Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **489,** 4389–4417 (2019) doi:10.1093/mnras/stz2311
Advance Access publication 2019 August 21

JINGLE – V. Dust properties of nearby galaxies derived from hierarchical Bayesian SED fitting

Isabella Lamperti [1]★ Amélie Saintonge [1] Ilse De Looze,[1,2] Gioacchino Accurso,[1] Christopher J. R. Clark [3] Matthew W. L. Smith [4] Christine D. Wilson [5] Elias Brinks,[6] Toby Brown,[5] Martin Bureau,[7] David L. Clements [8] Stephen Eales,[4] David H. W. Glass,[9] Ho Seong Hwang [10] Jong Chul Lee,[10] Lihwai Lin,[11] Michal J. Michalowski [12] Mark Sargent,[13] Thomas G. Williams,[4] Ting Xiao[14] and Chentao Yang [15,16,17]

Affiliations are listed at the end of the paper

Accepted 2019 August 15. Received 2019 August 1; in original form 2019 May 24

ABSTRACT

We study the dust properties of 192 nearby galaxies from the JINGLE survey using photometric data in the 22–850 μm range. We derive the total dust mass, temperature T, and emissivity index β of the galaxies through the fitting of their spectral energy distribution (SED) using a single modified blackbody model (SMBB). We apply a hierarchical Bayesian approach that reduces the known degeneracy between T and β. Applying the hierarchical approach, the strength of the T–β anticorrelation is reduced from a Pearson correlation coefficient $R = -0.79$ to $R = -0.52$. For the JINGLE galaxies we measure dust temperatures in the range 17−30 K and dust emissivity indices β in the range 0.6−2.2. We compare the SMBB model with the broken emissivity law modified blackbody (BMBB) and the two modified blackbody (TMBB) models. The results derived with the SMBB and TMBB are in good agreement, thus applying the SMBB, which comes with fewer free parameters, does not penalize the measurement of the cold dust properties in the JINGLE sample. We investigate the relation between T and β and other global galaxy properties in the JINGLE and *Herschel* Reference Survey (HRS) sample. We find that β correlates with the stellar mass surface density ($R = 0.62$) and anticorrelates with the H I mass fraction ($M_{\rm H\,I}/M_*$, $R = -0.65$), whereas the dust temperature correlates strongly with the star formation rate normalized by the dust mass ($R = 0.73$). These relations can be used to estimate T and β in galaxies with insufficient photometric data available to measure them directly through SED fitting.

Key words: dust, extinction – galaxies: evolution – galaxies: ISM – submillimetre: ISM.

1 INTRODUCTION

Interstellar dust plays an important role in galaxies: it helps to balance gas heating and cooling and the surface of dust grains provides a favourable place for chemical reactions to occur. Dust contributes only a small fraction of the mass of the interstellar medium (ISM), but in normal star-forming galaxies it can re-radiate up to ∼ 30 per cent of the stellar light in the infrared (e.g. Clements et al. 1996).

The two main places where dust is formed are in the ejecta of core-collapse supernovae and in the envelopes of asymptotic giant branch (AGB) stars (Galliano, Galametz & Jones 2018). These two production mechanisms alone however cannot account for the amount of dust observed in high-redshift galaxies (Bertoldi et al. 2003; Priddey et al. 2003; Rowlands et al. 2014; Michałowski 2015; Watson et al. 2015). Grain growth is another mechanism that can increase the dust content of a galaxy, but it is not well understood how much this process can contribute to the total dust production (Barlow 1978; Ferrara, Viti & Ceccarelli 2016; Ceccarelli et al. 2018). In order to resolve this tension, first we need to improve our understanding of all the mechanisms of dust production and growth. Secondly, it is necessary to have tools to accurately measure the dust content of distant galaxies and have a good understanding of the uncertainties on these measurements; this is the question this paper tackles.

Dust masses are measured by fitting the spectral energy distribution (SED) of galaxies in the far-infrared/sub-millimetre spectral

★ E-mail: isabellalamperti@gmail.com

© 2019 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

range. The standard model used is a modified blackbody function (MBB), which depends on the dust mass, temperature (T), and emissivity index β. An anticorrelation between temperature and β has been observed in galactic sources and luminous infrared galaxies (Dupac et al. 2003; Yang & Phillips 2007; Désert et al. 2008). However, it has been shown that noise in the data can introduce an artificial anticorrelation between T and β (e.g. Shetty et al. 2009a,b). An incorrect estimate of T and β would consequently bias the measurement of the dust mass. A way to overcome this problem and break the $T-\beta$ degeneracy is to use a hierarchical Bayesian approach (Kelly et al. 2012; Juvela et al. 2013; Veneziani et al. 2013; Galliano 2018). The hierarchical approach uses the information from the parameter distribution of the entire sample of galaxies to better constrain temperature and β for each single galaxy. The hierarchical method has the advantage that it does not require knowing the prior distribution of the parameters before the fitting, but can infer the parameters describing the prior directly during the fitting procedure, after assuming the shape of the distribution. The limitation of this is that the prior is only valid for the sample of galaxies under consideration, i.e. the prior depends on the population that one is considering.

The *Herschel Space Observatory*[1] (Pilbratt et al. 2010) has been key for the study of dust in nearby galaxies, providing photometric observations in the wavelength range 100–500 μm, that allowed to characterize the shape of their far-infrared SED. The *Herschel* Reference Survey (HRS; Boselli et al. 2010) is a guaranteed time program that measured the far-infrared SED of ∼300 nearby galaxies. Using HRS galaxies, Cortese et al. (2014) show that their far-infrared and sub-mm colours are inconsistent with a single modified blackbody (SMBB) model with the same emissivity index β for all galaxies.

Dust continuum observations can also be used to infer the molecular gas mass of a galaxy. It has been shown that the dust continuum luminosity of galaxies correlates with the CO luminosity (Hildebrand 1983; Eales et al. 2012; Magdis et al. 2012; Scoville et al. 2014; Groves et al. 2015) and this relation can be used to infer the molecular gas mass of a galaxy by applying a molecular gas-to-dust ratio. This method can be extremely useful for faint or high-redshift galaxies, since the dust emission is brighter and therefore easier to observe than the CO line emission. This method can therefore be beneficial for measuring the molecular gas content of large samples of galaxies.

The JINGLE (JCMT dust and gas In Nearby Galaxies Legacy Exploration), survey is a large program on the James Clerk Maxwell Telescope (JCMT) which aims to characterize the dust and molecular gas in nearby galaxies and study the relation between the two (Saintonge et al. 2018). JINGLE combines dust observations from the SCUBA-2 camera on the JCMT (and from *Herschel*), with the cold gas measurements obtained with the JCMT RxA instrument. With both measurements of the dust and cold gas properties for a statistical sample of nearby galaxies, we can study the variations in the dust-to-gas mass ratio as a function of galaxy and dust properties.

One of the objectives of the survey is to benchmark dust scaling relations with other galaxy properties such as stellar mass, metallicity, and star formation rate (SFR). These relations can be used to estimate the dust temperature and dust emissivity index in galaxies for which there are not enough photometric data available to measure them directly through SED fitting. This can be useful especially for high-redshift galaxies.

An excess of emission at wavelengths ≥500 μm with respect to the MBB model has been observed in numerous dwarf galaxies (e.g. Galametz et al. 2011; Rémy-Ruyer et al. 2013, 2015), in late-type galaxies (Dumke, Krause & Wielebinski 2004; Bendo et al. 2006; Galametz et al. 2009), in the Magellanic Clouds (Bot et al. 2010; Israel et al. 2010), and in M33 (Hermelo et al. 2016; Relaño et al. 2018). The origin of this 'sub-mm' excess is still an open question. The SCUBA-2 observations at 850 μm can help to place better constraints on the sub-mm slope and investigate the presence of this excess in the JINGLE sample.

In this paper we take advantage of the large and homogeneous JINGLE sample and apply a hierarchical Bayesian approach to reduce the $T-\beta$ degeneracy and obtain more accurate measurements of the dust parameters using MBB models. The hierarchical approach is crucial to disentangle dust temperature T and emissivity index β and allows us for the first time to study the independent relations of these two dust quantities with other galaxy global properties.

This paper is organized as follows. In Section 2, we present the sample and the data used in this paper. Then we describe the classical and hierarchical Bayesian SED fitting methods and compare the two methods using simulated SEDs (Section 3). Section 4 illustrates the results of the SED fitting of the JINGLE sample, the T–β relation, and comparison of different MBB models. In Section 5, we derive scaling relations between dust quantities and other global galaxy properties. Finally, in Section 6, we summarize the main results and our conclusions. Readers who are less interested in the statistical methods and tests of the fitting methods may wish to skip ahead to Section 4.

2 SAMPLE AND DATA

2.1 JINGLE sample

The 192 galaxies in the JINGLE sample have stellar masses in the range $\log M_*/\mathrm{M}_\odot = 9-11.3$ and are in the redshift range $0.01 < z < 0.05$. The targets were selected from the H-ATLAS survey (Eales et al. 2010; Maddox et al. 2018) with the requirement to have a detection $\geq 3\sigma$ in the 250 and 350 μm SPIRE bands. Additionally, they have been selected to have a flat logarithmic stellar mass distribution. Due to these requirements, they are mainly main-sequence star-forming galaxies with $-1.5 < \log \mathrm{SFR}/[\mathrm{M}_\odot\ \mathrm{yr}^{-1}] < 1.5$ (see Fig. 1). A detailed description of the selection criteria is provided in Saintonge et al. (2018). Most of the JINGLE objects are late-type galaxies, with only seven classified as early-type galaxies (Saintonge et al. 2018).

Properties of the JINGLE galaxies used in this work (such as SFR, metallicity, distances, ...) are taken from the JINGLE catalogue (Saintonge et al. 2018). In particular, we use the SFRs and stellar masses measured with MAGPHYS (da Cunha, Charlot & Elbaz 2008). In this paper, we refer to JINGLE galaxies using their corresponding JINGLE ID, as described in the JINGLE catalogue (Saintonge et al. 2018).

2.2 HRS sample

To extend our analysis to a larger range in galaxy properties, we include in our analysis also galaxies from the HRS (Boselli et al. 2010). The HRS is a volume-limited sample (15 Mpc ≤ D ≤ 25 Mpc) of 323 galaxies, with flux limits in the K band to minimize

[1] *Herschel* is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

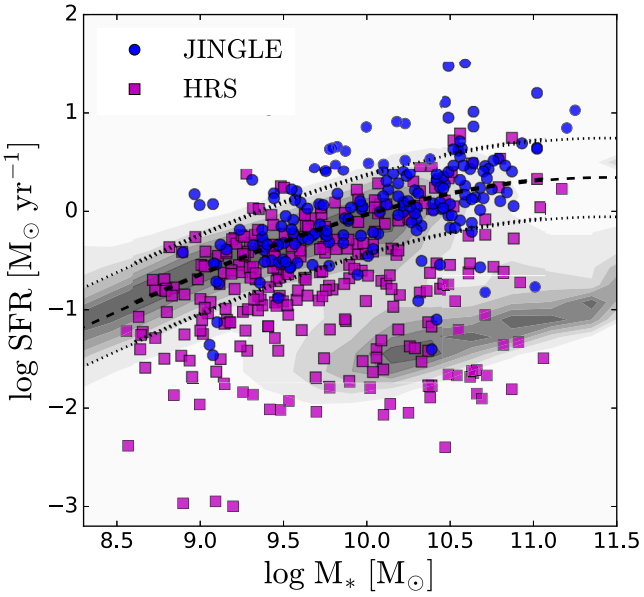

Figure 1. Distribution of the JINGLE and HRS sample in the SFR–M_* plane. The position of the star formation main sequence (Saintonge et al. 2016) is shown as a dashed line, the 0.4 dex dispersion is shown by dotted lines. The grey contours show the distribution of SDSS galaxies at redshift $z < 0.05$.

selection effects due to dust and young high-mass stars. A large fraction of HRS galaxies lie in clusters, with 47 per cent of the HRS galaxies listed in the Virgo Cluster Catalogue alone. They have stellar masses in the range $\log M_*/\mathrm{M}_\odot = 8.4-11.3$. Galaxies from the HRS have been observed in the five *Herschel* bands (at 100, 160, 250, 350, and 500 μm), but do not have observations at 850 μm. In our analysis we use the SFR and stellar masses measured with MAGPHYS by De Vis et al. (2017), to be consistent with the JINGLE measurements.

Fig. 1 shows the JINGLE and HRS galaxies on the SFR–M_* plane. With respect to the JINGLE galaxies, the HRS sample includes galaxies which are less massive ($\log M_* < 9$) and with lower SFR ($-2 < \log$ (SFR/[M$_\odot$ yr^{-1}]) < 0.6, mean $\log$ (SFR/[M$_\odot$ yr^{-1}]) $= -0.71$) compared to JINGLE, which has a mean $\log$ (SFR/[M$_\odot$ yr^{-1}]) $= 0.04$. HRS galaxies are also less dusty than JINGLE targets (De Looze et al., in preparation), since contrary to JINGLE they have not been selected based on detection in the infrared bands. The HRS sample includes also a large number of early-type galaxies (62/323; Smith et al. 2012a), which are not well represented in the JINGLE sample (7/192). Therefore by including this sample in our analysis, we can test whether the dust scaling relations that we find with the JINGLE sample hold also for other types of galaxies. Additionally, increasing the dynamical range of galaxy properties will help to constrain better the dust scaling relations.

2.3 Data

2.3.1 JINGLE

Our data set consists of photometric points at 22 μm (*WISE*), 60 μm (IRAS), 100, 160 μm (*Herschel*/PACS), 250, 350, 500 μm (*Herschel*/SPIRE), and 850 μm (SCUBA-2). A detailed description of the JINGLE photometric data set is given in Smith et al. (2019) and De Looze et al. (in preparation). Here, we summarize the most important points. The fluxes of the *WISE*, *Herschel*, and SCUBA-2 bands have been extracted from matched apertures based on the SPIRE 250 μm band. The flux extraction is described in detail by Smith et al. (2019). One galaxy (JINGLE 62) has been removed from the sample since it is not detected in the 250 μm band and therefore it is not listed in the release version of the *H*-ATLAS DR2 catalogue (Maddox et al. 2018). Thus, the sample analysed in this work consists of 192 galaxies.

We consider upper limits for fluxes with peak signal-to-noise ratio (S/N) < 3. Since the CO(3–2) 345.79 GHz line emits in the 850 μm band, we corrected the SCUBA-2 flux by subtracting the estimated contribution of the CO(3–2) line (for details see Smith et al. 2019). After subtracting the CO(3–2) emission, some of the fluxes become negative, due to the uncertainties in the 850 μm fluxes and in the CO(3–2) predictions. These fluxes are consistent with zero within the uncertainties and are considered as upper limits. In our sample, there are 66 galaxies with peak S/N < 3 and additionally four galaxies have negative 850 μm flux, even though their peak S/N > 3 before subtraction of the CO(3–2) contribution. For all these cases, we use conservative upper limits equal to five times the flux uncertainty in that band.

The IRAS 60 μm fluxes are derived using the Scan Processing and Integration Tool (SCANPI[2]), following the strategy of Sanders et al. (2003). In our sample, 69/192 galaxies have 5σ upper limits for the 60 μm flux and 22/192 do not have IRAS 60 μm observations.

2.3.2 HRS

For the HRS sample, we have flux measurements in the *Herschel*/PACS (Cortese et al. 2014) and *Herschel*/SPIRE bands (Ciesla et al. 2012), from 100 to 500 μm. We note that, contrary to JINGLE, this sample does not have observations at 850 μm, therefore the long-wavelength slope of the SED can be constrained only by the 500 μm point. In the case of non-detections, we consider upper limits equal to five times the flux uncertainties as we do for the JINGLE sample.

We exclude from the sample 39 galaxies which are not detected in all of the *Herschel* bands, and therefore do not have constraints on their dust properties. We also exclude four galaxies that do not have SFR and stellar mass measurements from De Vis et al. (2017). They were excluded from the sample because their SEDs show signs of contamination from dust heated by an active galactic nucleus or a hot X-ray halo or from synchrotron radiation emission (Eales et al. 2017). The final sample consists of 41 early-type and 239 late-type galaxies, for a total of 280 galaxies.

3 METHOD

3.1 Models

To describe the far-infrared and sub-millimetre SED we adopt the three models employed by Gordon et al. (2014) for the SED fit of the Magellanic Clouds: single modified blackbody (SMBB), broken emissivity law modified blackbody (BMBB), and two modified blackbody (TMBB). We describe below the analytic functions and the parameters used for the three models:

(i) **SMBB:** The SMBB model describes the dust emission F_λ (in units of W m^{-2} Hz^{-1} sr^{-1}) at each wavelength λ in the following way (Hildebrand 1983):

$$F_\lambda = \frac{M_{\rm dust}}{D^2}\kappa_\lambda B_\lambda(T), \quad (1)$$

[2]http://irsa.ipac.caltech.edu/applications/Scanpi/

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

where $M_{\rm dust}$ is the dust mass in the galaxy and D is the distance of the galaxy. $B_\lambda(T)$ is the Planck function for the emission of a blackbody with a dust temperature T given by

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\left(\frac{hc}{k_{\rm B}T\lambda}\right) - 1}. \quad (2)$$

The dust mass absorption coefficient κ describes which dust mass gives rise to an observed luminosity. The value of κ depends on the physical properties of the dust, such as the mass density of the constituent materials, the efficiency with which they emit, the grain surface-to-volume ratio, and the grain size distribution (Köhler, Ysard & Jones 2015; Ysard et al. 2018). The SMBB applies a dust emissivity power law to characterize the behaviour of κ as a function of wavelength:

$$\kappa_\lambda = \kappa_0 \left(\frac{\lambda_0}{\lambda}\right)^\beta. \quad (3)$$

where κ_0 is the reference dust mass absorption coefficient. Laboratory studies found that the absorption coefficient depends also on the dust temperature and dust emissivity index β, with higher κ values observed for higher temperatures and lower β values (Coupeaud et al. 2011). For simplicity, here we assume a constant value $\kappa_0 = \kappa(500\,\mu{\rm m}) = 0.051\,{\rm m^2\,kg^{-1}}$ from Clark et al. (2016). This model has three free parameters ($M_{\rm dust}$, T, and β), and assumes that the dust emission can be described by a dust component with a single temperature. At wavelengths shorter than 100 μm, a second warmer dust component can contribute to the FIR emission (e.g. Relaño et al. 2018). Therefore for this model, we use only the flux bands with wavelengths ≥100 μm. Additionally, we use the 60 μm point as an upper limit, in order to better constrain the dust temperature.

(ii) **BMBB:** When fitting the FIR SED with an SMBB model, some galaxies show an excess in the flux at wavelengths ≥500 μm, called 'sub-millimetre' excess (Lisenfeld et al. 2002; Galliano et al. 2003; Dumke et al. 2004; Bendo et al. 2006; Galametz et al. 2009; Bot et al. 2010; Israel et al. 2010; Hermelo et al. 2016). The BMBB model assumes that the sub-mm excess is due to variations in the wavelength dependence of the dust emissivity law. These variations are parametrized by a broken power law:

$$\kappa_\lambda = \begin{cases} \kappa_0 \left(\frac{\lambda_0}{\lambda}\right)^{\beta_1} & \text{if } \lambda < \lambda_{\rm b} \\ \kappa_0 \left(\frac{\lambda_0}{\lambda_{\rm b}}\right)^{\beta_1} \left(\frac{\lambda_{\rm b}}{\lambda}\right)^{\beta_2} & \text{if } \lambda > \lambda_{\rm b} \end{cases}, \quad (4)$$

where $\lambda_{\rm b}$ is the wavelength of the break. This model has five free parameters: $M_{\rm dust}$, T, β_1, β_2, and $\lambda_{\rm b}$. Also for this model, we use only the flux bands with wavelengths ≥100 μm. In order to have good constraints on the fitting parameters, it is crucial to have a detection of the 850 μm flux. If the SCUBA-2 point is not detected, an upper limit is not enough to constrain the parameters of this model. Without the 850 μm flux point, the 500 μm flux point is the only one that can be used to determine β_2 and $\lambda_{\rm b}$, leading to large uncertainties on their values.

