



Universiteit  
Leiden  
The Netherlands

## Dust temperature and mid-to-total infrared color distributions for star-forming galaxies at $0 < z < 4$

Schreiber, C.; Elbaz, D.; Pannella, M.; Ciesla, L.; Wang, T.; Franco, M.

### Citation

Schreiber, C., Elbaz, D., Pannella, M., Ciesla, L., Wang, T., & Franco, M. (2017). Dust temperature and mid-to-total infrared color distributions for star-forming galaxies at  $0 < z < 4$ . *Astronomy & Astrophysics (0004-6361)*, 609, A30. doi:10.1051/0004-6361/201731506

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/59170>

**Note:** To cite this publication please use the final published version (if applicable).

# Dust temperature and mid-to-total infrared color distributions for star-forming galaxies at $0 < z < 4$ <sup>★,★★</sup>

C. Schreiber<sup>1,2</sup>, D. Elbaz<sup>2</sup>, M. Pannella<sup>3</sup>, L. Ciesla<sup>2</sup>, T. Wang<sup>4,2</sup>, and M. Franco<sup>2</sup>

<sup>1</sup> Leiden Observatory, Leiden University, 2300 RA Leiden, The Netherlands  
e-mail: cschreib@strw.leidenuniv.nl

<sup>2</sup> Laboratoire AIM-Paris-Saclay, CEA/DSM/Irfu – CNRS – Université Paris Diderot, CEA-Saclay, pt courrier 131, 91191 Gif-sur-Yvette, France

<sup>3</sup> Faculty of Physics, Ludwig-Maximilians Universität, Scheinerstr. 1, 81679 Munich, Germany

<sup>4</sup> School of Astronomy and Astrophysics, Nanjing University, 210093 Nanjing, PR China

Received 4 July 2017 / Accepted 24 October 2017

## ABSTRACT

We present a new, publicly available library of dust spectral energy distributions (SEDs). These SEDs are characterized by only three parameters: the dust mass ( $M_{\text{dust}}$ ), the dust temperature ( $T_{\text{dust}}$ ), and the mid-to-total infrared color ( $\text{IR8} \equiv L_{\text{IR}}/L_8$ ). The latter measures the relative contribution of polycyclic aromatic hydrocarbon (PAH) molecules to the total infrared luminosity. We used this library to model star-forming galaxies at  $0.5 < z < 4$  in the deep CANDELS fields, using both individual detections and stacks of *Herschel* and ALMA imaging, and extending this sample to  $z = 0$  using the *Herschel* Reference Survey. At first order, the dust SED of a galaxy was observed to be independent of stellar mass, but evolving with redshift. We found trends of increasing  $T_{\text{dust}}$  and IR8 with redshift and distance from the SFR– $M_*$  main sequence, and quantified for the first time their intrinsic scatter. Half of the observed variations of these parameters was captured by the above empirical relations, and after subtracting the measurement errors we found residual scatters of  $\Delta T_{\text{dust}}/T_{\text{dust}} = 12\%$  and  $\Delta \log \text{IR8} = 0.18$  dex. We observed second order variations with stellar mass: massive galaxies ( $M_* > 10^{11} M_{\odot}$ ) at  $z \leq 1$  have slightly lower temperatures indicative of a reduced star formation efficiency, while low mass galaxies ( $M_* < 10^{10} M_{\odot}$ ) at  $z \geq 1$  showed reduced PAH emission, possibly linked to their lower metallicities. Building on these results, we constructed high-fidelity mock galaxy catalogs to predict the accuracy of infrared luminosities and dust masses determined using a single broadband measurement. Using a single *James Webb* Space Telescope (JWST) MIRI band, we found that  $L_{\text{IR}}$  is typically uncertain by 0.15 dex, with a maximum of 0.25 dex when probing the rest-frame  $8 \mu\text{m}$ , and this is not significantly impacted by typical redshift uncertainties. On the other hand, we found that ALMA bands 8 to 7 and 6 to 3 measured the dust mass at better than 0.2 and 0.15 dex, respectively, and independently of redshift, while bands 9 to 6 only measured  $L_{\text{IR}}$  at better than 0.2 dex at  $z > 1$ , 3.2, 3.8, and 5.7, respectively. Starburst galaxies had their  $L_{\text{IR}}$  significantly underestimated when measured by a single JWST or ALMA band, while their dust mass from a single ALMA band were moderately overestimated. This dust library and the results of this paper can be used immediately to improve the design of observing proposals, and interpret more accurately the large amount of archival data from *Spitzer*, *Herschel* and ALMA.

**Key words.** galaxies: evolution – galaxies: ISM – galaxies: statistics – infrared: galaxies – submillimeter: galaxies

## 1. Introduction

Properly accounting for the amount of stellar light absorbed by dust has proven to be a key ingredient to study star formation in galaxies. The most obvious breakthrough linked to deep infrared (IR) surveys was probably the exciting new outlook they provided on the cosmic history of star formation (e.g., Smail et al. 1997; Hughes et al. 1998; Barger et al. 1998; Blain et al. 1999; Elbaz et al. 1999, 2002, 2007, 2011; Flores et al. 1999; Lagache et al. 1999; Gispert et al. 2000; Franceschini et al. 2001; Papovich et al. 2004; Le Floc’h et al. 2005; Daddi et al. 2009; Magnelli et al. 2009; Gruppioni et al. 2010; Rodighiero et al. 2011; Magdis et al. 2012; Madau & Dickinson 2014, and references therein). In the meantime, the emission of dust in distant galaxies has also been used to study the dust itself, which turned out to be a valuable tool to learn

more about non-stellar baryonic matter, present in the interstellar medium (ISM) either in the form of dust grains or atomic and molecular gas (e.g., Chapman et al. 2003; Hwang et al. 2010; Elbaz et al. 2011; Magdis et al. 2012; Berta et al. 2013; Scoville et al. 2014; Santini et al. 2014; Béthermin et al. 2015; Genzel et al. 2015; Tacconi et al. 2017).

In the Local Universe, the large amount of infrared data acquired in the Milky Way and nearby galaxies has given birth to detailed models aiming to provide a description of the dust content from first principles (e.g., Zubko et al. 2004; Draine & Li 2007; Galliano et al. 2011; Jones et al. 2013, to only name a few of the most recent ones). These models typically contain three main components (see, e.g., Desert et al. 1990): big grains (BGs,  $> 0.01 \mu\text{m}$ ), very small grains (VSGs,  $< 0.01 \mu\text{m}$ ), and complex molecules (polycyclic aromatic hydrocarbon, or PAH). The most prominent one is the emission of big grains, which are at thermal equilibrium with the ambient interstellar radiation. These grains radiate like gray bodies with a typical temperature of  $T_{\text{dust}} \sim 20\text{--}40$  K, and therefore emit the bulk of their energy in the far-infrared (FIR) around the rest-frame  $100 \mu\text{m}$ . Smaller grains

\* The dust library described in this paper is available publicly at <http://cschreib.github.io/s17-irlib/>

\*\* Tables A.1–A.3 are also available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](http://cdsarc.u-strasbg.fr) (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/609/A30>

have a too small cross-section to be at equilibrium with the ambient radiation, and are instead only transiently heated to temperatures  $\sim 1000$  K. This produces continuum emission in the mid-infrared (MIR). Lastly, PAHs are large carbonated molecules which cool down through numerous rotational and vibrational modes, and thus produce a group of bright and broad emission lines between  $\lambda = 3.3$  and  $12.3 \mu\text{m}$  (Leger & Puget 1984; Allamandola et al. 1985).

To reproduce a set of observations in the IR, one can vary the total mass of dust encompassing all ISM components ( $M_{\text{dust}}$ ), the distribution of energy they receive from their surrounding medium ( $U$ ), coming mostly from stellar light, and the properties of each species of grains and molecules, including their size distribution, their chemical state and composition (neutral vs. ionized PAHs, silicate vs. carbonated grains). Given these parameters, models can output the expected infrared spectrum and interpret the observed data. However, most of these parameters are degenerate or unconstrained and the number of degrees of freedom is too large, hence assumptions have to be made when applying such models to photometric data. Typical approaches (e.g., Draine & Li 2007, hereafter DL07 or da Cunha et al. 2008) assume fixed grain distributions (motivated by observations from clouds of the Milky Way), and consider simplified geometries of dusty regions (e.g., birth clouds, diffuse ISM, hot torus around a super-massive black hole). This still allows much flexibility in the output spectrum, and can describe observations accurately (e.g., da Cunha et al. 2015; Gobat et al. 2017).

But even then, properly constraining the fit parameters (in particular the dust temperature) requires exquisite IR spectral energy distributions (SEDs) with good wavelength sampling, at a level of quality that can currently only be reached either in the Local Universe or at high-redshifts for the most extreme starbursts (e.g., Hwang et al. 2010; Magdis et al. 2010; Riechers et al. 2013), strongly lensed galaxies (e.g., Sklias et al. 2014), or on stacked samples (e.g., Magnelli et al. 2014; Béthermin et al. 2015). For the typical higher redshift galaxy, the available IR SED is limited to one or two photometric points (e.g., Elbaz et al. 2011), and even simpler approaches are often preferred. A number of empirical libraries have been constructed to this end, each composed of a reduced number of template SEDs. These templates are typically associated to different values of a single parameter, for example the 8 to  $1000 \mu\text{m}$  luminosity ( $L_{\text{IR}}$ ; Chary & Elbaz 2001, hereafter CE01), a FIR color (Dale & Helou 2002, hereafter DH02), or the average intensity of the interstellar radiation field,  $\langle U \rangle$  (Magdis et al. 2012; Béthermin et al. 2015). In all cases the number of free parameters is reduced to two: the normalization of the template (which can be linked either to the mass or luminosity of the dust) and its “shape” (essentially its average dust temperature, which defines the wavelength at which the template peaks). Despite their extreme simplicity, these models are sufficient to reproduce the observed IR features of the vast majority of distant galaxies, illustrating the fact that the dust SED of a galaxy taken as a whole is close to universal (see, e.g., Elbaz et al. 2010, 2011).

This universality echoes another key observation of the last decade: the main sequence of star-forming galaxies (Noeske et al. 2007; Elbaz et al. 2007). This tight correlation between the star formation rate (SFR) and the stellar mass ( $M_*$ ) has been observed across a broad range of redshifts up to  $z \sim 6$  (e.g., Daddi et al. 2007; Pannella et al. 2009, 2015; Rodighiero et al. 2011; Whitaker et al. 2012; Bouwens et al. 2012; Whitaker et al. 2014; Salmon et al. 2015; Schreiber et al. 2015, 2017b). Since its discovery, the main sequence has been used to put upper limits on the variability of the star formation histories, showing that

galaxies form their stars mostly through a unique and secular way, as opposed to random bursts (see, e.g., the discussion in Noeske et al. 2007).

Through observations of their molecular gas content, galaxies belonging to the main sequence have also been shown to form their stars with a roughly constant efficiency ( $SFE \equiv SFR/M_{\text{gas}}$ ) (e.g., Daddi et al. 2010; Tacconi et al. 2013; Genzel et al. 2015). The same conclusion can be drawn from the dust emission of these galaxies and the measurement of their average dust temperature (Magdis et al. 2012; Béthermin et al. 2015, but see however Schreiber et al. 2016). Indeed,  $T_{\text{dust}}$  is a proxy for  $L_{\text{IR}}/M_{\text{dust}}$ , which itself can be linked directly to  $SFE/Z$  (Magdis et al. 2012), where  $Z$  is the gas-phase metallicity. The universality of the dust SED therefore also suggests that star formation in galaxies is the product of a universal mechanism, which still remains to be fully understood (see, e.g., Dekel et al. 2013; or Tacchella et al. 2015).

Departures from this “universal” SED do exist however. Galaxy-to-galaxy variations of  $T_{\text{dust}}$  have been observed, with a first correlation identified with  $L_{\text{IR}}$  (e.g., Soifer et al. 1987, 1989; Dunne et al. 2000; Chapman et al. 2003; Chapin et al. 2009; Symeonidis et al. 2009; Amblard et al. 2010; Hwang et al. 2010). It was later argued that this correlation is not fundamental, but in fact consequential of two effects: on the one hand a global increase of the temperature with redshift (e.g., Magdis et al. 2012; Magnelli et al. 2014; Béthermin et al. 2015), and on the other hand an additional increase of temperature for galaxies that are offset from the main sequence (Elbaz et al. 2011; Magnelli et al. 2014; Béthermin et al. 2015), suggesting these galaxies form stars more efficiently than the average. Quantifying changes of the dust temperature can thus provide crucial information about the star formation efficiency in galaxies, and it is therefore an important ingredient in any library.

In addition, significant galaxy-to-galaxy variations have been observed in the MIR around the rest-frame 3 to  $12 \mu\text{m}$ . As written above, the dust emission in this wavelength domain is mostly produced by small grains and PAHs. The PAH emission lines are so bright that they typically contribute about 80% of the observed broadband MIR fluxes (e.g., Helou et al. 2000; Huang et al. 2007), however, their strength is strongly reduced in starbursts and active galactic nuclei (AGNs; e.g., Armus et al. 2007), in which hot dust takes over. Therefore, the observed diversity in the MIR can be expected to come mostly from a diversity of PAH properties, at least for galaxies without strong AGNs (e.g., Fritz et al. 2006). The interplay between the overall strength of PAHs and physical conditions inside the host galaxy is not yet fully understood. Two main trends are known at present: on the one hand an anti-correlation with metallicity (e.g., Madden et al. 2006; Wu et al. 2006; O’Halloran et al. 2006; Smith et al. 2007; Draine et al. 2007; Galliano et al. 2008; Ciesla et al. 2014; Rémy-Ruyer et al. 2015), and on the other hand a correlation with  $L_{\text{IR}}$  (e.g., Pope et al. 2008; Elbaz et al. 2011; Nordon et al. 2012). Although this latter correlation suffers from a significant scatter, it implies that PAH features or the  $8 \mu\text{m}$  luminosity can be used as a rough tracer of star formation rate (Pope et al. 2008; Shipley et al. 2016).

An interesting property of PAHs is that they are set aglow mostly in photo-dissociation regions, at the interface between the ionized and molecular interstellar medium (e.g., Tielens et al. 1993), whereas the FIR dust continuum is emitted from the whole volume of the dust clouds. Therefore, by relating the dust continuum to the PAH emission one can probe the geometry of star-forming regions, and in particular the filling factor of H II regions. Using this approach and combining *Spitzer* and

*Herschel* data, [Elbaz et al. \(2011\)](#) have used the  $IR8 = L_{IR}/L_8$  ratio as a tracer of compactness in distant galaxies: at fixed  $L_{IR}$ , a lower  $L_8$  indicates a higher filling factor of H II regions, hence a higher compactness. Main sequence galaxies have a constant  $IR8 \sim 4$ , while, as for the dust temperature,  $IR8$  increases as a function of the distance to the main sequence (see also, [Nordon et al. 2012](#); [Rujopakarn et al. 2013](#); [Murata et al. 2014](#)). These trends confirm that galaxies above the main sequence form their stars in a different way, with a higher efficiency and in more compact volumes.

While the study of the physical origin of the MIR emission is obviously of interest on its own, it is also important to the extragalactic community for practical observational reasons. Since the PAH emission is strong and found at low infrared wavelengths, the wavelength domain around the rest-frame  $8\mu\text{m}$  is easier to observe than than the FIR continuum. It is particularly the case for galaxies at  $z \sim 2$ , where the rest-frame  $8\mu\text{m}$  shifts into the very deep *Spitzer* MIPS  $24\mu\text{m}$  band, and allows the detection of galaxies significantly fainter than the detection limit of other infrared observatories like *Herschel*. However, these galaxies have by construction a very poorly constrained infrared SED, and extrapolating the total  $L_{IR}$  from the  $8\mu\text{m}$  alone is challenging (see [Daddi et al. 2007](#); [Elbaz et al. 2011](#); [Magdis et al. 2011](#); [Rujopakarn et al. 2013](#); [Shivaei et al. 2017](#)). Doing so requires an accurate understanding of the  $IR8$  ratio.

Another important practical interest for the rest-frame  $8\mu\text{m}$  is that it will be easily accessed by the *James Webb* Space Telescope (JWST) in the near future, for both local and distant galaxies. Once this satellite is launched, there will be a need for a properly calibrated library to exploit these data together with ancillary *Herschel* and *Spitzer* observations, and in particular to cope with their absence for the faintest objects.

Our goal in this paper is the following. We introduce in Sect. 3 a new SED library in which both  $T_{\text{dust}}$  and  $IR8$  are free parameters. This library provides an increased level of detail compared to standard libraries (e.g., [CE01](#), [DH02](#)), but still keeps the number of adjustable parameters low. In Sect. 4.1 we determine the redshift evolution of both  $T_{\text{dust}}$  and  $IR8$  using the MIR-to-FIR stacks introduced in [Schreiber et al. \(2015\)](#), to which we add stacks of  $16\mu\text{m}$  and ALMA  $870\mu\text{m}$  to better constrain the PAH features and the dust temperature at high redshifts. We then apply this model to individual *Herschel* detections in Sect. 4.2 to constrain the scatter on the model parameters, and also to quantify their enhancements for those galaxies that are offset from the main sequence. Using these results, we derive in Sect. 5 a set of recipes for optimal SED fitting in the IR, in particular when a single photometric band is available. Finally, we quantify the accuracy of such measurements using mock galaxy catalogs in Sect. 6, and provide in Sect. 6.2 conversion factors to determine dust masses and infrared luminosities from ALMA fluxes and JWST MIRI luminosities. These are valid for  $0 < z < 4$ , and are extrapolated to  $z = 8$  for ALMA.

In the following, we assume a  $\Lambda$ CDM cosmology with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.3$ ,  $\Omega_\Lambda = 0.7$  and a [Salpeter \(1955\)](#) initial mass function (IMF), to derive both star formation rates and stellar masses. All magnitudes are quoted in the AB system, such that  $M_{\text{AB}} = 23.9 - 2.5 \log_{10}(S_\nu [\mu\text{Jy}])$ .

## 2. Sample and observations

We based this analysis on the sample and data described in [Schreiber et al. \(2015, hereafter S15\)](#), which covers redshifts from  $z = 0.3$  to  $z = 4$ . We complemented this sample with  $z = 0$  galaxies from the *Herschel* Reference Survey (HRS;

[Boselli et al. 2010](#)), and  $z = 2$  to 4 galaxies in the Extended Chandra Deep Field South (ECDFS) observed by ALMA as part of the ALESS program ([Hodge et al. 2013](#)). In this section, we make a brief summary of these observations.

### 2.1. CANDELS

The catalogs we used in this work are based on the CANDELS ([Grogin et al. 2011](#); [Koekemoer et al. 2011](#)) *Hubble* Space Telescope (HST) WFC3 *H* band images in the fields covered by deep *Herschel* PACS and SPIRE observations, namely GOODS–South ([Guo et al. 2013](#)), UDS ([Galametz et al. 2013](#)) and COSMOS ([Nayyeri et al. 2017](#)). For the GOODS–North field the CANDELS catalog was not yet finalized, and we used instead the  $K_s$ -selected catalog of [Pannella et al. \(2015\)](#). Each of these fields is about  $150 \text{ arcsec}^2$  and they are evenly distributed on the sky to mitigate cosmic variance. We also used a catalog of the COSMOS  $2 \square^\circ$  field ([Muzzin et al. 2013](#)), which has overall shallower data but covers a much larger area; this field provides important statistics for the rarest and brightest objects.

The ancillary photometry varies from one field to another, being a combination of both space- and ground-based imaging from various facilities. The UV to near-IR wavelength coverage typically goes from the *U* band up the *Spitzer* IRAC  $8\mu\text{m}$ , including at least the HST bands *F606W*, *F814W*, *F125W*, and *F160W* in CANDELS, and a deep *K* (or  $K_s$ ) band. All these images are among the deepest available views of the sky. These catalogs therefore cover most of the important galaxy spectral features across a wide range of redshifts, even for intrinsically faint objects.

We augmented these catalogs with mid-IR photometry from *Spitzer* MIPS and far-IR photometry from *Herschel* PACS and SPIRE taken as part of the GOODS–*Herschel* ([Elbaz et al. 2011](#)), CANDELS–*Herschel* programs (PI: M. Dickinson), PEP ([Lutz et al. 2011](#)) and HerMES ([Oliver et al. 2010](#)).

Photometric redshifts and stellar masses were computed following [Pannella et al. \(2015\)](#) using EAZY ([Brammer et al. 2008](#)); for COSMOS  $2 \square^\circ$  we used the redshifts from [Muzzin et al. \(2013\)](#) which were computed the same way. For all catalogs, stellar masses were then computed using FAST ([Kriek et al. 2009](#)) by fixing the redshift to the best-fit photo- $z$  and fitting the observed photometry up to the IRAC  $4.5\mu\text{m}$  band<sup>1</sup> using the [Bruzual & Charlot \(2003\)](#) stellar population synthesis model, assuming a [Salpeter \(1955\)](#) IMF, a [Calzetti et al. \(2000\)](#) extinction law and a delayed exponentially-declining star formation history.

Galaxies with an uncertain photometric redshift (redshift odds less than 0.8) or bad SED fitting (reduced  $\chi^2$  larger than 10) were excluded from our sample. The resulting sample is the one we used for stacking the *Herschel* images in S15. In this previous work, we estimated the evolution of the stellar mass completeness (at the 90% level) of these catalogs at all redshifts, and found that all the stacked samples with significant signal were complete in mass. For example, at  $z = 1$  the completeness is as low as  $5 \times 10^8 M_\odot$ .

We estimated SFRs by summing the emerging UV light and the dust obscured component observed in the mid- to far-IR,

<sup>1</sup> The last two IRAC channels, at  $5.8$  and  $8\mu\text{m}$ , were not used to derive the stellar mass for two reasons. First, at low redshift these bands are contaminated by the dust emission and AGNs, which cannot be taken into account by FAST. Second, while the corresponding images are reasonably deep in the two GOODS fields, the observations in UDS and COSMOS are substantially shallower. Excluding these bands from the fit therefore prevents introducing field-to-field systematics.

following [Daddi et al. \(2007\)](#) and [Kennicutt \(1998\)](#) to convert the observed luminosities into star formation rates:

$$SFR = 2.17 \times 10^{-10} L_{UV} [L_{\odot}] + 1.72 \times 10^{-10} L_{IR} [L_{\odot}]. \quad (1)$$

UV luminosities ( $1500 \text{ \AA}$ ) were computed from the best-fit photo- $z$  template from EAZY, while IR luminosities were computed from the best-fit dust SED obtained with our new library (see Sect. 5). We applied this method to both the stacked samples and to individual galaxies with mid- or far-IR detections. For individual galaxies, we did not attempt to measure the SFRs of the rest of the sample without IR data since estimates based on the UV light alone are less reliable ([Goldader et al. 2002](#); [Buat et al. 2005](#); [Elbaz et al. 2007](#); [Rodighiero et al. 2011, 2014](#)). When working with individual galaxies (as opposed to stacking), we therefore only considered the sub-sample of IR-detected galaxies, which implicitly corresponds to a SFR threshold at each redshift (see [Elbaz et al. 2011](#)). The resulting selection biases are discussed in Sect. 4.2.2.

Lastly, the rest-frame  $U$ ,  $V$  and  $J$  magnitudes were computed for each galaxy using EAZY, by integrating the best-fit galaxy template from the photo- $z$  estimation. These colors were used, following [Williams et al. \(2009\)](#), to separate galaxies that are “quiescent” from those that are “star-forming”. We used the same selection criteria as those described in [S15](#), that is, a galaxy was deemed star-forming if its colors satisfy

$$UVJ_{SF} = \begin{cases} U - V < 1.3, \text{ or} \\ V - J > 1.6, \text{ or} \\ U - V < 0.88 \times (V - J) + 0.49; \end{cases} \quad (2)$$

otherwise the galaxy was considered as quiescent as was thus excluded from the present study. As shown in [S15](#), only a very small fraction of the IR-detected galaxies are classified as  $UVJ$  quiescent, therefore this selection is only important for stacking.

