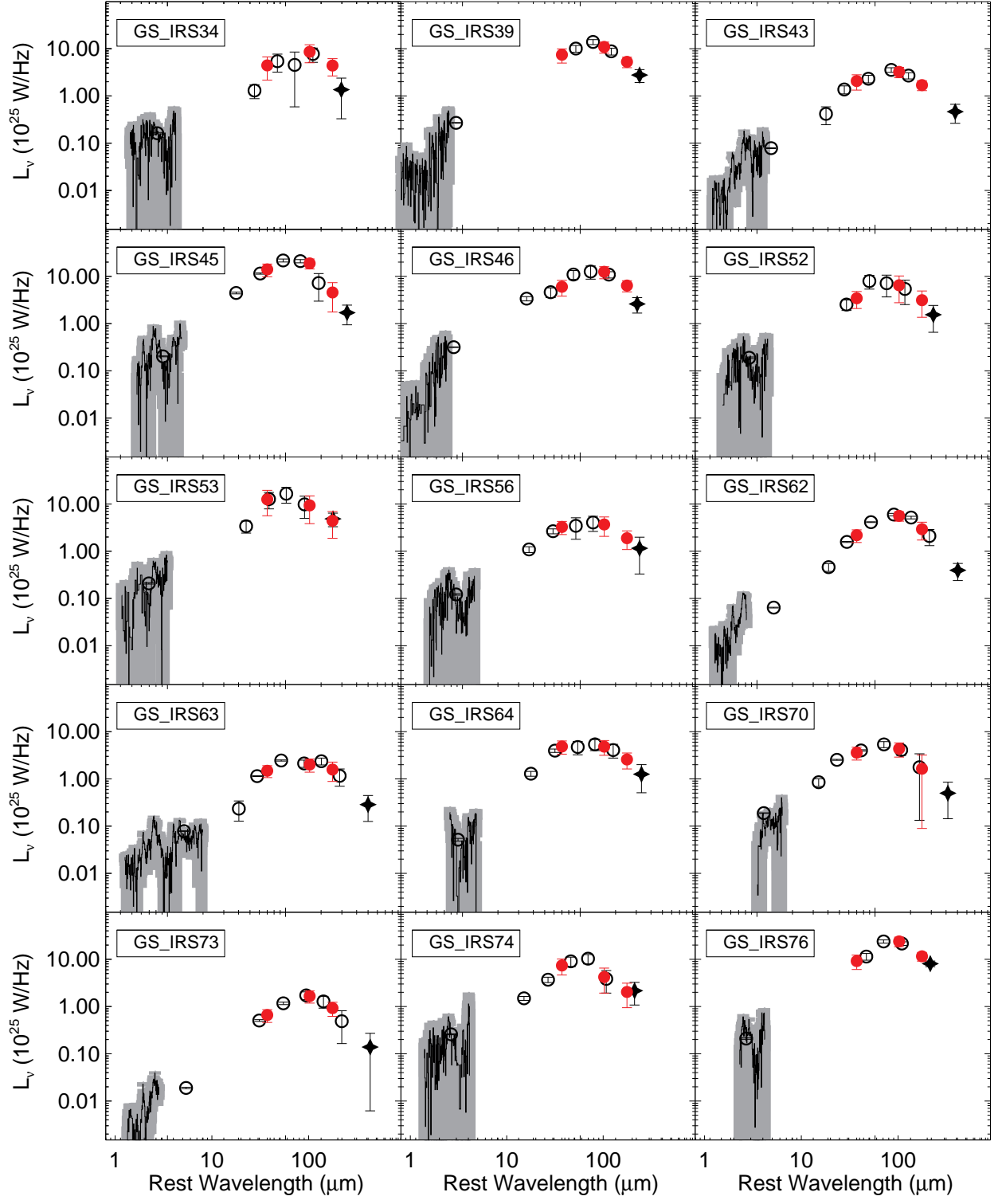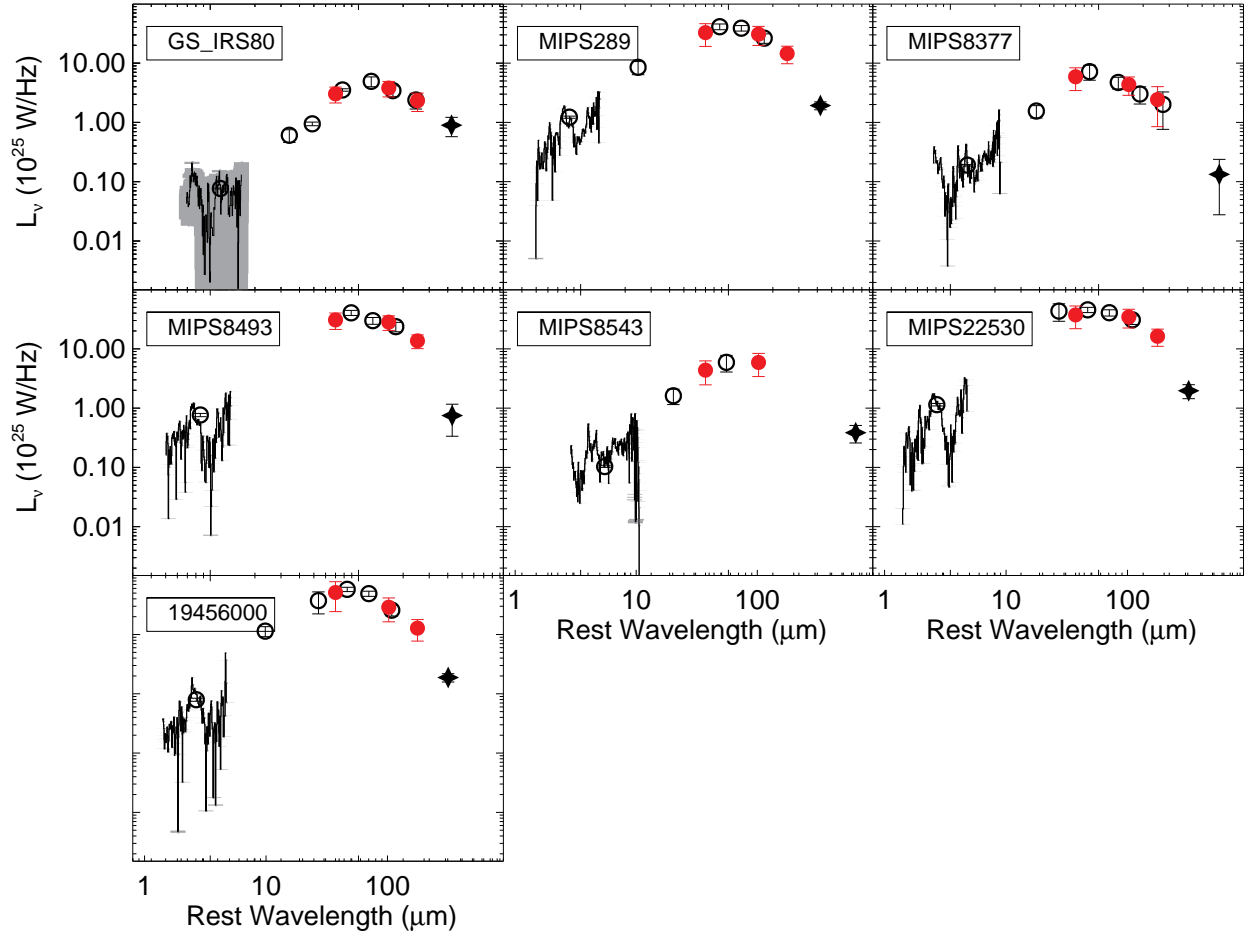

FIG. 10.— We demonstrate how accurately the estimated rest frame 70, 160, 250 μm photometry (red filled circles) matches the observed *Herschel* and *Spitzer* SEDs for the Supersample. We plot as the filled stars the ground based submm photometry used to calculate $M_{\rm dust}$.