(iii) **TMBB:** The TMBB model assumes that the FIR SED is emitted by two dust populations with different temperatures. The dust emission is parametrized by TMBB: one for the cold dust (indicatively $T < 40$ K) and one for the warm dust (indicatively $T > 40$ K):

$$F_\lambda = F_\lambda^{\rm SMBB_{cold}} + F_\lambda^{\rm SMBB_{warm}}, \quad (5)$$

where the two SMBB components are defined as above. In order to reduce the number of free parameters, we fix the β value of the warm component to 1.5 (Coupeaud et al. 2011; Boselli et al. 2012), while we leave the β value of the cold component as a free parameter. So in this model we have five free parameters: $M_{\rm cold}$, $T_{\rm cold}$, $\beta_{\rm cold}$, $M_{\rm warm}$, and $T_{\rm warm}$. For the fitting, we use the fluxes in all available bands from 22 to 850 μm.

All these models assume that dust grains are optically thin. According to dust models, this assumption holds for wavelengths ≥100 μm, while at shorter wavelengths it is possible that dust is optically thick (Draine & Li 2007). Casey (2012) modelled the SED of 65 luminous infrared galaxies from the GOALS survey (Armus et al. 2009) and found that even if the dust is optically thick, the difference in the SED shape at 22 μm would be small. Utomo et al. (2019) studied the dust emission at resolved scales in four nearby galaxies (Small and Large Magellanic Clouds, M31, and M33) and found that most of the dust emitting at wavelengths longer than 100 μm is optically thin. They observe that at wavelengths ∼20 μm some regions of the galaxies become optically thick, but on global galaxy scales we do not expect these regions to dominate the emission.

We apply the SMBB model to both the JINGLE and HRS sample, while we apply the BMBB and TMBB models only to the JINGLE sample. We make this decision because for the HRS sample we do not have the 850 μm flux point, and therefore we do not have enough flux points for models with a large number of free parameters. Additionally, for the BMBB model it is very important to have the 850 μm point to constrain the emissivity index β_2 after the break. Fig. 2 shows an example of the SED fitting of one galaxy from the JINGLE sample using the three models.

3.2 Introduction to the Bayesian SED fitting method

In this section, we briefly describe the Bayesian approach used for the SED fitting (we follow the same notation as in Galliano 2018). Readers who are less interested in the statistical methods may wish to go directly to the results presented in Section 4. The observed SED of a galaxy ($F^{\rm obs}$) can be described in the following way:

$$F^{\rm obs}(\lambda_j) = F^{\rm mod}(\lambda_j, \boldsymbol{\theta}) + \epsilon(\lambda_j) \cdot F^{\rm err}(\lambda_j), \quad (6)$$

where $F^{\rm obs}(\lambda_j)$ is the flux observed at the wavelength λ_j and $F^{\rm mod}(\lambda_j, \boldsymbol{\theta})$ is the flux described by our model with parameters $\boldsymbol{\theta}$. The last term describes the deviation of the observed flux from the model due to random noise: $F^{\rm err}(\lambda_j)$ is the amplitude of the noise and $\epsilon(\lambda_j)$ is a random variable with mean $\langle\epsilon\rangle = 0$ and standard deviation $\sigma(\epsilon) = 1$. We can reverse the previous formula to express $\epsilon(\lambda_j)$ as a function of the other quantities:

$$\epsilon(\lambda_j) = \frac{F^{\rm obs}(\lambda_j) - F^{\rm mod}(\lambda_j, \boldsymbol{\theta})}{F^{\rm err}(\lambda_j)}. \quad (7)$$

The goal is to find the best parameters to fit the data by minimizing the offset between the model and the data. From a Bayesian point of view, this is equivalent to maximizing the likelihood of the model, given the data. The probability of the data given the model parameters $\boldsymbol{\theta}$ can be expressed as

$$p(\boldsymbol{F}^{\rm obs}|\boldsymbol{\theta}) = \prod_{j=1}^{m} p(\epsilon(\lambda_j, \boldsymbol{\theta})), \quad (8)$$

where $\boldsymbol{F}^{\rm obs} = \left(F^{\rm obs}(\lambda_1), \ldots, F^{\rm obs}(\lambda_m)\right)$, is the vector containing the flux emission at each waveband $j = 1, \ldots, m$. We are interested in the probability of the model parameters, knowing the observations.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

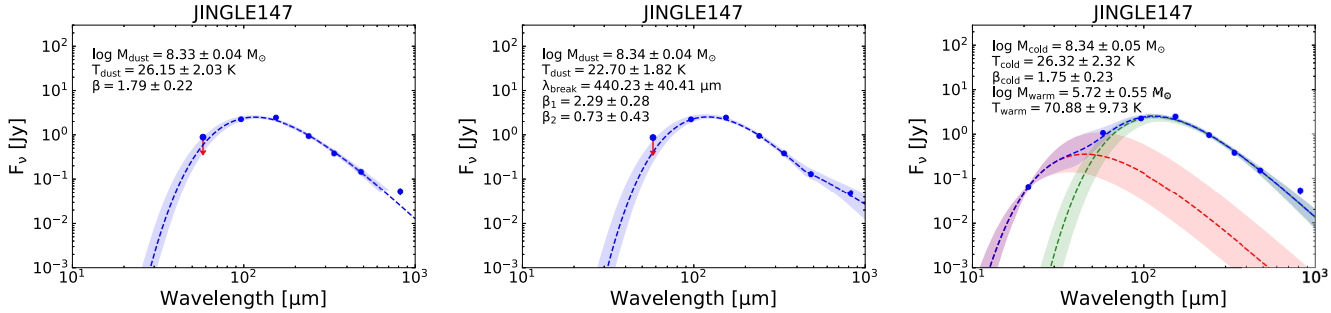

Figure 2. Example of FIR SED of one galaxy from the JINGLE sample, fitted with the non-hierarchical approach using the three models: SMBB (left-hand panel), BMBB (middle panel), and TMBB (right-hand panel). The shaded regions show the lower and upper 1σ uncertainties on the SED models, defined by taking the maximum and minimum flux values of the models with likelihood values in the highest 68th percentile.

Thus, we can use the Bayes' theorem to write the expression:

$$p(\boldsymbol{\theta}|\boldsymbol{F}^{\rm obs}) = \frac{p(\boldsymbol{F}^{\rm obs}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta})}{p(\boldsymbol{F}^{\rm obs})} \propto p(\boldsymbol{F}^{\rm obs}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}), \quad (9)$$

where $p(\boldsymbol{\theta})$ is the 'prior' distribution and $p(\boldsymbol{\theta}|\boldsymbol{F}_{\rm obs})$ is the 'posterior' distribution. The denominator $p(\boldsymbol{F}^{\rm obs})$ can be neglected since it is constant for a given set of observed fluxes. By sampling the posterior distribution in the parameter space we can construct the posterior probability density function (PDF). Examples of posterior PDFs are shown in the appendix (Fig. D1). The figure shows the PDFs obtained from the SED fit of one galaxy using the SMBB, BMBB, and TMBB models.

3.3 Hierarchical Bayesian method

The difference between the classical and hierarchical Bayesian method is that in the former the prior distribution is an assumption and in the latter it is defined by the data sample (e.g. Gelman et al. 2004; Galliano 2018). Hierarchical methods require therefore a population of objects, which are used to define the prior distributions. In the case of SED fitting, the sample can be formed by multiple spatially resolved regions of the same galaxy or by a sample of galaxies with similar properties. The entire sample is then fitted simultaneously, in order to extract both the information about the prior distribution of the sample and the posterior distribution of the single elements of the sample.

Kelly et al. (2012) showed that the hierarchical method can be used to reduce the degeneracy between T and β. This approach has subsequently been used in other studies to reduce the T–β degeneracy (Juvela et al. 2013; Veneziani et al. 2013; Galliano 2018). The key assumption behind the hierarchical approach is that the dust parameters of our sample of galaxies follow a common distribution. In our case we assume that they follow a Student's t-distribution. Thanks to this assumption, we are able to better constrain model parameters, especially for galaxies with low S/N, where a large range of combinations of T and β provide reasonably good fits to the data. In those cases, the prior helps to constrain the range of possible T and β. The key point of the hierarchical approach is that we do not need to specify the mean and standard deviation of the prior distribution before doing the fit, but they can be inferred by the data.

The new parameters describing the prior distribution of the parameters $\boldsymbol{\theta}$ are called *hyper-parameters*. The commonly used hyper-parameters are

(i) $\boldsymbol{\mu}$: the average of the parameter vector $\boldsymbol{\theta}$;

(ii) Σ: the covariance matrix describing the standard deviation and correlation of $\boldsymbol{\theta}$.

Using this formalism, the posterior distribution of the parameters given the data $p(\boldsymbol{\theta}|\boldsymbol{F}_{\rm obs})$ for the ith galaxy in the sample becomes

$$p(\boldsymbol{\theta_i}|\boldsymbol{F_i}^{\rm obs}, \boldsymbol{\mu}, \Sigma) \propto p(\boldsymbol{F_i}^{\rm obs}|\boldsymbol{\theta_i}) \cdot p(\boldsymbol{\theta_i}|\boldsymbol{\mu}, \Sigma). \quad (10)$$

This is the hierarchical equivalent of equation (9). The posterior distribution of the parameters and hyper-parameters for the entire sample of n galaxies is

$$p(\boldsymbol{\theta_1}, \ldots, \boldsymbol{\theta_n}, \boldsymbol{\mu}, \Sigma|\boldsymbol{F_1}^{\rm obs}, \ldots, \boldsymbol{F_n}^{\rm obs}) \propto \prod_{i=1}^{n} p(\boldsymbol{\theta_i}|\boldsymbol{F_i}^{\rm obs}, \boldsymbol{\mu}, \Sigma) \cdot p(\boldsymbol{\mu}) \cdot p(\Sigma)$$

$$\propto \prod_{i=1}^{n} p(\boldsymbol{F_i}^{\rm obs}|\boldsymbol{\theta_i}) \cdot p(\boldsymbol{\theta_i}|\boldsymbol{\mu}, \Sigma) \cdot p(\boldsymbol{\mu}) \cdot p(\Sigma), \quad (11)$$

where $p(\boldsymbol{\mu})$ and $p(\Sigma)$ are the prior distributions of the hyper-parameters. When compared to the classical Bayesian method, the hierarchical method is able to recover the distribution of parameters with better precision, especially if the noise in the data is high (Kelly et al. 2012; Galliano 2018). In that case, the hierarchical approach uses the information about the parameter distribution obtained from the rest of the sample to better constrain the parameters for the particular objects where the quality of the data is low. The hierarchical method will not necessarily perform better in measuring the parameter of a single object, but it will be less biased when measuring the distribution of parameters of the entire population.

3.4 Noise distribution

In this section, we describe the functions used to model the noise distribution for both the non-hierarchical and hierarchical approaches. The noise is usually modelled with a normal distribution or a Student's t-distribution. The Student's t-distribution has a higher probability in the tails with respect to the normal distribution, allowing for more outliers. Its shape is described by the number of degrees of freedom f: as f decreases, more probability will be in the tails of the distribution. The normal distribution is a special case of the t-distribution with the number of the degrees of freedom that goes to infinity, $f \rightarrow \infty$.

The probability density of a normal distribution is defined as

$$\text{Normal}(y|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right), \quad (12)$$

where μ is the mean and σ is the standard deviation. The multivariate normal distribution is the generalization of the one-dimensional

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

Table 1. Prior parameter ranges assumed for the Bayesian non-hierarchical SED modelling using the SMBB function.

Parameter	Range
$\log M_{\rm dust}/{\rm M}_\odot$	(5, 9)
T (K)	(5, 50)
β	(0.1, 3)

normal distribution to a higher dimension m:

$$\text{MultiNormal}(\boldsymbol{y}|\boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{m/2}} \frac{1}{\sqrt{|\Sigma|}} \times \exp\left(-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{y}-\boldsymbol{\mu})\right), \quad (13)$$

where m is the dimension of the vector $\vec{y}$, Σ is the $m \times m$ covariance matrix, and $(\boldsymbol{y}-\boldsymbol{\mu})^T$ indicates the transpose of the vector $(\boldsymbol{y}-\boldsymbol{\mu})$.

The Student's t-distribution is defined as

$$\text{Student}(y|\mu, \sigma, f) = \frac{\Gamma((f+1)/2)}{\Gamma(f/2)} \frac{1}{\sqrt{f\pi}\sigma} \left(1 + \frac{1}{f}\left(\frac{y-\mu}{\sigma}\right)^2\right)^{-\frac{f+1}{2}}, \quad (14)$$

where f is the number of degrees of freedom. The multivariate Student's t-distribution is the generalization of the one-dimensional distribution to a higher dimension m:

$$\text{MultiStudent}(\boldsymbol{y}|\boldsymbol{\mu}, \Sigma, f) = \frac{\Gamma((f+m)/2)}{\Gamma(f/2)} \frac{1}{(f\pi)^{m/2}} \frac{1}{\sqrt{|\Sigma|}} \times \left(1 + \frac{1}{f}(\boldsymbol{y}-\boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{y}-\boldsymbol{\mu})\right)^{-\frac{f+m}{2}}, \quad (15)$$

where m is the dimension of the vector $\boldsymbol{y}$.

We expect to observe a flux excess at 850 μm for some galaxies, given the fact that the sub-mm excess has been reported in numerous studies (e.g. Galametz et al. 2011; Rémy-Ruyer et al. 2013, 2015; Hermelo et al. 2016). Since the 850 μm fluxes have usually larger uncertainties than the other points, if we use a Student's t-distribution, the SMBB model will assume that every change in slope at 850 μm is due to the error being underestimated, rather than to a physical effect. The model will then 'ignore' the 850 μm point, and produce a fit considering only the *Herschel* points. Since we believe that there is information in the longer wavelength points, we therefore decide to use a normal distribution for the error. In Appendix A, we compare the results obtained using the Student and normal distribution.

In both the non-hierarchical and hierarchical case, we model the noise as

$$p(\boldsymbol{F}^{\rm obs}|\boldsymbol{F}^{\rm mod}(\boldsymbol{\theta}), C) = \text{MultiNormal}(\boldsymbol{F}^{\rm obs}|\boldsymbol{F}^{\rm mod}(\boldsymbol{\theta}), C), \quad (16)$$

where C is the covariance matrix, which describes the uncertainties associated with the flux densities in the different wavebands (see Section 3.6 for the definition of the covariance matrix).

3.5 Prior distributions

In this section, we describe the prior distributions assumed for the hierarchical and non-hierarchical method.

Non-hierarchical: For the prior distribution of the parameters $\boldsymbol{\theta}$, we assume uniformly distributed ('flat') priors, i.e. $p(\theta) = 1$, in the ranges described in Table 1.

Hierarchical: For the definition of the prior distributions in the hierarchical framework, we follow Kelly et al. (2012), Galliano (2018), and the STAN manual (Stan Development Team 2017).

(i) *Parameters*: For the definition of the prior distributions of the parameters given the hyper-parameters, we follow Kelly et al. (2012) and Galliano (2018). We assume a multivariate Student's t-distribution with $f = 8$ degrees of freedom:

$$p(\boldsymbol{\theta_i}|\boldsymbol{\mu}, \Sigma) = \text{MultiStudent}(\boldsymbol{\theta_i}|\boldsymbol{\mu}, \Sigma, f = 8). \quad (17)$$

We also tried to vary the number of degrees of freedom and did not see any differences in the results. Assuming a Student's t-distribution allows one to have more galaxies with dust parameters that are 'outliers' from the mean of the sample. In this way, we make sure that our assumption that the galaxies belong to the same population is not too stringent. We note that the parameters $\boldsymbol{\theta_i}$ are not constrained within a certain range but they are allowed to take any value. Their distribution is described by the prior distribution and we set some constraints on the allowed range of the hyper-priors (mean and standard deviation) that determine the shape of the priors (see next point).

Table 2. Ranges of the priors on the hyper-parameter $\boldsymbol{\mu}$ (sample mean) for the Bayesian hierarchical SED modelling using the SMBB, BMBB, and TMBB functions.

Hyper-parameter	Range
SMBB	
$\mu(\log M_{\rm dust}/{\rm M}_\odot)$	(6, 9)
$\mu(T)$ (K)	(15, 50)
$\mu(\beta)$	(0.5, 3)
BMBB	
$\mu(\log M_{\rm dust}/{\rm M}_\odot)$	(5, 9)
$\mu(T)$ (K)	(5, 50)
$\mu(\beta_1)$	(0, 5)
$\mu(\beta_2)$	(0, 5)
$\mu(\lambda_{\rm b})$ (μm)	(420, 500)
TMBB	
$\mu(\log M_{\rm cold}/{\rm M}_\odot)$	(6, 10)
$\mu(T_{\rm cold})$ (K)	(5, 40)
$\mu(\beta_{\rm cold})$	(0.5, 5)
$\mu(\log M_{\rm warm}/{\rm M}_\odot)$	(2, 7)
$\mu(T_{\rm warm})$ (K)	(50, 90)

(ii) *Hyper-parameters*: For the mean $\boldsymbol{\mu}$ of the parameters, we assume a uniform prior with a large parameter range. In this way we ensure that the prior is proper (i.e. $\int p(\theta){\rm d}\theta < \infty$), and at the same time we maintain the prior vague enough to not constrain the results (Gelman & Hill 2007; Tak, Ghosh & Ellis 2018). The prior ranges for $\boldsymbol{\mu}$ are shown in Table 2. We note that we set the prior range of $\mu(T_{\rm warm})$ to be >50 K, because we want the distribution of warm temperatures to be well separated from the distribution of cold temperatures. For the covariance matrix Σ, we use the *separation strategy* from Barnard, McCulloch & Meng (2000). This formalism ensures that the prior distributions of the correlations between parameters are uniform over the range $[-1, 1]$, meaning that all values of the correlations are equally likely. The separation strategy breaks down the covariance matrix in

$$\Sigma = SRS, \quad (18)$$

where S is a diagonal matrix with the values of the standard deviation and R is the correlation matrix. Both S and R have dimension $q \times q$, where q is the number of free parameters in the model. The prior distribution of the hyper-parameters is then

$$p(\boldsymbol{\mu}) \cdot p(\Sigma) \propto p(\boldsymbol{\mu}) \cdot p(S) \cdot p(R). \quad (19)$$

For the priors on the S and R, we follow the recommendations given by the STAN manual (Stan Development Team 2017). For the priors on the diagonal elements of S, we use a weakly informative prior,

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

parametrized by a half-Cauchy distribution with a small scale $\sigma = 2.5$ (Stan Development Team 2017):

$$p(S_{k,k}) = \text{Cauchy}(0, \sigma) = \frac{1}{\pi\sigma} \frac{1}{1 + \left(\frac{S_{k,k}}{\sigma}\right)^2}, \quad (20)$$

where $S_{k,k} > 0$, for $k = 1, \ldots, q$. For the priors on the correlation matrix R, we use an LKJ correlation distribution with shape $\nu = 2$:

$$p(R) = \text{LKJ Corr}(R, \nu) \propto \det(R)^{\nu-1} \quad (21)$$

(see Lewandowski, Kurowicka & Joe 2009 for definitions). The basic idea of the LKJ correlation distribution is that as ν increases, the prior increasingly concentrates around the identity matrix.