## 2.2. ALESS

To improve the statistics on the dust temperature of individual galaxies at  $z > 1$ , we complemented this sample with the 99 galaxies observed by ALMA in the ALESS program ([Hodge et al. 2013](#)). ALESS is a targeted program aiming to deblend the  $870 \mu\text{m}$  emission of sources detected in the single-dish LABOCA image of the ECDFS, which covers about  $0.3 \text{ deg}^2$  centered on the CANDELS GOODS–South field. The high resolution of the ALMA imaging allows a precise localization of the sub-millimeter source and avoids flux boosting from blended neighbors.

We used the photometric redshifts and stellar masses determined in [da Cunha et al. \(2015\)](#) using Magphys ([da Cunha et al. 2008](#)). We used the  $24 \mu\text{m}$  and PACS photometry from PEP, the SPIRE photometry as measured in [Swinbank et al. \(2014\)](#) using the ALMA detections to improve the deblending, and the ALMA photometry from [Hodge et al. \(2013\)](#). The galaxies were then treated the same way as those from CANDELS, and used only in Sect. 4.2.

## 2.3. HRS

To complement our sample toward the local Universe, we used the *Herschel* Reference Survey (HRS; [Boselli et al. 2010](#)). This is a volume-limited survey targeting a few hundred galaxies in and out of the Virgo cluster, obtaining in particular *Herschel* PACS and SPIRE photometry for a full sampling of their dust SEDs. All the galaxies in the HRS also have UV-to-NIR coverage to determine stellar masses and colors. Of this sample, we

only considered the  $UVJ$  star-forming galaxies which do not belong to the Virgo cluster, to avoid systematic effects caused by this peculiar environment, for a total of 131 galaxies with a minimum stellar mass of  $10^9 M_{\odot}$ . This same sample was studied in [Schreiber et al. \(2016\)](#); further informations can be found there.

We used the photometry from [Ciesla et al. \(2012\)](#) and modeled the HRS galaxies using CIGALE ([Noll et al. 2009](#); [Roehlly et al. 2014](#)), which fits simultaneously the stellar and dust emission. Using CIGALE proved necessary because the contribution of the stellar continuum to the  $8 \mu\text{m}$  luminosity (or more generally to the PAH domain in the mid-IR) can be non-negligible at  $z = 0$ , owing to the overall lower star-formation activity. We performed the fits using same SED libraries as for the CANDELS sample, that is, the dust SEDs introduced in this paper and the [Bruzual & Charlot 2003](#) templates with a delayed SFH. Our dust SEDs are made available to all CIGALE users in the official package.

## 3. A new far infrared template library

### 3.1. Description of the model

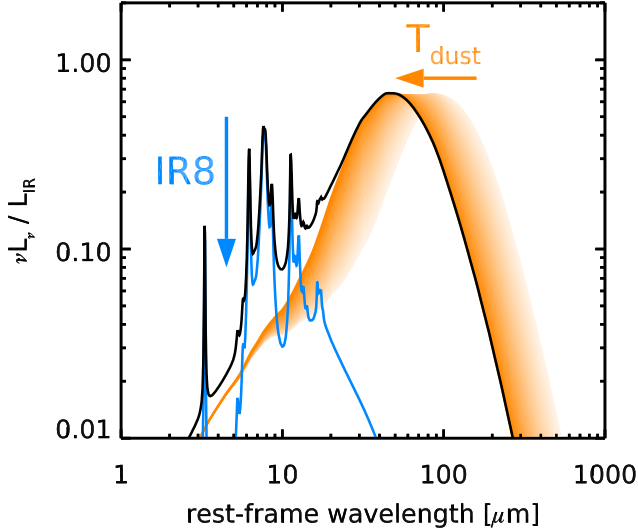
Since it was published, the CE01 library has been used routinely to derive infrared luminosities, and therefore star formation rates, for large samples of galaxies at various redshifts. In [S15](#), we found that, in spite of the relatively small number of different SEDs it contains, it is able to fit well our stacked FIR photometry (rest-frame 30-to- $500 \mu\text{m}$ ) at all  $z = 0.5$  to  $z = 4$ . However, once properly adjusted to the observed FIR data, the behavior of these SEDs, calibrated on local starbursts, may not adequately describe the observed MIR photometry. Indeed, the CE01 library enforces a unique relation between  $T_{\text{dust}}$ , IR8 and  $L_{\text{IR}}$  which was calibrated from Local Universe galaxies. The relevance of this assumption for distant galaxies was unknown at the time, and in fact it was shown to break when applied to  $24 \mu\text{m}$ -detected  $z = 2$  galaxies ([Papovich et al. 2007](#); [Daddi et al. 2007](#)). It was later understood that this was caused by an evolving  $L_{\text{IR}}$ –IR8 relation, and another more universal calibration was proposed where the IR8 varies as a function of the distance to the main sequence ([Elbaz et al. 2011](#); [Nordon et al. 2012](#); [Rujopakarn et al. 2013](#)) rather than with the absolute luminosity. Similar conclusions have been drawn for the dust temperature ([Elbaz et al. 2011](#); [Magnelli et al. 2014](#)).

To build a more up to date and versatile library, we started by making each of these observables independent of one another, that, we created a set of templates that allow us to vary  $T_{\text{dust}}$ , IR8 and  $L_{\text{IR}}$  simultaneously. To do so, we made the arbitrary choice of separating the IR emission into two components: on the one hand, the dust continuum of varying  $T_{\text{dust}}$  created by big and small grains, and on the other hand the MIR features emitted by PAH molecules (see Fig. 1):

$$S_{\nu} = M_{\text{dust}}^{\text{cont}} \bar{S}_{\nu}^{\text{cont}} + M_{\text{dust}}^{\text{PAH}} \bar{S}_{\nu}^{\text{PAH}}. \quad (3)$$

$M_{\text{dust}}^{\text{PAH}}$  and  $M_{\text{dust}}^{\text{cont}}$  are defined as the mass of dust grain found in the form of PAH molecules and silicate+carbonated grains, respectively, while  $\bar{S}_{\nu}^{\text{PAH}}$  and  $\bar{S}_{\nu}^{\text{cont}}$  are the spectra of each grain population normalized to unit dust mass (these are described in the next sections).

This decomposition implies that PAHs are almost exclusively responsible for the observed diversity in IR8. Indeed, decomposing the MIR emission into multiple components (PAHs, very small grains, and AGN torus) is a degenerate problem when only broadband photometry is used, and a choice needs to be



**Fig. 1.** Cartoon picture illustrating the two effective parameters impacting the shape of our FIR SEDs. The total SED, normalized to unit  $L_{\text{IR}}$ , is shown with a black solid line, while the dust continuum and PAH components are shown with solid orange and blue lines, respectively. We also show how the shape of the SED varies with dust temperature  $T_{\text{dust}}$  by displaying several templates of different  $T_{\text{dust}}$  in orange lines of varying intensities. The orange and blue arrows illustrate how the SED is modified by increasing  $T_{\text{dust}}$  and IR8, respectively.

made. Since our objective is mostly to study galaxies and not AGNs, we neglect the presence of AGNs torus emission, and assume a fixed fraction of small vs. big grains (see below). It is certainly possible to use our library in combination with an AGN template if sufficient MIR data is available, but this goes out of the scope of this work.

To create our templates, we used the amorphous carbon dust model of Galliano et al. (2011, hereafter G11). This model can output the mid- to far-IR spectrum emitted by a dust cloud of mass  $1 M_{\odot}$  under the influence of a uniform radiation field of integrated intensity  $U$  (taken here in units of the Mathis et al. 1983 interstellar radiation field in the solar neighborhood,  $U_{\odot}$ ). In the following, we call each spectrum generated by this model an “elementary” G11 template. The next two sections describe how our library is built from these templates, as well as the underlying assumptions.

### 3.1.1. Dust continuum

In the G11 model, the dust continuum elementary templates are produced by a combination of silicate and amorphous carbon grains of varying sizes, split into “big” (thermalized) and “small” (transiently heated) grains. The size distributions of these grains were taken from Zubko et al. (2004) and were assumed to be universal. Here we assumed the Milky Way (MW) mass-fraction of carbonated vs. silicate grains as derived by Zubko et al. for big grains, but fixed the mass-fraction of small silicate grains to zero (as in Compiègne et al. 2011) instead of 11%. While this is compatible with observational constraints (Li & Draine 2001), our motivation for this choice was purely empirical: reducing the emission of small grains in the mid-IR increased the range of IR8 values that our model can reach.

The only remaining free parameter is the radiation intensity  $U$ . This parameter controls the energy of the radiation field to which each grain is exposed. For grains big enough to be

thermalized, an increase of this energy implies an enhanced grain temperature (this is quantified later in Sect. 3.2), and affects the shape of their FIR spectrum. For smaller grains which are not thermalized, only the overall normalization of the spectrum is modified.

### 3.1.2. PAH emission

To produce the associated PAH emission, we assumed that these molecules are subject to the same  $U$  as the other dust grains, although in this case this choice has very little consequence since, as for small grains, the PAH molecules are not thermalized and therefore the effect of increasing  $U$  is essentially only to increase the normalization of the PAH SED at fixed dust mass. This will affect the absolute values of the PAH masses, in which we have limited interest here. The only free parameter in the G11 model regarding the PAH composition is the fraction of neutral vs. ionized molecules, which we chose here to be the MW value of 50% (Zubko et al. 2004). This parameter mostly influences the relative strengths of the 8 vs. 12  $\mu\text{m}$  PAH features, and we found that the MW value provided indeed a good match to the stacked *Spitzer* IRS spectra of  $z = 2$  ULIRGs (see Sect. 3.3), as well as to our stacked  $S_{24}/S_{16}$  broadband flux ratio at  $z = 1$  (see Sect. 4.1).

### 3.1.3. Radiation field distribution

The elementary templates introduced above are not well suited to describe an entire galaxy, since the hypothesis of a uniform  $U$  usually does not hold in such kind of extended systems. Instead, a “composite” template must be built by adding together the emission of different dusty regions, heated by different radiation intensities. As in Dale et al. (2001), G11 assumed that the distribution of  $U$  in a composite system follows a power law in  $dM_{\text{dust}}/dU \sim U^{-\alpha}$ , where  $dM_{\text{dust}}$  is the mass of dust associated to the elementary region illuminated with an intensity  $[U, U + dU]$ . This distribution is then integrated from  $U = U_{\text{min}}$  to  $U = U_{\text{max}}$  to form the final template. Contrary to DL07, they did not assume the presence of an additional component linked to photo-dissociation regions since it was shown not to provide significant improvement to the fits (as demonstrated in the appendix of G11).

The parameter  $\alpha$  is the slope of the mass distribution of  $U$  within a galaxy. This parameter affects the composite spectrum in a non-trivial way: large and small values of  $\alpha$  will accumulate most of the dust mass close to  $U_{\text{min}}$  and  $U_{\text{max}}$ , respectively ( $\alpha = 1$  and 2 give a uniform weighting in mass and luminosity, respectively). To avoid this complexity, we fixed  $\alpha = 2.6$  for all our templates. This value was chosen to reproduce the width of the CE01 SEDs between 15 and 500  $\mu\text{m}$ , since these templates are known to provide a good description of the FIR emission of distant galaxies (Elbaz et al. 2010). We also checked a posteriori that this value provided a satisfactory fit to our stacked FIR photometry (see Sect. 4.1).

With our adopted  $U$  distribution, the final SED is relatively insensitive to the precise choice of  $U_{\text{max}}$ , provided  $U_{\text{max}} \gg U_{\text{min}}$  (Draine et al. 2007). Hence the only remaining parameter that allows us to tune the SED shape is  $U_{\text{min}}$  or, equivalently, the mass-weighted intensity  $\langle U \rangle$ . In particular, we note in Sect. 3.2 that  $\langle U \rangle$  is related to the average dust temperature through a simple power law, which is one of the parameter our library aims to describe. Therefore, we generated a logarithmic grid of  $U_{\text{min}}$  ranging from 0.1 to 5000  $U_{\odot}$  with 250 samples, and took  $U_{\text{max}} = 10^6 U_{\odot}$  (Draine et al. 2007). This allows our library to

describe dust temperatures ranging from about 15 to 100 K with a roughly constant step of 0.3 K. The resulting SEDs can be obtained at <http://cschreib.github.io/s17-irlib/>.

### 3.1.4. Amorphous carbon or graphite?

Compared to more standard dust models (e.g., DL07), the one we used here assumes that carbonated grains are found exclusively in the form of amorphous carbon grains, rather than graphites. While this has no visible impact on the shape of the generated spectra, it systematically lowers the value of the measured dust masses by a factor of about 2.0 compared to graphite dust, owing to the different emissivities of these grain species. This was in fact the motivation for using amorphous carbon in G11: lowering the measured dust masses eases the tension between the observed dust-to-gas ratio and stellar abundances in the Large Magellanic Cloud (LMC). This conclusion was not only reached in the LMC, which has a particularly low metallicity, but also in more normal galaxies including the MW (e.g., Compiègne et al. 2011; Jones et al. 2013; Fanciullo et al. 2015; Planck Collaboration XXIX 2016). As stressed in G11, purely amorphous carbon is just one possibility to achieve higher emissivities. We therefore do not give much credit to carbon dust being truly amorphous, but since this type of grains does describe the content of the ISM in a more consistent way, we chose to favor it instead of graphite. The impact of this choice on gas-to-dust ratios is discussed in Sect. 6.3.

## 3.2. Basic usage of the library and useful relations

In this section we provide a set of simple relations to relate the internal parameters of the library, namely  $\langle U \rangle$ ,  $M_{\text{dust}}^{\text{cont}}$ , and  $M_{\text{dust}}^{\text{PAH}}$ , to more commonly used observables such as the total infrared luminosity, the dust mass, and the IR8. We then provide instructions for the most basic usage of this template library.

### 3.2.1. Dust temperature

The dust temperature of our model SEDs was computed by applying Wien's law to each elementary G11 template for the dust continuum:

$$T_{\text{dust}}[\text{K}] = 2.897 \times 10^3 / (\lambda_{\text{max}}[\mu\text{m}]), \quad (4)$$

determining  $\lambda_{\text{max}}$  as the wavelength corresponding to the peak of  $\lambda^\beta L_\nu$  (the term  $\lambda^\beta$  takes into account the effective emissivity of the templates,  $\beta \simeq 1.5$ ). We then weighted each value by the dust mass associated to the corresponding template ( $dM_{\text{dust}}$ ), therefore producing a mass-weighted average. We found that the following relation links together  $T_{\text{dust}}$  and the radiation field intensity:

$$\frac{\langle U \rangle}{U_\odot} = \left( \frac{T_{\text{dust}}}{18.2 \text{ K}} \right)^{5.57}. \quad (5)$$

As stated earlier, our library covers  $T_{\text{dust}}$  values ranging from 15 to 100 K. We also applied Wien's law (as above) to the peak of the final dust template, to obtain a light-weighted average  $T_{\text{dust}}^{\text{light}}$ . In practice, we find the difference between the two to be simply a constant factor, with

$$T_{\text{dust}} = 0.91 \times T_{\text{dust}}^{\text{light}}. \quad (6)$$

$T_{\text{dust}}^{\text{light}}$  is less stable because the summed dust template is broader, making it harder to accurately locate the position of the peak.

Its physical meaning is also less clear, since our templates do not have a single temperature, but it has the advantage of being less model-dependent and is essentially the temperature one would measure by using a modified blackbody model with a single temperature and an emissivity of  $\beta = 1.5$ . We therefore provide tabulated values for both temperatures in the library, and in the following, unless otherwise stated, we will refer to the dust temperature as the mass-weighted value. Likewise, using Eq. (5) we mapped a dust temperature to each value of  $\langle U \rangle$  and will refer to the two quantity interchangeably.

### 3.2.2. Total infrared and 8 $\mu\text{m}$ luminosities

Each template and component in our library, both for the continuum and PAH emission, is associated with a value of  $L_{\text{IR}}$  (8 to 1000  $\mu\text{m}$ ) and  $L_8$  (integrated with the response of the *Spitzer* IRAC channel 4, i.e., 6.4 to 9.3  $\mu\text{m}$ ). Both luminosities were computed as the integral of  $L_\lambda d\lambda$  within their respective wavelength interval, and are given in units of total solar luminosity ( $L_\odot = 3.839 \times 10^{26}$  W). The luminosities are proportional to the mass of dust, and depend linearly on  $\langle U \rangle$ :

$$L_{\text{IR}} = L_{\text{IR}}^{\text{cont}} + L_{\text{IR}}^{\text{PAH}}, \quad (7)$$

$$L_{\text{IR}}^{\text{cont}}[L_\odot] = 191 (M_{\text{dust}}^{\text{cont}}/M_\odot) (\langle U \rangle/U_\odot), \quad (8)$$

$$L_{\text{IR}}^{\text{PAH}}[L_\odot] = 325 (M_{\text{dust}}^{\text{PAH}}/M_\odot) (\langle U \rangle/U_\odot), \quad (9)$$

$$L_8 = L_8^{\text{cont}} + L_8^{\text{PAH}}, \quad (10)$$

$$L_8^{\text{cont}}[L_\odot] = 7.05 (M_{\text{dust}}^{\text{cont}}/M_\odot) (\langle U \rangle/U_\odot), \quad (11)$$

$$L_8^{\text{PAH}}[L_\odot] = 755 (M_{\text{dust}}^{\text{PAH}}/M_\odot) (\langle U \rangle/U_\odot). \quad (12)$$

Defining the mass fraction of PAHs as  $f_{\text{PAH}} \equiv M_{\text{dust}}^{\text{PAH}}/M_{\text{dust}}$ , we have

$$\text{IR8} \equiv \frac{L_{\text{IR}}}{L_8} = \frac{L_{\text{IR}}^{\text{cont}} (1 - f_{\text{PAH}}) + L_{\text{IR}}^{\text{PAH}} f_{\text{PAH}}}{L_8^{\text{cont}} (1 - f_{\text{PAH}}) + L_8^{\text{PAH}} f_{\text{PAH}}},$$

$$\frac{1}{f_{\text{PAH}}} = 1 - \frac{L_{\text{IR}}^{\text{PAH}} - L_8^{\text{PAH}} \text{IR8}}{L_{\text{IR}}^{\text{cont}} - L_8^{\text{cont}} \text{IR8}}. \quad (13)$$

Using the above approximate equations, this becomes:

$$\text{IR8} = \frac{191 + 134 f_{\text{PAH}}}{7.05 + 748 f_{\text{PAH}}} \quad \text{and} \quad f_{\text{PAH}} = \frac{191 - 7.05 \text{IR8}}{-134 + 748 \text{IR8}}. \quad (14)$$

By varying the relative contribution of PAHs to the total dust mass, our library can reach IR8 values in the range 0.48 to 27.7, which covers the vast majority of the observed parameter space (Elbaz et al. 2011).

### 3.2.3. Basic usage

Our library is composed of multiple templates, each corresponding to a given dust temperature  $T_{\text{dust}}$ . As illustrated in Fig. 1, each template is composed of two components: dust continuum on the one hand, and PAH emission on the other hand. The amplitude of each component is internally dictated by the corresponding mass of dust grains,  $M_{\text{dust}}^{\text{cont}}$  and  $M_{\text{dust}}^{\text{PAH}}$ , respectively (Eq. (3)), and both can be freely adjusted to match the observed data, effectively varying the dust mass (or  $L_{\text{IR}}$ ) and the IR8.

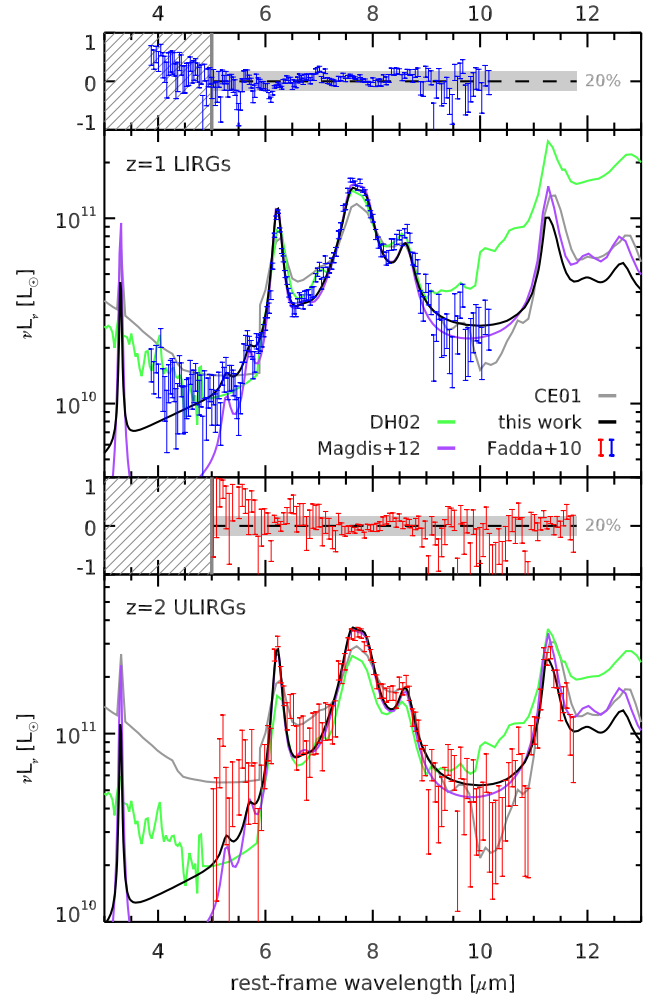
To perform the fit, both PAH and continuum components must be redshifted to the assumed redshift of the source, and the expected flux in each observed passband is computed by integrating the redshifted template multiplied with the corresponding filter response curve. At this stage, the expected fluxes can be fit to the observed ones through a linear solver, varying simultaneously  $M_{\text{dust}}^{\text{cont}}$  and  $M_{\text{dust}}^{\text{PAH}}$ . By performing such a fit for each value of  $T_{\text{dust}}$  in the library and picking the smallest  $\chi^2$ , one can find the optimal model (in the  $\chi^2$  sense) corresponding to the provided photometry. This is the simplest way to use the library, and it will work in most cases if enough photometry is available and if the redshift is precisely known. In other cases, a more careful approach should be used, and we describe it later in Sect. 5.

The immediate products of this fit are the dust masses,  $M_{\text{dust}}^{\text{cont}}$  and  $M_{\text{dust}}^{\text{PAH}}$  (expressed in solar mass), and the dust temperature  $T_{\text{dust}}$ . The total dust mass can be obtained by summing the two components,  $M_{\text{dust}} = M_{\text{dust}}^{\text{cont}} + M_{\text{dust}}^{\text{PAH}}$ , while  $L_{\text{IR}}$  and  $L_8$  are tabulated (per unit solar mass of dust) for each  $T_{\text{dust}}$  value in the library, or can be estimated using the above relations. In the next section, we check the accuracy of our PAH templates by fitting stacked MIR spectra from the *Spitzer* IRS spectrometer.

### 3.3. Comparison against stacked MIR spectroscopy

During the cryogenic phase of the *Spitzer* mission, the IRS spectrometer could observe in the MIR from  $5.3$  to  $38 \mu\text{m}$  at low ( $R = 90$ ) and medium ( $R = 600$ ) spectral resolutions. It has therefore provided valuable measurements of the PAH emission, both in local and distant galaxies. In particular, Fadda et al. (2010) have observed a sample of  $z = 1$  LIRGs and  $z = 2$  ULIRGs in the GOODS–South field (plus a few in the wider ECFDS). The galaxies in these two samples have been selected based on their redshift and *Spitzer* MIPS  $24 \mu\text{m}$  flux ( $0.2$ – $0.5$  mJy), therefore their measured properties cannot be straightforwardly compared to the mass-complete stacks that we will analyze in Sect. 4.1. Because of the  $24 \mu\text{m}$  flux limit, the sample of Fadda et al. is biased toward starburst galaxies located above the main sequence. With this caveat in mind, this dataset can still be used as a consistency check for our SED library. Indeed, although these samples may be biased, our library must be able to at least broadly reproduce their PAH spectra.