FIG. 1.— *Continued*

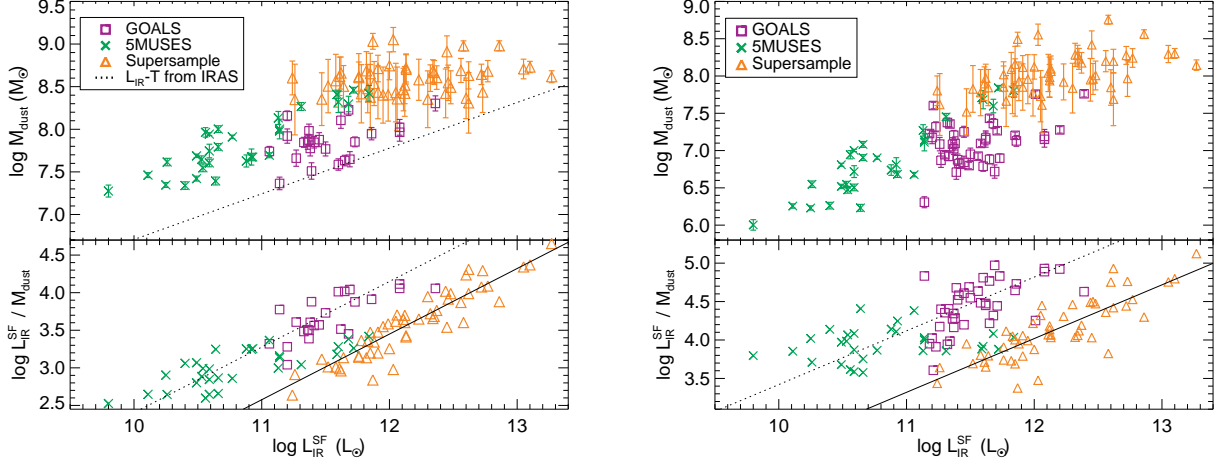FIG. 1.— *Continued*

Fig. 1.— *Continued*

FIG. 2.— $M_{\rm dust}$ v. $L_{\rm IR}^{\rm SF}$ with $T_{\rm dust}$ varying. *Left panel–* $T_{\rm dust}$ is approximated from a source's $L_{\rm IR}^{\rm SF}$. The dotted line shows how $M_{\rm dust}$ of the GOALS and 5MUSES galaxies will change when using the low z $L_{\rm IR} - T_{\rm dust}$ relation from IRAS BGS sources (Chapman et al. 2003). *Right panel–* $T_{\rm dust}$ and β (used to derive κ) come from the best fitting IR template to each source from the library of Dale et al. (2014). In both panels, there is a clear evolution of $L_{\rm IR}^{\rm SF}/M_{\rm dust}$ with redshift, indicating that the method used to derive $M_{\rm dust}$ is not responsible for this trend.

B. DUST MASS CALCULATIONS

We explore how trends between $L_{\rm IR}^{\rm SF}$ and $M_{\rm dust}$ will change if we use different $T_{\rm dust}$ to calculate $M_{\rm dust}$.

Determining a dust temperature from the SED involves fitting a modified blackbody to the far-IR data. This temperature is then a luminosity-weighted temperature that may not represent the true temperature of the bulk of the dust mass. Instead, the effective dust temperature probes the mean ionization parameter U "seen" by the dust (Draine & Li 2007). If a single modified blackbody is used, $T_{\rm dust}$ is measured using all far-IR photometry, but below $\lambda < 100\,\mu$m, dust is being heated primarily in star forming regions, as opposed to the diffuse dust in the ISM heated by the interstellar radiation field (Lonsdale Persson & Helou 1987; Dunne & Eales 2001). Dust masses calculated from a one temperature modified blackbody can underestimate the true dust mass by at least a factor of 2 (Magdis et al. 2012). The cold diffuse component is what makes up the bulk of the dust mass. $T_{\rm cold}$ is often measured by fitting a two temperature modified blackbody (e.g., Dunne & Eales 2001; Kirkpatrick et al. 2012; Lee et al. 2016), but this method requires full sampling of the far-IR SED or assumptions about β and the warm dust component, not to mention that the model is unphysical.

We now explore trends in dust mass when the dust temperature of the diffuse ISM is assumed to scale with increasing $L_{\rm IR}^{\rm SF}$, as this is consistent with what would be measured by fitting a one temperature modified blackbody to the SED (e.g., Blain et al. 2003; Chapman et al. 2003; Chapin et al. 2009; Casey et al. 2012; Lee et al. 2016). We calculate $T_{\rm dust}$ according to