3.6 Covariance matrix, beam, and filter corrections

In order to perform an accurate fit, it is important to take into account correctly the uncertainties associated with each flux measurement as well as the correlation between these uncertainties. The covariance matrix C describes the uncertainties associated with the flux densities in the different wavebands, and includes both calibration and measurement uncertainties. Calibration uncertainties can be correlated between bands observed with the same instrument. For the definition of the covariance matrix, we follow Gordon et al. (2014). The calibration covariance matrix is defined as

$$C^{\text{cal}}_{j,k} = [A_{\text{cor},j,k} + A_{\text{uncor},j,k}] = [\sigma^2_{\text{cor},j,k} + \delta_{j,k}\sigma^2_{\text{uncor},j,k}], \quad (22)$$

where A_{cor} is the matrix of the noise correlated between bands, A_{uncor} is the diagonal matrix of repeatability that is uncorrelated between bands. $\sigma_{\text{cor},j,k}$ and $\sigma_{\text{uncor},j,k}$ are thepercentage of correlated and uncorrelated uncertainties, respectively, between the jth and kth band, and $\delta_{j,k}$ is one for $j = k$ and zero otherwise. The calibration uncertainty values that we use are reported in Table 3, given inpercentage of the flux.

The total covariance matrix C is a combination of the calibration and measurement uncertainties:

$$C_{j,k} = C^{\text{cal}}_{j,k} \cdot F_j \cdot F_k + F^{\text{err}}_j \cdot F^{\text{err}}_k, \quad (23)$$

where F_j and F_k are the fluxes in the jth and kth waveband, and F^{err}_j and F^{err}_k are the corresponding measurement uncertainties.

The colour and beam corrections applied to our data are described in detail in De Looze et al. (in preparation).

Non-hierarchical: The filter corrections are applied to the model SED by convolving the model flux points with the appropriate filter response curve in each band. The *Herschel*/SPIRE fluxes were corrected also for the effective beam area, which depends on the shape of the spectrum due to the absolute SPIRE calibration in units of flux density per beam. The SED shape is described by the dust temperature T and the emissivity index β. At each step of the Markov chain Monte Carlo (MCMC) algorithm, the *Herschel*/SPIRE fluxes are corrected according to the two model parameters, before comparing them to the fluxes of the SED model. For the BMBB model, we applied the beam and colour corrections using β_1 or β_2 depending on the wavelength position of the break λ_b. For the TMBB model, we calculate which of the two components (warm or cold) contribute the most to the flux in every band. Then we calculate the corrections using the temperature T and β values of the dominant component in each band.

Hierarchical: The beam and filter corrections make it more difficult for the code to converge, since in every MCMC step the fluxes are slightly modified. This is more problematic for the hierarchical approach, because it has a larger number of free parameters. Therefore, in order to achieve convergence in a reasonable amount of time, we apply a slightly different approach to implement the beam and filter corrections in the hierarchical case. We first do the hierarchical fit without beam and filter corrections. Then we apply the beam and filter corrections on the fluxes based on the values of T and β measured from the fit with no corrections, and finally we repeat the hierarchical fit using the 'corrected' fluxes. The beam and filter corrections are generally small compared to the flux uncertainties, therefore this approximation of the corrections does not affect the results significantly.

3.7 Implementation of the SED fitting

Non-hierarchical method: For the implementation of the classical Bayesian SED fitting method, we employ the affine-invariant ensemble sampler for MCMC (Metropolis et al. 1953) code EMCEE (Goodman & Weare 2010; Foreman-Mackey et al. 2013). The MCMC algorithm is designed to sample the posterior distribution of the unknown parameters, i.e. the probability of the parameters given the data. The values of the parameters with the corresponding uncertainties can then be inferred from the posterior distribution. We consider as results the median values of the marginalized posterior probability distributions, and we estimate the uncertainties from the values corresponding to the 16th and 84th percentiles.

To monitor the convergence we look at the effective sample size (N_{eff}), which is defined as the number of iterations divided by the integrated autocorrelation time $N_{\text{eff}} = N_{\text{iter}}/\tau_{\text{int}}$. The autocorrelation time τ_{int} measures the number of steps after which the drawings are truly independent (Foreman-Mackey et al. 2013). It is recommended to have at least $N_{\text{eff}} > 10$, to ensure that the sequence has converged (Gelman et al. 2004).

Hierarchical method: For the implementation of the hierarchical Bayesian fitting we use STAN (Carpenter et al. 2017, http://mc-stan.org/), a software for Bayesian inference which employs the No-U-Turn sampler, a variant of Hamiltonian Monte Carlo (HMC) sampler. The HMC sampling (Duane et al. 1987; Neal 1994, 2011) is a form of MCMC sampling which uses the gradient of the logarithmic probability function to accelerate the parameter exploration and the convergence to the stationary distribution (Stan Development Team 2017). The HMC algorithm is more efficient than other MCMC algorithms (as for example the Metropolis–Hastings algorithm) in sampling the parameter space and in finding the region of high likelihood, because it samples the probability distribution with fewer samples. Therefore, it is particularly well suited for problems with high dimension, as is the case for hierarchical models. For example, for the hierarchical fit of 100 galaxies using the SMBB model, which has three free parameters, the dimension is of the order of ∼300. Another advantage of STAN is that it can sample simultaneously the posterior distribution of parameters and hyper-parameters. STAN allows to define the model by specifying the probability distribution of each parameter (or hyper-parameter) independently, without the need of computing the full posterior distribution. For the practical implementation, we used PYSTAN,[3] which is the PYTHON interface to STAN (Stan Development Team 2018).

The recommended method for monitoring the convergence of the MCMC chains in STAN is computing the potential scale reduction statistics $\hat{R}$ (Gelman & Rubin 1992), which gives an estimate of the factor by which the scale of the posterior distribution may be

[3] http://pystan.readthedocs.io/en/latest/http://mc-stan.org

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

Table 3. Percentage of correlated and uncorrelated uncertainties for the different instruments.

Instrument	Waveband (μm)	Correlated uncertainty (per cent)	Uncorrelated uncertainty (per cent)	Reference
WISE	22	–	5.7	Jarrett et al. (2011)
IRAS	60	–	20	Sanders et al. (2003); Miville-Deschênes & Lagache (2005)
PACS	100, 160	5	2	Balog et al. (2014), Decin & Eriksson (2007)
SPIRE	250, 350, 500	4	1.5	Bendo et al. (2013)
SCUBA	850	–	10	Smith et al. (2019)

reduced as the number of iterations goes to infinity. If $\hat{R}$ is large, it means that increasing the number of iterations is likely to improve the inference. If $\hat{R} \sim 1$, then we can be confident that the number of iterations that we are using is large enough. Thus, we set the requirement that for our runs $\hat{R} < 1.15$. We also check that the effective sample size $N_{\rm eff}$ is always larger than 10.

3.8 Validation of the method with simulations of mock SEDs

We test our fitting methods using simulated FIR SEDs. For the mock SEDs, we know the input parameter values, thus we can assess how well our fitting procedure is able to recover them. The simulation code takes as input parameters the dust mass ($\log M_{\rm dust}$), temperature T, and emissivity index β, and it uses these parameters to generate an SED assuming an SMBB model. Then it extracts the flux density in the selected wavebands and it adds random noise at each flux point. We assume the noise to be Gaussian distributed around zero, with amplitude equal to the noise level. We assume a different noise level in every band. For the wavebands (100, 160, 250, 350, 500, 850) μm, we use the following noise levels, given aspercentages of the flux: (20, 10, 5, 10, 20, 25) per cent, respectively. We estimate these values by taking the mean of the error fraction in each band from our data.

The goal of the test is to assess how well the non-hierarchical Bayesian approach can measure the values of temperature and β. We simulate 100 SEDs with the same input parameters ($\log M_{\rm dust} = 8\ {\rm M}_{\odot}$, $T = 30$ K, $\beta = 1.5$), adding to every SED random noise in every band as explained above. Fig. 3 shows the results in the T–β plane. As we can see from the figure, an artificial anticorrelation is generated only from the effect of adding noise to the fluxes. This suggests that the non-hierarchical Bayesian approach will always measure a T–β anticorrelation, even if it is not present in the data. Thus, in order to asses if the T–β anticorrelation is indeed present in our sample, we need a more sophisticated fitting method.

We run the same simulation, but this time we use the hierarchical code to fit the SEDs. The results are in better agreement with the input value, and do not show any artificial correlation or anticorrelation between T and β. The non-hierarchical method measures a large range of temperatures ($T = 22-42$ K) and β values ($\beta = 0.8-2.3$). The hierarchical method measures smaller ranges of $T = 27-30$ K and $\beta = 1.50-1.55$, which are closer to the input values. Consequently, also the dust masses are better measured with the hierarchical method. The dust masses measured with the non-hierarchical method are in the range $\log M_{\rm dust}/{\rm M}_{\odot} = 7.87-8.23$, with typical uncertainties of ∼0.13 dex, while the ones measured with the hierarchical method are in the range $\log M_{dust}/{\rm M}_{\odot} = 8.06-8.09$, with typical uncertainties 0.02 dex.

We also test whether the codes can recover a positive or negative T–β correlation. In both cases, the hierarchical method perform equally or better than the non-hierarchical code. Details of these simulations can be found in Appendix C.

4 RESULTS

4.1 JINGLE sample: non-hierarchical versus hierarchical results

In the previous section we have demonstrated, using simulated SEDs, that the hierarchical method works better than the non-hierarchical approach. Here, we apply both methods to the 192 galaxies of the JINGLE sample and we show the advantages of using the hierarchical method.

We start by using the simplest model, the SMBB. Fig. 4 shows the comparison of the dust masses, dust temperatures and β derived with the two approaches. In general, dust masses agree quite well between the two methods (median difference $= 0.07$ dex). The dust masses derived using the hierarchical method are slightly smaller, and this is probably due to the variations in dust temperatures. For a given constant flux, higher dust temperatures correspond to lower dust masses. In the range 15−25 K the dust temperatures from the hierarchical approach are indeed slightly higher. At high temperatures, the differences between the two methods are larger and the non-hierarchical method measures much higher temperatures ($T > 30$ K) than the hierarchical method. This is because as the dust temperature increases, the peak of the SED moves to shorter wavelengths. If the SED peaks at wavelengths shorter than 100 μm, it is not sampled by the flux bands considered in the fit, since for the SMBB we are considering the 60 μm point as an upper limit. Therefore, it is more difficult to constrain the temperature. If we were to include flux points at shorter wavelengths we would need to consider a second MBB component with a warmer temperature, because the assumption of a single temperature MBB does not hold over such a large wavelength range. Instead, in the hierarchical framework, the code uses the information from the temperature distribution of the galaxy population to constrain T, and it will consider more likely for the galaxy to have a temperature close to the population mean temperature than an extreme value. Therefore, the hierarchical method can better constrain the dust temperature.

The range of temperatures is smaller in the hierarchical case ($T = 17-30$ K), than in the non-hierarchical case ($T = 15-48$ K). The same is true for the range of β: in the hierarchical case $\beta = 0.6-2.2$, while in the non-hierarchical case $\beta = 0.0-2.5$. In the hierarchical approach, we assume that the population follows a common distribution, thus the fitting is less likely to return extreme values of β. However, the hierarchical code can accommodate some outliers, since we do not define a priori the standard deviation of the prior distribution. Thus, if the data require it, the standard deviation can be large, allowing for more 'extreme' values of β.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

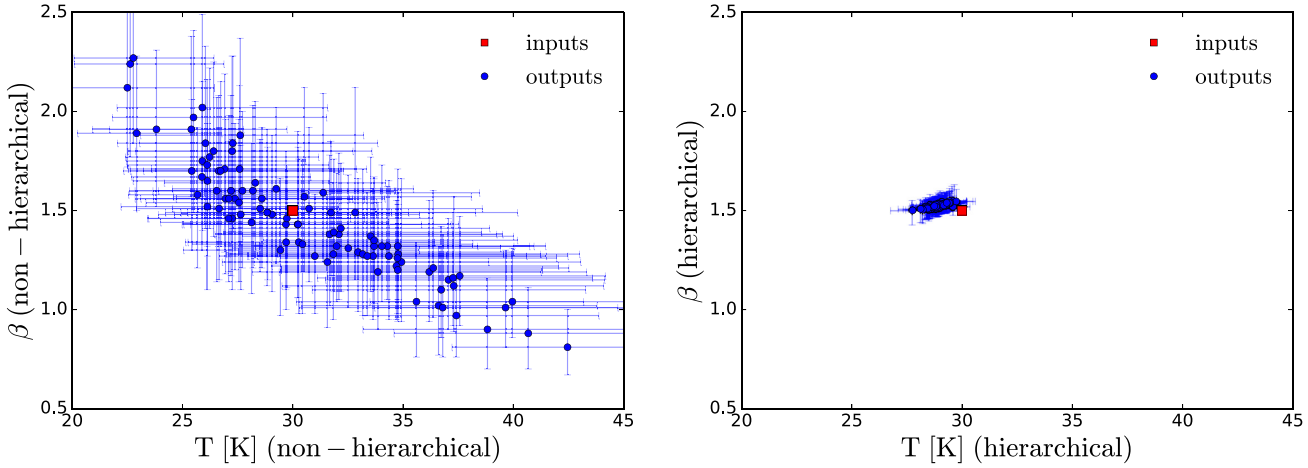

Figure 3. Results of temperature and β from the fit of 100 simulated SMBB SEDs with the same input parameters ($\log M_{\rm dust}/M_\odot = 8$, $T = 30$ K, and $\beta = 1.5$) and 10 per cent added noise. The output values are derived with the non-hierarchical (left-hand panel) and hierarchical (right-hand panel) SED fitting method. In red is shown the input value and in blue are the measured values.

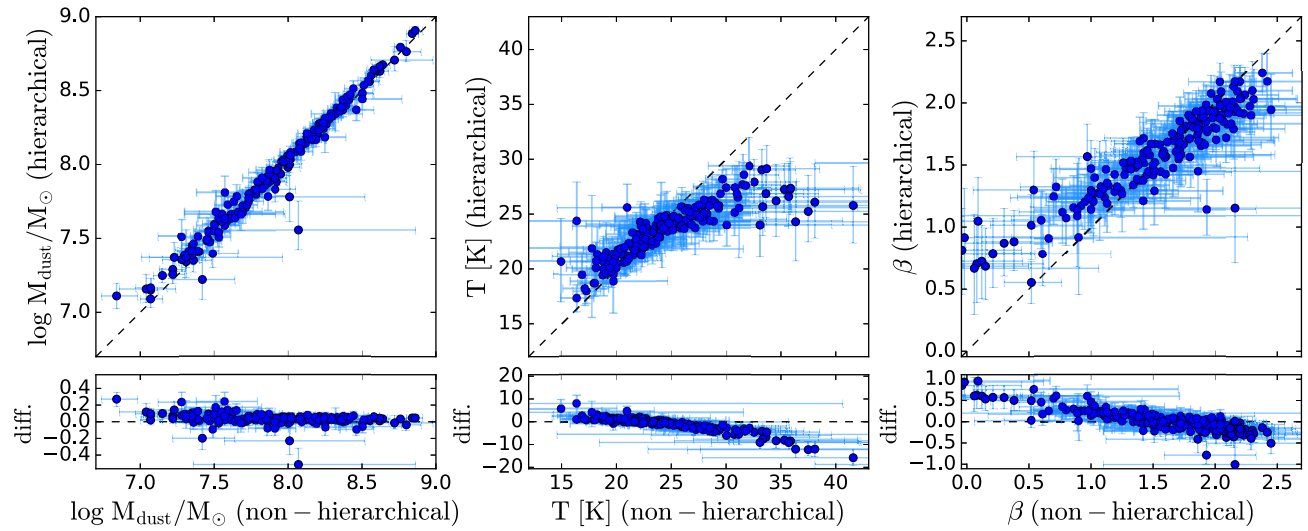

Figure 4. Comparison of dust properties of the JINGLE sample obtained through the fit of an SMBB using the non-hierarchical and hierarchical approaches. The lower panels show the difference between the hierarchical and non-hierarchical fit in each of the derived properties.

But if the extreme objects have large noise on the flux values, then the hierarchical method considers more likely that they are not 'true outliers', but that their extreme SED shape is only due to the noise in the data points. If we believe that the hierarchical approach gives more accurate results for the cases with high noise level, we conclude that the extreme values found with the non-hierarchical approach are likely not reliable, but only due to the noise in the data. The results of the hierarchical fit using the SMBB model are given in Table E1.

4.2 *T*–β relation in the JINGLE sample

We use the results of the SED fitting using the SMBB model to investigate whether there is a relation between dust temperature and β in our sample of galaxies. An anticorrelation between T and β has been observed in many studies (e.g. Dupac et al. 2003; Désert et al. 2008), but it has been demonstrated that it can be attributed to the degeneracy between the two parameters and the effect of noise on the data (Shetty et al. 2009a,b).

Fig. 5 shows the results from the non-hierarchical and hierarchical approach applied to our sample of 192 galaxies. The results from the non-hierarchical method show a significant anticorrelation between T and β. The Pearson correlation coefficient is $R_{\rm pear} = -0.79$ (p-value $= 1.19 \times 10^{-41}$). The results from the hierarchical method shows a weaker anticorrelation ($R_{\rm pear} = -0.52$, p-value $= 9.79 \times 10^{-15}$). This shows that the choice of the method used is really important and can deeply influence the results. This result confirms previous findings (Shetty et al. 2009a,b; Kelly et al. 2012; Juvela et al. 2013; Veneziani et al. 2013) that the observed $T-\beta$ anticorrelation is mainly driven by the fact that they are

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

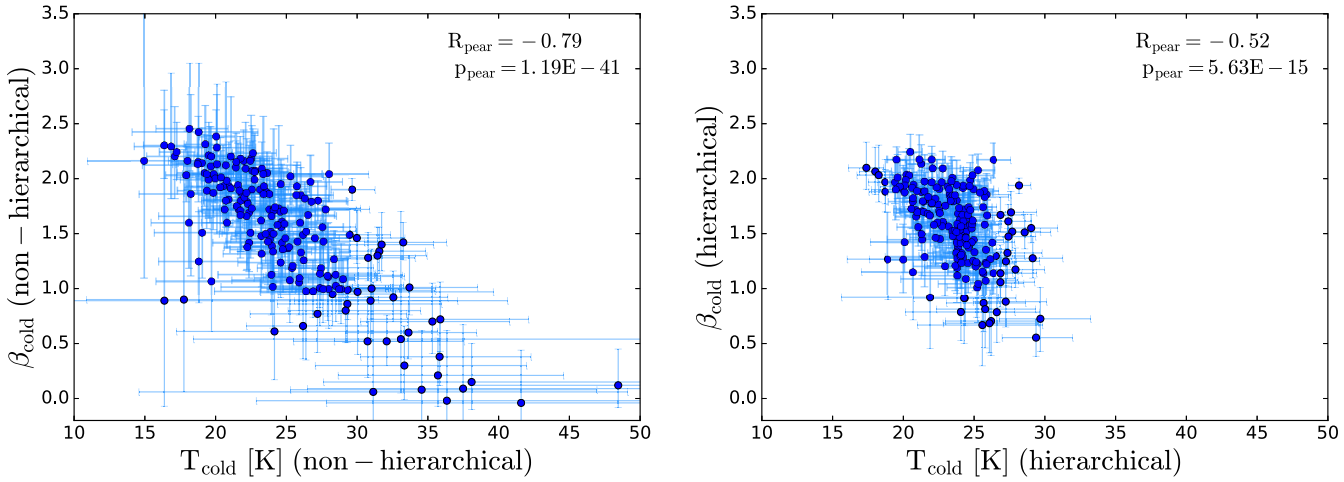

Figure 5. Relation between the dust temperature and dust emissivity index (T–β relation) for the JINGLE sample derived with non-hierarchical (left-hand panel) and hierarchical (right-hand panel) Bayesian methods. In both cases, we fit the SED using an SMBB model and we include the 850 μm flux point in the fit.

degenerate parameters, and by the noise on the data. There is still an anticorrelation between T and β even using the hierarchical approach ($R_{\rm pear} = -0.52$). This could mean that there is indeed a physical relation between these two quantities. However, it is also possible that the hierarchical method is not able to remove completely the $T-\beta$ degeneracy, leaving a residual anticorrelation. With our current data we are not able to distinguish whether the observed relation is a physical effect or whether it is due to a residual degeneracy.