We therefore applied the fitting method described in the previous section to the stacked spectra of both  $z = 1$  and  $z = 2$  samples. Since the dust temperature cannot be constrained from IRS spectroscopy alone, we fixed it to its redshift-average value of  $T_{\text{dust}} = 28$  and  $33$  K, respectively (see Sect. 4.1), although this choice had no impact on the quality of the fit. We also did not attempt to fit rest-frame wavelengths  $\lambda < 5 \mu\text{m}$  to avoid contamination from stellar continuum. The result is shown in Fig. 2, together with fits from a number of other libraries from the literature. It can be seen from this figure that we are able to fit these spectra with good accuracy (typically better than 20%) and obtain a significantly improved match compared to the CE01 or DH02 libraries. The Magdis et al. (2012) library (which uses the DL07 models) yields a similarly good fit as ours, with only subtle differences. This should not come as a surprise, since the physical properties of the PAHs in both the Magdis et al. library and ours are adapted from DL07.



**Fig. 2.** Comparison of our templates (black solid line) against stacked *Spitzer* IRS spectra of  $z = 1$  LIRGs (blue, top) and  $z = 2$  ULIRGs (red, bottom) from Fadda et al. (2010). The relative residuals of the fits are shown above the plot for each sample; the region of perfect agreement is shown with a dashed line, surrounded by a  $\pm 20\%$  confidence interval. We also show the best fit using other models: CE01 (gray), DH02 (green), and Magdis et al. (2012; purple). In all cases the fit only uses observations at  $\lambda > 5 \mu\text{m}$ , since shorter wavelength can be contaminated by the stellar continuum, as illustrated with the hashed region in the residual plots.

## 4. Observed dust temperatures and IR8

### 4.1. Average values from stacked photometry

#### 4.1.1. Description of the stacked data

We now proceed to apply our new library to model the stacked *Spitzer* and *Herschel* photometry in the CANDELS fields. These stacks were presented in detail in S15, and we briefly recall here the essential information. We selected all galaxies in the GOODS–North, GOODS–South, UDS and COSMOS CANDELS fields that are brighter than  $H = 26$  and  $UVJ$  star-forming (Eq. (2)). These galaxies were then binned according to their redshift and stellar mass, for a total of 24 bins (six redshift bins from  $z = 0.3$  to 5, and four mass bins from  $M_* = 3 \times 10^9$  to  $3 \times 10^{11} M_\odot$ ). The number of stacked galaxies in each bin is given in Fig. 4 of S15. The smallest number of stacked galaxy were 53 and 28, in the highest mass bins at  $z \sim 0.5$  and  $z \sim 4$ ,



respectively. Otherwise, all the bins had at least 100 galaxies (median of 628).

In the bins where we estimated our stellar-mass completeness to be above 90%, we stacked the *Spitzer* and *Herschel* images at the positions of all the galaxies in the bin. The fluxes in the resulting stacked images were measured through point spread function (PSF) fitting with a free background. The measured fluxes were corrected for clustering a posteriori using an empirical recipe calibrated on simulated images, and the method is described fully in the appendix of S15. Briefly, we simulated the *Herschel* maps using the positions of the real galaxies of CANDELS and the prescriptions described in Schreiber et al. (2017a) to predict (statistically) their  $L_{\text{IR}}$  and FIR fluxes. The resulting images have the same statistical properties (pixel distribution and number counts) as the real images. We then applied the same stacking method and compared the measured fluxes to the true flux averages in the stacked sample. We found that the flux boosting caused by clustering is roughly constant in a given band, hence we assumed a fixed correction in each *Herschel* band throughout, at all redshifts and masses (see also Béthermin et al. 2015). The largest correction was a reduction of the fluxes by 25% for the SPIRE 500  $\mu\text{m}$  band.

The uncertainty on each flux measurement was finally computed by taking the maximum value of two independent estimates: first by doing bootstrapping, that is, repeatedly removing half of the sample from the stack, and second from the RMS of the residual image after subtraction of the best-fit PSF model.

In S15, we only stacked the *Spitzer* MIPS 24  $\mu\text{m}$  band as well as the *Herschel* bands redward of 100  $\mu\text{m}$ , which are the only bands covered in all the four CANDELS fields. For the present work, we extended these stacks to also include the *Spitzer* IRS 16  $\mu\text{m}$  imaging (acquired in the GOODS fields only; Teplitz et al. 2011) as well as the *Herschel* PACS 70  $\mu\text{m}$  (acquired in the GOODS–South field only). Because these images only cover some of the CANDELS fields, the stacked fluxes can be affected by field-to-field variations. To correct for this effect, we first computed the  $S_{16}/S_{24}$  flux ratio observed when stacking only the two GOODS fields. We then multiplied the stacked 24  $\mu\text{m}$  flux obtained with the four fields by this ratio to estimate the corresponding 16  $\mu\text{m}$  flux. The same procedure is used for the 70  $\mu\text{m}$  flux, based on the  $S_{70}/S_{100}$  flux ratio observed in GOODS–South. Lastly, to better constrain the dust temperature for our  $z = 4$  bin we added the average ALMA 890  $\mu\text{m}$  of our galaxies as measured in Schreiber et al. (2017b; in all fields but GOODS–North).

#### 4.1.2. Fitting of the stacked SEDs

Using the fitting method described in Sect. 3.2, we used our new library to model the stacked fluxes from S15 (avoiding again  $\lambda < 5 \mu\text{m}$ ). We used the mean redshift of the stacked sample to shift the SED in the observed frame. Uncertainties on the redshifts are of the order of 5% and are thus much smaller than the bins (which have a constant width of 25%), we therefore ignored them, but we did broaden the templates by the width of the redshift bins. To derive accurate error estimates on the fitting parameters, we bootstrapped the stacked sample and applied the fitting procedure on each bootstrapped SED. This allows a better treatment of correlated noise (flux fluctuations affecting multiple bands simultaneously due, e.g., to contamination from neighboring sources on the map). The resulting models are compared to the measured photometry in Fig. 3.

The first thing to note is the excellent agreement between our templates and the photometry. With only three free parameters,

no clear tension is observed, reinforcing the idea of a universal SED. Compared to our previous fits with the CE01 library, we found very similar values of  $L_{\text{IR}}$ . The most extreme difference arose in the lowest redshift bin ( $0.3 < z < 0.7$ ) where we obtained values that are systematically 0.1 dex lower with the new library. This difference is caused by a peculiar feature of the previously adopted best-fit CE01 template. This particular SED (ID 40) shows an enhanced flux around the rest-frame 30  $\mu\text{m}$  compared to our library. Without any data to constrain this feature, we cannot say whether it is real or not, although we tend to favor the result of our new SED library which has a consistent shape at all  $T_{\text{dust}}$ .

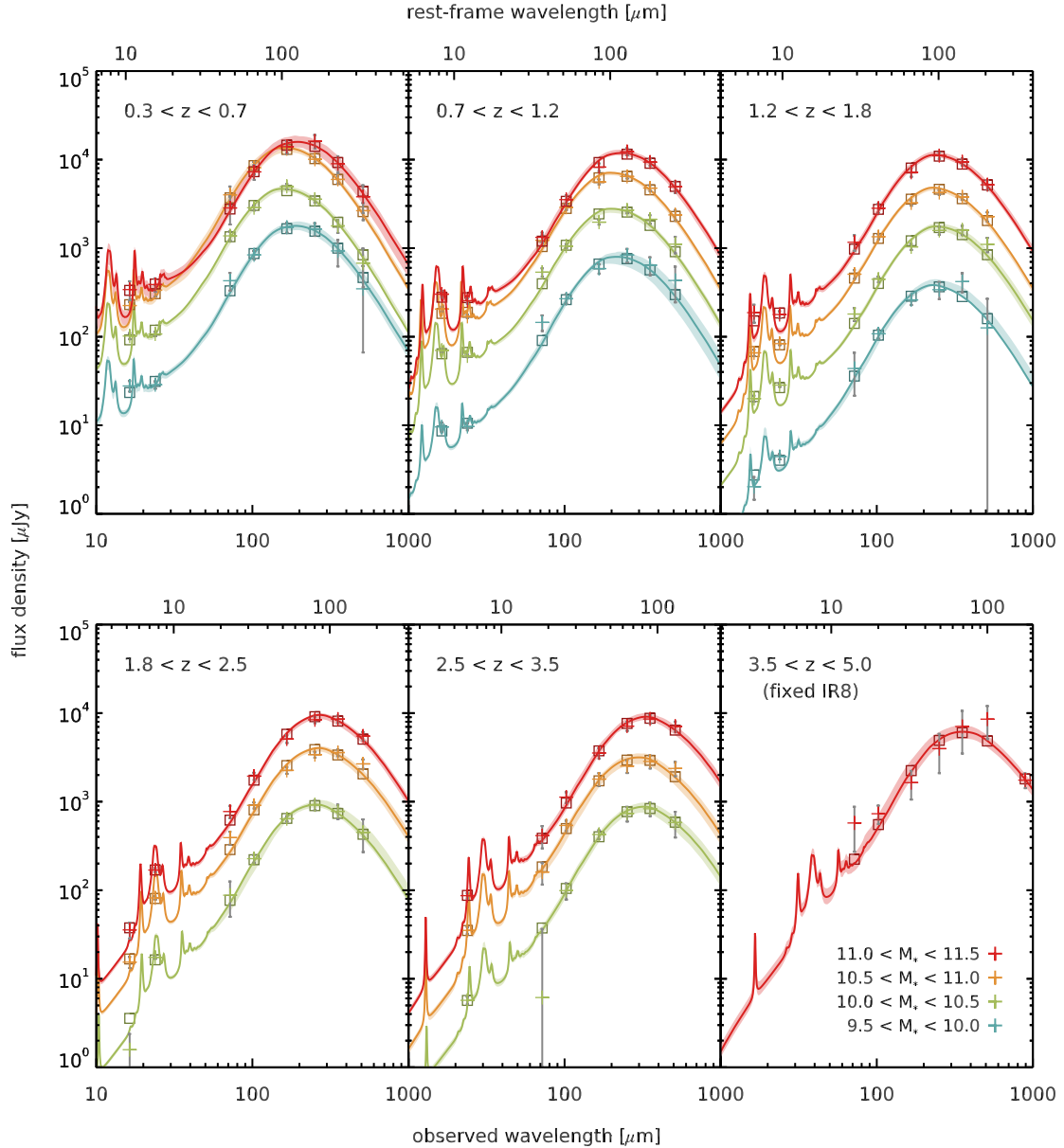
#### 4.1.3. Best fit parameters and systematic bias corrections

One major issue when interpreting stacked photometry is that the stacked SED is the flux-weighted average in each band, which is not necessarily equal to the SED of the average galaxy of the sample. This is particularly true if the brightest galaxies have a different SED than the average galaxy, and indeed such situation is expected in our case: starburst galaxies, which have the highest SFRs in a given bin of mass, have an increased temperature and IR8. Therefore our stacked SED will be biased towards higher temperatures and higher IR8 compared to their true average, and these biases need to be accounted for before going further.

To quantify these biases, we used the Empirical Galaxy Generator (EGG; Schreiber et al. 2017a). Using a set of empirical prescriptions, this tool can generate mock galaxy catalogs matching exactly the observed stellar mass functions at  $0 < z < 6$  and the galaxy main sequence with its scatter and starburst population (S15). We also implemented the relations derived in the present paper for  $T_{\text{dust}}$  and IR8 and their dependence on  $R_{\text{SB}}$ , namely Eqs. (15) to (19), adding the observed residual scatter (see Sects. 4.2.3 and 4.2.5) as a random Gaussian perturbation to match the observed distribution of both quantities. Thanks to these prescriptions, the tool can produce realistic mock catalogs of galaxies with a full infrared SED, and we showed in Schreiber et al. (2017a) that this empirical model reproduces faithfully the observed number counts in all FIR bands.

Generating a large mock catalog of about a million galaxies with  $M_* > 3 \times 10^9 M_{\odot}$  at  $0.3 < z < 5$ , we simulated our stacked SEDs by computing the average flux in each band for each of the stellar mass and redshift bins displayed in Fig. 3. The size of the mock catalog was chosen so that each stacked bin contained at least 1000 galaxies. No noise was added to the stacked fluxes, since we only looked for systematic biases. We then applied exactly the same fitting method to these synthetic stacks as used for the real stacks, and compared the resulting fit parameters to their true average. Since the recipes for  $T_{\text{dust}}$  and IR8 used in EGG depend to some extent on the bias correction we are now discussing, we proceeded iteratively: we derived a first estimate of Eqs. (15) to (19) without applying any bias correction, implemented these relations in EGG, and determined a first value of the corrections. We then applied these corrections to the observed values, re-evaluated the relations, updated EGG and the mock catalog, then determined the final corrections. The obtained values were not significantly different from the first estimates, so we only did two iterations of this procedure.

We found that the “raw” stacked values (before correction) were on average larger than their true average by  $1.5 \pm 0.3 \text{ K}$  for  $T_{\text{dust}}$ , and by a factor of  $10 \pm 1\%$  for IR8. These are relatively small corrections which do not impact our conclusions, but we applied them nevertheless to our best fit values.



**Fig. 3.** *Spitzer* and *Herschel* stacks of S15 (crosses) of main sequence galaxies at different redshifts (from left to right) and for different stellar masses (colors, see legend). These also include new stacks of the *Spitzer*  $16\mu\text{m}$  and *Herschel*  $70\mu\text{m}$  images in the GOODS fields, as well as ALMA  $890\mu\text{m}$  from Schreiber et al. (2017b) for the  $z = 4$  bin. We overplot the best fit template from our library with colored solid lines. Open boxes in the background show the modeled broadband flux from the best-fit template, to illustrate any offset with the observations. The colored regions in the background display the scatter of the best-fit model SED among all the bootstrap realizations.

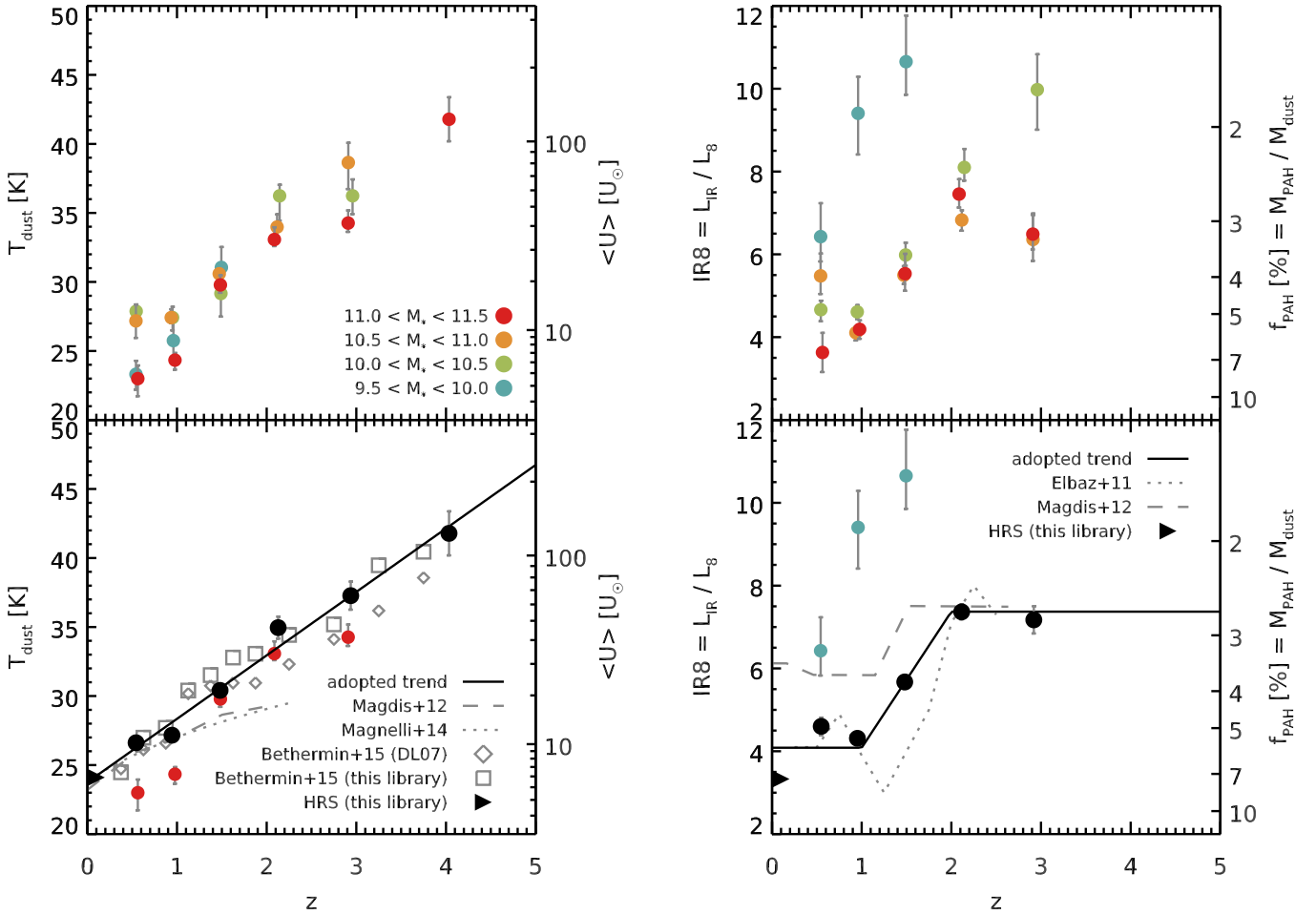
In Fig. 4 (top left and top right) we show the best-fit values we obtain for  $T_{\text{dust}}$  and IR8 in all bins of mass and redshifts where they could be measured. These are also tabulated in Table 1. To complement our stacks toward the local Universe, we also compute the average  $T_{\text{dust}}$  and IR8 for the UVJ star-forming galaxies of the HRS. The evolution with redshift and mass of  $T_{\text{dust}}$  and IR8 are quantified in the following sections.

#### 4.1.4. Evolution of the dust temperature

In most cases, varying the stellar mass has no influence on the dust temperature. The most massive galaxies ( $M_* > 10^{11} M_{\odot}$ ) tend to have colder temperatures by 2.3 K on average (see also Matsuki et al. 2017, who report a similar trend at  $z = 0$ ), but this is only significant at  $z < 1$ , in the domains where the main

sequence departs from a linear relation (e.g., Whitaker et al. 2015; Schreiber et al. 2015). This reduced temperature was already interpreted in our previous work as a sign that massive star-forming galaxies at low redshifts are in the process of shutting down star-formation, with a slowly declining efficiency (Schreiber et al. 2016). Since this is a minor effect compared to the overall redshift evolution, and since it only affects the few most massive galaxies, we do not discuss it further here.

Averaging the dust temperatures of the mass bins unaffected by the above effect, we recovered a previously reported trend for the dust temperature to increase with redshift (e.g., Magdis et al. 2012; Magnelli et al. 2013; Béthermin et al. 2015). Contrary to what Magdis et al. (2012) claimed, we observed a continuous rise of the temperature up to  $z = 4$ , confirming the results of Béthermin et al. (2015). Fitting the evolution of  $T_{\text{dust}}$  with



**Fig. 4.** *Left:* evolution of the effective dust temperature  $T_{\text{dust}}$  with redshift. The  $T_{\text{dust}}$  estimated from each stacked SED at different stellar masses are shown on the *top* plot with circles of different colors (see legend). The *bottom* plot shows the weighted mean of all mass bins as black circles (excluding the most massive bin, shown in red, or the least massive bin, shown in blue). The average of the galaxies from the HRS (Schreiber et al. 2016; Ciesla et al. 2014) is given with a solid triangle. The trend we adopt in this paper is shown with a solid black line. We also show the  $T_{\text{dust}}$  evolution of Magdis et al. (2012) and Béthermin et al. (2015); both converted from  $\langle U \rangle$  to  $T_{\text{dust}}$  using Eq. (5) as well as Magnelli et al. (2014; corrected from light-weighted to mass-weighted using Eq. (6)). We also show the best fitting temperature to the Béthermin et al. (2015) SEDs modeled with our own library as open squares. *Right:* evolution of the IR8 ratio with redshift. The legends are the same as for the plots on the left, except that here we show the values obtained by Magdis et al. (2012, computed from their best-fit template SEDs) and Elbaz et al. (2011).

redshift as an empirical power law, we obtained

$$T_{\text{dust}}^{\text{MS}} [\text{K}] = (32.9 \pm 2.4) + (4.60 \pm 0.35) \times (z - 2). \quad (15)$$

This relation is displayed in Fig. 4 (bottom left), where we compare it to results from the literature. In particular Magnelli et al. (2014) found  $T_{\text{dust}} = 26.5 \times (1 + z)^{0.18}$ . The normalization of this relation is higher than the one we report here, but the evolution with redshift is milder. This higher normalization is linked to the fact that Magnelli et al. (2014) measured  $T_{\text{dust}}$  by fitting modified blackbodies, making their dust temperatures light-weighted. Correcting for this difference using Eq. (6) (as was done in Fig. 4), their  $z = 0$  value is fully consistent with ours, but their measurement at  $z = 2$  falls short of ours by 5 K. The same is true for the stacks of Magdis et al. (2012). This could be caused by the absence of clustering correction in these two studies, since it affects preferentially the long wavelength *Herschel* bands which are crucial to determine the temperature. Magnelli et al. (2014) does exclude the bands for which they predict the flux is doubled by the effect of clustering, however this threshold is extreme and will leave substantial clustering signal in their photometry.

To further check for systematic issues in our stacked fluxes, we applied our model to the stacked SEDs of Béthermin et al. (2015), which were obtained for a different sample, in a single mass bin, and include longer wavelength data from LABOCA 870  $\mu\text{m}$  and AzTEC 1.1 mm at all redshifts. The correction for clustering is also performed with another method. Despite these differences, the evolution of  $T_{\text{dust}}$  in these data is in excellent agreement with the above relation, suggesting that our stacked fluxes are robustly measured. We also compare the  $\langle U \rangle$  value as reported by Béthermin et al. (2015) for reference; the DL07 model used in that work assumes a different functional form for the  $U$  distribution, so the comparison is limited in scope, but we nevertheless recover a similar slope for the redshift dependence of the temperature.