$$T_{\rm dust} = 0.47 \times (L_{\rm IR}^{\rm SF})^{0.144} \tag{B1}$$

determined from 767 SPIRE 250 μm selected sources spanning $z = 0 - 5$ where $T_{\rm dust}$ was calculated from the peak wavelength of the SED (Casey et al. 2012); we ignore the redshift evolution of the normalization for the time being. The left panel of Figure 2 shows $M_{\rm dust}$ as a function of $L_{\rm IR}^{\rm SF}$. There is an overall trend towards increasing $M_{\rm dust}$ with increasing $L_{\rm IR}^{\rm SF}$, although for the the Supersample, $M_{\rm dust}$ is nearly constant over two orders of magnitude in $L_{\rm IR}^{\rm SF}$. The Supersample is offset from the GOALS and 5MUSES sample, with higher $M_{\rm dust}$ at similar $L_{\rm IR}^{\rm SF}$, confirming the trend in Figure 5.

The GOALS and 5MUSES galaxies lie decidedly below the Supersample, and this is not explained by evolution in the $L_{\rm IR} - T_{\rm dust}$ relation with redshift. We can test for the effect of redshift by instead using the $L_{\rm IR} - T_{\rm dust}$ relation derived from local IRAS BGS sources, which we approximate as $T_{\rm dust} \propto (L_{\rm IR}^{\rm SF})^{0.137}$ (Chapman et al. 2003; Casey et al. 2012). When we apply the IRAS BGS relation to the 5MUSES and GOALS galaxies, they lie around the dotted line in Figure 2. Rather than bringing the low z and high z samples into better agreement, using two $L - T$ relations will make the dust masses *more* discrepant.

Alternately, we can calculate $T_{\rm dust}$ and β through template fitting. The Dale et al. (2014) library, which we use to approximate the far-IR colors of our galaxies in Figure 3, has an effective $T_{\rm dust}$ and β associated with each template. We fit the Dale et al. (2014) library to all sources using a χ^2 minimization technique. We then use the best fit template β to scale κ as

$$\kappa_\lambda = \kappa_{850} \times \left(\frac{850}{\lambda}\right)^\beta \tag{B2}$$

where $\kappa_{850} = 0.15\,{\rm m}^2/{\rm kg}$ (Weingartner & Draine 2001) and λ is the submm wavelength we calculate dust mass at for each galaxy. We use κ_λ and $T_{\rm dust}$ associated with the best fit template to calculate $M_{\rm dust}$. The results are shown in the right panel of Figure 2. Again, the same increase in $M_{\rm dust}$ with redshift for a given $L_{\rm IR}^{\rm SF}$ is seen. The only

difference is that now, at a given $L_{\mathrm{IR}}^{\mathrm{SF}}$, the low z galaxies have $\log L_{\mathrm{IR}}^{\mathrm{SF}}/M_{\mathrm{dust}}$ that is 0.8 dex below the high z galaxies, slightly larger than the 0.7 dex offset in Figure 5. We conclude that the observed evolution of $L_{\mathrm{IR}}^{\mathrm{SF}}/M_{\mathrm{dust}}$ with redshift is real and not a product of how we are deriving dust masses.

<div align="center">REFERENCES</div>

Armus, L., Mazzarella, J. M., Evans, A. S., et al. 2009, PASP, 121, 559

Barcos-Muñoz, L., Leroy, A. K., Evans, A. S. et al. 2015, ApJ, 799, 10

Battisit, A., Calzetti, D., Johnson, B. D. , Elbaz, D. 2015, ApJ, 800, 143

Behroozi, P. S., Wechsler, R. H., Conroy, C. 2013, ApJ, 770, 57

Bergin, E. A., et al. 2000, ApJ, 539, L147

Blain, A. W., Barnard, V. E., Chapman, S. C. 2003, MNRAS, 338, 733

Bluck, A. F. L., Conselice, C. J., Buitrage, F., et al. 2012, ApJ, 747, 34

Bothwell, M. S., Maiolino, R., Cicone, C., Peng, Y., Wagg, J. 2016, A&A, accepted

Bouché, N., Dekel, A., Genzel, R., et al. 2010, ApJ, 718, 1001

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Casey, C. M., Narayanan, D., Cooray, A. 2014, PhR, 541, 45