We also compare the results obtained with and without including the 850 μm flux point in the fit using the hierarchical approach (see Fig. 6). In general, the emissivity indices β measured with the 850 μm flux point are equal or lower than the ones measured without the 850 μm point. This means that without the SCUBA-2 flux, the fits of the *Herschel* points alone have steeper slopes. This suggests that there is indeed a 'sub-mm' excess visible at 850 μm, at least in some galaxies. This is visible especially for low values of $\beta < 1$. We note that not all galaxies show this behaviour: for some galaxies the β values measured with and without SCUBA-2 flux are in good agreement, or they show a small deficit at 850 μm. Consequently, the dust temperatures show the opposite trend: they are in general larger when the 850 μm point is included in the fit, because they have to compensate for the lower β values. The mass measurements are only slightly affected by the presence of the SCUBA-2 flux point (median difference: 0.002 dex). The largest difference in the dust masses measured with and without the SCUBA-2 flux point is 0.07 dex. The fact that the dust masses do not show a larger variation depends on the fact that we assumed a constant absorption coefficient κ_0. Laboratory studies show that κ changes with dust temperature T and β (Coupeaud et al. 2011; Demyk et al. 2017a,b). Therefore, by keeping κ constant we erase the difference in dust masses that would arise from the different temperature and β values. A certain value of κ_0 will give an accurate dust mass only if the β value used for the fit is the same that was used to measure κ_0 (Bianchi 2013). However, a recent laboratory study by Demyk et al. (2017a) shows that variations in κ_0 are more prominent for high temperatures ($T > 30$ K) than for low temperatures. For the temperature range considered in this study (10−30 K) they do not observe variations in κ_0. A possible approach to account for variations in κ_0 would be to change the value of κ_0 according to the value of T and β used for the fitting in an iterative way. We plan to investigate this in the future.

4.3 Comparison of models: SMBB, BMBB, TMBB

In many cases, the SMBB model is not enough to fit the FIR/sub-mm SED accurately. Especially at long wavelengths, the SED often shows a change in the slope. Therefore, we consider also two other models: the BMBB and TMBB models, described in Section 3.1. In this section, we compare the results obtained applying these models to the SED fit of the JINGLE sample. The results of the hierarchical fit using the BMBB and TMBB models are given in Tables E2 and E3.

4.3.1 BMBB

The BMBB model (Gordon et al. 2014) allows for a variation in the wavelength dependence of the dust emissivity law, to account for a sub-mm excess. This is parametrized by using two emissivity indices for shorter and longer wavelengths. The break wavelength is a free parameter in our model. For the JINGLE sample we find values in the range 480–488 μm. The emissivity index at wavelengths shorter than $\lambda_{\rm break}$ (β_1) is in the range 0.6–2.2. The range of the second emissivity index at wavelengths $> \lambda_{\rm break}$ (β_2) is larger (0.1–3.3). We compared the results obtained using the BMBB model with the results from the SMBB model (Fig. 7). The dust masses measured with the BMBB model are in agreement with the ones measured with the SMBB model, with a maximum difference of 0.1 dex. The BMBB model measures generally slightly lower temperatures than the SMBB model (median difference of 1 K). In the case of a shallower slope of the sub-mm SED, the SMBB model fits it by using a lower value of β and a higher T. The BMBB can correct using a smaller value of β_2, without affecting the temperature measurement. Thus, T does not depend anymore on the longer wavelength points and can have a lower value. We compare also the emissivity index β from the SMBB model, with the parameter β_1 that describes the slope of the BMBB model before the break. β_1 tends to be larger than β from the SMBB for low values of β. This is due to the fact that any excess at longer wavelength can be modelled by a second index β_2, while in the case of the SMBB the excess needs to be taken into account by β.

The results from the BMBB model are more similar to the SMBB fit without the 850 μm point. This is due to the fact that the BMBB model fits the fluxes at longer wavelengths (500 and 850 μm point)

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

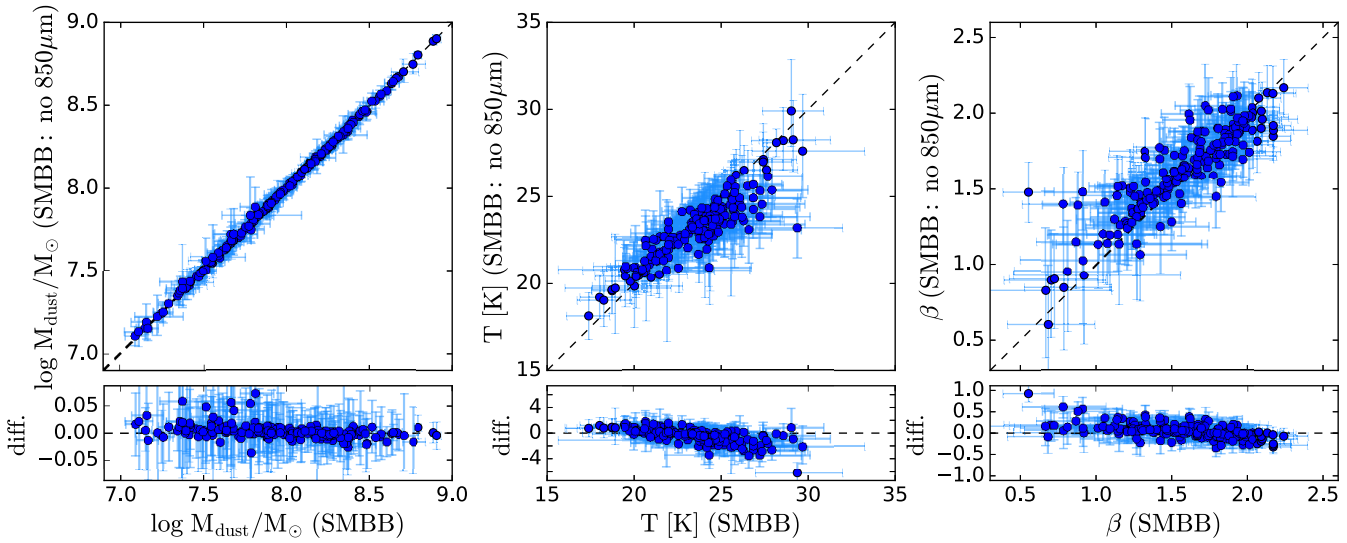

Figure 6. Comparison of the dust masses, temperatures, and emissivity indices obtained through the fit of an SMBB using the hierarchical approach, with and without the SCUBA-2 flux point at 850 μm. The lower panels show the difference between fit with and without the SCUBA-2 flux in each of the derived properties.

using a second emissivity index β_2, thus the measurements of T and β_1 are not sensitive to the flux measurement at 500 μm and 850 μm. Fig. 8 shows an example of the SMBB and BMBB fit of one galaxy for which the difference in temperature is more evident (JINGLE 1). This model is especially useful to quantify the possible sub-mm excess, given by the difference between the two emissivity indices β_1 and β_2. Further discussion on the sub-mm excess is presented in Section 5.2.

4.3.2 TMBB

The bottom panels of Fig. 7 show the comparison of the SMBB and the two MBB models (TMBB). The dust masses are in good agreement, with the cold dust masses derived from the TMBB being slightly higher (median offset: 0.03 dex).

The dust temperatures of the cold component obtained with the TMBB model tend to be lower than the ones measured from the SMBB model by about 3 per cent (or 0.8 K). This is expected, since the warm component is contributing to the fit of the 100 μm flux, allowing the cold component to shift to longer wavelengths, corresponding to colder temperatures. Consequently, the $\beta_{\rm cold}$ values from the TMBB are also slightly higher (median offset: 0.05). The outlier is JINGLE 33 (Fig. 9). This galaxy has a high 60 μm flux, compared to the 100 μm flux, which results in the warm dust component (with $T_{\rm warm} = 52.3$ K) reproducing most of the emission, and skewing the cold dust component to a lower temperature ($T_{\rm cold} = 17.2$ K) and a higher dust mass.

The warm dust component does not contribute much to the entire dust mass. Warm dust masses are in the range $10^{3.4}$–$10^{6.6}$ M$_\odot$, which correspond to only 0.01–4.4 per cent of the total dust mass of the galaxies. Nevertheless, it is important to take into account this component because, as we have shown, it can affect the measurement of the temperature and emissivity β of the cold component. The temperatures of the warm component are in the range 66−76 K, with the exception of JINGLE 33 which has a lower temperature (52.3 K).

If we compare the total dust masses ($M_{\rm dust,tot} = M_{\rm cold} + M_{\rm warm}$) from the TMBB with the cold dust masses $M_{\rm cold}$ from the SMBB, the latter are smaller by 10 per cent (∼0.08 dex) on average. Other studies found that fitting the SED using the TMBB model will result in higher cold dust masses. For example Gordon et al. (2014) found that the dust masses of the Small and Large Magellanic Clouds are 6–15 times larger when estimated using a TMBB model instead of the SMBB model. Clark et al. (2015) found that the warm dust mass can contribute up to 38 per cent of the total dust mass of galaxies in the *Herschel*-ATLAS survey. The disagreement with our findings is probably due the fact that these studies do not include the 22 μm flux point in their fit. Consequently, their warm component is shifted to longer wavelength and has lower temperature than ours, thus contributing more to the total dust mass. The cold dust temperature of the TMBB will also be smaller than in the SMBB case, thus resulting in higher cold dust masses.

4.4 Model comparison with information criterion

In order to decide which of the models provides a better fit to the data, we applied a criterion based on the comparison of the likelihoods. We consider the Bayesian Information Criterion (BIC) (Schwarz 1978) that takes into account not only the likelihood of the fit, but also the number of free parameters of the models. The latter point is important, since increasing the number of free parameters would generally lead to better fits. The BIC (Schwarz 1978) is defined as

$$\mathrm{BIC} = -2 \cdot \ln(L) + q \cdot \ln(m), \tag{24}$$

where L is the likelihood (i.e. the probability of the data given the parameter $p(\boldsymbol{F}|\boldsymbol{\theta})$), q is the number of free parameters of the model, and m is the number of data points (wavebands). The model with the lowest BIC value is the preferred model according to this criterion. To calculate the likelihood L_i for the ith galaxy we consider the product of the likelihood $p(F^{\rm obs}_{i,j}|\boldsymbol{\theta_i}, F^{\rm err}_{i,j}, \delta_j)$ in all wavebands $j =$

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

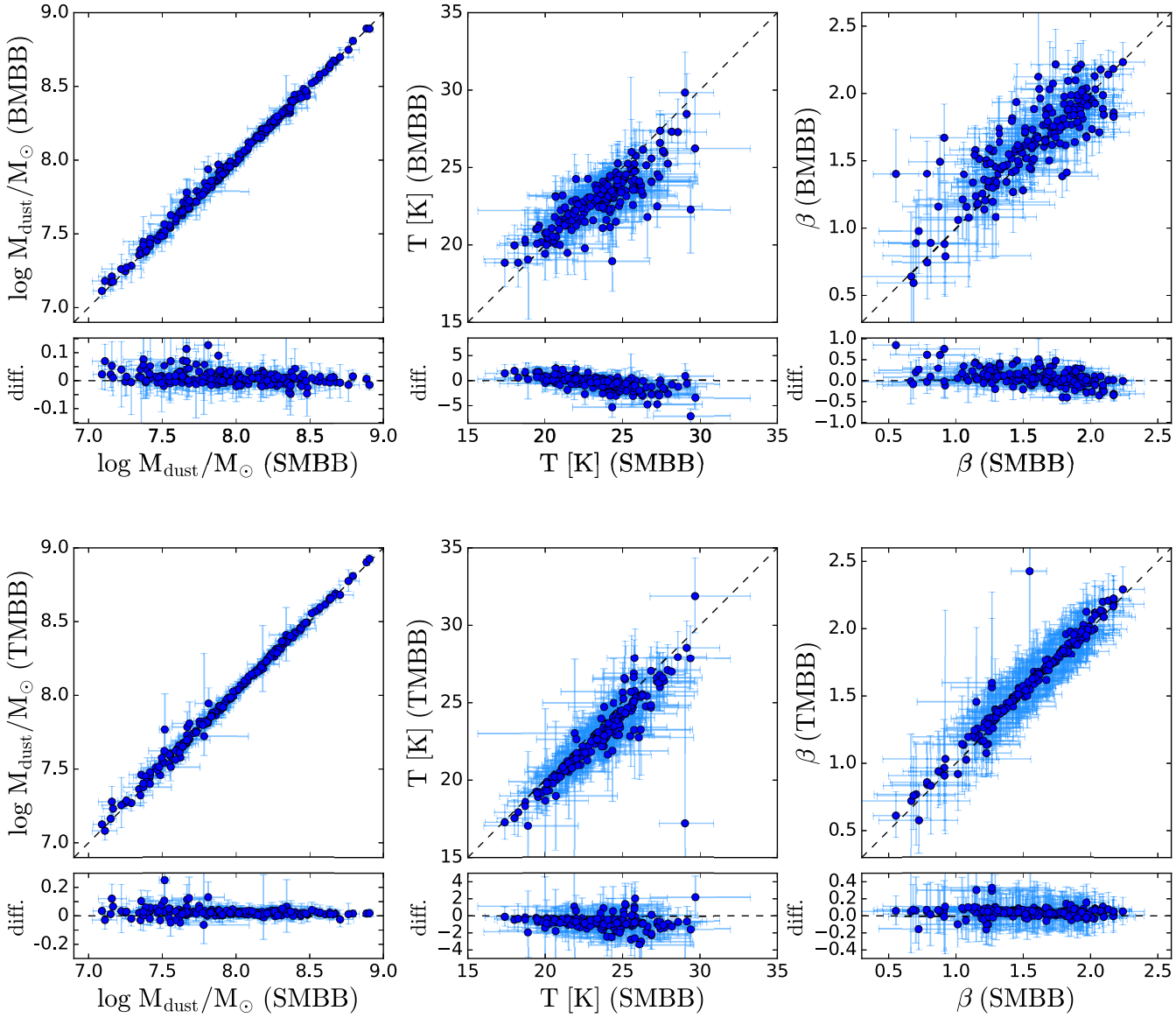

Figure 7. *Upper panels*: Comparison of the cold dust masses, temperatures, and emissivity index obtained through the fit of an SMBB and a broken emissivity power-law MBB model (BMBB). For the BMBB model, the β value shown in the plot is β_1, i.e. the emissivity index at wavelength $< \lambda_{\rm break}$. The lower sub-panels show the difference between the two models in each of the derived properties. *Bottom panels:* Comparison of the results from the SMBB and TMBB models. For the TMBB model, the values shown in the plot are the parameters of the cold component ($\log M_{\rm cold}$, $T_{\rm cold}$, $\beta_{\rm cold}$).

$1, \ldots, m$.

$$L_i = \prod_{j=1}^{m} p(F_{i,j}^{\rm obs} | \boldsymbol{\theta_i}, F_{i,j}^{\rm err}, \delta_j). \quad (25)$$

Fig. 10 shows the BIC values for the BMBB and TMBB models compared to the SMBB model. For most of the galaxies (180/192, 94 per cent), the TMBB model is preferred. This is probably due to the fact that the additional warm component can help to improve the fit at 100 μm, without affecting the fit of the points at longer wavelengths.

For seven galaxies the preferred model is the BMBB model (JINGLE ID: 35, 56, 77, 101, 118, 133, and 147). In all these galaxies there is a clear sub-mm excess at 850 μm. The BIC criterion does not identify all galaxies for which the 850 μm flux is enhanced with respect to the SMBB model, but selects the ones for which the discrepancy cannot be attributed to flux uncertainties or uncertainties in the model.

There are five galaxies that are best modelled with the SMBB model (JINGLE ID 83, 110, 142, 159, and 186). The TMBB model is not able to fit well the 60 and 100 μm flux points of these galaxies. For JINGLE 83 and JINGLE 159, the 60 μm flux is too low and is not well fitted by the TMBB model. For JINGLE 110 the 60 μm, flux is instead too high compared to the 100 μm flux. For JINGLE 186, the uncertainty on the 60 μm flux is very small, and therefore even a small deviation from the perfect fit of that data point results in a low likelihood. In JINGLE 142, the 500 μm point is enhanced with respect to the 350 μm flux point and the 850 μm upper limit. In general neither the SMBB and TMBB models are able to produce a good fit for this galaxy. The SED fits with the BMBB and TMBB models for all galaxies are shown in Fig. D2.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

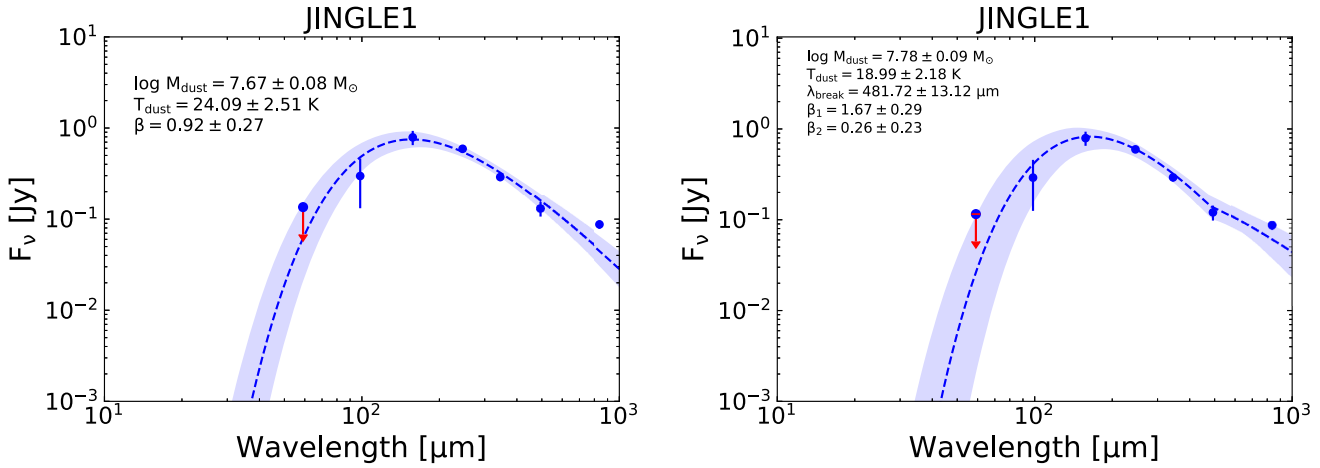

Figure 8. SMBB and BMBB fit for the galaxy JINGLE 1, where there is a clear difference in the dust temperature measured with the two different models.