We therefore confirm that the dust temperature increases continuously from about 25 K at  $z = 0$  up to more than 40 K at  $z = 4$ , with little to no dependence on stellar mass (hence  $L_{\text{IR}}$ ) at fixed redshift. This is important to take into account when only limited FIR photometry is available and  $T_{\text{dust}}$  needs to be assumed to extrapolate the total IR luminosity, particularly for

**Table 1.** Best fit dust parameters for the stacked SEDs.

$z$	$\log_{10} M_*$ $\log_{10} M_\odot$	$L_{\text{IR}}$ $10^{10} L_\odot$	$M_{\text{dust}}$ $10^7 M_\odot$	$T_{\text{dust}}$ K	$f_{\text{PAH}}$ %	IR8
0.3–0.7	9.5–10.0	$1.49^{+0.08}_{-0.09}$	$0.136^{+0.034}_{-0.025}$	$23.3^{+1.0}_{-1.1}$	$3.08^{+0.40}_{-0.46}$	$6.4^{+0.8}_{-0.6}$
	10.0–10.5	$4.61^{+0.20}_{-0.20}$	$0.180^{+0.019}_{-0.017}$	$27.9^{+0.5}_{-0.4}$	$4.66^{+0.36}_{-0.28}$	$4.66^{+0.22}_{-0.28}$
	10.5–11.0	$12.9^{+0.9}_{-0.9}$	$0.57^{+0.10}_{-0.07}$	$27.2^{+1.2}_{-1.2}$	$3.8^{+0.5}_{-0.5}$	$5.5^{+0.5}_{-0.4}$
	11.0–11.5	$14.9^{+1.5}_{-1.5}$	$1.43^{+0.38}_{-0.26}$	$23.0^{+1.0}_{-1.3}$	$6.4^{+0.9}_{-0.9}$	$3.6^{+0.5}_{-0.5}$
0.7–1.2	9.5–10.0	$2.24^{+0.13}_{-0.18}$	$0.130^{+0.064}_{-0.035}$	$25.7^{+1.8}_{-1.7}$	$1.76^{+0.35}_{-0.24}$	$9.4^{+0.9}_{-1.0}$
	10.0–10.5	$8.48^{+0.23}_{-0.35}$	$0.354^{+0.061}_{-0.031}$	$27.4^{+0.8}_{-0.9}$	$4.73^{+0.26}_{-0.25}$	$4.61^{+0.16}_{-0.17}$
	10.5–11.0	$21.2^{+0.8}_{-0.7}$	$0.88^{+0.08}_{-0.08}$	$27.4^{+0.6}_{-0.4}$	$5.47^{+0.33}_{-0.26}$	$4.11^{+0.15}_{-0.19}$
	11.0–11.5	$35.2^{+1.8}_{-2.3}$	$2.58^{+0.31}_{-0.31}$	$24.3^{+0.5}_{-0.7}$	$5.36^{+0.33}_{-0.31}$	$4.19^{+0.22}_{-0.23}$
1.2–1.8	9.5–10.0	$2.91^{+0.22}_{-0.14}$	$0.063^{+0.010}_{-0.016}$	$31.1^{+1.5}_{-0.3}$	$1.40^{+0.23}_{-0.27}$	$10.7^{+1.1}_{-0.8}$
	10.0–10.5	$13.0^{+0.6}_{-0.5}$	$0.40^{+0.10}_{-0.06}$	$29.2^{+0.3}_{-1.7}$	$3.36^{+0.20}_{-0.21}$	$5.98^{+0.30}_{-0.26}$
	10.5–11.0	$36.3^{+1.3}_{-1.5}$	$0.85^{+0.07}_{-0.07}$	$30.6^{+0.3}_{-0.5}$	$3.75^{+0.19}_{-0.21}$	$5.50^{+0.22}_{-0.21}$
	11.0–11.5	$84^{+4}_{-4}$	$2.35^{+0.37}_{-0.24}$	$29.8^{+0.7}_{-0.6}$	$3.72^{+0.38}_{-0.38}$	$5.5^{+0.5}_{-0.4}$
1.8–2.5	10.0–10.5	$16.3^{+0.7}_{-0.8}$	$0.131^{+0.069}_{-0.024}$	$36.2^{+0.8}_{-1.8}$	$2.13^{+0.13}_{-0.19}$	$8.10^{+0.45}_{-0.32}$
	10.5–11.0	$65.4^{+2.3}_{-2.8}$	$0.76^{+0.16}_{-0.13}$	$34.0^{+0.9}_{-1.1}$	$2.75^{+0.12}_{-0.14}$	$6.83^{+0.24}_{-0.26}$
	11.0–11.5	$144^{+6}_{-7}$	$2.04^{+0.28}_{-0.28}$	$33.1^{+0.9}_{-0.5}$	$2.44^{+0.15}_{-0.20}$	$7.46^{+0.36}_{-0.32}$
2.5–3.5	10.0–10.5	$26.5^{+2.2}_{-2.0}$	$0.213^{+0.069}_{-0.033}$	$36.2^{+1.2}_{-1.3}$	$1.52^{+0.30}_{-0.21}$	$10.0^{+0.9}_{-1.0}$
	10.5–11.0	$102^{+7}_{-10}$	$0.58^{+0.14}_{-0.13}$	$38.6^{+1.4}_{-1.9}$	$2.98^{+0.36}_{-0.37}$	$6.4^{+0.6}_{-0.5}$
	11.0–11.5	$257^{+17}_{-16}$	$2.9^{+0.5}_{-0.5}$	$34.3^{+0.9}_{-0.6}$	$2.96^{+0.25}_{-0.25}$	$6.49^{+0.46}_{-0.37}$
3.5–5.0	11.0–11.5	$357^{+37}_{-61}$	$1.41^{+0.26}_{-0.18}$	$41.8^{+1.6}_{-1.6}$	–	–

sub-millimeter samples which do not probe the emission blueward of the peak of the dust emission (see Sect. 6).

#### 4.1.5. Evolution of the IR8

Elbaz et al. (2011) proposed that a unique value of  $\text{IR8} = 4.9$  holds for all main sequence galaxies, however it could be seen already from their stacked data (see their Fig. 7) that the average IR8 is closer to 8 at  $z = 2$  (see also Reddy et al. 2012).

After applying the correction described in Sect. 4.1.3, we found  $\text{IR8} \sim 7$  at  $z = 2$ , mildly but significantly larger than the value first reported in Elbaz et al. (2011) at  $z = 1$ . On the other hand, we found the  $z = 0$  galaxies from the HRS have a marginally smaller IR8 of about 3.5, very similar to our  $z = 1$  value of 4. This implies that the average value has evolved continuously between  $z = 2$  and  $z = 1$  only, and our stacks at  $z = 3$  suggest that this evolution stops at higher redshifts. We thus fit the evolution of the average IR8 as a broken linear relation and obtained

$$\text{IR8}^{\text{MS}} = 4.08 \pm 0.29 + (3.29 \pm 0.24) \times \begin{cases} 0 & \text{if } z < 1, \\ (z - 1) & \text{if } 1 \leq z \leq 2, \\ 1 & \text{if } z > 2. \end{cases} \quad (16)$$

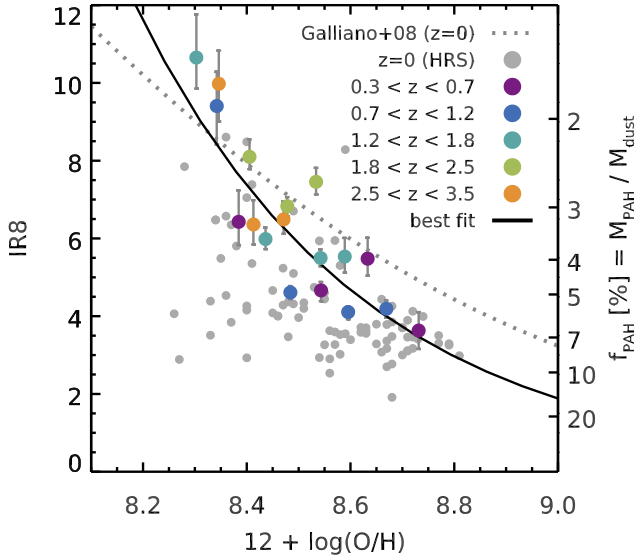
This relation is displayed in Fig. 4 (bottom right) and compared to the values obtained by Elbaz et al. (2011, Fig. 7) and Magdis et al. (2012), which are in rough agreement. The IR8 value of Magdis et al. (2012) at  $z = 0$  is significantly higher than ours, probably because their local sample was flux-limited (da Cunha et al. 2010), hence is biased toward starbursts which have a systematically higher IR8 (see Sect. 4.2.2).

Interestingly, we found that low mass galaxies ( $M_* < 10^{10} M_\odot$ ) have systematically higher IR8 values, implying that

their PAH emission is reduced. Shivaie et al. (2017) observed a similar trend in  $z \sim 2$  galaxies with spectroscopic redshifts. Shivaie et al. also observe that this trend of reduced PAH emission can be linked to decreasing metallicity, as observed in the local Universe (e.g., Galliano et al. 2003; Ciesla et al. 2014). The physical origin of this trend is still debated. One plausible explanation is that a metal-poor interstellar medium blocks less efficiently the UV radiation of young stars, and makes it harder for PAH molecules to survive (e.g., Galliano et al. 2003). Other scenarios have been put forward, suggesting either that low metallicity objects are simply too young to host enough carbon grains to form PAH complexes (Galliano et al. 2008), or that this is instead caused by a different filling factor of molecular clouds in metal poor environments (Sandstrom et al. 2012). Metallicity, in turn, is positively correlated with the stellar mass through the mass-metallicity relation (Lequeux et al. 1979; Tremonti et al. 2004), and this relation has been found to evolve with time, so that galaxies were more metal-poor in the past (e.g., Erb et al. 2006). If metallicity was the main driver of PAH emission, one would expect to find the strongest PAH features (and the lowest IR8) within massive low-redshift galaxies, which is indeed what we found.

To test this hypothesis quantitatively, we used the Fundamental Metallicity Relation (FMR; Mannucci et al. 2010) to estimate the average metallicity of our stacked galaxies from their stellar masses and SFRs. The resulting relation between metallicity and IR8 is shown in Fig. 5. We found a clear anti-correlation between the two quantities, quantitatively matching the trend observed in the local Universe by Galliano et al. (2008) and confirming the results of Shivaie et al. (2017). The best fitting power law is

$$\text{IR8} = (3.5 \pm 0.3) \times (Z/Z_\odot)^{-0.99 \pm 0.15}, \quad (17)$$



**Fig. 5.** Relation between the IR8 observed in stacked *Spitzer* and *Herschel* photometry, and the gas-phase metallicity. The metallicity, expressed in terms of oxygen abundance  $12 + \log_{10}(\text{O}/\text{H})$  (the solar value is 8.73, [Asplund et al. 2009](#)), was estimated using the Fundamental Metallicity Relation (FMR, [Mannucci et al. 2010](#)). Data points are colored with redshift as indicated in the legend. We also show the galaxies from the HRS as gray circles, using their measured metallicities. The black line shows the best fitting power law to the stacked data. The  $z = 0$  relation obtained by [Galliano et al. \(2008\)](#) is shown for reference with a dotted gray line, converting their measured PAH mass fractions to IR8 using Eq. (14).

where  $Z$  is the metallicity (which can be substituted for the oxygen abundance  $\text{O}/\text{H}$  under the assumption that  $\text{O}/\text{H}$  scales linearly with  $Z$ ). The galaxies of the HRS, for which we have individual metallicity measurements ([Boselli et al. 2010](#)), also follow a similar trend.

Unfortunately, the power of the FMR in predicting metallicities is limited (particularly at high redshifts, e.g., [Béthermin et al. 2015](#)), and measuring metallicities from emission lines for individual galaxies is both expensive and prone to systematics ([Kewley & Ellison 2008](#)). Excluding the masses below  $10^{10} M_{\odot}$ , the model of Eq. (16) with a simple dependence on redshift provides a fit of equal quality to the data. Since Eq. (16) is more easily applicable to large samples, we chose to adopt it as the fiducial model and caution that the relations derived below for IR8 only apply to galaxies more massive than  $10^{10} M_{\odot}$ .

## 4.2. Values for individual galaxies

In this section we describe the scatter on both  $T_{\text{dust}}$  and IR8 about the average “main sequence” values we obtained in Sect. 4.1. We also describe how these quantities are modified for starburst galaxies, that is, those galaxies that have an excess SFR at a given stellar mass compared to the main sequence. To quantify this latter excess, we used the “starburstiness” ([Elbaz et al. 2011](#)) which is defined as  $R_{\text{SB}} \equiv \text{SFR}/\text{SFR}_{\text{ms}}$ ; galaxies with  $R_{\text{SB}} = 1$  are on the main sequence, and those with  $R_{\text{SB}} > 1$  are located above the sequence.

### 4.2.1. Fitting individual galaxies

We used our library to fit the FIR-detected galaxies in the CANDELS fields and COSMOS  $2 \square^{\circ}$ . To ensure reliable fits, we

selected galaxies with a good enough wavelength coverage and robust photometry. In particular we only used the *Herschel* photometry for clean sources ([Elbaz et al. 2011](#)) to avoid biases toward low  $T_{\text{dust}}$  because of blending, and excluded all sources for which the matching of the  $24 \mu\text{m}$  emission to a  $H$  or  $K_s$ -band counterpart was ambiguous (as defined in [Schreiber et al. 2016](#)). We also excluded fits of poor quality by rejecting galaxies with  $\chi^2$  larger than 10 (less than 10% of our sample), indicative of uncaught issues in the photometry and counterpart identification.

For the measurement of  $T_{\text{dust}}$ , following [Hwang et al. \(2010\)](#) we required at least one photometric measurement with significance greater than  $5\sigma$  on either side of the peak of the dust continuum. This produced a sample of 438 galaxies, which are shown on the left panel of Fig. 6.

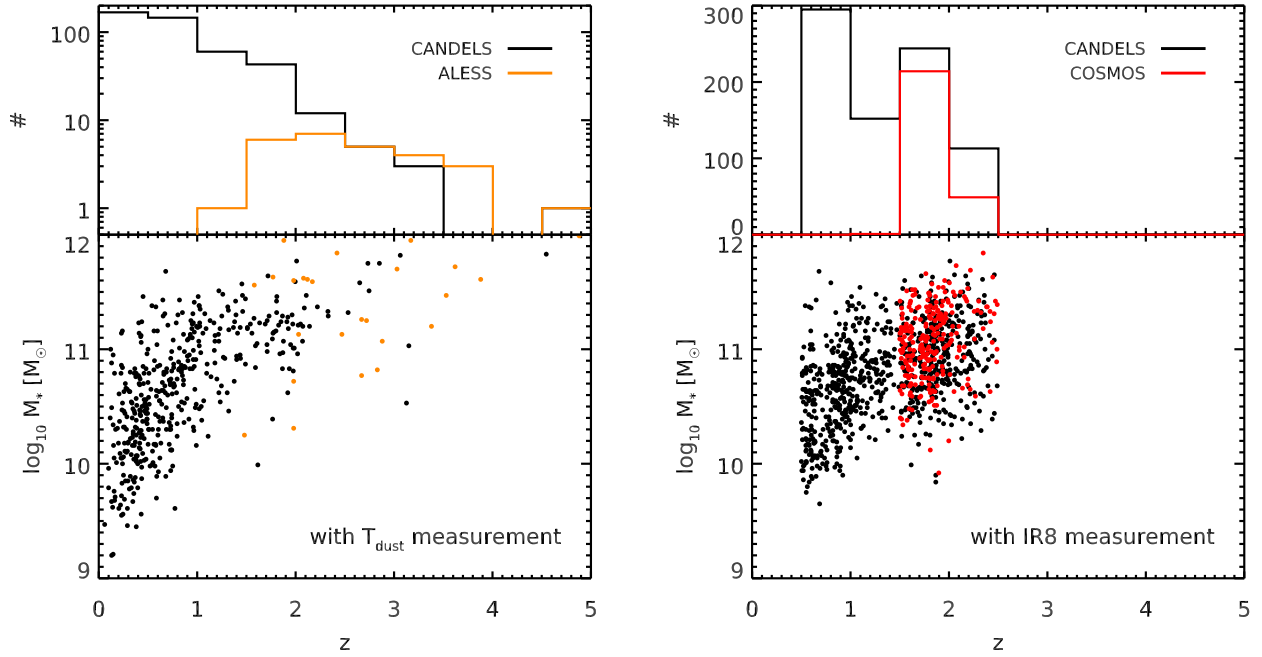
For the measurement of IR8, we selected only the galaxies at  $0.5 < z < 1.5$  with at least a  $3\sigma$  detection in *Spitzer* IRS  $16 \mu\text{m}$  and the galaxies at  $1.5 < z < 2.5$  with at least a  $3\sigma$  detection in *Spitzer* MIPS  $24 \mu\text{m}$  to ensure the rest-frame  $8 \mu\text{m}$  emission is well constrained. Of these, we only kept the galaxies for which  $L_{\text{IR}}$  could be independently measured using longer wavelength photometry at better than  $5\sigma$ . This produced a sample of 1068 galaxies, shown on the right panel of Fig. 6, and 264 of these are in COSMOS  $2 \square^{\circ}$ .

### 4.2.2. Completeness and selection biases

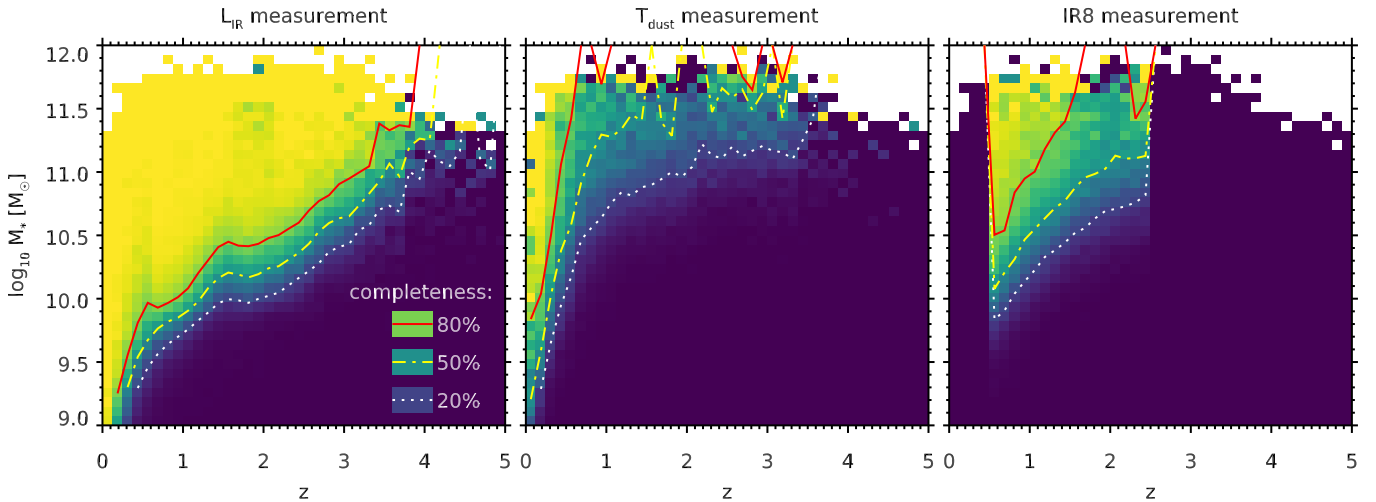
To quantify the selection biases introduced by the numerous criteria listed above, we created a new mock catalog with EGG (see Sect. 4.1.3) over  $10 \text{ deg}^2$  and down to a *Spitzer* IRAC  $3.6 \mu\text{m}$  magnitude of 25 (slightly deeper than the typical  $5\sigma$  depth in CANDELS). We then perturbed the generated fluxes within the uncertainties typical for CANDELS, and mimicked the *Spitzer* and *Herschel* flux extraction procedure described in [Elbaz et al. \(2011\)](#): we only kept the  $16$  and  $24 \mu\text{m}$  fluxes for galaxies with a  $3.6 \mu\text{m}$  flux larger than  $0.5 \mu\text{Jy}$  ([Magnelli et al. 2011](#)), we only kept the  $70$  and  $100 \mu\text{m}$  fluxes for galaxies with a  $24 \mu\text{m}$  flux larger than  $21 \mu\text{Jy}$  ( $3\sigma$ ), we only kept the  $160$  to  $500 \mu\text{m}$  fluxes for galaxies with a  $24 \mu\text{m}$  flux larger than  $35 \mu\text{Jy}$  ( $5\sigma$ ), and finally we only kept the  $350$  and  $500 \mu\text{m}$  fluxes for galaxies with a  $250 \mu\text{m}$  flux larger than  $3.8 \text{ mJy}$  ( $2\sigma$ ). Since most of our sample has no ALMA data, we did not include ALMA photometry in the mock catalog.

We then ran the same fitting procedure as for the real galaxies, and applied the same selection criteria to produce the two samples with robust  $T_{\text{dust}}$  and IR8. Normalizing the number of objects in the mock catalog over an area equal to that of CANDELS (and accounting for the loss of effective area caused by the rejection of the non-clean *Herschel* sources), the mock catalog predicts 396 galaxies with a robust  $T_{\text{dust}}$  measurement, and 1101 galaxies with a robust IR8. These numbers are in excellent agreement with the observed ones, and confirm the validity of the mock catalog.

We display in Fig. 7 the completeness of various samples: selecting galaxies with an  $L_{\text{IR}}$  measured at better than  $3\sigma$ , selecting galaxies with a robust  $T_{\text{dust}}$ , and selecting galaxies with a robust IR8. It is clear from this plot that one can only obtain robust measurements of  $T_{\text{dust}}$  or IR8 for a fraction of the IR-detected galaxies. For the dust temperature, our sample reaches 80% completeness only at  $z < 0.6$ , however it can still probe galaxies with  $M_* > 10^{11} M_{\odot}$  at much higher redshifts (up to  $z \sim 3$ ) albeit with a lower completeness of 20%; this implies that selection effects are important and have to be studied carefully. The situation for the IR8 is not as dramatic, as the sample



**Fig. 6.** Distribution of galaxies with a  $T_{\text{dust}}$  (left) or IR8 (right) measurement in the  $M_*$ -redshift plane. The galaxies from CANDELS are shown with black circles, that from the larger COSMOS field in red, and the smaller sample of ALESS galaxies is shown in orange. At the top we display the redshift distribution of each sample.



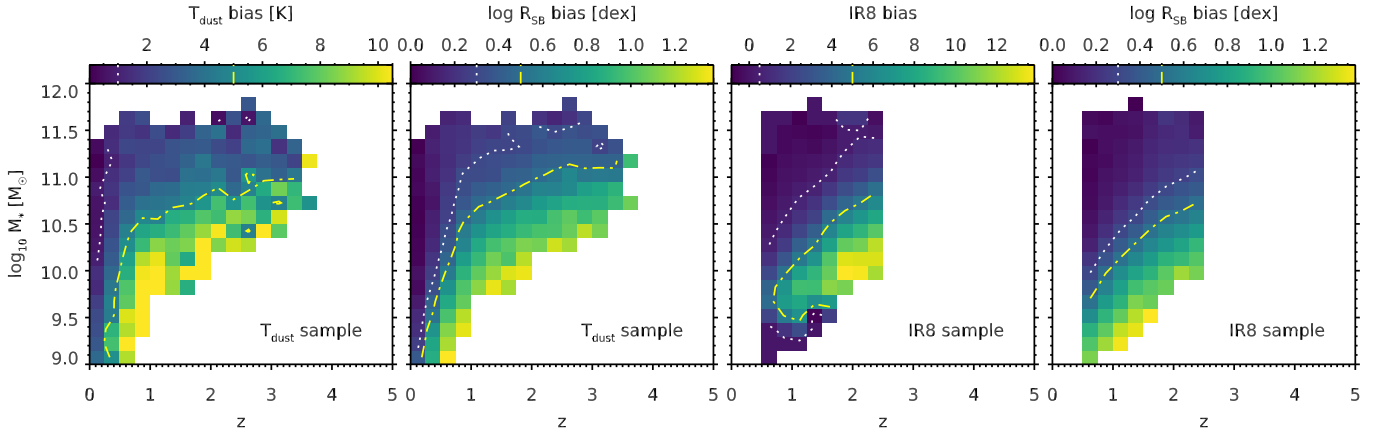
**Fig. 7.** Left: completeness of a CANDELS-like sample computed from the mock catalog, limited to galaxies with a measurement of  $L_{\text{IR}}$  with  $S/N > 3$ , from either *Spitzer* MIPS or *Herschel*. The completeness is shown as a function of redshift and stellar mass: each cell of the plot displays the completeness in the corresponding bin of redshift and mass. The 80, 50 and 20% completeness levels are indicated with red, yellow and white lines, respectively. Center: same as left, but limited to the sample with a robust  $T_{\text{dust}}$  measurement. Right: same as left, but limited to the sample with a robust IR8 measurement.

reaches 80% completeness at  $z = 1$  above  $M_* > 8 \times 10^{10} M_{\odot}$ , and 50% completeness at  $z = 2$  above the same mass.

We display in Fig. 8 the resulting selection biases on  $T_{\text{dust}}$ , IR8 and  $R_{\text{SB}}$ . We found that our selection criteria bias our samples toward galaxies with higher SFR at fixed mass, hence to galaxies offset from the main sequence. For example, our sample with  $T_{\text{dust}}$  measurement will only probe galaxies a factor three above the main sequence for masses less than  $10^{11} M_{\odot}$  at  $1 < z < 3$ . Because  $T_{\text{dust}}$  and IR8 both correlate with  $R_{\text{SB}}$  (Elbaz et al. 2011; Magnelli et al. 2014), this translates into a positive bias on both quantities: at a given mass and redshift, the observed average  $T_{\text{dust}}$  and IR8 of our sample are higher than

their true average. This is particularly the case for  $T_{\text{dust}}$ , and is apparent also on the real data set (Fig. 9). Interestingly, although our  $T_{\text{dust}}$  measurements require a detection in the SPIRE bands, which could bias our sample toward colder temperatures, we find that this is a negligible effect compared to the bias toward starbursts: on no occasion is the average measured  $T_{\text{dust}}$  lower than the true average.