Casey, C. M., Berta, S., Béthermin, M., et al. 2012, ApJ, 761, 140

Chabrier, G. 2003, PASP, 115, 763

Chakrabarti, S. & McKee, C. 2008, ApJ, 683, 693

Chakrabarti, S. & McKee, C. 2005, ApJ, 631, 792

Chanial, P., Flores, H., Guiderdoni, B., et al. 2007, A&A, 462, 81

Chapin, E. L., Hughes, D. H., Aretxaga, I. 2009, MNRAS, 393, 653

Chapman, S. C., Helou, G., Lewis, G. F., Dale, D. A. 2003, ApJ, 588, 186

Chapman, S. C., Smail, I., Ivison, R. J., et al. 2002, ApJ, 573, 66

Chary, R. & Elbaz, D. 2001, ApJ, 556, 562

Chu, J. K., Sanders, D. B., Larson, K. L., et al. 2017, ApJ, submitted

Clements, D. L., Dunne, L., Eales, S. 2010, MNRAS, 403, 274

Conselice, C. J., Yang, C., Bluck, A. F. L. 2009, MNRAS, 394, 1956

Conselice, C. J., Rajgor, S., Myers, R. 2008, MNRAS, 386, 909

da Cunha, E., Charlot, S., Elbaz, D. 2008, MNRAS, 388, 1595

Dalcanton, J. J., Fouesneau, M., Hogg, D. W., et al. 2015, ApJ, 814, 3

Dale, D. A., Helou, G., Magdis, G. E., et al. 2014, ApJ, 784, 83

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

Davé, R., Finlator, K., & Oppenheimer, B. D. 2012, MNRAS, 421, 98

Díaz-Santos, T., Charmandaris, V., Armus, L., et al. 2010, ApJ, 723, 993

Díaz-Santos, T., Charmandaris, V., Armus, L., et al. 2010, ApJ, 741, 32

Draine, B. T. 2003, ARA&A, 41, 241

Draine, B. T. & Li, A. 2007, ApJ, 657, 810

Dwek, E., Staguhn, J., Arendt, R. G., et al. 2014, ApJ, 788, L30

Dwek, E. 1998, ApJ, 501, 643

Dunne, L., and Eales, S. A. 2001, MNRAS, 327, 697

Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, 119

Galametz, M., Madden, S., Galliano, F., et al. 2011, A&A, 532, A56

Galametz, M., Kennicutt, R. C., Albrecht, M., et al. 2012, MNRAS, 425, 763

Galametz, M., Albrecht., M., Kennicutt, R. et al. 2014, MNRAS, 439, 2542

Geach, J. E., Smail, I., Moran, S. M., et al. 2011, ApJ, 730, L19

Genzel, R., Tacconi, L. J., Lutz, D., et al. 2015, ApJ, 800, 20

González-Alfonso, E., Smith, H. A., Fischer, J., Cernicharo, J. 2004, ApJ, 613, 247