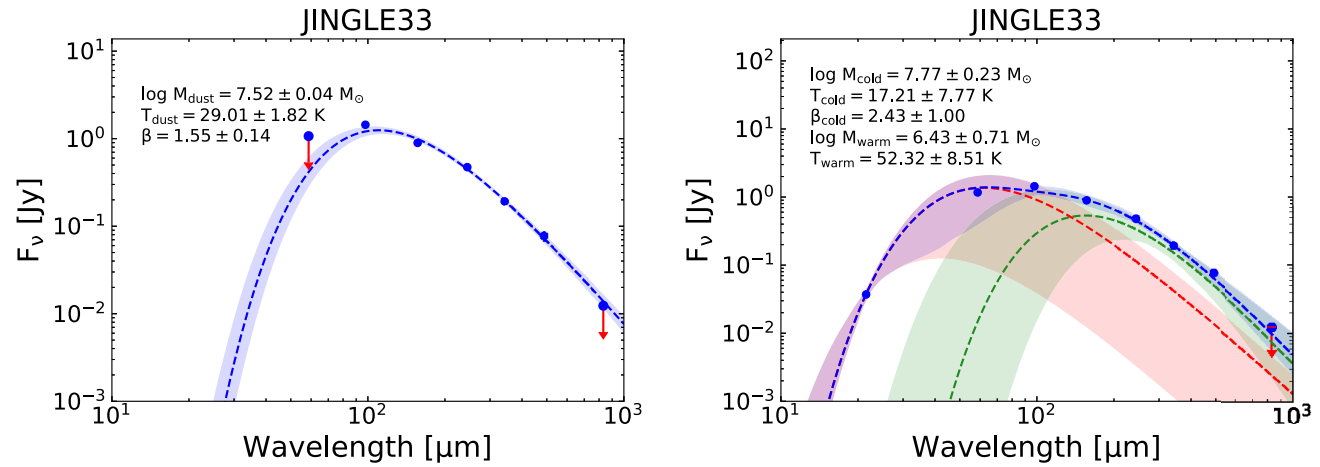

Figure 9. SMBB and TMBB fit for the galaxy JINGLE 33, which shows a clear difference in the cold dust mass measured with the two different methods. The warm component has a large contribution to the total dust emission in this galaxy.

We conclude that the TMBB model produces the best fit of the FIR SED for most of the galaxies. Additionally, the comparison of the BIC of the SMBB and BMBB model can be used to identify galaxies which show a strong sub-mm excess or deficit.

5 RELATION BETWEEN DUST PROPERTIES AND GALAXY PROPERTIES

In this section, we investigate how dust properties correlate with global galaxy properties. We use the results obtained using the SMBB model, even though the TMBB model is preferred according to the BIC. We decide to use the SMBB model because one of the goals of this analysis is to provide prescriptions to estimate T and β from other galaxy quantities. These prescriptions can be useful in those cases where only a few photometric data points are available and in such cases it is preferred to use the model with the smallest number of free parameter (i.e. the SMBB model). Additionally, as we have shown in the previous section, the differences in T and β derived from the SMBB and the TMBB models are not very large and they are mainly systematic shifts, that can be accounted for.

We include in this analysis also the galaxies from the HRS (Boselli et al. 2010), which allow us to extend the parameter range to lower SFR and specific SFR (SSFR), since a large fraction of the HRS sample are galaxies that lie below the star formation main sequence (see Fig. 1). In this case, the total sample of galaxies consists of two populations: star-forming galaxies (main-sequence galaxies) and passive galaxies (below main-sequence). Therefore, the basic assumption for the use of the hierarchical method that all galaxies belong to the same population does not hold any more. We therefore divide the 'total' sample (JINGLE + HRS) into two sub-samples according to their position in the SFR–M_* plane and fit each separately. In this way, the assumption that the galaxies in one sub-sample belong to the same population is still valid. We define the two sub-samples as follows:

(i) *Main-sequence galaxies/star-forming sample*: Galaxies belonging to the SF main sequence or laying above it. This sample consists of all galaxies which fall above the lower limit of the SF main sequence, defined as 0.4 dex below the SF main sequence from Saintonge et al. (2016).

(ii) *Below main-sequence sample/passive sample*: Galaxies laying below the SF main sequence. These are the galaxies which lie more than 0.4 dex below the SF main sequence defined by Saintonge et al. (2016).

The star-forming sample consists of 313 galaxies (177 from JINGLE and 136 from HRS) and the passive sample of 159 galaxies

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

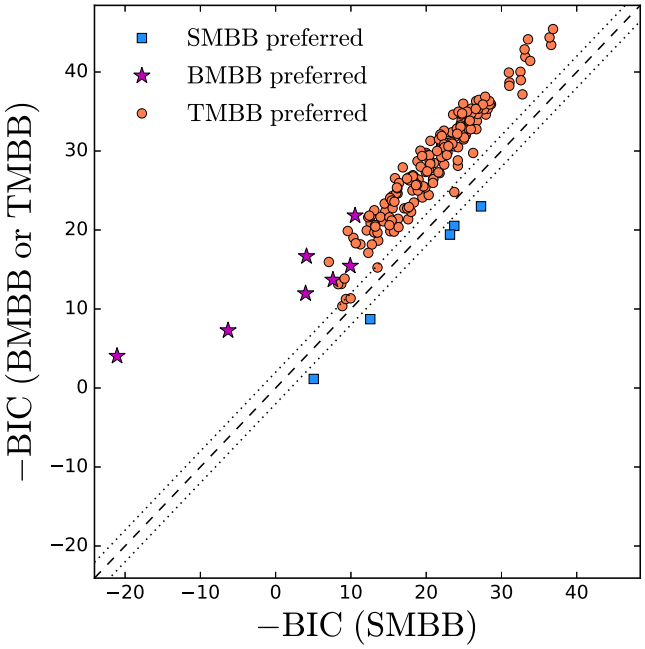

Figure 10. Comparison of the negative Bayesian Information Criterion (−BIC) for the fit using the three models: SMBB, BMBB, TMBB. The model with the largest value of −BIC is the preferred model. If the difference between the BICs is smaller than 2 (shown by the dotted lines), there is little evidence to prefer one model over another.

(15 from JINGLE and 144 from HRS). We did a test fitting galaxies belonging to the two sub-samples together. This test confirms that it is necessary to separate the sample in two, to avoid to force the two sub-samples to move towards a common mean, introducing systematic biases in the results.

Fig. 11 shows the galaxies on the SFR–M_* plane, colour-coded by dust temperature T and emissivity index β. The dust temperature increases when moving from the bottom-right corner (high M_*, low SFR) to the upper-left corner (low M_*, high SFR). The emissivity β instead tends to increase with M_*. From this figure we can already see that T and β are related to different galaxy properties, with T varying depending on the SSFR and β on the stellar mass.

We quantify the strength of these relations by calculating the correlation coefficients between T, β and the following quantities: stellar mass, stellar mass surface density ($\mu_* = M_*/(2\pi R_{50}^2)$, where R_{50} is the optical half-light radius in the i band from SDSS), metallicity (12 + log(O/H), using the O3N2 calibration of Pettini & Pagel 2004), H I mass fraction ($M_{\rm H\,I}/M_*$), SFR, SSFR, SFR surface density ($\Sigma_{\rm SFR}$), and SFR divided by dust mass. We consider all quantities in log space.

We calculate the Pearson correlation coefficient R and perform a linear fit when the absolute value of the correlation coefficient is higher than 0.4, both for the total sample and for the JINGLE and HRS samples separately. We did the fit also for the two samples separately to see whether there are differences in the correlations derived using JINGLE or HRS. We apply a correction to account for the fact that the stellar mass distribution of our sample does not exactly represent the stellar mass distribution in the local Universe, using the method developed for the xCOLD GASS survey (Saintonge et al. 2017). We compare the mass distribution of our sample, in bins of 0.1 dex in $\log M_*$, to the expected mass distribution of a volume-limited sample based on the stellar mass function from Baldry et al. (2012). For each mass bin, we calculate the ratio between the normalized number of galaxies in our sample and in the mass distribution from Baldry et al. (2012). We apply this ratio as a statistical weight when we fit the dust scaling relations. The correlation coefficients and parameters of the linear fits are summarized in Table 4.

We find that the emissivity β shows a positive correlation with $\log M_*$ (Pearson correlation coefficient $R = 0.58$), $\log \mu_*$ ($R = 0.62$), and metallicity ($R = 0.58$). Since these galaxy properties are all correlated with each other, it is not surprising that they all correlate with β. These trends were already observed by Cortese et al. (2014) in the HRS sample. They also observed negative correlations of these quantities with dust temperature T, due to the fact that they used a non-hierarchical method for the fitting and therefore they could not break the degeneracy between T and β. Thus, they were not able to distinguish whether the fundamental physical correlations were driven by the temperature or by the emissivity index. In our analysis, these three quantities do not show a strong anticorrelation with temperature ($-0.29 \leq R \leq -0.19$). We note that for the JINGLE galaxies the metallicities are measured from the SDSS fibre spectra and therefore represent only the metallicities in the central 3 arcsec of the galaxies. For the HRS sample, metallicities are measured from long-slit integrated optical spectra (Boselli et al. 2013; Hughes et al. 2013), and thus represent better the global metallicities of the galaxies. Indeed we find that the correlation between β and metallicity is higher ($R = 0.67$ if we consider only the HRS sample. We also find an anticorrelation between β and the H I mass fraction ($R = -0.65$), that was already observed in Cortese et al. (2014). In this case, the H I mass fraction shows a weaker correlation with dust temperature ($R = 0.41$). The H I mass fraction is known to correlate with the inverse of the stellar mass surface density and with SSFR (Catinella et al. 2013). Thus, it is expected to see an anticorrelation with β and a positive correlation with T, due to the correlation of SSFR with T.

The dust temperature correlates with log SSFR ($R = 0.54$), $\log \Sigma_{\rm SFR}$ ($R = 0.49$), and $\log {\rm SFR}/M_{\rm dust}$ ($R = 0.73$). These correlations have already been observed by Clemens et al. (2013) and Cortese et al. (2014). As stated in Clemens et al. (2013), the fact the cold dust temperature correlates with SFR surface density but not with stellar mass surface density suggests that the cold dust is heated more by ongoing star formation or by young stars. Also Kirkpatrick et al. (2014) observed a correlation between cold dust temperature and SFR normalized by the 500 μm luminosity, that is a proxy for the dust mass, on spatially resolved scales in galaxies from the KINGFISH sample (Kennicutt et al. 2011). According to their work, this correlation suggests that the number of photons from young stars relative to the amount of dust has an important heating effect on the diffuse cold dust component. Moreover, Galametz et al. (2012) studied a sub-sample of galaxies from the KINGFISH sample and observed that the higher dust temperatures coincide with the centre of star-forming regions, showing a connection between dust temperature and star formation.

The temperature of the dust is regulated by the radiation from star formation, weighted by the amount of dust present in the galaxy. The relation between T and SSFR shows more scatter at low SSFR. This may be related in part to the fact that SFR measurements are less accurate for low SSFR ($\log {\rm SSFR} < -10.6$; Hunt et al. 2019). Also it is likely that the contribution of the older stellar population to the dust heating is higher in low SSFR galaxies, since the star formation is weak and the contribution from old stars can be more significant.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

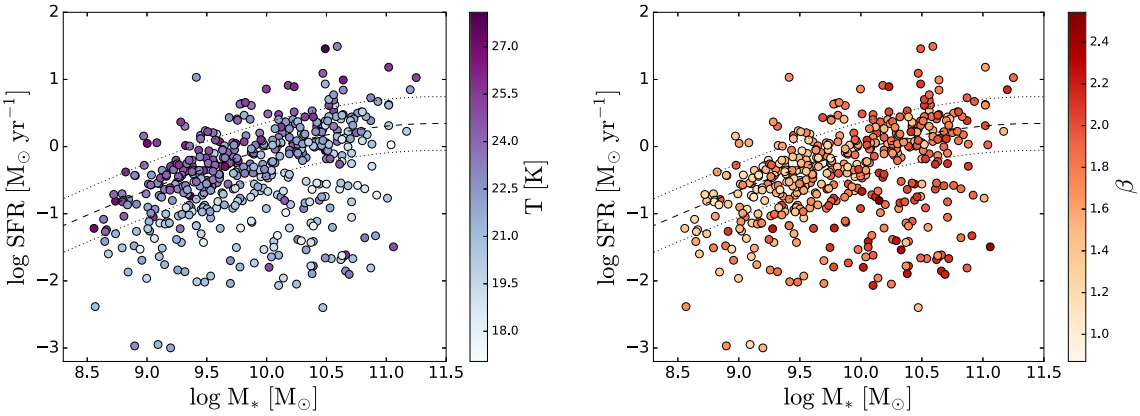

Figure 11. Distribution of the JINGLE and HRS sample in the SFR–M_* plane, colour-coded by dust temperature (left) and emissivity index β (right). Dust temperatures and β are measured using the SMBB model and the hierarchical SED fitting approach. The position of the star formation main sequence (Saintonge et al. 2016) is shown as a dashed lines, the 0.4 dex dispersion is shown by dotted lines.

5.1 Primary correlation analysis

In this section, we investigate which are the primary parameters driving the correlation with dust properties. This analysis has two goals: (1) to provide prescriptions to estimate the temperature T and the emissivity index β of the dust from other galaxy properties and (2) to understand which are the physical quantities that influence and set T and β in a galaxy.

We perform a Bayesian inference analysis to find the best combination of parameters that can be used to estimate the dust properties. We consider the galaxy parameters which, alone or combined, show some correlation with β and T: stellar mass M_*, SFR, dust mass, metallicity, and surface area ($A = 2\pi R_{50}^2$, where R_{50}^2 is the optical half-light radius in the i band from SDSS in kpc). The surface area is used to calculate, for example, the SFR and stellar mass 'surface density'. We fit first-order polynomial models with a different number of parameters, exploring all possible combinations of parameters. The number of possible combination of k parameters selected from a total sample of n parameters is $C_{n,k} = \frac{n!}{k!(n-k)!}$. We use a first-order polynomial model in log space:

$$Q_{\rm model}(x_1, \ldots, x_k) = \sum_{j=1}^{k} a_j \log(x_j) + b, \tag{26}$$

where k is the number of galaxy properties x_j considered, and $Q_{\rm model}$ is the value of the dust quantity (T or β) approximated by the model. We use a Bayesian inference method to determine the optimal number of parameters needed to fit the data and the best-fitting relations. We model the probability of observing our data, given the model and the uncertainties, as a normal distribution:

$$p(Q_i|Q_{{\rm model},i}(x_{1,i}, \ldots, x_{k,i}), Q_{err,i}) = w_i \cdot {\rm Normal}(Q_{{\rm model},i}, Q_{{\rm err},i}), \tag{27}$$

for each galaxy i in our sample, where w_i is the weight correcting for the flat M_* distribution (see Section 5). We consider only the uncertainties on the dust quantity Q_i, but not on the galaxy properties $x_{j,i}$. We make this choice because we want to minimize the difference between Q_i and $Q_{{\rm model},i}$, given the quantities $x_{j,i}$. We perform an MCMC fit using STAN to find the best-fitting parameters and measure the likelihood of the different models. Then we apply the BIC to find the optimal number of parameters and the best model.

We consider first the models to estimate β. According to the BIC, the preferred model has five parameters: stellar mass, surface area, metallicity, SFR, and H I mass. The best-fitting relation is given by

$$\begin{aligned}\beta_{\rm model} = {} & 0.26^{+0.03}_{-0.03} \cdot \log M_* - 0.27^{+0.03}_{-0.03} \cdot \log {\rm area} \\ & + 0.60^{+0.09}_{-0.09} \cdot \left[12 + \log({\rm O/H})\right] + 0.18^{+0.03}_{-0.03} \cdot \log {\rm SFR} \\ & - 0.23^{+0.03}_{-0.03} \cdot \log M_{\rm HI} - 3.54^{+0.82}_{-0.84}. \end{aligned} \tag{28}$$

This model includes five parameters, several of which are known to be correlated, therefore it is difficult to know which one is more fundamentally related to β. To assess this, we measure the increase in R^2 that each parameter produces when it is added to a model that contains already all other parameters. This change represents the amount of variance that can be explained by each parameter and that is not explained by the other variables. We measure R^2 ($0 < R^2 < 1$) as the squared Pearson correlation coefficient between the dust parameter (β or T) and the 'modelled' parameter ($\beta_{\rm model}$, $T_{\rm model}$), i.e. the parameter estimated by the linear combination of galaxy properties. Table 5 shows the results. From the analysis of the increase of the R^2, we can see that the most fundamental parameter determining β is the stellar mass (increase in R^2: $\Delta R^2 = 11.2$ per cent). The second one is the surface area ($\Delta R^2 = 8.0$ per cent). Since they have opposite coefficients in the fit with almost the same magnitude (0.26 ± 0.03 for M_* and -0.27 ± 0.03 for the surface area), this can be interpreted as the stellar mass surface density correlating with β. If we consider stellar mass and surface area combined as a single parameter in the analysis, the increase in R^2 due to stellar mass surface density is $\Delta R^2 = 17.9$ per cent. The following parameter in order of importance is the metallicity ($\Delta R^2 = 7.1$ per cent). SFR and H I mass cause a smaller increase in R^2 ($\Delta R^2 = 5.0$ per cent and 5.7 per cent, respectively), and the dust mass has a negligible contribution ($\Delta R^2 = 0.5$ per cent).