This analysis implies that the bias toward higher  $R_{\text{SB}}$  is the dominant source of bias on the measured dust properties. In the following, we focus on the trend of both  $T_{\text{dust}}$  and IR8 with  $R_{\text{SB}}$ , and therefore we will not be affected by this bias. Yet it is important to keep in mind that our results are based almost entirely on



**Fig. 8.** *Left:* bias in the average  $T_{\text{dust}}$  as a function of redshift and mass, computed from the mock catalog and for the “robust  $T_{\text{dust}}$ ” sample. Each cell of the plot shows the difference between the observed average  $T_{\text{dust}}$  in the mock and the true average of the mock in the corresponding bin of redshift and mass. Biases of 1 and 5K are shown with white and yellow lines, respectively. *Center-left:* same plot, but showing instead the bias in “starburstiness” ( $R_{\text{SB}}$ ), or offset from the main sequence. Biases of a factor two and three are shown with white and yellow lines, respectively. *Center-right:* same as *left*, but showing the bias in IR8 for the “robust IR8” sample. Biases of 0.5 and 5 are shown with white and yellow lines, respectively. *Right:* same as *center-left*, but for the “robust IR8” sample.

galaxies with masses larger than  $3 \times 10^{10} M_{\odot}$ ; studying galaxies at lower masses will require next generation instruments such as ALMA and JWST.

#### 4.2.3. Dust temperatures

In Fig. 9 (top left) we show the measured dust temperatures for individual *Herschel* detections, ALESS galaxies, and galaxies from the HRS. These temperatures match broadly the trend observed in the stacked SEDs, but with a tendency to show systematically larger values. As discussed in the previous section, this is to be expected: the requirement of a well-measured IR SED biases this sample towards strongly star-forming starburst galaxies (particularly at high redshifts,  $z > 0.5$ ), which have higher temperatures. Subtracting our redshift-dependent average from the measured  $T_{\text{dust}}$  values (Eq. (15)), we then observed how the residuals correlate with the offset from the main sequence. The result is shown in Fig. 9 (top right). We found a positive correlation and parametrized it with a linear relation (obtained as the bisector of the data):

$$T_{\text{dust}}[\text{K}] = T_{\text{dust}}^{\text{MS}} - (0.77 \pm 0.04) + (10.1 \pm 0.6) \times \log_{10}(R_{\text{SB}}); \quad (18)$$

where  $T_{\text{dust}}^{\text{MS}}$  is defined in Eq. (15). The error bars were determined by bootstrapping. Such a trend was first observed in Elbaz et al. (2011), and later quantified by Magnelli et al. (2014) who stacked galaxies at various locations on the SFR– $M_{*}$  plane (see also Matsuki et al. 2017). They reported a linear relation between  $T_{\text{dust}}$  and  $\log_{10}(R_{\text{SB}})$  – which they call  $\Delta \log(s\text{SFR})$  – with a slope of 6.5 K, which is shallower than the one we measure here but still roughly fits our data.

The residual intrinsic scatter of the temperatures is presented in Fig. 11 as a function of redshift, after subtracting the measurement and redshift uncertainties (assuming  $\Delta z/(1+z) = 3\%$ ; Pannella et al. 2015). The absolute scatter (in Kelvins) increases mildly with redshift, and this evolution is consistent with a constant relative scatter of 12%, relative to  $T_{\text{dust}}^{\text{MS}}$  as given in Eq. (15). The evolution of this scatter beyond  $z = 2$  is poorly constrained, and larger samples with 870 or 1200  $\mu\text{m}$  coverage would be required to determine whether it keeps increasing at higher redshifts. Over the whole sample, the scatter in temperature was

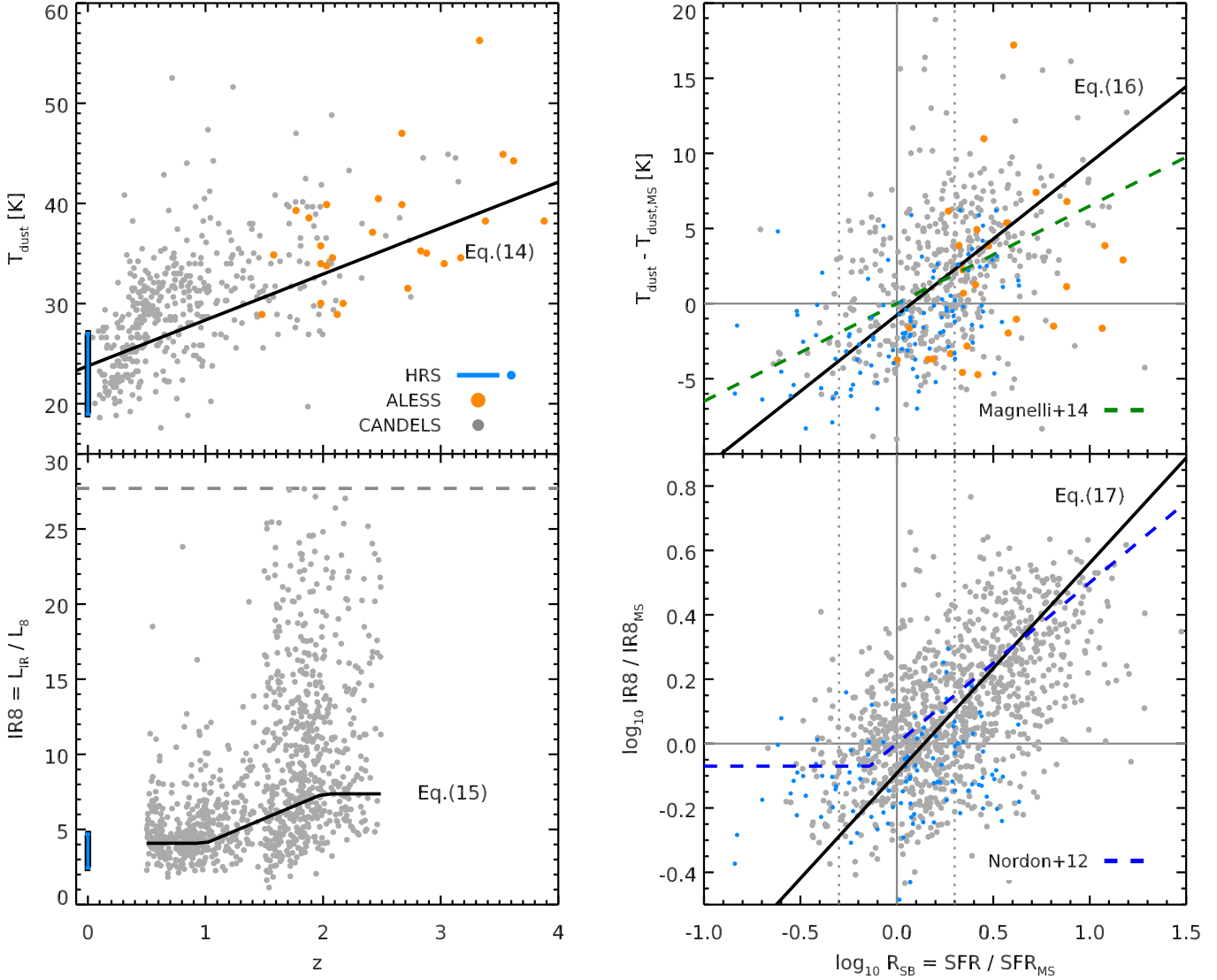
initially 17% (relative to the median). After removing the evolution of the main sequence temperature with redshift, this scatter dropped to 15% (relative to  $T_{\text{dust}}^{\text{MS}}$ ), and removing the starburstiness trend further reduced the scatter to the final value of 12%.

#### 4.2.4. Temperature–luminosity relation

Historically, the first studied correlation was that between the dust temperature and the infrared luminosity (see references in Sects. 1 and 7). Whether this correlation or that involving the redshift and the starburstiness provides the best description of the observations has been repeatedly debated in the literature (see, e.g., Casey et al. 2012; Symeonidis et al. 2013; Magnelli et al. 2014). Unfortunately, owing to the strong selection effects of flux-limited FIR samples, it is generally difficult to de-correlate the effect of redshift and luminosity. As shown in Fig. 10 (left), our sample is affected by this issue: while our galaxies do show a clear correlation between  $T_{\text{dust}}$  and  $L_{\text{IR}}$ , with  $T_{\text{dust}} = 5.57 \times L_{\text{IR}}^{0.0638}$ ,  $L_{\text{IR}}$  also strongly correlates with the redshift. The residual scatter around the  $L_{\text{IR}}-T_{\text{dust}}$  relation is 13%, which is comparable to that found in the previous section; individual galaxies do not favor one description over the other.

At present, only stacking has enough discriminative power to address this question. For example, Magnelli et al. (2014) found, at fixed redshift, a tighter correlation between  $T_{\text{dust}}$  and  $R_{\text{SB}}$  than with  $L_{\text{IR}}$  in their stacks. In our stacks in bins of stellar mass and redshift, we found no evidence for a correlation between  $T_{\text{dust}}$  and  $L_{\text{IR}}$  at fixed redshift; for example at  $z = 1.5$ ,  $T_{\text{dust}}$  remains mostly constant while  $L_{\text{IR}}$  spans almost two orders of magnitude, as shown in Fig. 10 (left). This suggests that there exists no fundamental correlation between  $L_{\text{IR}}$  and  $T_{\text{dust}}$  (at a given redshift) beyond that induced by the starburstiness, which gets averaged-out in our stacks, and selection effects.

To further demonstrate this, we show in Fig. 10 (right) the  $L_{\text{IR}}-T_{\text{dust}}$  relation arising in the mock catalog produced with EGG (see Sect. 4.2.2). This relation is very similar to that observed in the real catalog, albeit with a slightly higher scatter (15%). This  $L_{\text{IR}}-T_{\text{dust}}$  relation was not imposed when generating the mock, instead it emerges naturally as a combination of selection effects and the relations discussed in the previous section.



**Fig. 9.** *Left:* evolution of the dust temperature ( $T_{\text{dust}}$ , *top*) and  $\text{IR8} \equiv L_{\text{IR}}/L_8$  (*bottom*) of galaxies individually detected with *Herschel* in the CANDELS fields (gray dots), from ALESS (orange circles) and the HRS (blue circles or range). We overplot the trends found in stacking (Sect. 4.1) with solid black lines. The dashed horizontal line indicates the maximum IR8 value that our library can reach. *Right:* evolution of  $T_{\text{dust}}$  (*top*) and IR8 (*bottom*) with the starburstiness ( $R_{\text{SB}}$ , see text). The legend is the same as for the plots on the left, except that here the black solid line shows our best-fit relation to the data. For  $T_{\text{dust}}$  we show the relation previously found by Magnelli et al. (2014) with a dashed green line, and for IR8 we show the relation of Nordon et al. (2012) with a dashed blue line.

#### 4.2.5. IR8

We applied the same procedure to IR8, and the results are shown in Fig. 9 (bottom left and right). Consistently with the results of Elbaz et al. (2011) and Nordon et al. (2012), we found a correlation between IR8 and  $\log_{10}(R_{\text{SB}})$ , meaning that starburst galaxies have depressed PAH emission, which Elbaz et al. interpreted as a sign of increased compactness of the star-forming regions. We modeled this dependence with a linear relation (obtained as the bisector of the data):

$$\text{IR8} = \text{IR8}^{\text{MS}} \times (0.81 \pm 0.02) \times R_{\text{SB}}^{0.66 \pm 0.05}. \quad (19)$$

This relation is consistent with that derived by Nordon et al. (2012), although we do not find the need for a separate regime for galaxies below the main sequence. Over the whole sample, the scatter in  $\log_{10}(\text{IR8})$  is 0.28 dex. This is reduced to 0.22 and 0.18 dex after subtracting the redshift and starburstiness dependences, respectively.

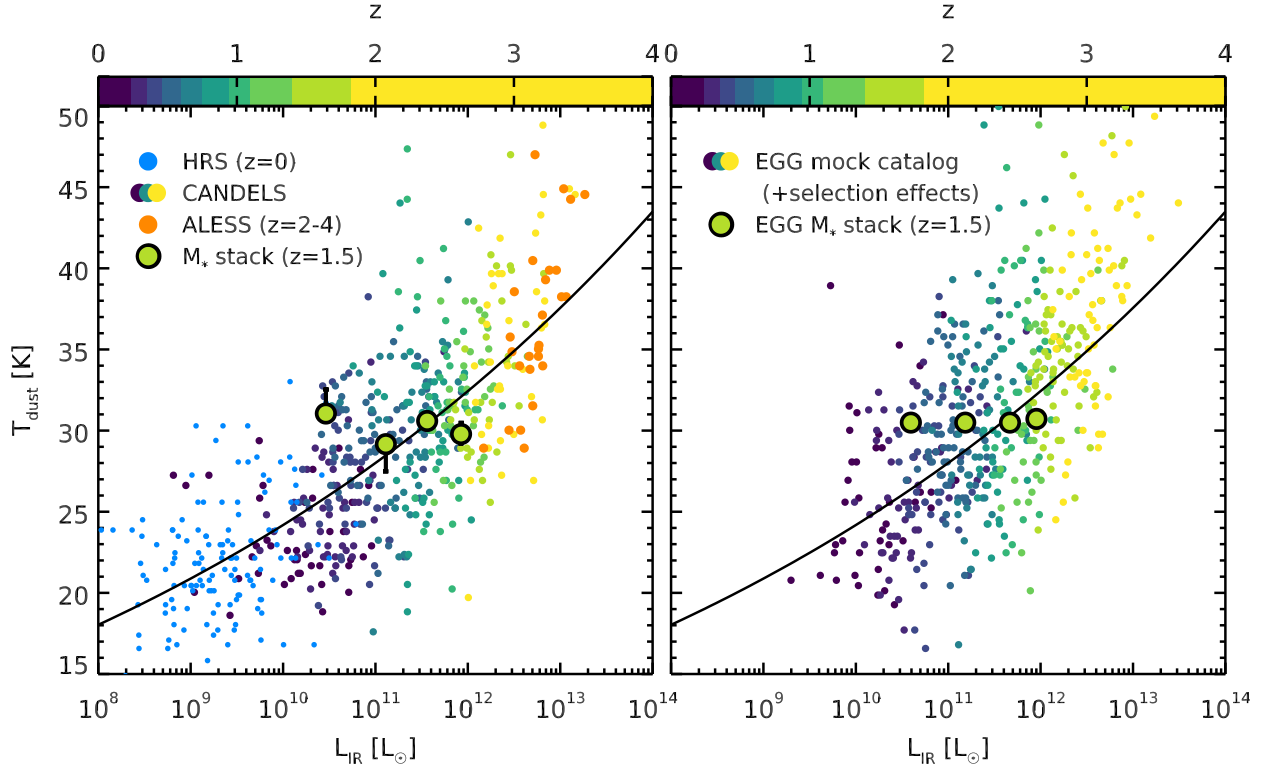
## 5. Optimal $L_{\text{IR}}$ and $M_{\text{dust}}$ measurements

Contrary to the standard FIR libraries from the literature (e.g., CE01), ours has three degrees of freedom: the normalization (either  $L_{\text{IR}}$  or  $M_{\text{dust}}$ ), the dust temperature, and the IR8 (or  $f_{\text{PAH}}$ ). This requires particular care when the observed SEDs are poor and some of these parameters are unconstrained. To obtain a robust determination of  $L_{\text{IR}}$  for galaxies with variable wavelength coverage, the procedure we recommend is described in the following subsections. In the next section we quantify the accuracy of monochromatic  $L_{\text{IR}}$  or  $M_{\text{dust}}$  measurements (i.e., when only a single broadband flux is available for a galaxy) when this procedure is applied, and make predictions for JWST and ALMA.

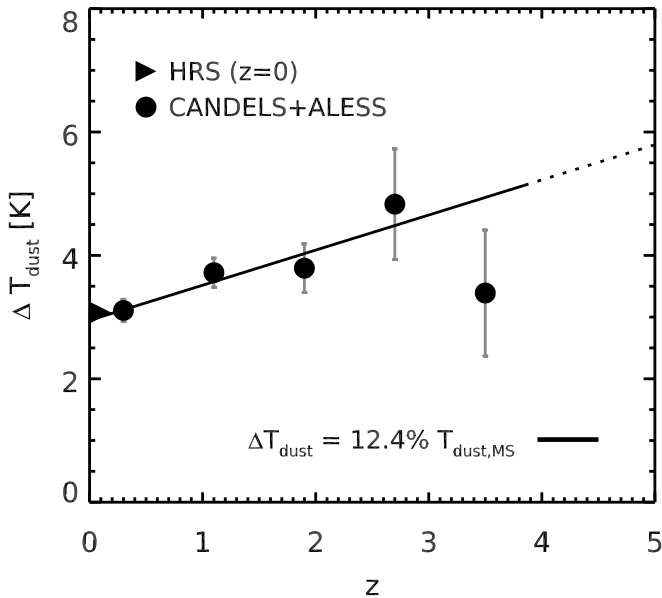
### 5.1. Selection of the free parameters

The dust temperature must be fixed if the peak of the FIR emission is not constrained. In practice, this requires at least





**Fig. 10.** *Left:* relation between the dust temperature ( $T_{\text{dust}}$ ) and the total infrared luminosity ( $L_{\text{IR}}$ ) for galaxies individually detected with *Herschel* in the CANDELS fields (colored circles, dark purple to yellow from low to high redshift), from ALESS (orange circles) and the HRS (light blue circles). We overplot the best-fit power law (see text) with a black solid line, as well as the values obtained at  $z = 1.5$  by stacking galaxies in different bins of mass (see Sect. 4.1). *Right:* same as left, but for the mock CANDELS catalog produced with EGG (defined in Sect. 4.2.2).



**Fig. 11.** Residual intrinsic scatter of  $T_{\text{dust}}$ , obtained after removing the redshift evolution (Eq. (15)) and starburstiness dependence (Eq. (18)), and subtracting statistically the uncertainty on the  $T_{\text{dust}}$  measurements and on the redshift. Measurements of the scatter in CANDELS and ALESS are shown with solid circles, and the HRS is shown with a solid triangle. The trend with redshift is modeled as a constant relative uncertainty (black line).

one measurement at more than  $3\sigma$  on either side of the peak (Hwang et al. 2010), and within the rest-frame  $15\mu\text{m}$  to 3 mm

to avoid contribution from PAH or free-free emission. Here we defined the “peak” by first fitting the galaxy with a free  $T_{\text{dust}}$ , and measuring  $\lambda_{\text{max}}$  from the resulting best-fit template. To fix the temperature, one will use Eq. (15) evaluated at the redshift of the galaxy. The library is then reduced to a single dust continuum and a single PAH template.

Likewise, the IR8 must be fixed if no measurement probes the rest-frame 5 to  $15\mu\text{m}$  (to constrain  $L_8$ ), or if no observation is available for wavelengths greater than  $15\mu\text{m}$  (to constrain  $L_{\text{IR}}$ ). In this case the value of IR8 should be taken from Eq. (16) and evaluated at the redshift of the galaxy. Using Eq. (13), IR8 can then be translated to  $f_{\text{PAH}}$  and fix the relative amplitude of the PAH and continuum templates.

## 5.2. Fitting the observed photometry

For each value of  $T_{\text{dust}}$  allowed in the fit, the templates (provided in rest-frame quantities) must be translated to the observer frame and integrated under the filter response curves of the available photometry. Using a linear solver, one can then fit the convolved templates to the observed photometry as a linear combination of the dust continuum and PAH emission:

$$S_{\nu}^{\text{model}} = M_{\text{dust}}^{\text{cont}} S_{\nu}^{\text{cont}} + M_{\text{dust}}^{\text{PAH}} S_{\nu}^{\text{PAH}}, \quad (20)$$

where  $M_{\text{dust}}^{\text{cont}}$  and  $M_{\text{dust}}^{\text{PAH}}$  are free parameters, and compute the  $\chi^2$ . Among all the templates in the library, one can then pick as the best-fit solution the  $T_{\text{dust}}$  value which produced the smallest  $\chi^2$  (or the only available  $T_{\text{dust}}$  value if it is kept fixed). If the fit is performed with a standard linear solver, the amplitude of either component can become negative. This typically happens when the observed photometry requires an IR8 value larger than what

the library has to offer (which is rare by construction), and the fit uses a PAH component with a negative amplitude to reduce the mid-IR emission. In such cases, one could fix  $f_{\text{PAH}} = 0$  (i.e., no PAH emission) or exclude the rest-frame  $8\ \mu\text{m}$  photometry altogether. Such situations can hint at the presence of obscured AGNs (Donley et al. 2012), or galaxies with strong silicate absorption (Magdis et al. 2011).

If IR8 is fixed, the two components of the library are merged into one single template for each value of  $T_{\text{dust}}$ . Equation (20) becomes

$$S_{\nu}^{\text{model}} = M_{\text{dust}} \left[ (1 - f_{\text{PAH}}) \times S_{\nu}^{\text{cont}} + f_{\text{PAH}} \times S_{\nu}^{\text{PAH}} \right], \quad (21)$$

where  $f_{\text{PAH}}$  is computed from IR8 using Eq. (13), and the only free parameter is  $M_{\text{dust}}$ . Similarly to the procedure above, one can then vary  $T_{\text{dust}}$  and choose as the best-fit solution the value of  $T_{\text{dust}}$  that produced the smallest  $\chi^2$ .

The dust mass is computed as  $M_{\text{dust}} = M_{\text{dust}}^{\text{cont}} + M_{\text{dust}}^{\text{PAH}}$ , while the other parameters ( $L_{\text{IR}}$  and  $L_8$ ) can be obtained from the tabulated values corresponding to the best-fit  $T_{\text{dust}}$ .

### 5.3. Computing uncertainties on best-fit parameters

The most straightforward and secure way to determine uncertainties is to perform a Monte Carlo simulation for each galaxy, where the observed fluxes are randomly perturbed with a Gaussian scatter of amplitude set by the flux uncertainties. This needs to be repeated at least 100 times for reliable  $1\sigma$  error bars. For each parameter, the probability distribution function can be determined from the distribution of best fit values among all random realizations.

## 6. Accuracy of monochromatic measurements

While our library provides three degrees of freedom, in most cases the observed SED of a galaxy will consist of only a couple of data points. The procedure we described in the previous section provides a simplified fit procedure where the number of degrees of freedom is progressively reduced until only one is left: the normalization of the template SED. At fixed  $T_{\text{dust}}$ , this normalization can be translated either into  $L_{\text{IR}}$  or  $M_{\text{dust}}$ , which are usually the quantities observers are after. Often, only one observed data point is available – typically either from *Spitzer* MIPS  $24\ \mu\text{m}$ , ALMA  $870\ \mu\text{m}$ , or  $1.1\ \text{mm}$ , but also in the future from *James Webb*. In this case we dub the measurement of  $L_{\text{IR}}$  or  $M_{\text{dust}}$  as “monochromatic”.

In the case of such monochromatic measurements, since the shape of the SED depends on  $T_{\text{dust}}$  in a strongly non-linear way, the uncertainty on  $T_{\text{dust}}$  will propagate into different uncertainties on  $L_{\text{IR}}$  and  $M_{\text{dust}}$  depending on which band is used to determine the normalization of the template (the case of the many ALMA bands in the millimeter regime is discussed in Sect. 6.2.2). We have shown in Sect. 4.2.3 that fixing  $T_{\text{dust}}$  to the average value for galaxies at a given redshift (Eq. (15)) provides the correct temperature within 15%. In this section, we discuss the uncertainties on  $L_{\text{IR}}$  and  $M_{\text{dust}}$  resulting from this uncertainty on  $T_{\text{dust}}$ , and show which bands are best suited for monochromatic measurements or either quantities. We also discuss how well the rest-frame  $8\ \mu\text{m}$  luminosity can be used to measure  $L_{\text{IR}}$ , which is a situation arising for  $z \sim 1$  to 2 galaxies too faint to be detected with *Herschel* but seen by *Spitzer* MIPS at  $24\ \mu\text{m}$  (or in the future with JWST, which is discussed more thoroughly in Sects. 6.2 and 6.2.1).