Gordon, K. D., Misselt, K. A., Witt, A. N., Clayton, G. C. 2001, ApJ, 551, 269

Haan, S., Surace, J. A., Armus, L., et al. 2011, AJ, 141, 100

Hayward, C. C., Kereš, D., Jonsson, P., et al. 2011, ApJ, 743, 159

Hayward, C. C., Jonsson, P., Kereš, D., et al. 2012, MNRAS, 424, 951

Hayward, C. C., Narayanan, D., Kereš, D., et al. 2013, MNRAS, 428, 2529

Helou, G., Roussel, H., Appleton, P., et al. 2005, ApJS, 154, 253

Hopkins, P. F., Croton, D., Bundy, K., et al. 2010, ApJ, 724, 915

Hopkins, P. F., Bundy, K., Murray, N., et al. 2009, MNRAS, 398, 898

Houck, J. R., Roellig, T. L., van Cleve, J., et al. 2004, ApJS, 154, 18

Howell, J. H., Armus, L., Mazzarella, J. M., et al. 2010, ApJ, 715, 572

Huang, M.-L. & Kauffmann, G. 2014, MNRAS, 443, 1329

Kartaltepe, J. S., Dickinson, M., Alexander, D. M., et al. 2012, ApJ, 757, 23

Kartaltepe, J. S., Mozena, M., Kocevski, D. D., et sl. 2015, ApJS, 221, 11

Kennicutt, R. C., Jr. 1998, ApJ, 498, 541

Kirkpatrick, A., Pope, A., Alexander, D. M., et al. 2012, ApJ, 759, 139

Kirkpatrick, A., Calzetti, D., Galametz, M., et al. 2013b, ApJ, 778, 51

Kirkpatrick, A., Calzetti, D., Kennicutt, R. C., et al. 2014a, ApJ, 789, 130

Kirkpatrick, A., Pope, A., Aretxaga, I., et al. 2014b, ApJ, 796, 135

Kirkpatrick, A., Pope, A., Sajina, A., et al. 2013, ApJ, 814, 9

Larson, K. L., Sanders, D. B., Barnes, J. E., et al. 2016, ApJ, 825, 128

Lanz, L., Hayward, C. C., Zezas, A., et al. 2014, ApJ, 785, 39

Lee, B., Giavalisco, M., Williams, C. C., et al. 2013, ApJ, 774, 47

Lee, N., Sanders, D. B., Casey, C. M., et al. 2015, ApJ, 801,80

Lee, J. C., Hwang, H. S., Lee, G.-H. 2016, ApJ, submitted, arXiv:1610.07265

Lonsdale Persson, C. J. & Helou, G. 1987, ApJ, 314, L513

Lutz, D., Yan, L., Armus, L., et al. 2005, ApJ, 632, L13

Nelson, E. J., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2016 ApJ, 828, 27

Nelson, E. J., van Dokkum, P., Franx, M., et al. 2014, Nature, 513, 394

Madau, P. & Dickinson, M. 2014 ARA&A, 52, 415

Magdis, G. E., Rigopoulou, D., Hopwood, R., et al. 2014, ApJ, 796, 63

Magdis, G. E., Rigopoulou, D., Helou, G., et al. 2013, A&A, 558, 136

Magdis, G. E., Daddi, E., Sargent, M., et al. 2012, ApJ, 758, L9

Magnelli, B., Lutz, D., Saintonge, A., et al. 2014, A&A, 561, 86

Magnelli, B., Elbaz, D., Chary, R. R., et al. 2011, A&A, 528, A35

Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., Gnerucci, A. 2010, MNRAS, 408, 2115

Martínez-Galarza, J. R., Smith, H. A., Lanz, L., et al. 2016, ApJ, 817, 76

Martínez-Sansigre, A., Karim, A., Schinnerer, E., et al. 2009, ApJ, 706, 184

McKinnon, R., Torrey, P., Vogelsberger, M. 2016a, MNRAS, 457, 3775

McKinnon, R., Torrey, P., Vogelsberger, M., Hayward, C. C., Marinacci, F. 2016b, MNRAS, submitted, arXiv:1606.02714

Misselt, K. A., Gordon, K. D., Clayton, G. C., Wolff, M. J. 2001, ApJ, 551, 277

Murphy, E. J., Chary, R.-R., Dickinson, M., et al. 2011, ApJ, 732, 126

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ, 660, 43

Nordon, R., Lutz, D., Shao, L., et al. 2010, A&A, 518, L24

Oliver, S. J.,Wang, L., Smith, A. J., et al. 2010, A&A, 518, L21

Oliver, S. J., Bock, J., Altieri, B., et al. 2012, MNRAS, 424, 1614

Pannella, M., Elbaz, D., Daddi, E., et al. 2015, ApJ, 807, 141

Peeters, E., Spoon, H. W. W., & Tielens, A. G. G. M. 2004, ApJ, 613, 986

Penner, K., Pope, A., Chapin, E. L., et al. 2011, MNRAS, 410, 2749

Petric, A. O., Armus, L., Howell, J., et al. 2011, ApJ, 730, 28

Planck Collaboration 2016, A&A, 586, 132

Pollack, J. B., Hollenbach, D., Beckwith, S., et al. 1994, ApJ, 421 615

Pope, A., Scott, D., Dickinson, M., et al. 2006, MNRAS, 370, 1185

Pope, A., Chary, R.-R., Alexander, D. M., et al. 2008, ApJ, 675, 1171

Pope, A., Wagg, J., Frayer, D., et al. 2013, ApJ, 772, 92

Popping, G., Somerville, R. S., Galametz, M. MNRAS, accepted

Popping, G., Somerville, R. S., Trager, S. C. 2014a, MNRAS, 442, 2398

Popping, G., Pérez-Beaupuits, J. P., Spaans, M., Trager, S. C., Somerville, R. 2014b, MNRAS, 444, 1301