Smith et al. (2012b) studied the variation of β in M31 (Andromeda). They found that β decreases with galactocentric radius. Since also the stellar mass surface density, μ_*, in M31 decreases with radius (Tamm et al. 2012), their result is consistent with a correlation between β and μ_*. Köhler et al. (2015) found that the emissivity index of grains evolve from lower to higher β values when transitioning from diffuse to denser ISM due to grain coagulations. If the stellar mass density is related to the density of the ISM, this could explain the relation between β and the stellar mass surface density.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

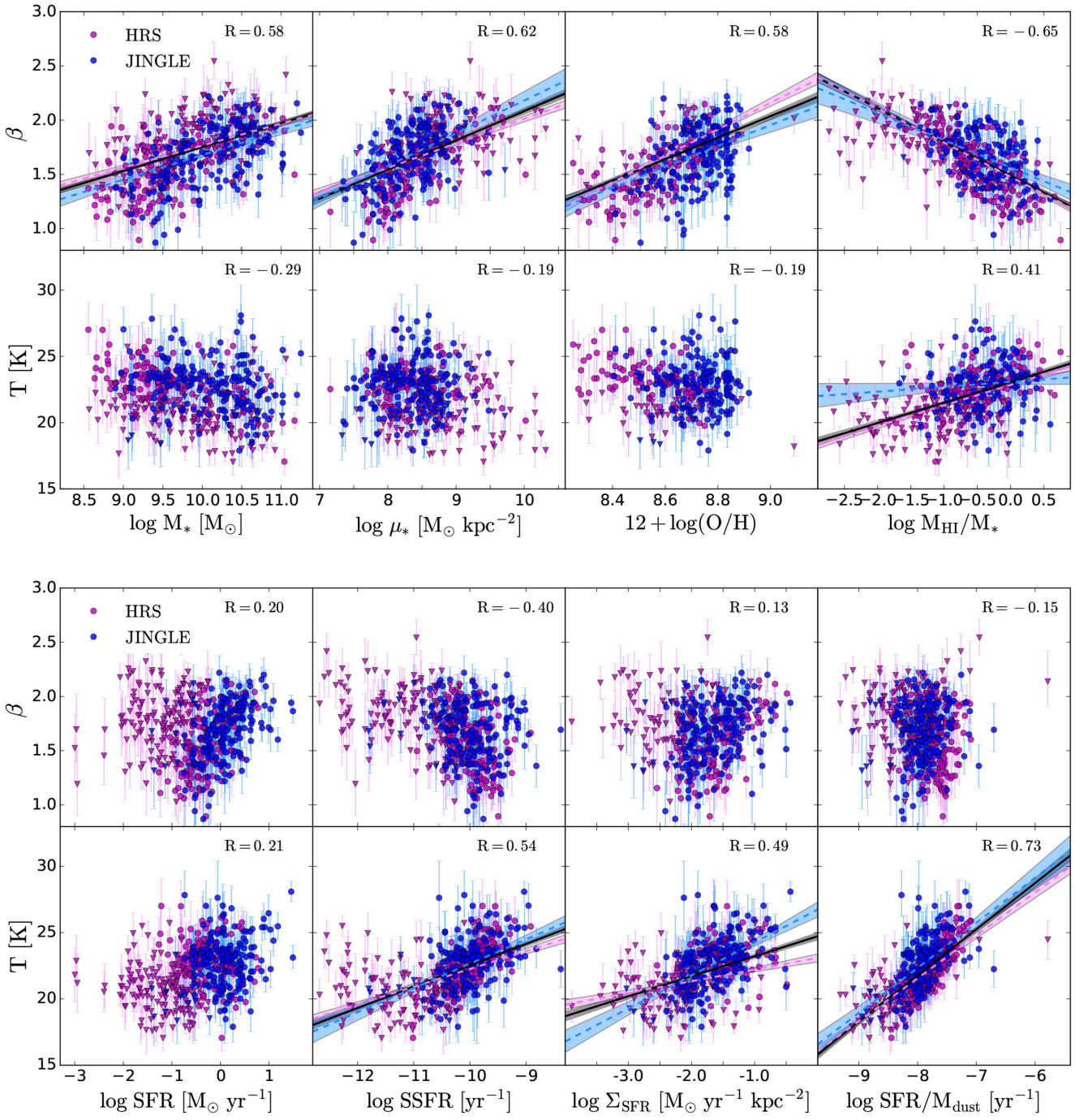

Figure 12. Dust scaling relations: correlation of dust temperature T and effective β with other global galaxy properties: stellar mass (M_*), stellar mass surface density ($\mu_* = M_*/(2\pi R_{50}^2)$, where R_{50}^2 is the optical half-light radius in the i band from SDSS in kpc), metallicity (12 + log(O/H), O3N2 calibration of Pettini & Pagel 2004), H I mass fraction ($M_{\rm H\,I}/M_*$), SFR, SSFR, SFR surface density ($\Sigma_{\rm SFR}$), and SFR over dust mass (SFR/$M_{\rm dust}$). Dust temperatures and β are measured using the SMBB model and the hierarchical SED fitting approach. The JINGLE sample is shown in blue and the HRS sample in magenta. Galaxies of the 'main-sequence' sample are shown with circles and galaxies of the 'below main-sequence' sample are shown with triangles. In every panel we show the Pearson correlation coefficient R. For the cases where $R > 0.4$, the plot shows the linear fit to the JINGLE sample (in blue), to the HRS sample (in magenta), and to the two samples together (in black).

As we have seen in the previous section, β correlates also with metallicity and with the inverse of the H I mass fraction. This indicates a relation between β and the state of evolution of a galaxy: more evolved galaxies tend to have higher metallicity and lower H I fraction. A possible interpretation of the variation of β with metallicity and H I mass fraction is related to the structure and composition of dust grains. Crystalline or carbonaceous dust is characterized by a lower β with respect to amorphous or silicate dust (Désert, Boulanger & Puget 1990; Jones et al. 2013). We expect less evolved (metal-poor) galaxies undergoing an elevated

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

Table 4. The table shows the Pearson correlation coefficient R between dust properties (dust emissivity index β and dust temperature T) and global galaxy properties. If $|R| > 0.4$ we provide the best-fitting relation (slope and intercept) between the selected galaxy property (p) and T (or β).

Properties p	Correlation with β			Correlation with T		
	R	Slope	Intercept	R	Slope	Intercept
$\log M_*$	0.58	0.23 ± 0.02	-0.60 ± 0.22	-0.29		
$\log \mu_*$	0.62	0.30 ± 0.03	-0.84 ± 0.27	-0.19		
$12 + \log(\mathrm{O/H})$	0.58	0.95 ± 0.13	-6.64 ± 1.16	-0.19		
$\log M_{\mathrm{H\,I}}/M_*$	-0.65	-0.25 ± 0.04	1.56 ± 0.02	0.41	0.38 ± 0.23	23.07 ± 0.15
log SFR	0.20			0.21		
log SSFR	-0.40			0.54	1.83 ± 0.19	41.02 ± 1.90
$\log \Sigma_{\mathrm{SFR}}$	0.13			0.49	2.49 ± 0.23	26.74 ± 0.38
$\log \mathrm{SFR}/M_{\mathrm{dust}}$	-0.15			0.73	3.40 ± 0.29	49.52 ± 2.32

Table 5. Increase in R^2 when the parameter is added to a model that already contains the other parameters.

Parameter	β Increase in ΔR^2 (per cent)	T Increase in ΔR^2 (per cent)
$\log M_*$	11.2	0.5
log SFR	5.0	80.0
log area	8.0	2.4
	7.1	1.5
$12 + \log(\mathrm{O/H})$		
$\log M_{\mathrm{dust}}$	0.5	13.6
$\log M_{\mathrm{H\,I}}$	5.7	0.5

period of star formation activity to produce a lot of dust in stars (Zhukovska 2014), and this dust has a more crystalline structure at the beginning (Waelkens et al. 1996; Waters et al. 1996; de Vries et al. 2010) and tends to become more 'amorphous' with time (e.g. Demyk et al. 2001). Therefore, more evolved galaxies can be expected to have more amorphous dust and higher β. Additionally, silicate dust is thought to survive for a longer time compared to carbon dust (e.g. Jones & Nuth 2011). Thus, we expect dust in a more evolved galaxy to have a larger fraction of silicate grains that are associated with higher values of β. Another possible explanation for the relation between β and metallicity is the observation that the abundance of carbon stars, which produce carbon dust, decreases at high metallicities (Boyer et al. 2019). Thus, we can expect high-metallicity galaxies to have less carbonaceous dust and consequently a higher β.

Another possibility is that the low β values are due to temperature mixing. In our analysis we are not measuring directly the emissivity of dust grains but we are measuring an 'effective β', which includes both the actual emissivity of the dust and the effect of temperature mixing (e.g. Hunt et al. 2015). It has been shown that variations of the dust temperatures along the line of sight can broaden the SED and mimic the effect of a low β value (Shetty et al. 2009a). Rémy-Ruyer et al. (2015) find the SED of low-metallicity dwarf galaxies to be broader than the one of higher metallicity galaxies, consistent with our finding of lower β in low-metallicity galaxies. They explain this effect with the fact that dwarf galaxies have a clumpier ISM that produces a wider distribution of dust temperatures.

Since the preferred relation to approximate β needs a large number of parameters, we also provide the best relation with two parameters (stellar mass and surface area) and with three parameters (stellar mass, surface area, and metallicity), that are more practical to use:

$$\beta_{\mathrm{model}} = 0.42^{+0.02}_{-0.02} \cdot \log M_* - 0.37^{+0.03}_{-0.03} \cdot \log \mathrm{area} - 1.97^{+0.18}_{-0.18}, \quad (29)$$

$$\beta_{\mathrm{model}} = 0.28^{+0.03}_{-0.03} \cdot \log M_* - 0.38^{+0.03}_{-0.03} \cdot \log \mathrm{area} + 0.80^{+0.09}_{-0.09} \cdot \left[12 + \log(\mathrm{O/H})\right] - 7.48^{+0.64}_{-0.67}. \quad (30)$$

A summary with the best relations for every number k of parameters can be found in Table 6.

We perform a similar analysis to investigate which combination of parameters gives the better approximation of the dust temperature T. According to the BIC, the preferred model has three parameters: SFR, dust mass, and metallicity (BIC = 848.8). Also the two-parameter model with SFR and dust mass has a similar BIC (BIC = 849.6), meaning that adding the metallicity parameter has only a small effect on improving the correlation. This confirms our previous finding that dust temperature correlates strongly with SFR per unit dust mass. The R^2 analysis gives the same result: the most important parameter is clearly the SFR ($\Delta R^2 = 87.9$ per cent), with a secondary dependence on the dust mass ($\Delta R^2 = 16.6$ per cent). The other four parameters have a very small effect ($\Delta R^2 < 3$ per cent).

This relation is however of limited practical interest since it requires prior knowledge of the dust mass. Therefore, we consider also the two-parameter model with the best BIC that do not include $\log M_{\mathrm{dust}}$ as a parameter. The two-parameter model uses SFR and stellar mass ($R = 0.50$):

$$T_{\mathrm{model}} = 2.50^{+0.22}_{-0.22} \cdot \log \mathrm{SFR} - 2.14^{+0.20}_{-0.19} \cdot \log M_* + 44.24^{+1.93}_{-2.02}. \quad (31)$$

Tables for T and β with all the relations with two or three parameters are in the appendix (Tables E4 and E5).

5.2 Sub-mm excess

In this section, we discuss the behaviour of the SED at long wavelengths ($\lambda > 500\,\mu$m). In particular, we are interested in galaxies that show a so-called sub-mm excess. An excess at sub-mm wavelength has been observed in dwarf galaxies (Lisenfeld et al. 2002; Galliano et al. 2003), in late-type galaxies (Dumke et al. 2004; Bendo et al. 2006; Galametz et al. 2009), and in the Magellanic Clouds (Bot et al. 2010; Israel et al. 2010). The most significant excesses cannot be explained by contribution from synchrotron,

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

Table 6. Results of the correlation analysis to derive an expression to approximate the emissivity β and the dust temperature using global galaxy properties. The table shows the coefficients a_j of the best polynomial expression ($Q_{\rm model}(x_1, \ldots, x_k) = \sum_{j=1}^{k} a_j \log(x_j) + b$) to estimate β and T using a different number of parameters k. The table also shows the BIC and the Pearson correlation coefficient R between the dust parameter (β or T) and the 'modelled' parameter ($\beta_{\rm model}$, $T_{\rm model}$), i.e. the parameter estimated by the linear combination of galaxy properties.

Emissivity index β									
Parameters	$\log M_*$ ($M_\odot$)	log SFR ($M_\odot$ yr^{-1})	log area (kpc^2)	12 + log(O/H)	$\log M_{\rm dust}$ ($M_\odot$)	$\log M_{\rm H\,I}$ ($M_\odot$)	Intercept	BIC	R
$k = 1$				0.98 ± 0.06			-6.77 ± 0.59	170.56	0.61
$k = 2$	0.42 ± 0.02		-0.37 ± 0.03				-1.97 ± 0.18	53.19	0.64
$k = 3$	0.28 ± 0.03		-0.38 ± 0.03	0.80 ± 0.09			-7.48 ± 0.64	-14.37	0.70
$k = 4$	0.33 ± 0.03		-0.29 ± 0.03	0.69 ± 0.10		-0.13 ± 0.03	-5.92 ± 0.69	-27.87	0.71
$k = 5$	0.26 ± 0.03	0.18 ± 0.03	-0.27 ± 0.03	0.60 ± 0.09		-0.23 ± 0.03	-3.54 ± 0.82	-54.40	0.73
$k = 6$	0.31 ± 0.04	0.23 ± 0.04	-0.25 ± 0.04	0.66 ± 0.10	-0.13 ± 0.08	-0.20 ± 0.04	-3.84 ± 0.89	-51.19	0.73
Temperature									
Parameters	$\log M_*$ ($M_\odot$)	log SFR ($M_\odot$ yr^{-1})	log area (kpc^2)	12 + log(O/H)	$\log M_{\rm dust}$ ($M_\odot$)	$\log M_{\rm H\,I}$ ($M_\odot$)	Intercept	BIC	R
$k = 1$		0.65 ± 0.13					22.93 ± 0.08	1024.78	0.15
$k = 2$		4.19 ± 0.29			-3.73 ± 0.30		51.88 ± 2.20	849.60	0.68
$k = 3$		4.06 ± 0.29		-1.85 ± 0.75	-3.31 ± 0.31		64.7 ± 5.44	848.76	0.68
$k = 4$		3.93 ± 0.31	-0.66 ± 0.28	-2.24 ± 0.79	-2.71 ± 0.41		64.13 ± 5.67	849.08	0.69
$k = 5$	0.36 ± 0.35	3.99 ± 0.32	-0.63 ± 0.29	-2.36 ± 0.76	-3.08 ± 0.57		64.86 ± 5.82	853.77	0.69
$k = 6$	0.29 ± 0.39	4.01 ± 0.33	-0.58 ± 0.30	-2.59 ± 0.81	-2.86 ± 0.64	-0.23 ± 0.29	67.87 ± 7.28	859.06	0.70

free–free or molecular line emission (e.g. Galliano et al. 2003). Different explanations proposed to explain this phenomenon are for example the presence of a very cold dust component, a temperature-dependent emissivity (Meny et al. 2007), and the presence of rotating or magnetic grains (Draine & Hensley 2012).

We identify the galaxies with an excess at 850 μm with respect to the SMBB model, taking into account uncertainties on the SCUBA-2 fluxes and on the SMBB model:

$$F^{\rm obs} - F^{\rm model} > F^{\rm obs}_{\rm err} + F^{\rm model}_{\rm err}. \quad (32)$$

There are 27/192 galaxies (14 per cent) that satisfy this criterion. If we adopt a more stringent criterion, requiring the galaxy to have an excess above 2σ (i.e. $(F^{\rm obs} - F^{\rm model}) > 2 \cdot F^{\rm obs}_{\rm err}$), we find that 24 galaxies (12 per cent) satisfy this criterion. From a normal distribution, we would expect to find only 2.5 per cent of the galaxies with an excess above 2σ, thus we think that it is a statistically significant result. The galaxies with sub-mm excess do not appear to be in a particular region of the SFR–M_* plane (see Fig. 13). There also some galaxies that show a deficit at 850 μm.

A weak point of this analysis is that the sub-mm excess is determined only by a single point, the 850 μm SCUBA-2 flux. Therefore, the presence of an excess can also be due to a number of factors including measurement errors, uncertainties on the apertures, contamination by other sources, and uncertainties on the CO(3–2) contribution. In order to better characterize and quantify the sub-mm excess, additional flux points at longer wavelengths are needed. We plan to investigate this in the future. We have an accepted proposal to observe 18 JINGLE targets at 1 and 2 mm with NIKA-2 on the IRAM-30 m telescope. With two additional flux points we will be able to characterize better the sub-mm excess and to test different models proposed to account for it.

6 CONCLUSIONS

In this paper, we analyse a sample of 192 star-forming galaxies from the JINGLE survey. We also include in the analysis 323 galaxies from the HRS to expand our analysis to galaxies with lower SSFR. We fit their far-infrared/sub-mm SED with MBB models using a hierarchical Bayesian approach that allows to reduce the degeneracy between parameters, especially between dust temperature and emissivity index β. We consider three models: SMBB, TMBB, and MBB with a broken emissivity law (BMBB).

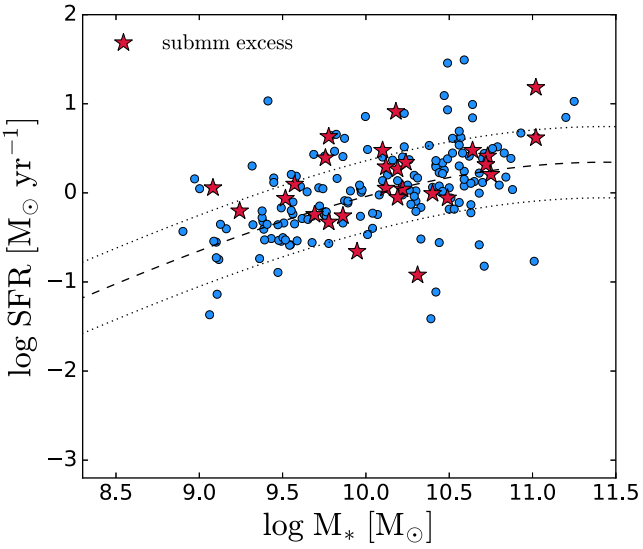

Figure 13. JINGLE galaxies with a sub-mm excess are shown by red stars symbols, while the JINGLE sample is shown in light blue. The position of the star formation main sequence (Saintonge et al. 2016) is shown as a dashed lines, the 0.4 dex dispersion is shown by dotted lines.

The main results of our study are as follows:

(i) *Dust masses*: The choice of the model (SMBB, BMBB, or TMBB) has only a small effect on the dust mass estimates. The cold dust masses measured with the TMBB are larger than the ones measured by the SMBB by only 0.04 dex on average, and the dust masses measured with the BMBB model agree very well with the SMBB results.

(ii) T–β relation: The use of the hierarchical Bayesian approach to fit the FIR SED is crucial to infer the intrinsic relation between

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

dust temperature and dust emissivity index β. In the JINGLE sample, the anticorrelation between T and β is reduced when we use the hierarchical approach ($R = -0.52$) with respect to the non-hierarchical result ($R = -0.79$). Using the hierarchical approach, both T and β span smaller ranges ($17\,\mathrm{K} < T < 30$ K, $0.6 < \beta < 2.2$) with fewer outliers.

(iii) *Dust scaling relations:* The hierarchical approach is able to reduce the degeneracy between T and β and to separate their relations with other galaxy properties. We find that the dust emissivity index β correlates with stellar mass surface density, metallicity and anticorrelates with H I mass fraction ($M_{\rm H\,I}/M_*$). The strongest relation is with stellar mass surface density. The dust temperature correlates with H I mass fraction, SSFR, SFR surface density, and SFR per unit dust mass. The strongest relation is with SFR per unit dust mass. These relations can be used to estimate the dust temperature or emissivity index in galaxies where insufficient data prevents determining them directly through SED fitting.

(iv) *Sub-mm excess*: We observe an excess at 850 μm with respect to the flux predicted from the SMBB fit in 26/192 (14 per cent) galaxies, but we do not find these galaxies to lie in a particular region in the stellar mass–SFR plane. Additional flux points at longer wavelengths are needed to better characterize the sub-mm excess and to investigate its origin.

The dust scaling relations derived in this work based on low-redshift galaxies show that dust properties correlate with global galaxy properties. After calibrating these relations with data at higher redshift, they could be applied to the study of high-redshift galaxies. Thanks to ALMA it is now possible to detect dust emission in galaxies at redshifts as high as $z > 7$ (e.g. Watson et al. 2015; Laporte et al. 2017), but the measurement of dust masses in these objects is difficult due to the scarcity of photometric points. The possibility to use scaling relations to predict what dust properties to apply in the SED modelling will increase the precision of the dust mass measurements in the early Universe, and consequently will help our understanding of dust evolution over cosmic time.

ACKNOWLEDGEMENTS

We thank Boris Leistedt, Frédéric Galliano, Luca Cortese, and Lorne Whiteway for useful discussions. AS acknowledges support from the Royal Society through the award of a University Research Fellowship. IDL gratefully acknowledges the support of the Research Foundation Flanders (FWO). CDW acknowledges support from the Natural Science and Engineering Research Council of Canada and the Canada Research Chairs program. EB acknowledges support from the UK Science and Technology Facilities Council [grant number ST/M001008/1]. MJM acknowledges the support of the National Science Centre, Poland, through the SONATA BIS grant 2018/30/E/ST9/00208. The James Clerk Maxwell Telescope is operated by the East Asian Observatory on behalf of The National Astronomical Observatory of Japan; Academia Sinica Institute of Astronomy and Astrophysics; the Korea Astronomy and Space Science Institute; the Operation, Maintenance and Upgrading Fund for Astronomical Telescopes and Facility Instruments, budgeted from the Ministry of Finance (MOF) of China and administrated by the Chinese Academy of Sciences (CAS), as well as the National Key R&D Program of China (No. 2017YFA0402700). Additional funding support is provided by the Science and Technology Facilities Council of the United Kingdom and participating universities in the United Kingdom and Canada. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

This research has made use of data from HRS project. HRS is a Herschel Key Programme utilizing Guaranteed Time from the SPIRE instrument team, ESAC scientists, and a mission scientist. The HRS data was accessed through the Herschel Database in Marseille (HeDaM – http://hedam.lam.fr) operated by CeSAM and hosted by the Laboratoire d'Astrophysique de Marseille.