### 6.1. Mock catalog

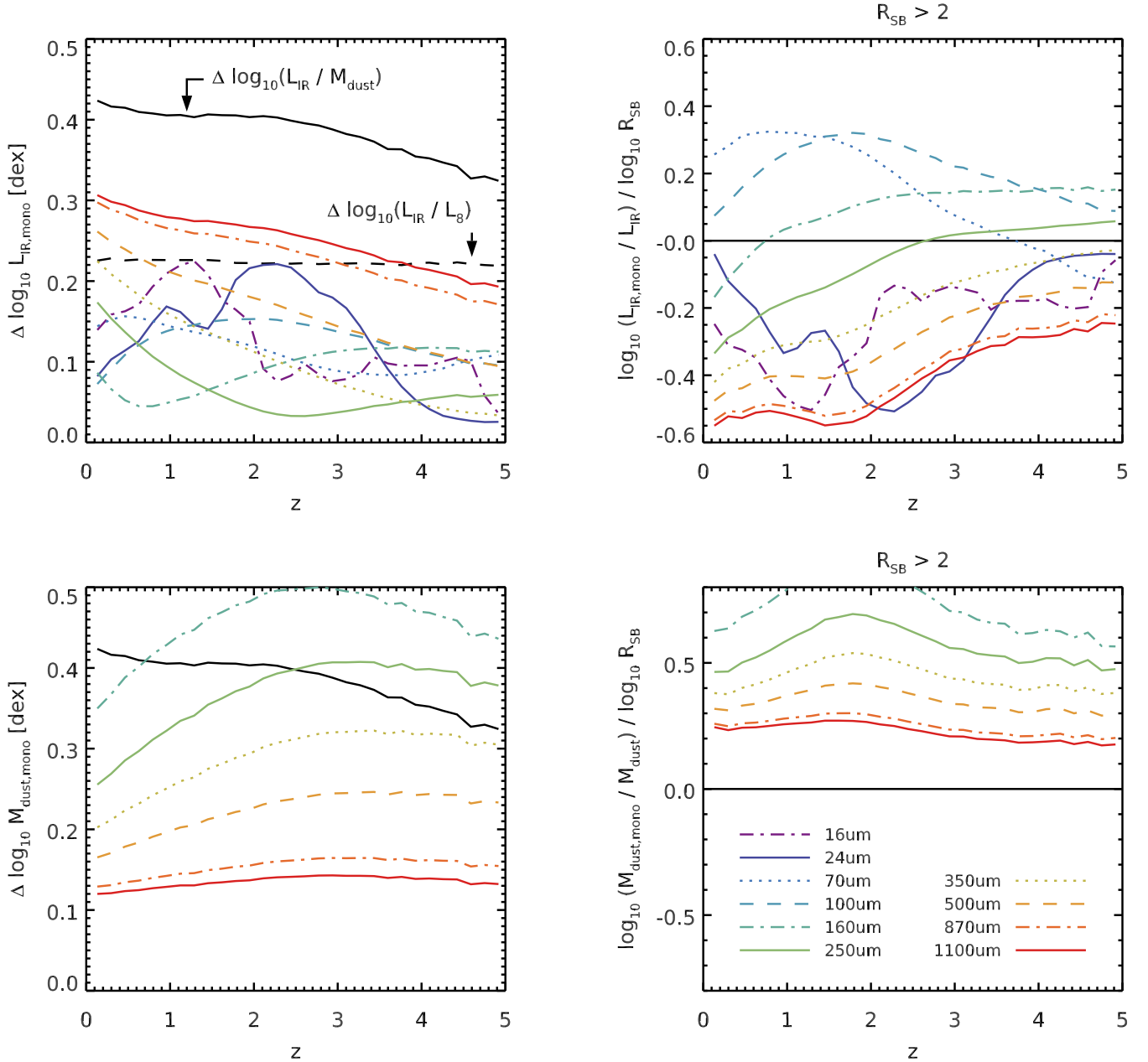
Using a mock catalog built with EGG (see Sect. 4.1.3), we simulated monochromatic measurements and compared them to the true values of  $L_{\text{IR}}$  and  $M_{\text{dust}}$ . We produced a mock catalog spanning  $10\ \text{deg}^2$  and selected all the star-forming galaxies more massive than  $M_* = 10^{10}\ M_{\odot}$ . We then created a flux catalog containing the fluxes of all galaxies in the simulation in all *Spitzer* and *Herschel* bands, as well as JWST MIRI and ALMA. For each galaxy in the mock catalog, we built its “average” expected SED from Eqs. (15) and (16) and used this SED to convert each flux into a value of  $L_{\text{IR}}$  and  $M_{\text{dust}}$ . We note that we did not simulate the stellar continuum for this experiment; we assumed it can be properly constrained and subtracted using the shorter wavelengths. This will be particularly important for the  $16$  and  $24\ \mu\text{m}$  bands at high redshifts. Likewise, the impact of AGNs was also ignored, and these will increase the uncertainty in these same bands at all redshifts (e.g., Mullaney et al. 2011). Therefore the uncertainties we derived should be considered as lower limits.

The resulting uncertainties are shown in Fig. 12 as a function of redshift, for all *Spitzer* and *Herschel* bands, as well as two ALMA bands for illustration (band 7 at  $870\ \mu\text{m}$ , and band 6 at  $1100\ \mu\text{m}$ ); the case of JWST and ALMA are discussed further in Sect. 6.2. Several interesting features come out of this figure and we describe them in the following sections.

#### 6.1.1. Infrared luminosity

When measuring fluxes on the dust continuum, we found  $L_{\text{IR}}$  is best measured when the photometric measurement is close to the peak of the FIR SED: the optimal uncertainty is of the order of 0.05 dex, using  $100\ \mu\text{m}$  at  $z < 0.2$ ,  $160\ \mu\text{m}$  at  $0.2 < z < 1.5$ ,  $250\ \mu\text{m}$  at  $1.5 < z < 3.9$ , and  $350\ \mu\text{m}$  at  $3.9 < z < 5$ . These correspond respectively to rest-frame wavelengths of 90, 86, 68 and  $64\ \mu\text{m}$ , which are precisely the peak wavelengths of the dust SED at each redshift (see also Schreiber et al. 2015, Fig. 9). Rightward of the peak, the uncertainty rises continuously as the rest wavelength increases, since the emission beyond the rest-frame  $250\ \mu\text{m}$  is rather tracing the dust mass (see Sects. 6.1.2, 6.2.2, and Scoville et al. 2014), and fluctuations in the  $L_{\text{IR}}/M_{\text{dust}}$  ratio (displayed in Fig. 12, top and bottom left) are driven by the adopted scatter in  $T_{\text{dust}}$ .

Leftward of the peak, the uncertainties rise significantly up to 0.22 dex when probing the rest-frame  $5$  to  $10\ \mu\text{m}$ , that is, for  $16\ \mu\text{m}$  at  $0.5 < z < 2$  and  $24\ \mu\text{m}$  at  $1.5 < z < 3$ . This, in turn, is caused by variations of IR8 (displayed in Fig. 12, top left), and shows that fluxes in this wavelength domain should be interpreted carefully. Outside of these ranges dominated by PAH emission, our model suggests that the short wavelengths can be excellent tracers of  $L_{\text{IR}}$ , for example with  $24\ \mu\text{m}$  reaching the same accuracy as  $350\ \mu\text{m}$  at  $z > 3.9$ . This is linked to our assumption of a constant fraction of very small grains. Little data can back up this assumption at present; in the local Universe Chary & Elbaz (2001) reported a scatter of 0.15 dex between  $L_{\text{IR}}$  and the  $12$  or  $16\ \mu\text{m}$  luminosities, which is consistent with the values we obtained with our model and suggest the small grain population does not vary strongly from one galaxy to the next. On the other hand, the hotter  $T_{\text{dust}}$  and reduced PAH emission observed in distant galaxies suggest that small grains could get destroyed more efficiently at this epoch. In the near future, JWST will be able to detect distant galaxies at rest wavelengths less than  $12\ \mu\text{m}$  and check this assumption. Another source of uncertainty will be the subtraction of the stellar continuum which can start to dominate below  $5\ \mu\text{m}$ , although this should be less of

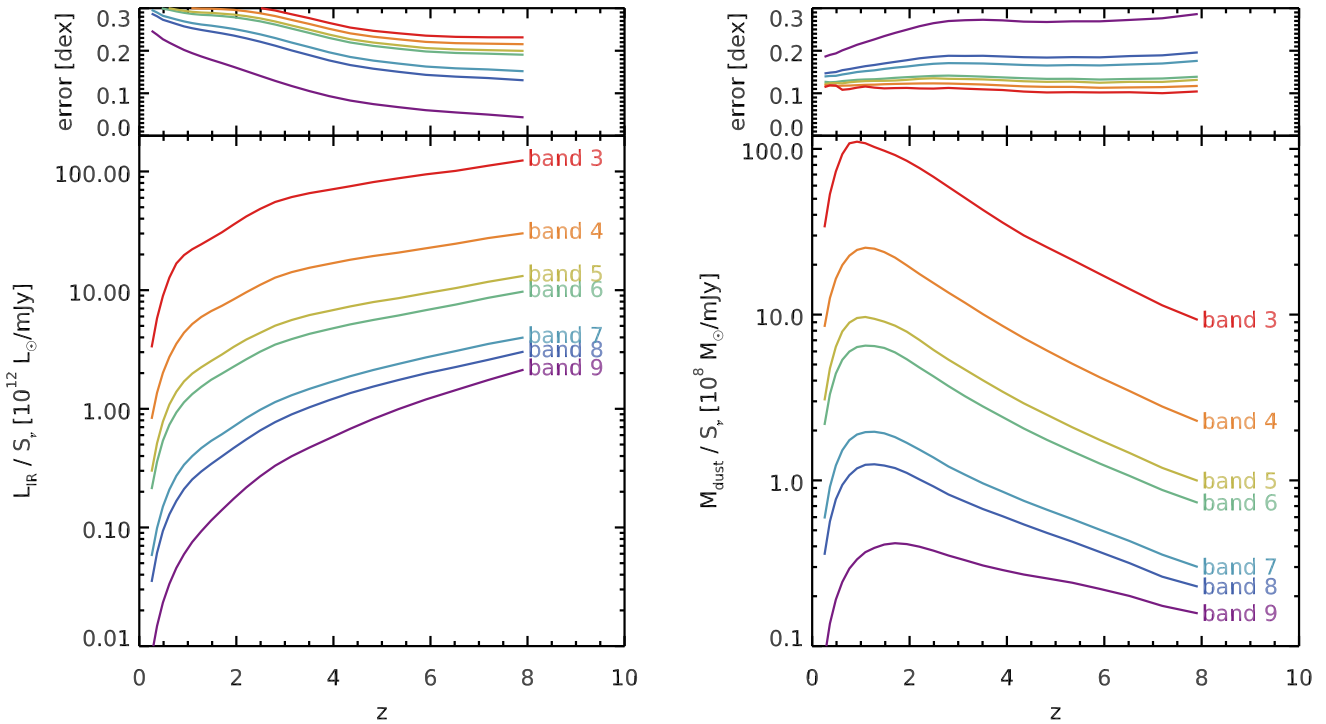


**Fig. 12.** *Top left:* predicted evolution of the uncertainty in  $L_{\text{IR}}^{\text{mono}}$ , that is, the  $L_{\text{IR}}$  inferred from a single broadband photometric measurement, each band corresponding to a different color and line style (see legend). This uncertainty is derived by measuring the standard deviation of the difference between the true  $L_{\text{IR}}$  that was put in the simulated catalog and the observed monochromatic rest-frame luminosity. This is an optimal uncertainty, assuming 1) no error on the measured flux, 2) knowledge of the best average  $L_{\text{IR}}/L_{\nu}$  conversion factor (i.e., the best average SED), 3) perfect subtraction of the stellar component (which only matters for 16 and 24  $\mu\text{m}$  at high-redshift), and 4) no contamination from AGNs. For comparison, we show with a black solid and dashed lines the logarithmic scatter in  $L_{\text{IR}}/M_{\text{dust}}$  and  $L_{\text{IR}}/L_8$  in the sample at each redshift, which are the main drivers of SED variations at  $\lambda > 30$  and  $\lambda < 30 \mu\text{m}$  in our model, respectively. *Top right:* predicted systematic error on the  $L_{\text{IR}}$  of starburst galaxies (selected here with  $R_{\text{SB}} > 2$ ) normalized to the galaxies' offset from the main sequence. In other words, a value of  $x$  on this plot means that the  $L_{\text{IR}}$  will be wrong on average by a factor  $R_{\text{SB}}^x$ . *Bottom left and right:* same as top, but for  $M_{\text{dust}}^{\text{mono}}$ .

a problem at high redshifts where the specific SFRs are higher. Last but not least, we caution that this whole discussion is ignoring AGNs, which seem to be very common at least among massive galaxies at  $z > 3$  (Marsan et al. 2017).

Using a single photometric point, and therefore fixing IR8 and  $T_{\text{dust}}$  to their redshift average, one will be systematically biased for those galaxies which have unusual IR8 or  $T_{\text{dust}}$  values. As shown in Sect. 4.2, this is the case for starburst galaxies. For this reason, we also display in Fig. 12 (top right) the predicted value of this systematic bias for galaxies with  $R_{\text{SB}} > 2$  (i.e., at

least one sigma away from the main sequence). The trend is for measurements leftward of the peak to barely overestimate the  $L_{\text{IR}}$  by no more than  $R_{\text{SB}}^{0.3}$ . In other words, a galaxy which is truly a factor five above the main sequence will be observed instead at a factor eight. On the other hand, measurements rightward of the peak or those dominated by PAH emission can reach systematic underestimations by a factor of  $R_{\text{SB}}^{0.5}$ , so a galaxy a factor five above the MS will be seen at only a factor 2.2. These systematic errors will tend to bring starburst galaxies closer to the main sequence than what they are in reality, to the point where they can



**Fig. 13.** *Left:* conversion factor from observed flux (in mJy) to  $L_{\text{IR}}$  (in units of  $10^{12} L_{\odot}$ ). *Right:* conversion factor from observed flux (in mJy) to a dust mass (in units of  $10^8 M_{\odot}$ ). The *top panels* give the associated relative uncertainty on the conversion. These data are tabulated for easier access in Tables A.1 and A.2, respectively. These factors are provided for all ALMA bands from band 8 to band 3, assuming the standard frequency.

no longer be identified as such. Therefore, SFRs measured only from a single  $24\mu\text{m}$  or ALMA flux at  $z = 2$  should not be used to study starburst galaxies.

### 6.1.2. Dust mass

We present similar figures for  $M_{\text{dust}}$  in Fig. 12 (bottom left and right) for submillimeter bands. The most striking fact to take out of this plot is that no band provides a measurement of  $M_{\text{dust}}$  at better than 0.1 dex (as observed in the local Universe in Groves et al. 2015). This uncertainty rises steadily with redshift for the SPIRE bands, as rest-wavelengths get closer to the peak of the dust emission. This is partly compensated by the increase of the dust temperature with redshift (Eq. (15)), which shifts the peak toward shorter wavelengths. As a consequence we found that the  $870\mu\text{m}$  or  $1.1\text{mm}$  bands suffer roughly equivalent uncertainties, with only a minimal increase with redshift. Scoville et al. (2014) recommended using rest-wavelengths larger than  $250\mu\text{m}$  to measure the dust mass, which should rule out the use of  $870\mu\text{m}$  beyond  $z = 2.5$ . However they have assumed a constant and relatively cold average dust temperature of 25 K: for  $\langle T_{\text{dust}} \rangle = 42\text{K}$  as we observed at  $z = 4$ , the equivalent wavelength limit (in terms of distance to the peak) becomes  $145\mu\text{m}$ , which corresponds to  $725\mu\text{m}$  at  $z = 4$ . This implies that high frequency ALMA bands can still be used to trace the dust mass at high redshifts with a similar accuracy as  $500\mu\text{m}$  at  $z < 1$  (see also Sects. 6.2.2 where this is developed further).

The systematic bias on the dust mass of starburst galaxies is relatively constant with redshift, and most importantly never reaches zero. For these galaxies, we predict a systematic overestimation of the dust masses by a factor of at least  $R_{\text{SB}}^{0.2}$ , so a galaxy a factor five above the main sequence will have its dust

mass overestimated by at least 40%. When the dust mass is converted to a gas mass (see Sect. 6.3), this implies that the gas fraction and depletion time of starbursts will be overestimated by a similar factor.

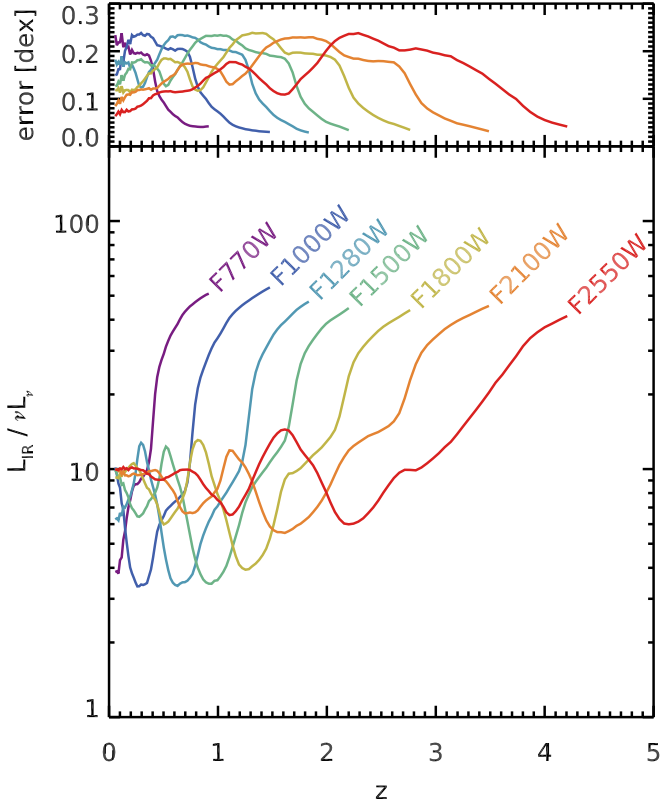
## 6.2. Monochromatic conversion factors for ALMA and JWST

We provide numerical factors to convert an observed ALMA flux into a dust mass and infrared luminosity in Tables A.1 and A.2, respectively, and conversion from a JWST MIRI luminosity into  $L_{\text{IR}}$  in Table A.3. The later is truncated beyond the redshift where the JWST bands probe the rest-frame  $\lambda < 4\mu\text{m}$ , where the stellar continuum starts to dominate the emission. These tables also include the conversion uncertainty (in dex), as shown in Fig. 12 for the *Spitzer* and *Herschel* bands. These data are also displayed in Figs. 13 and 14.

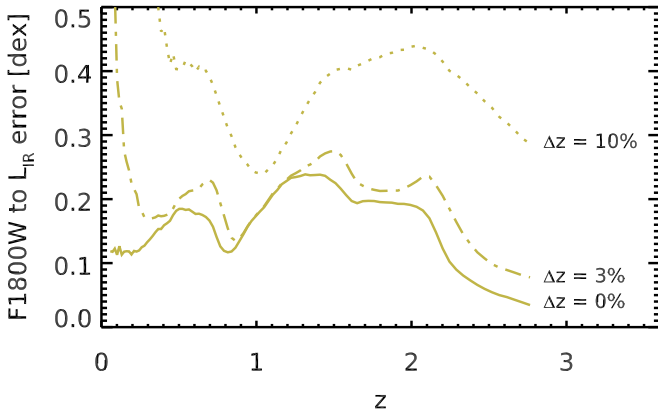
It is clear from Fig. 14 that the accuracy of a JWST band in measuring  $L_{\text{IR}}$  will strongly depend on the redshift, as was already apparent for *Spitzer* 16 and  $24\mu\text{m}$ . Depending on the redshift range of interest, it will therefore be more profitable to observe with one band or the other: for example, using *F2550W* instead of *F1800W* at  $z = 1.5$  can reduce the uncertainty on  $L_{\text{IR}}$  from 0.23 to 0.12 dex. This has to be considered together with the telescope’s sensitivity in each band in determining the optimal observational setup (given, for example, that *F2550W* is expected to be seven time less sensitive than *F1800W*).

### 6.2.1. Impact of redshift uncertainties for JWST

A potentially important consideration related to JWST broadbands is the effect of redshift uncertainties. Because a large fraction of the flux in these bands comes from relatively narrow PAH



**Fig. 14.** Conversion factor from observed monochromatic luminosity to  $L_{\text{IR}}$ . These factors are provided for all JWST MIRI bands, at all redshifts where the band probes the rest-frame  $>3\mu\text{m}$  (i.e., where it may not be dominated by stellar continuum). The *top panel* gives the associated relative uncertainty on the conversion. These data are tabulated for easier access in Table A.3.



**Fig. 15.** Uncertainty on the conversion of an observed JWST F1800W luminosity into  $L_{\text{IR}}$ . This uncertainty is shown with three cases: no uncertainty on the redshift (solid line), 3% uncertainty (dot-dashed line), and 10% uncertainty (dotted line).

emission lines, the conversion factor from flux to  $L_{\text{IR}}$  depends steeply on redshift as lines fall in and out of the bandpass. This can be precisely modeled if the spectroscopic redshift is known, but the case of photometric redshifts requires more care. To explore the impact of photometric redshift uncertainties, we have created two sets of “observed redshifts” for the galaxies in our mock catalog, obtained by randomly perturbing the true redshift with a Gaussian distribution of width  $\Delta z$ . We picked  $\Delta z/(1+z) = 3\%$  and  $10\%$ ;  $3\%$  is a typical (if not a conservative) value of

the uncertainty in deep fields (Muzzin et al. 2013; Skelton et al. 2014; Pannella et al. 2015; Straatman et al. 2016), while  $10\%$  is a more extreme value which applies only to rare outliers.

The result is shown in Fig. 15, and cannot be described as a simple increase in quadrature from the case with no redshift uncertainty. In the case of  $\Delta z/(1+z) = 3\%$ , the impact of the redshift uncertainty is null at  $0.9 < z < 1.2$ , and shows a maximal increase (in quadrature) of  $0.16$  dex at  $z = 0.7, 1.5$  and  $2.2$ . This has the net effect of shifting the domains of best and worst accuracy toward slightly higher redshifts. This can be explained through the so-called “negative  $k$ -correction”: when increasing the redshift, the decrease in flux caused by the larger distance can be compensated, partly or fully, if the intrinsic flux is higher at shorter rest-wavelengths. This happens when the broadband filter falls on the long-wavelength side of a PAH line. But globally, the uncertainty on the conversion is not dramatically increased, and the impact of the redshift uncertainty can be mostly ignored. In contrast, when  $\Delta z/(1+z) = 10\%$  the situation is much worse, with a minimal uncertainty of  $0.25$  dex around  $z = 1$ , and  $0.4$  dex at other redshifts (i.e., essentially not a measurement). This highlights the sensitivity of the PAH region to redshift outliers, and suggests that multiple MIRI bands may be required to properly characterize the galaxies with the most uncertain redshifts (e.g., extremely dusty starbursts).

### 6.2.2. ALMA: dust masses or infrared luminosities?

The dust mass and the infrared luminosity are the two main quantities one can measure from a set of ALMA fluxes; typically to infer gas masses and star formation rates, respectively. However, the conversion from an ALMA flux to either of these quantities depends on other factors, in particular on the dust temperature (see Eq. (8)). The most secure approach would thus be to fit for the dust temperature, and then derive  $L_{\text{IR}}$  and  $M_{\text{dust}}$ . But when only a single flux measurement is available, the question arises: does this photometric point measure  $L_{\text{IR}}$ , or  $M_{\text{dust}}$ ? Because both depend on the dust temperature, the correct answer is “neither”, which is not so helpful. However, it is clear that one quantity will be constrained better than the other depending on the observed frequency (see previous sections and Fig. 12): one expects high frequency bands to better trace  $L_{\text{IR}}$ , and low frequency bands to better trace  $M_{\text{dust}}$ .