Preibisch, T., Zinnecker, H., & Schmitt, J. H. M. M. 1993, A&A, 279, L33

Rieke, G. H., Alonso-Herrero, A., Weiner, B. J., et al. 2009, ApJ, 692, 556

Rosario, D. J., McIntosh, D. H., van der Wel, A. et al. 2015, A&A, 573, A85

Rujopakarn, W., Rieke, G. H., Weiner, B. J., et al. 2013, ApJ, 767, 73

Rujopakarn, W., Rieke, G. H., Eisenstein, D. J., Juneau, S. 2011, ApJ, 726, 93

Safarzadeh, M., Hayward, C. C., Ferguson, H. C., Somerville, R. S. 2016, ApJ, 818, 62

Saintonge, A., Kauffmann, G., Wang, J., et al. 2011, MNRAS, 415, 61

Sajina, A., Yan, L., et al. 2012, ApJ, 757, 13

Sajina, A., Spoon, H., Yan, L., et al. 2009, ApJ, 703, 270

Sajina, A., Yan, L., Lutz, D., et al. 2008, ApJ, 683, 659

Sajina, A., Scott, D., Dennefeld, M., et al. 2006, MNRAS, 369, 939

Sakamoto, K., Wang, J., Wiedner, M. C., et al. 2008, ApJ, 684, 957

Salpeter, E. E. 1955, ApJ, 121, 161

Sanders, D. B. & Mirabel, I. F. 1996, ARA&A, 34, 749

Sanders, D. B., Mazzarella, J. M., Kim, D.-C., Surace, J. A., & Soifer, B. T. 2003, AJ, 126, 1607

Santini, P., Maiolino, R., Magnelli, B., et al. 2014, A&A, 562, 30

Sargent, M. T., Daddi, E., Béthermin, M., et al. 2014, ApJ, 793, 19

Schinnerer, E., Groves, B., Sargent, M. T., et al. 2016, ApJ, 833, 112

Scoville, N., Lee, N., Vanden Bout, P., et al. 2017, ApJ, 837, 150

Scoville, N., Sheth, K., Aussel, H., et al. 2016, ApJ, 820, 83

Scoville, N., Faisst, A., Capak, P., et al. 2015, ApJ, 800, 108

Shi, Y., Helou, G., Yan, L., et al. 2011, ApJ, 733, 87

Speagle, J. S., Steinhardt, C. L., Capak, P. L., Silverman, J. D. ApJS, 214, 15

Spoon, H. W. W., Keane, J. V., Tielens, A. G. G. M, et al. 2002, A&A, 385, 1022

Stierwalt, S., Armus, L., Charmandaris, V., et al. 2014, ApJ, 790, 124

Stierwalt, S., Armus, L., Surace, J. A., et al. 2013, ApJS, 206, 1

Suzuki, T. K., Kodama, T., Tadaki, K., et al. 2015, ApJ, 806, 208

Symeonidis, M., Page, M. J., Seymour, N. et al. 2011, MNRAS, 411, 983

Symeonidis, M., Page, M. J., Seymour, N. et al. 2009, MNRAS, 397, 1728

Tacconi, L. J., Neri, R., Genzel, R., et al. 2013, ApJ, 768, 74

Tielens, A. G. G. M., et al. 1984, ApJ, 287, 697

Trujillo, I., Conselice, C. J., Bundy, K., et al. 2007, MNRAS, 382, 109

U, V., Sanders, D. B., Mazzarella, J. M., et al. 2012, ApJS, 203, 9

van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, ApJ, 677, 5

Weingartner, J. C., & Draine, B. T. 2001, ApJ, 548, 296

Weiß, A., Kovács, A., Coppin, K., et al. 2009, ApJ, 707, 1201

Whitaker, K. E., van Dokkum, P. G., Bezanson, R., et al. 2012, ApJ, 754, L29

Williams, C. C., Giavalisco, M., Cassata, P., et al. 2014, ApJ, 780, 1

Witt, A. N. & Gordon, K. D. 2000, ApJ, 528, 799

Wu, Y., Helou, G., Armus, L., et al. 2010, ApJ, 723, 895

Yan, L., Tacconi, L. J., Fiolet, N., et al., 2010, ApJ, 714, 100

Zahid, H. J., Dima, G. I., Kewley, L. J., Erb, D. K., Davé, R. 2012, ApJ, 757, 54

Zamojski, M., Yan, L., Dasyra, K., et al. 2011, ApJ, 730, 125