This research made use of ASTROPY, a community-developed core PYTHON package for Astronomy (The Astropy Collaboration 2013), MATPLOTLIB (Hunter 2007) and NUMPY (Van Der Walt, Colbert & Varoquaux 2011). This research used the TOPCAT tool for catalogue cross-matching (Taylor 2005). This research used the STAN interface for PYTHON PYSTAN (Stan Development Team 2018). This research used the CORNER PYTHON package (Foreman-Mackey 2016).

REFERENCES

Armus L. et al., 2009, PASP, 121, 559
Baldry I. K. et al., 2012, MNRAS, 421, 621
Balog Z. et al., 2014, Exp. Astron., 37, 129
Barlow M. J., 1978, MNRAS, 183, 417
Barnard J., McCulloch R., Meng X.-L., 2000, Stat. Sin., 10
Bendo G. J. et al., 2006, ApJ, 652, 283
Bendo G. J. et al., 2013, MNRAS, 433, 3062
Bertoldi F., Carilli C. L., Cox P., Fan X., Strauss M. A., Beelen A., Omont A., Zylka R., 2003, A&A, 406, L55
Bianchi S., 2013, A&A, 552, A89
Boselli A. et al., 2010, PASP, 122, 261
Boselli A. et al., 2012, A&A, 540, A54
Boselli A., Hughes T. M., Cortese L., Gavazzi G., Buat V., 2013, A&A, 550, A114
Bot C., Ysard N., Paradis D., Bernard J. P., Lagache G., Israel F. P., Wall W. F., 2010, A&A, 523, A20
Boyer M. L. et al., 2019, ApJ, 879, 109
Carpenter B. et al., 2017, J. Stat. Softw., 76, 1
Casey C. M., 2012, MNRAS, 425, 3094
Catinella B. et al., 2013, MNRAS, 436, 34
Ceccarelli C., Viti S., Balucani N., Taquet V., 2018, MNRAS, 476, 1371
Ciesla L. et al., 2012, A&A, 543, A161
Clark C. J. R. et al., 2015, MNRAS, 452, 397
Clark C. J. R., Schofield S. P., Gomez H. L., Davies J. I., 2016, MNRAS, 459, 1646
Clemens M. S. et al., 2013, MNRAS, 433, 695
Clements D. L., Sutherland W. J., McMahon R. G., Saunders W., 1996, MNRAS, 279, 477
Cortese L. et al., 2014, MNRAS, 440, 942
Coupeaud A. et al., 2011, A&A, 535, A124
da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595
Decin L., Eriksson K., 2007, A&A, 472, 1041
Demyk K. et al., 2001, A&A, 368, L38
Demyk K. et al., 2017a, A&A, 600, A123
Demyk K. et al., 2017b, A&A, 606, A50
Desert F.-X., Boulanger F., Puget J. L., 1990, A&A, 500, 313
Désert F. X. et al., 2008, A&A, 481, 411
De Vis P. et al., 2017, MNRAS, 464, 4680
de Vries B. L., Min M., Waters L. B. F. M., Blommaert J. A. D. L., Kemper F., 2010, A&A, 516, A86
Draine B. T., Hensley B., 2012, ApJ, 757, 103
Draine B. T., Li A., 2007, ApJ, 657, 810
Duane S., Kennedy A. D., Pendleton B. J., Roweth D., 1987, Phys. Lett. B, 195, 216
Dumke M., Krause M., Wielebinski R., 2004, A&A, 414, 475
Dupac X. et al., 2003, A&A, 404, L11
Eales S. et al., 2010, PASP, 122, 499
Eales S. et al., 2012, ApJ, 761, 168

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

Eales S., de Vis P., Smith M. W. L., Appah K., Ciesla L., Duffield C., Schofield S., 2017, MNRAS, 465, 3125
Ferrara A., Viti S., Ceccarelli C., 2016, MNRAS, 463, L112
Foreman-Mackey D., 2016, J. Open Source Softw., 24,
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Galametz M. et al., 2009, A&A, 508, 645
Galametz M., Madden S. C., Galliano F., Hony S., Bendo G. J., Sauvage M., 2011, A&A, 532, A56
Galametz M. et al., 2012, MNRAS, 425, 763
Galliano F., 2018, MNRAS, 476, 1445
Galliano F., Madden S. C., Jones A. P., Wilson C. D., Bernard J. P., Le Peintre F., 2003, A&A, 407, 159
Galliano F., Galametz M., Jones A. P., 2018, ARA&A, 56, 673
Gelman A., Hill J., 2007, Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge Univ. Press, Cambridge
Gelman A., Rubin D. B., 1992, Stat. Sci., 7, 457
Gelman A., Carlin J. B., Ster H. S., Dunson D. B., Vehtari A., Rubin D. B., 2004, Bayesian Data Analysis. Chapman and Hall/CRC, Boca Raton, FL
Goodman J., Weare J., 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Gordon K. D. et al., 2014, ApJ, 797, 85
Groves B. A. et al., 2015, ApJ, 799, 96
Hermelo I. et al., 2016, A&A, 590, A56
Hildebrand R. H., 1983, Q. J. R. Astron. Soc., 24, 267
Hughes T. M., Cortese L., Boselli A., Gavazzi G., Davies J. I., 2013, A&A, 550, A115
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Hunt L. K. et al., 2015, A&A, 576, A33
Hunt L. K. et al., 2019, A&A, 621, A51
Israel F. P., Wall W. F., Raban D., Reach W. T., Bot C., Oonk J. B. R., Ysard N., Bernard J. P., 2010, A&A, 519, A67
Jarrett T. H. et al., 2011, ApJ, 735, 112
Jones A. P., Nuth J. A., 2011, A&A, 530, A44
Jones A. P., Fanciullo L., Köhler M., Verstraete L., Guillet V., Bocchio M., Ysard N., 2013, A&A, 558, A62
Juvela M., Montillaud J., Ysard N., Lunttila T., 2013, A&A, 556, A63
Kelly B. C., Shetty R., Stutz A. M., Kauffmann J., Goodman A. A., Launhardt R., 2012, ApJ, 752, 55
Kennicutt R. C. et al., 2011, PASP, 123, 1347
Kirkpatrick A. et al., 2014, ApJ, 789, 130
Köhler M., Ysard N., Jones A. P., 2015, A&A, 579, A15
Laporte N. et al., 2017, ApJ, 837, L21
Lewandowski D., Kurowicka D., Joe H., 2009, J. Multivariate Anal., 100, 556
Lisenfeld U., Israel F. P., Stil J. M., Sievers A., 2002, A&A, 382, 860
Maddox S. J. et al., 2018, ApJS, 236, 30
Magdis G. E. et al., 2012, ApJ, 760, 6
Meny C., Gromov V., Boudet N., Bernard J. P., Paradis D., Nayral C., 2007, A&A, 468, 171
Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, J. Chem. Phys., 21, 1087
Michałowski M. J., 2015, A&A, 577, A80
Miville-Deschênes M.-A., Lagache G., 2005, ApJS, 157, 302
Neal R., 2011, in Brooks S., AGelman A., Jones G. L., Meng X. L., eds, Handbook of Markov Chain Monte Carlo. Chapman and Hall/CRC, Boca Raton, FL, p. 116
Neal R. M., 1994, J. Comput. Phys., 111, 194
Pettini M., Pagel B. E. J., 2004, MNRAS, 348, L59
Pilbratt G. L. et al., 2010, A&A, 518, L1
Priddey R. S., Isaak K. G., McMahon R. G., Robson E. I., Pearson C. P., 2003, MNRAS, 344, L74
Relaño M. et al., 2018, A&A, 613, A43
Rémy-Ruyer A. et al., 2013, A&A, 557, A95
Rémy-Ruyer A. et al., 2015, A&A, 582, A121
Rowlands K., Gomez H. L., Dunne L., Aragón-Salamanca A., Dye S., Maddox S., da Cunha E., van der Werf P., 2014, MNRAS, 441, 1040
Saintonge A. et al., 2016, MNRAS, 462, 1749
Saintonge A. et al., 2017, ApJS, 233, 22
Saintonge A. et al., 2018, MNRAS, 481, 3497
Sanders D. B., Mazzarella J. M., Kim D. C., Surace J. A., Soifer B. T., 2003, AJ, 126, 1607
Sawicki A., 2012, J. Phys. A: Math. Theor., 45, 505202
Schwarz G., 1978, Ann. Stat., 6, 461
Scoville N. et al., 2014, ApJ, 783, 84
Shetty R., Kauffmann J., Schnee S., Goodman A. A., 2009a, ApJ, 696, 676
Shetty R., Kauffmann J., Schnee S., Goodman A. A., Ercolano B., 2009b, ApJ, 696, 2234
Smith M. W. L. et al., 2012a, ApJ, 748, 123
Smith M. W. L. et al., 2012b, ApJ, 756, 40
Smith M. W. L. et al., 2019, MNRAS, 486, 4166
Stan Development Team, 2017, Stan Modeling Language: User's Guide and Reference Manual. Version 2.17.0, http://mc-stan.org
Stan Development Team, 2018, PyStan: The Python Interface to Stan, Version 2.17.1.0. http://mc-stan.org
Tak H., Ghosh S. K., Ellis J. A., 2018, MNRAS, 481, 277
Tamm A., Tempel E., Tenjes P., Tihhonova O., Tuvikene T., 2012, A&A, 546, A4
Taylor M. B., 2005 , in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV. Astron. Soc. Pac., San Francisco, p. 29
The Astropy Collaboration, 2013, A&A, 558, A33
Utomo D., Chiang I. D., Leroy A. K., Sandstrom K. M., Chastenet J., 2019, ApJ, 874, 141
Van Der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Veneziani M., Piacentini F., Noriega-Crespo A., Carey S., Paladini R., Paradis D., 2013, ApJ, 772, 56
Waelkens C. et al., 1996, A&A, 315, L245
Waters L. B. F. M. et al., 1996, A&A, 315, L361
Watson D., Christensen L., Knudsen K. K., Richard J., Gallazzi A., Michałowski M. J., 2015, Nature, 519, 327
Yang M., Phillips T., 2007, ApJ, 662, 284
Ysard N., Jones A. P., Demyk K., Boutéraon T., Koehler M., 2018, A&A, 617, A124
Zhukovska S., 2014, A&A, 562, A76

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Figure D2. FIR SED of the galaxies of the JINGLE sample, fitted with the hierarchical approach using the three models: SMBB (left-hand panel), BMBB (middle panel), and TMBB (right-hand panel).
Table E1. Result parameters from the hierarchical SED fitting using the SMBB model.
Table E2. Result parameters from the hierarchical SED fitting using the BMBB model.
Table E3. Result parameters from the hierarchical SED fitting using the TMBB model.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: NORMAL DISTRIBUTION VERSUS STUDENT'S *t*-DISTRIBUTION

In this section, we investigate how the choice of the prior distributions affect the results. In particular, the distribution of the parameter population $p(\boldsymbol{\theta}|\boldsymbol{\mu}, \Sigma)$ and the noise distribution $p(F_j^{\rm obs}|\boldsymbol{\theta}, F_j^{\rm err}, \delta_j)$.

The Student's *t*-distribution is appropriate for robust statistical models (Gelman et al. 2004; Kelly et al. 2012) and it is recommended when the measurement errors are assumed to be Gaussian,

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

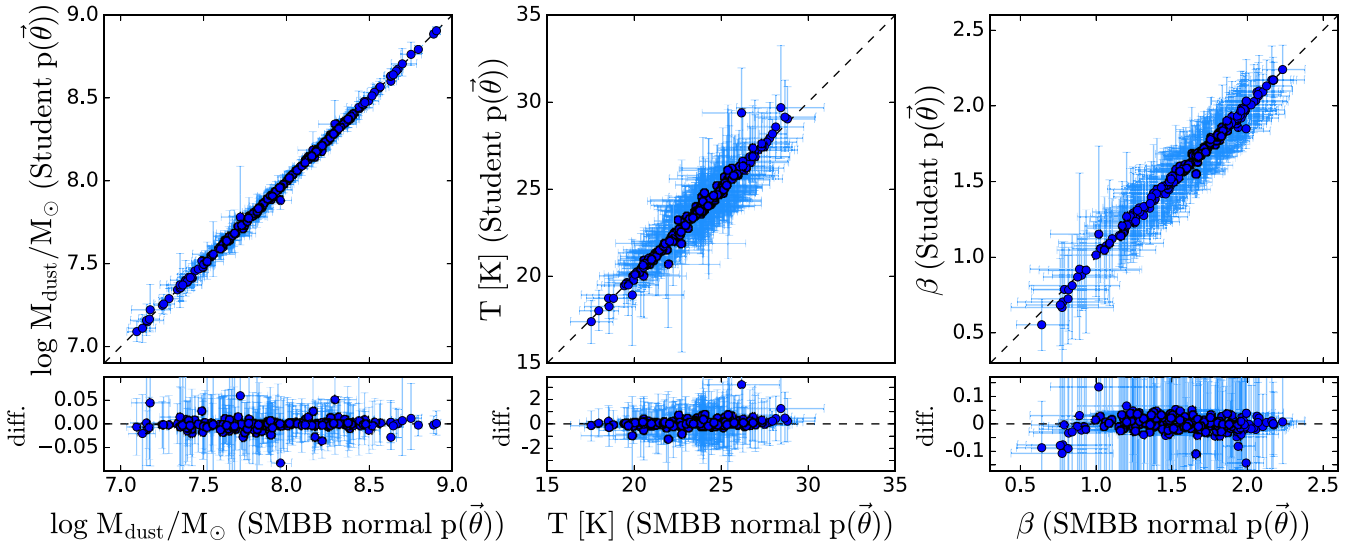

Figure A1. Comparison of the dust masses, temperatures, and emissivity index obtained through the fit of an SMBB using the hierarchical approach, assuming a normal or a Student's *t*-distribution with 8 degrees of freedom for the distribution of the parameters given the hyper-parameters $p(\boldsymbol{\theta}|\boldsymbol{\mu}, \Sigma)$.

but their standard deviation is not known but only estimated. If we assume that the true variance σ^2 follows a Scaled Inverse-χ^2 distribution with scale parameter σ'^2 (the estimated variance), then modelling the noise as a Student's *t*-distribution with standard deviation σ' is equivalent to the assumption that the noise is normal distributed with a standard deviation σ (Gelman et al. 2004). For the choice of the degrees of freedom we follow Kelly et al. (2012) and used $f = 3$, since it is the smallest value for which mean and variance of the distribution are finite. The results do not depends strongly on change on f which are less than an order of magnitude and $f < 10$ is a typical choice for robust models (e.g. Gelman et al. 2004). For a Student's *t*-distribution with $f = 3$, (61.5 per cent, 86.5 per cent, 94.6 per cent) of the distribution lie within (1σ, 2σ, 3σ) from the mean, respectively. In comparison, for a Gaussian distribution thepercentages are (68.3 per cent, 95.4 per cent, 99.7 per cent).

First, we focus on the population distribution of the parameters given the hyper-parameters (mean and standard deviation). We consider two distributions: normal and Student's *t*-distribution, which compared to the normal distribution allows for more outliers in the tail of the distribution. For the Student's *t*-distribution we use $f = 8$ degrees of freedom, which is the value we use for the analysis in this paper. As we can see from Fig. A1, the results do not change much. The dust masses do not vary depending on the choice of the sample distribution. Temperature and β show small differences, within the uncertainties, and no systematic offset. We conclude that the choice of the distribution does not affect the results critically.

The second assumption on the priors is about the noise distribution. We consider also in this case a normal and a Student's *t*-distribution with three degrees of freedom (see description in Section 3.4). The Student's *t*-distribution is less sensitive to flux points which may be outliers, due to large uncertainties or to the

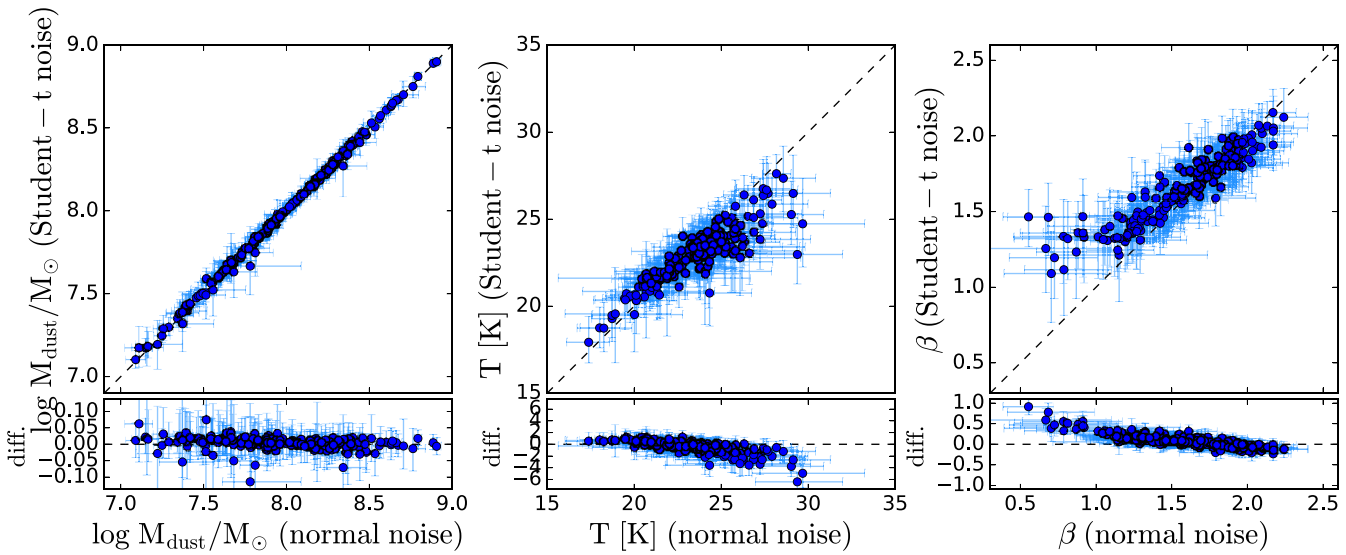

Figure A2. Comparison of the dust masses, temperatures, and emissivity index obtained through the fit of an SMBB using the hierarchical approach, assuming a normal distribution or a Student's *t*-distribution with three degrees of freedom for the noise.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

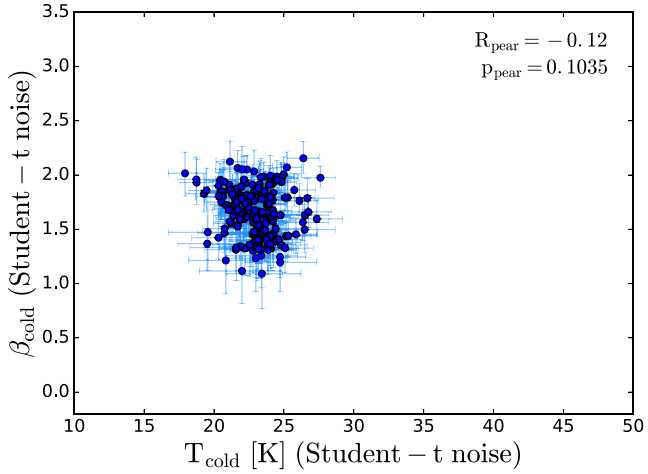

Figure A3. Relation between the dust temperature and dust emissivity index (T–β relation) from the SMBB hierarchical fit of the JINGLE sample assuming a Student's t-distribution with three degrees of freedom for the noise.

noise being underestimate. Fig. A2 shows the comparison plots. Again, the dust masses are robust with respect to the choice of the noise distribution. Temperature and β, on the other hand, show some variations, with the results obtained using the Student's t-distribution covering a smaller range of T and β values with respect to the results from the normal distribution ($T = 18-27$ K, $\beta = 1.1-2.1$ for the Student's t-distribution, $T = 17-30$ K, $\beta = 0.6-2.2$ for the normal distribution). The values of T or β which differ more from the mean values are determined mostly by the flux points at long (850 μm) or short wavelengths (100 μm). These are also the flux points which on average have the largest measurement uncertainties. With the assumption of a Student's t-distribution, we imply that the deviation of the SED shape from an SMBB with mean parameter values is not due to a change in T or β, but is more likely due to uncertainties in the flux measurements, which lead to 'outlier' flux points. Therefore, the measured T and β will cover a smaller range of values.