We quantified this question in two ways using our mock catalog. We first determined, for each band, the redshift ranges in which  $L_{\text{IR}}$  or  $M_{\text{dust}}$  are measured at better than  $0.2$  dex. We found that  $M_{\text{dust}}$  is measured at better than  $0.2$  dex at all redshifts, and for all bands except band 9 ( $440\mu\text{m}$ , which is always worse than  $0.2$  dex). In contrast, we found that  $L_{\text{IR}}$  can only be measured at this level of accuracy in band 9, 8, 7 and 6 and at  $z > 1, 3.2, 3.8$  and  $5.7$ , respectively. Therefore there are domains where either  $L_{\text{IR}}$  or  $M_{\text{dust}}$  can be measured at better than  $0.2$  dex from the same ALMA measurement. We note however that only one of the two quantities can really be “measured”, the other is then fully determined by the assumed  $T_{\text{dust}}$ .

Alternatively, for each redshift and band, we determined which of  $L_{\text{IR}}$  or  $M_{\text{dust}}$  is constrained with the lowest uncertainty. The answer is always  $M_{\text{dust}}$ , except for band 9, 8 and 7 at  $z > 0.8, 3.6$  and  $5.7$ , respectively, where  $L_{\text{IR}}$  takes over. Therefore, while band 7 ( $870\mu\text{m}$ ) can be used starting from  $z = 3.8$  to measure  $L_{\text{IR}}$ , it is only at  $z > 5.7$  that it traces  $L_{\text{IR}}$  better than it traces  $M_{\text{dust}}$ .

These results depend strongly on the evolution of the average dust temperature with redshift. If it had remained constant at the  $z = 0$  value, high frequency ALMA bands would have traced

$L_{\text{IR}}$  better at lower redshifts. This highlights that a proper knowledge of the dust temperature is crucial to interpret sub-millimeter fluxes, especially in cases where a single band is used at multiple redshifts to determine, for example, luminosity functions or cosmic SFR densities, or to build scaling laws.

### 6.3. Gas masses from dust masses

While the dust mass in itself is only of moderate interest, it is often used as a proxy for determining gas masses, though the assumption of a gas-to-dust ratio. In Schreiber et al. (2016), we have derived gas-to-dust ratios appropriate for our library: since systematic uncertainties on the dust masses are significant (see Sect. 3.1.4), it is crucial to use a consistent calibration of gas-to-dust ratios, derived using the same dust model. We obtained

$$\frac{M_{\text{gas}}}{M_{\text{dust}}} = (155 \pm 23) \times \frac{Z_{\odot}}{Z}, \quad (22)$$

where  $M_{\text{gas}}$  is the total mass of atomic and molecular hydrogen, including helium, and  $Z$  is the metallicity. This relation assumes that the metallicity is inferred from the oxygen abundance  $12 + \log(\text{O}/\text{H})$ , measured using the Pettini & Pagel (2004) calibration (see Magdis et al. 2012), a solar metallicity of  $Z_{\odot} = 0.0134$  and a solar oxygen abundance of 8.73 (Asplund et al. 2009). The Pettini & Pagel calibration was used by Mannucci et al. (2010) to build the Fundamental Metallicity Relation (FMR), therefore the above formula can be applied directly to metallicities estimated using the FMR.

The uncertainty in the above formula is only statistical. Using galaxies in the HRS with well measured dust masses and independent measurement of gas masses from CO and HI, we determined that gas masses derived from dust masses were accurate at the level of 0.2 dex. We emphasize that this gas-to-dust ratio was empirically calibrated using measured dust masses of nearby galaxies, and therefore the value of this ratio depends entirely on the model used to infer the dust masses. Using the DL07 model, for example, would produce higher dust masses by a factor two, hence the gas-to-dust ratio would have to be reduced by the same factor to remain consistent. Ultimately, the gas masses derived from dust masses of any model should be the same, provided the gas-to-dust ratios are properly calibrated.

## 7. Conclusions

We have introduced a new library of infrared SEDs, publicly available on-line, with three degrees of freedom: the dust mass or infrared luminosity, the dust temperature  $T_{\text{dust}}$ , and the mid-to-total infrared color  $\text{IR}8 \equiv L_{\text{IR}}/L_8$ .

Using this library, we fit stacked SEDs of complete samples of main sequence galaxies in the CANDELS fields and derived the redshift evolution of the average  $T_{\text{dust}}$  and IR8, recovering and extending the trends previously identified in the literature. We observed that the most massive galaxies ( $M_* > 10^{11} M_{\odot}$ ) at  $z < 1$  have a reduced  $T_{\text{dust}}$ , sign of a reduced star formation efficiency, and found that low mass galaxies ( $M_* < 10^{10} M_{\odot}$ ) have an increased IR8, probably because of their lower metallicities. Aside from these two mass domains, we found the infrared SED of the stacked galaxies only depend on the redshift, confirming the existence of a universal dust SED for main sequence galaxies at each epoch of the Universe.

We then used our new library to model galaxies individually detected in the *Herschel* images to determine how  $T_{\text{dust}}$  and IR8 vary for galaxies located above the main sequence, that is, the

starbursts, and measure for the first time the scatter of both quantities. We recovered previous claims of a positive correlation of  $T_{\text{dust}}$  and IR8 with the offset from the main sequence. Both trends hint that the star forming regions in starbursts are more compact than in the typical main sequence galaxy. We observed a low residual intrinsic scatter of 12% in  $T_{\text{dust}}$  and 0.18 dex for IR8, confirming that most of the observed variations of these two parameters are captured by the relations we derived with redshift and offset from the main sequence.

We have implemented these relations and scatters in the Empirical Galaxy Generator (EGG) to predict the accuracy of monochromatic measurements of  $L_{\text{IR}}$  and  $M_{\text{dust}}$ , as provided now by *Spitzer* MIPS 24  $\mu\text{m}$  and ALMA, and in the future with JWST. We found that  $L_{\text{IR}}$  is best measured by wavelengths close to the peak of the dust emission, with a minimal uncertainty of 0.05 dex, while mid-IR bands such as those of JWST have a typical uncertainty of 0.1 to 0.25 dex. The highest frequency bands of ALMA can also be used to determine  $L_{\text{IR}}$ , with an uncertainty of 0.2 dex or less at  $z > 0.9, 3.2, 3.8,$  and  $5.7$  for bands 9, 8, 7, and 6, respectively. Using randomly perturbed redshifts, we found these values to be only moderately increased in case of a typical redshift uncertainty of 3%. When measuring flux leftward of the peak (in wavelength),  $L_{\text{IR}}$  is only barely biased for starburst galaxies, however measurements rightward of the peak or using the rest-frame 8  $\mu\text{m}$  will underestimate the  $L_{\text{IR}}$  of starburst galaxies to the point of artificially bringing them back within the upper envelope of the main sequence.

Using the same mock catalog, we determined that the dust masses are best determined from the longest wavelength bands with an uncertainty of less than 0.15 dex. High-frequency ALMA bands such as band 8 and 7 can also be used with an uncertainty of 0.2 dex, however these bands will also tend to overestimate the dust (and gas) masses of starburst galaxies.

Finally, we tabulated the coefficients to convert observed ALMA fluxes into  $M_{\text{dust}}$  and  $L_{\text{IR}}$  and JWST luminosities into  $L_{\text{IR}}$ , and provided estimates of the uncertainty associated to this conversion.

These results and our library can be used immediately to interpret the many observations in the ALMA archive, which most often consist of a single measurement per galaxy. Furthermore, we expect this will also be most useful for future proposals, either for ALMA or JWST, by providing accurate predictions for the expected flux range of individual galaxies at various epochs.

*Acknowledgements.* The authors want to thank the anonymous referee for his/her comments that improved the consistency and overall quality of this paper. Most of the numerical analysis conducted in this work have been performed using `phy++`, a free and open source C++ library for fast and robust numerical astrophysics (<http://cschreib.github.io/phypp/>). This work is based on observations taken by the CANDELS Multi-Cycle Treasury Program with the NASA/ESA HST, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. This research was supported by the French Agence Nationale de la Recherche (ANR) project ANR-09-BLAN-0224 and by the European Commission through the FP7 SPACE project ASTRODEEP (Ref. No.: 312725).

## References

- Allamandola, L. J., Tielens, A. G. G. M., & Barker, J. R. 1985, *ApJ*, **290**, L25
- Amblard, A., Cooray, A., Serra, P., et al. 2010, *A&A*, **518**, L9
- Armus, L., Charmandaris, V., Bernard-Salas, J., et al. 2007, *ApJ*, **656**, 148
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, **47**, 481
- Barger, A. J., Cowie, L. L., Sanders, D. B., et al. 1998, *Nature*, **394**, 248
- Berta, S., Lutz, D., Santini, P., et al. 2013, *A&A*, **551**, A100
- B  thermin, M., Daddi, E., Magdis, G., et al. 2015, *A&A*, **573**, A113
- Blain, A. W., Smail, I., Ivison, R. J., & Kneib, J. 1999, *MNRAS*, **302**, 632
- Boselli, A., Eales, S., Cortese, L., et al. 2010, *PASP*, **122**, 261

- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012, *ApJ*, 754, 83
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
- Buat, V., Iglesias-Páramo, J., Seibert, M., et al. 2005, *ApJ*, 619, L51
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682
- Casey, C. M., Berta, S., Béthermin, M., et al. 2012, *ApJ*, 761, 140
- Chapin, E. L., Pope, A., Scott, D., et al. 2009, *MNRAS*, 398, 1793
- Chapman, S. C., Helou, G., Lewis, G. F., & Dale, D. A. 2003, *ApJ*, 588, 186
- Chary, R., & Elbaz, D. 2001, *ApJ*, 556, 562
- Ciesla, L., Boselli, A., Smith, M. W. L., et al. 2012, *A&A*, 543, A161
- Ciesla, L., Boquien, M., Boselli, A., et al. 2014, *A&A*, 565, A128
- Compiègne, M., Verstraete, L., Jones, A., et al. 2011, *A&A*, 525, A103
- da Cunha, E., Charlot, S., & Elbaz, D. 2008, *MNRAS*, 388, 1595
- da Cunha, E., Charmandaris, V., Díaz-Santos, T., et al. 2010, *A&A*, 523, A78
- da Cunha, E., Walter, F., Smail, I. R., et al. 2015, *ApJ*, 806, 110
- Daddi, E., Dickinson, M., Morrison, G., et al. 2007, *ApJ*, 670, 156
- Daddi, E., Dannerbauer, H., Stern, D., et al. 2009, *ApJ*, 694, 1517
- Daddi, E., Elbaz, D., Walter, F., et al. 2010, *ApJ*, 714, L118
- 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
- Dekel, A., Zolotov, A., Tweed, D., et al. 2013, *MNRAS*, 435, 999
- Desert, F., Boulanger, F., & Puget, J. L. 1990, *A&A*, 237, 215
- Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, *ApJ*, 748, 142
- Draine, B. T., & Li, A. 2007, *ApJ*, 657, 810
- Draine, B. T., Dale, D. A., Bendo, G., et al. 2007, *ApJ*, 663, 866
- Dunne, L., Eales, S., Edmunds, M., et al. 2000, *MNRAS*, 315, 115
- Elbaz, D., Cesarsky, C. J., Fadda, D., et al. 1999, *A&A*, 351, L37
- Elbaz, D., Cesarsky, C. J., Chantal, P., et al. 2002, *A&A*, 384, 848
- Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, *A&A*, 468, 33
- Elbaz, D., Hwang, H. S., Magnelli, B., et al. 2010, *A&A*, 518, L29
- Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, *A&A*, 533, A119
- Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006, *ApJ*, 646, 107
- Fadda, D., Yan, L., Lagache, G., et al. 2010, *ApJ*, 719, 425
- Fanciullo, L., Guillet, V., Aniano, G., et al. 2015, *A&A*, 580, A136
- Flores, H., Hammer, F., Thuan, T. X., et al. 1999, *ApJ*, 517, 148
- Franceschini, A., Aussel, H., Cesarsky, C. J., Elbaz, D., & Fadda, D. 2001, *A&A*, 378, 1
- Fritz, J., Franceschini, A., & Hatziminaoglou, E. 2006, *MNRAS*, 366, 767
- Galametz, A., Grazian, A., Fontana, A., et al. 2013, *ApJS*, 206, 10
- Galliano, F., Madden, S. C., Jones, A. P., et al. 2003, *A&A*, 407, 159
- Galliano, F., Dwek, E., & Chantal, P. 2008, *ApJ*, 672, 214
- Galliano, F., Hony, S., Bernard, J., et al. 2011, *A&A*, 536, A88
- Genzel, R., Tacconi, L. J., Lutz, D., et al. 2015, *ApJ*, 800, 20
- Gispert, R., Lagache, G., & Puget, J. L. 2000, *A&A*, 360, 1
- Gobat, R., Daddi, E., Magdis, G., et al. 2017, *Nat. Astron.*, submitted [[arXiv:1703.02207](https://arxiv.org/abs/1703.02207)]
- Goldader, J. D., Meurer, G., Heckman, T. M., et al. 2002, *ApJ*, 568, 651
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, 197, 35
- Groves, B. A., Schinnerer, E., Leroy, A., et al. 2015, *ApJ*, 799, 96
- Gruppioni, C., Pozzi, F., Andreani, P., et al. 2010, *A&A*, 518, L27
- Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, 207, 24
- Helou, G., Lu, N. Y., Werner, M. W., Malhotra, S., & Silbermann, N. 2000, *ApJ*, 532, L21
- Hodge, J. A., Karim, A., Smail, I., et al. 2013, *ApJ*, 768, 91
- Huang, J., Ashby, M. L. N., Barmby, P., et al. 2007, *ApJ*, 664, 840
- Hughes, D. H., Serjeant, S., Dunlop, J., et al. 1998, *Nature*, 394, 241
- Hwang, H. S., Elbaz, D., Magdis, G., et al. 2010, *MNRAS*, 409, 75
- Jones, A. P., Fanciullo, L., Köhler, M., et al. 2013, *A&A*, 558, A62
- Kennicutt, R. C. J. 1998, *ARA&A*, 36, 189
- Kewley, L. J., & Ellison, S. L. 2008, *ApJ*, 681, 1183
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36
- Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, 700, 221
- Lagache, G., Abergel, A., Boulanger, F., Désert, F. X., & Puget, J. 1999, *A&A*, 344, 322
- Le Floc'h, E., Papovich, C., Dole, H., et al. 2005, *ApJ*, 632, 169
- Leger, A., & Puget, J. L. 1984, *A&A*, 137, L5
- Lequeux, J., Peimbert, M., Rayo, J. F., Serrano, A., & Torres-Peimbert, S. 1979, *A&A*, 80, 155
- Li, A., & Draine, B. T. 2001, *ApJ*, 554, 778
- Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, *A&A*, 532, A90
- Madau, P., & Dickinson, M. 2014, *ARA&A*, 52, 415
- Madden, S. C., Galliano, F., Jones, A. P., & Sauvage, M. 2006, *A&A*, 446, 877
- Magdis, G. E., Elbaz, D., Hwang, H. S., et al. 2010, *MNRAS*, 409, 22
- Magdis, G. E., Elbaz, D., Dickinson, M., et al. 2011, *A&A*, 534, A15
- Magdis, G. E., Daddi, E., Béthermin, M., et al. 2012, *ApJ*, 760, 6
- Magnelli, B., Elbaz, D., Chary, R. R., et al. 2009, *A&A*, 496, 57
- Magnelli, B., Elbaz, D., Chary, R. R., et al. 2011, *A&A*, 528, A35
- Magnelli, B., Popesso, P., Berta, S., et al. 2013, *A&A*, 553, A132
- Magnelli, B., Lutz, D., Saintonge, A., et al. 2014, *A&A*, 561, A86
- Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, *MNRAS*, 408, 2115
- Marsan, Z. C., Marchesini, D., Brammer, G. B., et al. 2017, *ApJ*, 842, 21
- Mathis, J. S., Mezger, P. G., & Panagia, N. 1983, *A&A*, 128, 212
- Matsuki, Y., Koyama, Y., Nakagawa, T., & Takita, S. 2017, *MNRAS*, 466, 2517
- Mullaney, J. R., Alexander, D. M., Goulding, A. D., & Hickox, R. C. 2011, *MNRAS*, 414, 1082
- Murata, K., Matsuhara, H., Inami, H., et al. 2014, *A&A*, 566, A136
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJS*, 206, 8
- Nayyeri, H., Hemmati, S., Mobasher, B., et al. 2017, *ApJS*, 228, 7
- Noeske, K. G., Faber, S. M., Weiner, B. J., et al. 2007, *ApJ*, 660, L47
- Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, *A&A*, 507, 1793
- Nordon, R., Lutz, D., Genzel, R., et al. 2012, *ApJ*, 745, 182
- O'Halloran, B., Satyapal, S., & Dudik, R. P. 2006, *ApJ*, 641, 795
- Oliver, S. J., Wang, L., Smith, A. J., et al. 2010, *A&A*, 518, L21
- Pannella, M., Carilli, C. L., Daddi, E., et al. 2009, *ApJ*, 698, L116
- Pannella, M., Elbaz, D., Daddi, E., et al. 2015, *ApJ*, 807, 141
- Papovich, C., Dole, H., Egami, E., et al. 2004, *ApJS*, 154, 70
- Papovich, C., Rudnick, G., Le Floc'h, E., et al. 2007, *ApJ*, 668, 45
- Pettini, M., & Pagel, B. E. J. 2004, *MNRAS*, 348, L59
- Planck Collaboration XXIX. 2016, *A&A*, 586, A132
- Pope, A., Chary, R., Alexander, D. M., et al. 2008, *ApJ*, 675, 1171
- Reddy, N., Dickinson, M., Elbaz, D., et al. 2012, *ApJ*, 744, 154
- Rémy-Ruyer, A., Madden, S. C., Galliano, F., et al. 2015, *A&A*, 582, A121
- Riechers, D. A., Bradford, C. M., Clements, D. L., et al. 2013, *Nature*, 496, 329
- Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, *ApJ*, 739, L40
- Rodighiero, G., Renzini, A., Daddi, E., et al. 2014, *MNRAS*, 443, 19
- Roehly, Y., Burgarella, D., Buat, V., et al. 2014, in *Astronomical Data Analysis Software and Systems XXIII, ASP Conf. Ser.*, 347, 485
- Rujopakarn, W., Rieke, G. H., Weiner, B. J., et al. 2013, *ApJ*, 767, 73
- Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, *ApJ*, 799, 183
- Salpeter, E. E. 1955, *ApJ*, 121, 161
- Sandstrom, K. M., Bolatto, A. D., Bot, C., et al. 2012, *ApJ*, 744, 20
- Santini, P., Maiolino, R., Magnelli, B., et al. 2014, *A&A*, 562, A30
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, *A&A*, 575, A74
- Schreiber, C., Elbaz, D., Pannella, M., et al. 2016, *A&A*, 589, A35
- Schreiber, C., Elbaz, D., Pannella, M., et al. 2017a, *A&A*, 602, A96
- Schreiber, C., Pannella, M., Leiton, R., et al. 2017b, *A&A*, 599, A134
- Scoville, N., Aussel, H., Sheth, K., et al. 2014, *ApJ*, 783, 84
- Shipley, H. V., Papovich, C., Rieke, G. H., Brown, M. J. I., & Moustakas, J. 2016, *ApJ*, 818, 60
- Shivaei, I., Reddy, N. A., Shapley, A. E., et al. 2017, *ApJ*, 837, 157
- Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, *ApJS*, 214, 24
- Sklia, P., Zamojski, M., Schaefer, D., et al. 2014, *A&A*, 561, A149
- Smail, I., Ivison, R. J., & Blain, A. W. 1997, *ApJ*, 490, L5
- Smith, J. D. T., Draine, B. T., Dale, D. A., et al. 2007, *ApJ*, 656, 770
- Soifer, B. T., Sanders, D. B., Madore, B. F., et al. 1987, *ApJ*, 320, 238
- Soifer, B. T., Bohmer, L., Neugebauer, G., & Sanders, D. B. 1989, *AJ*, 98, 766
- Straatman, C. M. S., Spitler, L. R., Quadri, R. F., et al. 2016, *ApJ*, 830, 51
- Swinbank, A. M., Simpson, J. M., Smail, I., et al. 2014, *MNRAS*, 438, 1267
- Symeonidis, M., Page, M. J., Seymour, N., et al. 2009, *MNRAS*, 397, 1728
- Symeonidis, M., Vaccari, M., Berta, S., et al. 2013, *MNRAS*, 431, 2317
- Tacchella, S., Carollo, C. M., Renzini, A., et al. 2015, *Science*, 348, 314
- Tacconi, L. J., Neri, R., Genzel, R., et al. 2013, *ApJ*, 768, 74
- Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2017, *ApJ*, submitted [[arXiv:1702.01140](https://arxiv.org/abs/1702.01140)]
- Teplitz, H. I., Chary, R., Elbaz, D., et al. 2011, *AJ*, 141, 1
- Tielens, A. G. G. M., Meixner, M. M., van der Werf, P. P., et al. 1993, *Science*, 262, 86
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, 613, 898
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, *ApJ*, 754, L29
- Whitaker, K. E., Franx, M., Leja, J., et al. 2014, *ApJ*, 795, 104
- Whitaker, K. E., Franx, M., Bezanson, R., et al. 2015, *ApJ*, 811, L12
- Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, *ApJ*, 691, 1879
- Wu, Y., Charmandaris, V., Hao, L., et al. 2006, *ApJ*, 639, 157
- Zubko, V., Dwek, E., & Arendt, R. G. 2004, *ApJS*, 152, 211

**Appendix A: Tabulated conversion factors for  $L_{\text{ir}}$  and  $M_{\text{dust}}$** **Table A.1.** Conversion from flux to infrared luminosity (in  $10^{12} L_{\odot}/\text{mJy}$ ) and uncertainty (in dex).

$z$	Band 9 678 GHz	Band 8 404 GHz	Band 7 343 GHz	Band 6 229 GHz	Band 5 202 GHz	Band 4 149 GHz	Band 3 96.3 GHz
0.25	0.008 (0.25)	0.035 (0.29)	0.057 (0.30)	0.210 (0.31)	0.295 (0.31)	0.821 (0.32)	3.290 (0.33)
0.37	0.015 (0.24)	0.061 (0.28)	0.099 (0.29)	0.358 (0.31)	0.512 (0.31)	1.364 (0.32)	5.797 (0.32)
0.49	0.023 (0.23)	0.094 (0.27)	0.150 (0.28)	0.539 (0.30)	0.780 (0.30)	2.009 (0.31)	8.936 (0.32)
0.62	0.033 (0.22)	0.129 (0.27)	0.209 (0.28)	0.734 (0.30)	1.087 (0.30)	2.735 (0.31)	12.75 (0.31)
0.76	0.046 (0.21)	0.167 (0.26)	0.272 (0.27)	0.931 (0.29)	1.392 (0.30)	3.530 (0.30)	16.79 (0.31)
0.92	0.059 (0.20)	0.211 (0.26)	0.337 (0.27)	1.136 (0.29)	1.701 (0.29)	4.379 (0.30)	19.72 (0.31)
1.09	0.075 (0.19)	0.254 (0.25)	0.399 (0.26)	1.330 (0.29)	1.978 (0.29)	5.182 (0.30)	22.12 (0.31)
1.28	0.093 (0.19)	0.297 (0.25)	0.466 (0.26)	1.532 (0.28)	2.238 (0.29)	5.937 (0.30)	24.36 (0.31)
1.48	0.115 (0.18)	0.343 (0.25)	0.536 (0.26)	1.757 (0.28)	2.530 (0.29)	6.657 (0.30)	27.21 (0.31)
1.70	0.140 (0.17)	0.397 (0.24)	0.610 (0.26)	1.979 (0.28)	2.865 (0.29)	7.378 (0.30)	30.70 (0.31)
1.94	0.174 (0.16)	0.466 (0.24)	0.708 (0.25)	2.261 (0.28)	3.319 (0.29)	8.346 (0.30)	35.68 (0.31)
2.20	0.217 (0.15)	0.552 (0.23)	0.836 (0.25)	2.615 (0.27)	3.832 (0.28)	9.595 (0.29)	41.80 (0.30)
2.49	0.270 (0.14)	0.656 (0.22)	0.980 (0.24)	3.026 (0.27)	4.367 (0.28)	11.06 (0.29)	48.33 (0.30)
2.80	0.332 (0.13)	0.772 (0.21)	1.140 (0.23)	3.477 (0.26)	5.018 (0.27)	12.74 (0.28)	55.34 (0.29)
3.14	0.398 (0.12)	0.898 (0.20)	1.302 (0.22)	3.867 (0.25)	5.604 (0.26)	14.18 (0.27)	60.88 (0.29)
3.51	0.472 (0.10)	1.031 (0.19)	1.473 (0.21)	4.284 (0.24)	6.164 (0.25)	15.45 (0.26)	65.69 (0.28)
3.91	0.561 (0.09)	1.181 (0.18)	1.662 (0.20)	4.687 (0.23)	6.674 (0.24)	16.64 (0.25)	69.96 (0.27)
4.35	0.678 (0.08)	1.356 (0.17)	1.881 (0.19)	5.141 (0.22)	7.290 (0.23)	18.05 (0.24)	75.17 (0.25)
4.82	0.822 (0.07)	1.540 (0.16)	2.125 (0.18)	5.608 (0.21)	7.931 (0.22)	19.39 (0.23)	81.47 (0.25)
5.34	0.994 (0.07)	1.748 (0.15)	2.389 (0.17)	6.117 (0.20)	8.553 (0.21)	20.70 (0.23)	87.66 (0.24)
5.91	1.199 (0.06)	1.989 (0.14)	2.708 (0.16)	6.781 (0.20)	9.419 (0.21)	22.53 (0.22)	94.57 (0.24)
6.52	1.438 (0.05)	2.249 (0.14)	3.065 (0.16)	7.553 (0.19)	10.40 (0.20)	24.61 (0.22)	101.3 (0.23)
7.19	1.748 (0.05)	2.587 (0.14)	3.519 (0.16)	8.614 (0.19)	11.77 (0.20)	27.49 (0.22)	111.9 (0.23)
7.92	2.135 (0.04)	3.032 (0.13)	3.990 (0.15)	9.739 (0.19)	13.20 (0.20)	30.18 (0.22)	124.3 (0.23)

**Table A.2.** Conversion from flux to dust mass (in  $10^8 M_{\odot}/\text{mJy}$ ) and uncertainty (in dex).

$z$	Band 9 678 GHz	Band 8 404 GHz	Band 7 343 GHz	Band 6 229 GHz	Band 5 202 GHz	Band 4 149 GHz	Band 3 96.3 GHz
0.25	0.085 (0.19)	0.355 (0.15)	0.588 (0.14)	2.157 (0.13)	3.036 (0.12)	8.430 (0.12)	33.62 (0.11)
0.37	0.137 (0.19)	0.566 (0.15)	0.914 (0.14)	3.311 (0.13)	4.741 (0.12)	12.61 (0.12)	53.35 (0.12)
0.49	0.191 (0.19)	0.768 (0.15)	1.230 (0.14)	4.414 (0.13)	6.398 (0.12)	16.47 (0.12)	73.09 (0.12)
0.62	0.244 (0.20)	0.939 (0.15)	1.520 (0.14)	5.336 (0.13)	7.918 (0.12)	19.90 (0.12)	92.96 (0.11)
0.76	0.294 (0.21)	1.075 (0.16)	1.749 (0.15)	5.987 (0.13)	8.951 (0.13)	22.69 (0.12)	108.0 (0.11)
0.92	0.334 (0.21)	1.182 (0.16)	1.891 (0.15)	6.371 (0.13)	9.545 (0.13)	24.56 (0.12)	110.6 (0.11)
1.09	0.369 (0.22)	1.241 (0.16)	1.954 (0.15)	6.504 (0.13)	9.679 (0.13)	25.34 (0.12)	108.2 (0.12)
1.28	0.391 (0.23)	1.250 (0.17)	1.966 (0.15)	6.458 (0.13)	9.430 (0.13)	25.01 (0.12)	102.6 (0.11)
1.48	0.410 (0.23)	1.228 (0.17)	1.922 (0.16)	6.296 (0.13)	9.067 (0.13)	23.85 (0.12)	97.50 (0.11)
1.70	0.418 (0.24)	1.187 (0.17)	1.824 (0.16)	5.921 (0.14)	8.579 (0.13)	22.09 (0.12)	91.84 (0.11)
1.94	0.413 (0.25)	1.109 (0.18)	1.685 (0.16)	5.380 (0.14)	7.893 (0.13)	19.85 (0.12)	84.81 (0.11)
2.20	0.398 (0.26)	1.013 (0.18)	1.533 (0.17)	4.797 (0.14)	7.029 (0.13)	17.59 (0.12)	76.64 (0.11)
2.49	0.377 (0.26)	0.920 (0.19)	1.375 (0.17)	4.247 (0.14)	6.131 (0.14)	15.53 (0.12)	67.83 (0.11)
2.80	0.352 (0.27)	0.822 (0.19)	1.213 (0.17)	3.706 (0.14)	5.347 (0.13)	13.58 (0.12)	58.99 (0.11)
3.14	0.330 (0.27)	0.746 (0.19)	1.083 (0.17)	3.219 (0.14)	4.663 (0.13)	11.80 (0.12)	50.66 (0.11)
3.51	0.307 (0.27)	0.672 (0.19)	0.960 (0.17)	2.792 (0.14)	4.017 (0.13)	10.07 (0.12)	42.83 (0.11)
3.91	0.288 (0.27)	0.608 (0.19)	0.856 (0.17)	2.413 (0.14)	3.434 (0.13)	8.560 (0.12)	35.99 (0.11)
4.35	0.270 (0.27)	0.542 (0.18)	0.753 (0.17)	2.057 (0.14)	2.917 (0.13)	7.217 (0.12)	30.07 (0.10)
4.82	0.256 (0.27)	0.482 (0.18)	0.665 (0.17)	1.756 (0.13)	2.487 (0.13)	6.076 (0.12)	25.53 (0.10)
5.34	0.241 (0.27)	0.427 (0.18)	0.584 (0.17)	1.496 (0.13)	2.091 (0.13)	5.060 (0.11)	21.43 (0.10)
5.91	0.221 (0.27)	0.370 (0.18)	0.504 (0.17)	1.263 (0.13)	1.753 (0.12)	4.191 (0.11)	17.60 (0.10)
6.52	0.201 (0.27)	0.317 (0.19)	0.431 (0.17)	1.064 (0.13)	1.463 (0.13)	3.457 (0.11)	14.27 (0.10)
7.19	0.175 (0.28)	0.262 (0.19)	0.356 (0.17)	0.876 (0.13)	1.190 (0.13)	2.790 (0.11)	11.38 (0.10)
7.92	0.158 (0.29)	0.229 (0.20)	0.300 (0.18)	0.733 (0.14)	0.993 (0.13)	2.274 (0.12)	9.301 (0.10)



**Table A.3.** Conversion from monochromatic luminosity to  $L_{\text{IR}}$  and uncertainty (in dex).

$z$	$F777W$ 7.66 $\mu\text{m}$	$F1000W$ 9.97 $\mu\text{m}$	$F1280W$ 12.8 $\mu\text{m}$	$F1500W$ 15.1 $\mu\text{m}$	$F1800W$ 18.0 $\mu\text{m}$	$F2100W$ 20.8 $\mu\text{m}$	$F2550W$ 25.4 $\mu\text{m}$
0.06	3.849 (0.23)	10.15 (0.15)	6.267 (0.18)	9.897 (0.13)	9.247 (0.12)	9.947 (0.08)	9.938 (0.06)
0.07	3.870 (0.21)	9.508 (0.15)	6.331 (0.17)	9.568 (0.13)	9.245 (0.12)	9.969 (0.09)	9.946 (0.07)
0.09	3.788 (0.22)	8.741 (0.16)	6.213 (0.18)	9.206 (0.14)	9.191 (0.12)	9.822 (0.10)	9.911 (0.07)
0.10	4.293 (0.22)	8.415 (0.16)	6.653 (0.17)	9.256 (0.13)	9.576 (0.11)	9.973 (0.09)	10.05 (0.07)
0.12	4.403 (0.24)	7.280 (0.19)	6.473 (0.19)	8.640 (0.15)	9.316 (0.13)	9.520 (0.10)	9.974 (0.08)
0.13	5.192 (0.21)	6.863 (0.18)	7.059 (0.17)	8.741 (0.14)	9.815 (0.11)	9.734 (0.10)	10.18 (0.07)
0.15	5.610 (0.22)	5.802 (0.20)	6.997 (0.18)	8.233 (0.15)	9.722 (0.12)	9.427 (0.10)	9.985 (0.07)
0.16	6.333 (0.21)	5.074 (0.21)	7.216 (0.17)	8.006 (0.15)	9.944 (0.12)	9.435 (0.11)	10.08 (0.08)
0.18	7.003 (0.21)	4.469 (0.22)	7.527 (0.18)	7.697 (0.16)	10.10 (0.12)	9.349 (0.11)	9.907 (0.07)
0.19	7.470 (0.19)	4.015 (0.21)	7.727 (0.17)	7.223 (0.16)	10.31 (0.11)	9.393 (0.11)	10.09 (0.07)
0.21	8.125 (0.20)	3.799 (0.23)	8.214 (0.17)	6.953 (0.18)	10.48 (0.12)	9.523 (0.11)	10.16 (0.08)
0.23	8.563 (0.20)	3.651 (0.23)	8.744 (0.16)	6.814 (0.18)	10.56 (0.12)	9.594 (0.11)	10.06 (0.08)
0.24	8.684 (0.19)	3.492 (0.23)	9.610 (0.15)	6.636 (0.18)	10.33 (0.12)	9.532 (0.11)	10.06 (0.08)
0.26	8.694 (0.20)	3.364 (0.23)	10.98 (0.14)	6.443 (0.18)	10.08 (0.12)	9.442 (0.12)	10.03 (0.08)
0.28	8.859 (0.20)	3.370 (0.24)	12.36 (0.13)	6.420 (0.18)	9.884 (0.13)	9.463 (0.12)	9.975 (0.08)
0.29	9.153 (0.20)	3.425 (0.24)	12.76 (0.12)	6.510 (0.18)	9.566 (0.13)	9.468 (0.12)	9.910 (0.09)
0.31	9.309 (0.19)	3.408 (0.23)	12.52 (0.13)	6.624 (0.18)	9.136 (0.14)	9.543 (0.12)	9.892 (0.09)
0.33	9.584 (0.19)	3.432 (0.23)	11.79 (0.13)	6.864 (0.18)	8.754 (0.14)	9.611 (0.12)	9.768 (0.09)
0.35	9.891 (0.19)	3.473 (0.24)	10.69 (0.15)	7.084 (0.18)	8.404 (0.15)	9.628 (0.12)	9.571 (0.10)
0.37	10.67 (0.19)	3.700 (0.24)	9.885 (0.16)	7.431 (0.18)	8.283 (0.15)	9.821 (0.12)	9.548 (0.10)
0.39	12.06 (0.18)	3.937 (0.23)	9.077 (0.16)	7.565 (0.17)	7.962 (0.15)	9.809 (0.12)	9.333 (0.10)
0.40	15.62 (0.16)	4.396 (0.23)	8.450 (0.17)	7.779 (0.17)	7.724 (0.16)	9.869 (0.12)	9.214 (0.11)
0.42	20.09 (0.14)	4.994 (0.22)	7.885 (0.18)	8.034 (0.17)	7.530 (0.16)	9.908 (0.12)	9.086 (0.11)
0.44	23.49 (0.12)	5.598 (0.21)	7.199 (0.18)	8.424 (0.16)	7.195 (0.17)	9.838 (0.12)	9.066 (0.11)
0.46	25.81 (0.11)	5.963 (0.21)	6.410 (0.19)	9.179 (0.16)	6.702 (0.17)	9.619 (0.13)	9.046 (0.11)
0.48	27.55 (0.10)	6.221 (0.21)	5.542 (0.20)	10.52 (0.14)	6.199 (0.18)	9.314 (0.13)	9.057 (0.11)
0.50	29.15 (0.09)	6.468 (0.21)	4.832 (0.21)	11.74 (0.13)	5.977 (0.18)	9.067 (0.14)	9.072 (0.12)
0.52	31.14 (0.08)	6.794 (0.21)	4.326 (0.22)	12.32 (0.13)	6.040 (0.18)	8.934 (0.14)	9.240 (0.12)
0.54	32.91 (0.08)	6.964 (0.20)	3.914 (0.22)	12.06 (0.13)	6.138 (0.18)	8.696 (0.14)	9.282 (0.12)
0.56	34.50 (0.07)	7.103 (0.21)	3.628 (0.23)	11.28 (0.14)	6.315 (0.18)	8.446 (0.15)	9.398 (0.12)
0.59	35.85 (0.07)	7.259 (0.20)	3.483 (0.23)	10.47 (0.15)	6.516 (0.18)	8.210 (0.15)	9.518 (0.11)
0.61	37.09 (0.06)	7.403 (0.20)	3.412 (0.23)	9.770 (0.15)	6.724 (0.18)	7.912 (0.15)	9.660 (0.11)
0.63	38.44 (0.06)	7.512 (0.20)	3.380 (0.23)	9.072 (0.16)	6.924 (0.18)	7.423 (0.16)	9.740 (0.11)
0.65	40.17 (0.06)	7.724 (0.21)	3.451 (0.23)	8.486 (0.17)	7.254 (0.18)	7.022 (0.17)	9.896 (0.11)
0.67	41.80 (0.05)	7.902 (0.20)	3.493 (0.23)	7.779 (0.18)	7.592 (0.17)	6.754 (0.17)	9.939 (0.12)
0.70	43.04 (0.05)	8.144 (0.20)	3.500 (0.23)	6.950 (0.19)	8.054 (0.17)	6.632 (0.17)	9.942 (0.12)
0.72	44.20 (0.05)	8.832 (0.20)	3.548 (0.23)	6.147 (0.20)	9.048 (0.16)	6.643 (0.17)	9.950 (0.12)
0.74	45.26 (0.04)	10.59 (0.19)	3.644 (0.23)	5.449 (0.20)	10.63 (0.14)	6.679 (0.17)	9.915 (0.12)
0.77	46.12 (0.04)	14.02 (0.17)	3.772 (0.23)	4.812 (0.21)	12.01 (0.13)	6.645 (0.17)	9.765 (0.12)
0.79	47.07 (0.04)	18.20 (0.14)	4.078 (0.23)	4.339 (0.22)	12.83 (0.12)	6.693 (0.18)	9.637 (0.12)
0.81	47.94 (0.04)	21.21 (0.12)	4.497 (0.22)	3.998 (0.22)	13.10 (0.12)	6.822 (0.17)	9.401 (0.13)
0.84	48.82 (0.04)	23.39 (0.11)	4.988 (0.22)	3.767 (0.23)	12.99 (0.12)	7.058 (0.17)	9.062 (0.13)
0.86	49.63 (0.04)	25.18 (0.11)	5.402 (0.22)	3.590 (0.23)	12.50 (0.12)	7.285 (0.17)	8.688 (0.14)
0.89	50.37 (0.04)	26.93 (0.10)	5.769 (0.21)	3.495 (0.23)	11.67 (0.13)	7.464 (0.17)	8.390 (0.14)
0.92	50.96 (0.04)	28.86 (0.09)	6.138 (0.21)	3.451 (0.23)	10.51 (0.15)	7.615 (0.17)	8.151 (0.15)
0.94	–	30.67 (0.08)	6.449 (0.21)	3.449 (0.23)	9.363 (0.16)	7.756 (0.17)	7.890 (0.15)
0.97	–	32.37 (0.08)	6.851 (0.21)	3.544 (0.23)	8.623 (0.17)	8.059 (0.17)	7.757 (0.15)
0.99	–	33.63 (0.07)	7.072 (0.20)	3.562 (0.23)	7.941 (0.17)	8.323 (0.16)	7.520 (0.16)
1.02	–	35.15 (0.07)	7.403 (0.20)	3.693 (0.23)	7.451 (0.18)	8.986 (0.16)	7.351 (0.16)
1.05	–	36.91 (0.06)	7.713 (0.20)	3.852 (0.23)	6.880 (0.19)	10.05 (0.15)	7.046 (0.17)
1.08	–	38.82 (0.06)	8.033 (0.20)	4.070 (0.23)	6.217 (0.20)	11.20 (0.14)	6.692 (0.17)
1.11	–	40.68 (0.05)	8.432 (0.20)	4.417 (0.23)	5.536 (0.21)	11.90 (0.13)	6.522 (0.18)
1.13	–	42.14 (0.05)	8.845 (0.20)	4.853 (0.22)	4.927 (0.21)	11.85 (0.13)	6.560 (0.18)
1.16	–	43.44 (0.04)	9.297 (0.20)	5.358 (0.22)	4.454 (0.22)	11.39 (0.14)	6.762 (0.18)
1.19	–	44.60 (0.04)	9.808 (0.20)	5.933 (0.22)	4.125 (0.23)	10.94 (0.14)	7.054 (0.17)
1.22	–	45.76 (0.04)	10.70 (0.19)	6.664 (0.21)	3.989 (0.23)	10.62 (0.15)	7.470 (0.17)
1.25	–	46.80 (0.04)	12.50 (0.18)	7.341 (0.21)	3.928 (0.23)	10.15 (0.15)	7.867 (0.17)
1.28	–	47.81 (0.03)	16.26 (0.16)	7.901 (0.20)	3.949 (0.24)	9.554 (0.16)	8.331 (0.16)
1.31	–	48.89 (0.03)	21.01 (0.13)	8.452 (0.20)	4.057 (0.24)	8.912 (0.17)	8.990 (0.16)
1.35	–	49.92 (0.03)	24.54 (0.11)	8.886 (0.20)	4.159 (0.24)	8.112 (0.18)	9.707 (0.15)

Table A.3. continued.

$z$	$F777W$ 7.66 $\mu\text{m}$	$F1000W$ 9.97 $\mu\text{m}$	$F1280W$ 12.8 $\mu\text{m}$	$F1500W$ 15.1 $\mu\text{m}$	$F1800W$ 18.0 $\mu\text{m}$	$F2100W$ 20.8 $\mu\text{m}$	$F2550W$ 25.4 $\mu\text{m}$
1.38	–	50.98 (0.03)	27.08 (0.10)	9.296 (0.20)	4.306 (0.24)	7.359 (0.19)	10.47 (0.14)
1.41	–	52.01 (0.03)	29.06 (0.09)	9.701 (0.20)	4.501 (0.24)	6.730 (0.20)	11.19 (0.14)
1.44	–	53.00 (0.03)	30.87 (0.08)	10.13 (0.19)	4.778 (0.24)	6.214 (0.21)	11.91 (0.13)
1.48	–	54.01 (0.03)	32.84 (0.07)	10.66 (0.19)	5.248 (0.23)	5.865 (0.21)	12.66 (0.12)
1.51	–	–	34.84 (0.07)	11.18 (0.19)	5.946 (0.22)	5.668 (0.22)	13.38 (0.12)
1.54	–	–	36.56 (0.06)	11.69 (0.19)	6.897 (0.22)	5.570 (0.22)	13.93 (0.11)
1.58	–	–	38.07 (0.06)	12.29 (0.19)	8.108 (0.21)	5.564 (0.22)	14.34 (0.11)
1.61	–	–	39.32 (0.05)	13.03 (0.18)	9.081 (0.20)	5.526 (0.23)	14.44 (0.11)
1.65	–	–	40.76 (0.05)	14.81 (0.17)	9.582 (0.19)	5.620 (0.23)	14.31 (0.11)
1.69	–	–	42.31 (0.04)	18.25 (0.16)	9.630 (0.20)	5.734 (0.23)	13.70 (0.12)
1.72	–	–	43.85 (0.04)	22.77 (0.13)	9.822 (0.20)	5.904 (0.23)	12.84 (0.13)
1.76	–	–	45.10 (0.03)	26.51 (0.10)	10.08 (0.20)	6.021 (0.23)	12.00 (0.14)
1.80	–	–	46.28 (0.03)	29.31 (0.09)	10.49 (0.20)	6.208 (0.23)	11.41 (0.15)
1.83	–	–	47.34 (0.03)	31.37 (0.08)	10.92 (0.19)	6.407 (0.23)	10.89 (0.16)
1.87	–	–	–	33.23 (0.07)	11.38 (0.19)	6.700 (0.23)	10.35 (0.17)
1.91	–	–	–	35.04 (0.06)	11.78 (0.19)	7.038 (0.23)	9.644 (0.18)
1.95	–	–	–	36.86 (0.06)	12.22 (0.19)	7.456 (0.22)	8.833 (0.19)
1.99	–	–	–	38.50 (0.05)	12.82 (0.19)	7.995 (0.22)	8.092 (0.20)
2.03	–	–	–	39.78 (0.05)	13.47 (0.19)	8.659 (0.22)	7.342 (0.21)
2.07	–	–	–	40.85 (0.04)	14.38 (0.18)	9.493 (0.21)	6.713 (0.22)
2.12	–	–	–	41.98 (0.04)	16.29 (0.17)	10.40 (0.20)	6.292 (0.23)
2.16	–	–	–	43.21 (0.04)	19.60 (0.15)	11.18 (0.19)	6.032 (0.23)
2.20	–	–	–	44.58 (0.03)	23.68 (0.12)	11.90 (0.19)	5.988 (0.24)
2.24	–	–	–	–	26.92 (0.10)	12.39 (0.18)	6.027 (0.24)
2.29	–	–	–	–	29.31 (0.09)	12.79 (0.18)	6.137 (0.24)
2.33	–	–	–	–	31.12 (0.08)	13.17 (0.18)	6.330 (0.23)
2.38	–	–	–	–	32.80 (0.07)	13.58 (0.18)	6.634 (0.23)
2.43	–	–	–	–	34.44 (0.07)	13.85 (0.18)	6.995 (0.23)
2.47	–	–	–	–	36.05 (0.06)	14.11 (0.18)	7.460 (0.22)
2.52	–	–	–	–	37.51 (0.05)	14.49 (0.18)	8.036 (0.22)
2.57	–	–	–	–	38.74 (0.05)	14.90 (0.18)	8.615 (0.21)
2.62	–	–	–	–	39.87 (0.05)	15.57 (0.18)	9.198 (0.21)
2.66	–	–	–	–	41.08 (0.04)	16.79 (0.17)	9.663 (0.20)
2.71	–	–	–	–	42.44 (0.04)	19.28 (0.15)	9.927 (0.20)
2.76	–	–	–	–	43.74 (0.03)	22.84 (0.13)	9.923 (0.20)
2.82	–	–	–	–	–	26.47 (0.11)	9.887 (0.21)
2.87	–	–	–	–	–	29.45 (0.09)	10.15 (0.20)
2.92	–	–	–	–	–	31.60 (0.08)	10.55 (0.20)
2.97	–	–	–	–	–	33.33 (0.07)	11.03 (0.20)
3.03	–	–	–	–	–	34.99 (0.07)	11.58 (0.19)
3.08	–	–	–	–	–	36.68 (0.06)	12.25 (0.19)
3.14	–	–	–	–	–	38.13 (0.05)	12.88 (0.19)
3.19	–	–	–	–	–	39.47 (0.05)	13.80 (0.18)
3.25	–	–	–	–	–	40.67 (0.05)	14.88 (0.18)
3.31	–	–	–	–	–	41.79 (0.04)	16.04 (0.17)
3.37	–	–	–	–	–	43.00 (0.04)	17.30 (0.16)
3.43	–	–	–	–	–	44.29 (0.04)	18.73 (0.15)
3.49	–	–	–	–	–	45.48 (0.03)	20.35 (0.14)
3.55	–	–	–	–	–	–	22.31 (0.13)
3.61	–	–	–	–	–	–	24.01 (0.12)
3.67	–	–	–	–	–	–	26.16 (0.11)
3.74	–	–	–	–	–	–	28.49 (0.10)
3.80	–	–	–	–	–	–	31.31 (0.08)
3.87	–	–	–	–	–	–	33.72 (0.07)
3.93	–	–	–	–	–	–	35.62 (0.06)
4.00	–	–	–	–	–	–	37.31 (0.06)
4.07	–	–	–	–	–	–	38.60 (0.05)
4.14	–	–	–	–	–	–	40.00 (0.05)
4.21	–	–	–	–	–	–	41.38 (0.04)