The choice of the noise distribution affects consequently also the derived relation between T and β, shown in Fig. A3. Assuming a Student's t-distribution for the noise, the results show a weak anticorrelation ($R = -0.12$) between T and β. This is similar to the result obtained from the fit without the 850 μm point. The fit with Student noise assumes that the variations at 850 μm are due to larger uncertainties on the estimate of the 850 μm uncertainties, rather than to a real variations in the sub-mm slope of the SED. Therefore, the fit tends to 'ignore' the extreme 850 μm flux points. In some cases the 850 μm point does not follow the same SED slope as the other points, but it shows an excess or a deficit. We have two ways to model this type of SED. One possibility is to assume that the true uncertainties on the 850 μm point are larger than the estimated ones, and therefore model the SED assuming a Student's t-distribution. If instead we believe that the different behaviour of the SED at wavelengths longer than 500 μm is real, we can model the noise using a normal distribution. In this paper, we decide to model the noise using a normal distribution.

APPENDIX B: UPPER LIMITS FORMALISM

In the case of a non-detection in one of the bands, the likelihood needs to be modified to include an upper limit for the non-detection. Following the formalism described in Sawicki (2012), the upper limit of an observation provides a limit on the evaluation of a definite integral. For a measured flux, F_j, which is clearly detected, the probability of observing our data, given the true value of the observables $F_j^{\rm true}$ and the measurement uncertainties $F_j^{\rm err}$, is

$$p(F_j|F_j^{\rm true}, F_j^{\rm err}). \tag{B1}$$

In the case of a single non-detection, we consider the upper limit $F_{{\rm lim},j}$, and the probability is

$$p(F_{{\rm lim},j}|F_j^{\rm true}, F_j^{\rm err}) = \int_{-\infty}^{F_{{\rm lim},j}} p(F_j|F_j^{\rm true}, F_j^{\rm err}){\rm d}F_j. \tag{B2}$$

Non-hierarchical: In the non-hierarchical approach, the likelihood in case of a non-detection on the jth flux measurement is given by

$$p(\boldsymbol{F}^{\rm obs}|\boldsymbol{\theta}) = \int_{-\infty}^{F_{{\rm lim},j}} {\rm MultiNormal}(\boldsymbol{F}^{\rm obs}|\boldsymbol{F}^{\rm mod}(\boldsymbol{\theta}), C){\rm d}F_j^{\rm obs}. \tag{B3}$$

Since the likelihood evaluation in case of upper limits includes the computation of integrals, the use of upper limit is computationally expensive. Thus, we allow our code to perform the SED fit with one flux point as upper limit at most, to avoid that the code has to calculate too many integrals. If more than one band has an upper limit, we consider only the upper limit in one band and we neglect the other flux point. We prefer to keep the 850 μm point, if it is an upper limit, since it is the longest wavelength point and it is the one that places more constraints on the SED slope.

Hierarchical: Similarly, for the hierarchical method the likelihood for the ith galaxy in case of a non-detection in the jth band is

$$p(\boldsymbol{F_i}^{\rm obs}|\boldsymbol{\theta_i}) = \int_{-\infty}^{F_{{\rm lim},i,j}} {\rm MultiNormal}(\boldsymbol{F_i}^{\rm obs}|\boldsymbol{F_i}^{\rm mod}(\boldsymbol{\theta_i}), C_i){\rm d}F_{i,j}^{\rm obs}. \tag{B4}$$

If the upper limit is in a band whose uncertainties are not correlated with other bands (i.e. the SCUBA-2 850 μm band or the IRAS 60 μm band), the expression for the upper limit can be divided in two parts, and the part that does not depend on $F_{i,j}$ can be taken out of the integral:

$$\begin{aligned} p(\boldsymbol{F_i}^{\rm obs}|\boldsymbol{\theta_i}) &= {\rm MultiNormal}(\boldsymbol{F_i'}^{\rm obs}|\boldsymbol{F_i'}^{\rm mod}(\boldsymbol{\theta_i}), C_i') \\ &\quad \cdot \int_{-\infty}^{F_{{\rm lim},i,j}} p(F_{i,j}^{\rm obs}|F_{i,j}^{\rm mod}(\boldsymbol{\theta_i}), F_{i,j}^{\rm err}){\rm d}F_{i,j}^{\rm obs} \\ &= {\rm MultiNormal}(\boldsymbol{F_i'}^{\rm obs}|\boldsymbol{F_i'}^{\rm mod}(\boldsymbol{\theta_i}), C_i') \\ &\quad \cdot \int_{-\infty}^{F_{{\rm lim},i,j}} {\rm Normal}\left(F_{i,j}^{\rm obs}|F_{i,j}^{\rm mod}(\boldsymbol{\theta_i}), \sqrt{C_{i,jj}}\right){\rm d}F_{i,j}^{\rm obs}, \end{aligned} \tag{B5}$$

where $\boldsymbol{F_i'}$ is the $(m-1)$-dimensional vector equal to the vector $\boldsymbol{F_i}$ but without the jth component. Similarly, C_i' is equal to the covariance matrix C_i, but without the jth component. $C_{i,jj}$ is the jj component of the covariance matrix C_i for the ith galaxy.

The integral of the univariate normal distribution can then be computed analytically:

$$\begin{aligned} \int_{-\infty}^{y_{\rm lim}} {\rm Normal}(y|\mu,\sigma){\rm d}y &= \frac{1}{\sqrt{2\pi}\sigma} \cdot \int_{-\infty}^{y_{\rm lim}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right){\rm d}y \\ &= \frac{1}{2}\left[{\rm erf}\left(\frac{y_{\rm lim}-\mu}{\sqrt{2}\sigma}\right)+1\right], \end{aligned} \tag{B6}$$

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

where 'erf' is the error function. If the upper limit is in one of the *Herschel* bands, the integral can also be computed analytically, but it requires more computations and it slows down code. Therefore, we decide to ignore the points with non-detections in the *Herschel* bands.

APPENDIX C: ADDITIONAL SIMULATIONS

***T*−*β* anticorrelated**: The second test we did was to see whether the hierarchical code can recover a *T*–*β* anticorrelation. We simulated 100 SEDs with temperatures uniformly distributed in the range 20–30 K and the corresponding β given by the relation:

$$\beta = -0.121 \cdot T + 4.595. \quad \text{(C1)}$$

The slope and intercept of this relation are derived from the results of the non-hierarchical fit to the real data. We also added some scatter to the $T-\beta$ anticorrelation. As before, we kept the dust mass constant ($\log M = 8$ M$_\odot$).

Fig. C1 shows the results from the hierarchical and non-hierarchical method. The non-hierarchical method points move in the $T-\beta$ anticorrelation direction. Thus, even if the $T-\beta$ anticorrelation is maintained, the differences between input values and measured values can be up to 8.6 K in temperature and 0.7 in β. Also the hierarchical code is able to recover the $T-\beta$ anticorrelation. The difference between input and measured are a bit smaller than in the non-hierarchical case (<5.8 K in temperature and 0.5 in β).

Comparing directly the input and output parameters, we see that the largest discrepancies between input and output temperatures happen for high-temperature values. This is due to the fact that the FIR SED moves to lower wavelengths with increasing temperature. Thus, for the high-temperature models ($T > 30$ K), the peak of

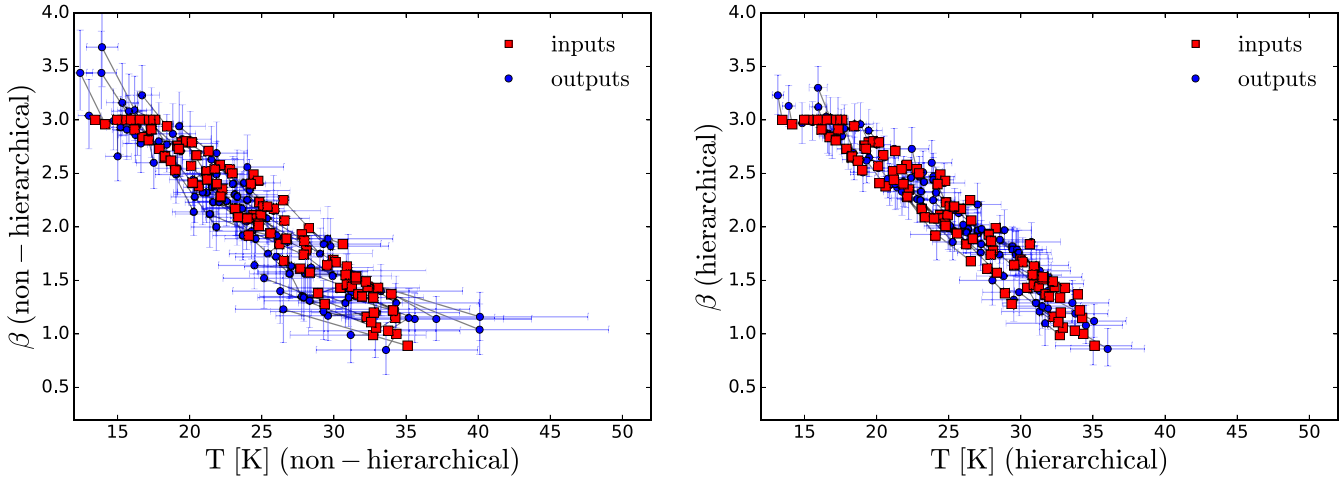

Figure C1. Results of temperature and β derived from the fit of 100 simulated SMBB SEDs with *T*–*β* anticorrelated and the same dust mass ($\log M_{\rm dust} = 8$ M$_\odot$). The temperature are linearly distributed and some scatter is added around a linear $T-\beta$ relation. We added to every band Gaussian noise with an amplitude proportional to the level of noise present in our data in that band. The output values are derived with the non-hierarchical (left-hand panel) and hierarchical (right-hand panel) fitting approach. In red are shown the input values and in blue are the measured values (outputs), the grey lines connect the corresponding inputs and outputs.

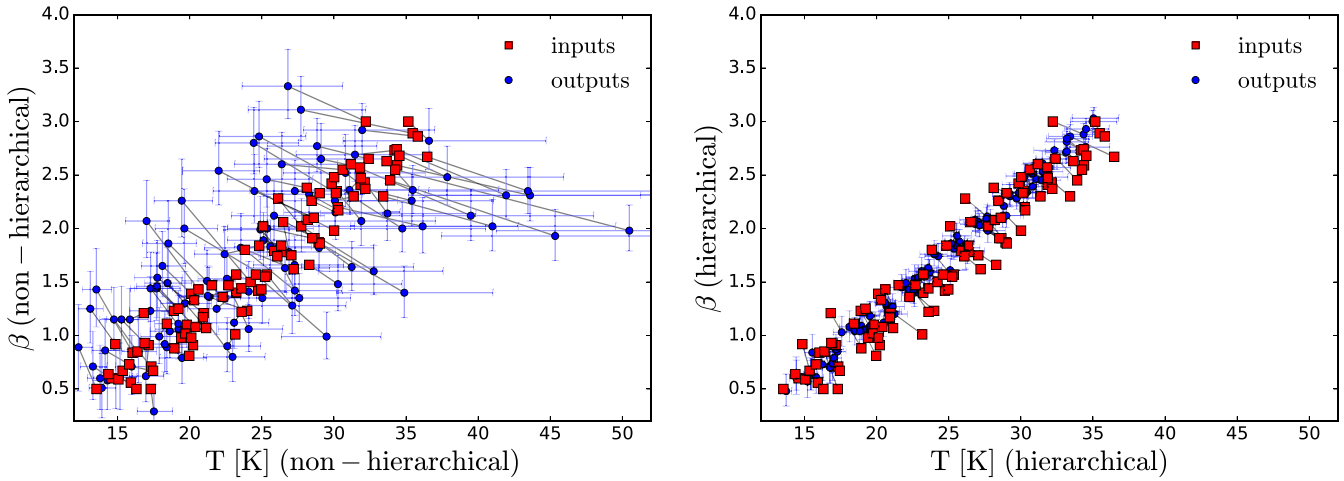

Figure C2. Results of temperature and β derived from the fit of 100 simulated SMBB SEDs with *T*–*β* correlated and the same dust mass ($\log M_{\rm dust} = 8$ M$_\odot$). The temperature are linearly distributed and some scatter is added around a linear $T-\beta$ relation. We added to every band Gaussian noise with an amplitude proportional to the level of noise present in our data in that band. The output values are derived with the non-hierarchical (left-hand panel) and hierarchical (right-hand panel) fitting approach. In red are shown the input values and in blue are the measured values (outputs), the grey lines connect the corresponding inputs and outputs.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

the SED is at wavelengths $<100\,\mu$m, which are not sampled by our data points/bands. This problem affects also the measurements of β: if the temperature is not well constrained, also β will not be determined with high precision, due to the degeneracy between the two parameters. Additionally, due to the assumed T–β anticorrelation, high T values correspond to low β values, i.e. shallower slopes of the SED. This will also contribute to the difficulties of accurately measure T and β.

***T–β* correlated:** We did the same test for positive correlation between T and β, parametrized by the relation:

$$\beta = 0.121 \cdot T - 1.325. \quad \text{(C2)}$$

As we can see from the left-hand panel of Fig. C2, the non-hierarchical method is not able to recover the positive correlation. The results of the fitting move away from the input values along diagonal lines in the $T-\beta$ plane, following the anticorrelation line. The right-hand panel of Fig. C2 shows the results from the hierarchical SED fitting. The code can recover the input values and the trend quite well. We note that the difference between input and output is often larger than the error bars. The points tend to move along diagonal lines in the $T-\beta$ plane, following the anticorrelation line. Therefore some points move outside the input correlation. However, the difference between input and output are small enough, that the $T-\beta$ positive correlation is visible also in the outputs value.

From these tests we can conclude that the hierarchical approach performs better than the non-hierarchical approach in all three cases of single input, $T-\beta$ correlation, and anticorrelation. In the case of a positive correlation, we note that even in the hierarchical approach the difference between input and output values can sometimes be larger than our errorbars. The differences in temperature are <3 K, and the difference in β are <0.3. For comparison, in the non-hierarchical case, the differences in temperature are <16 K, and the difference in β are <0.8.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

APPENDIX D: PLOTS OF THE FITTED SED

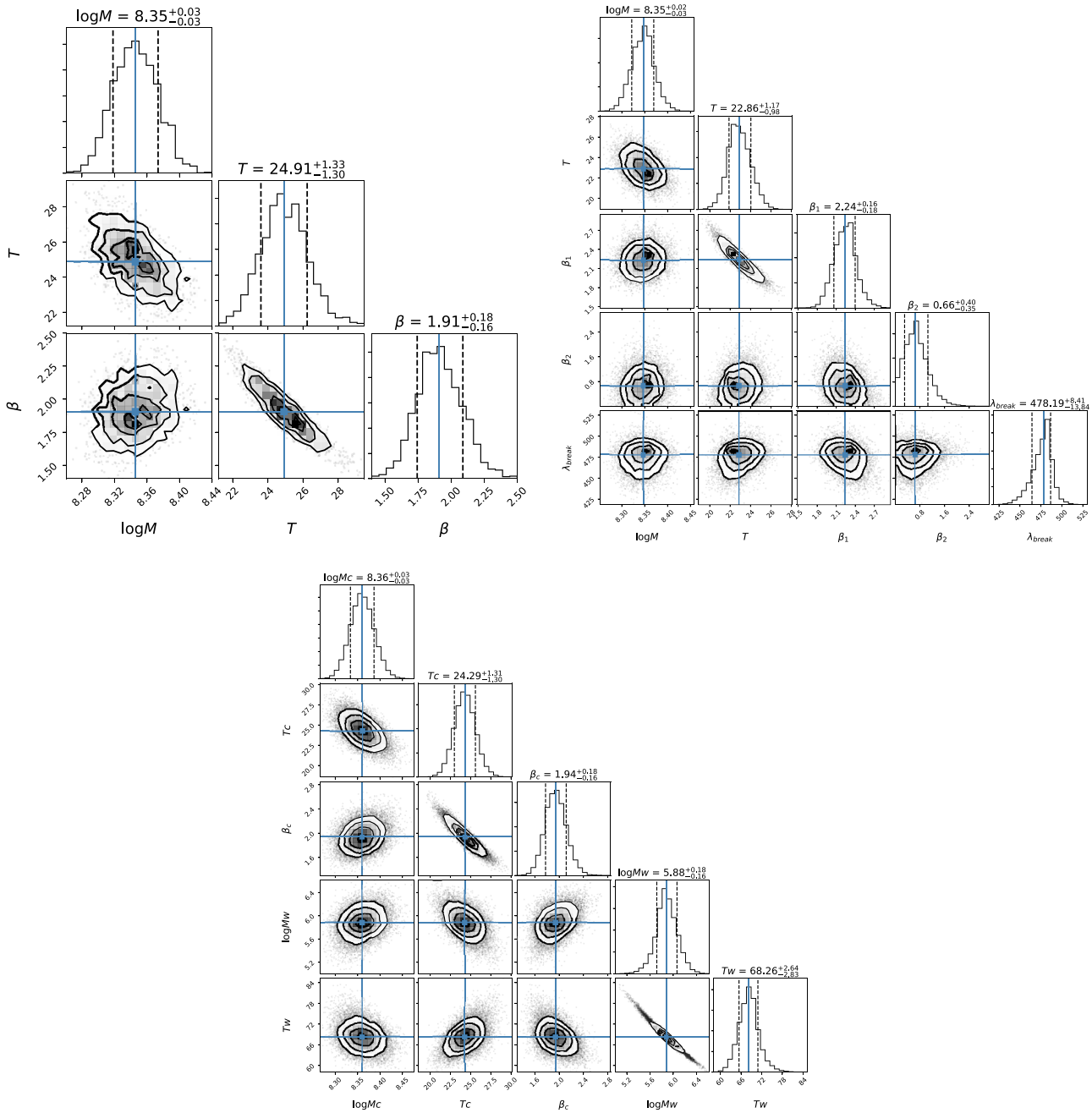

Figure D1. Example of the posterior PDFs of the model parameters obtained using the hierarchical method for the fit of one galaxy (JINGLE 147). The three panels show the results of fit using the SMBB (upper left), BMBB (upper right), and TMBB (bottom) models. The blue line indicates the median values, the dotted lines show the 16th and 84th percentiles that indicate the 1σ uncertainties on the parameters.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

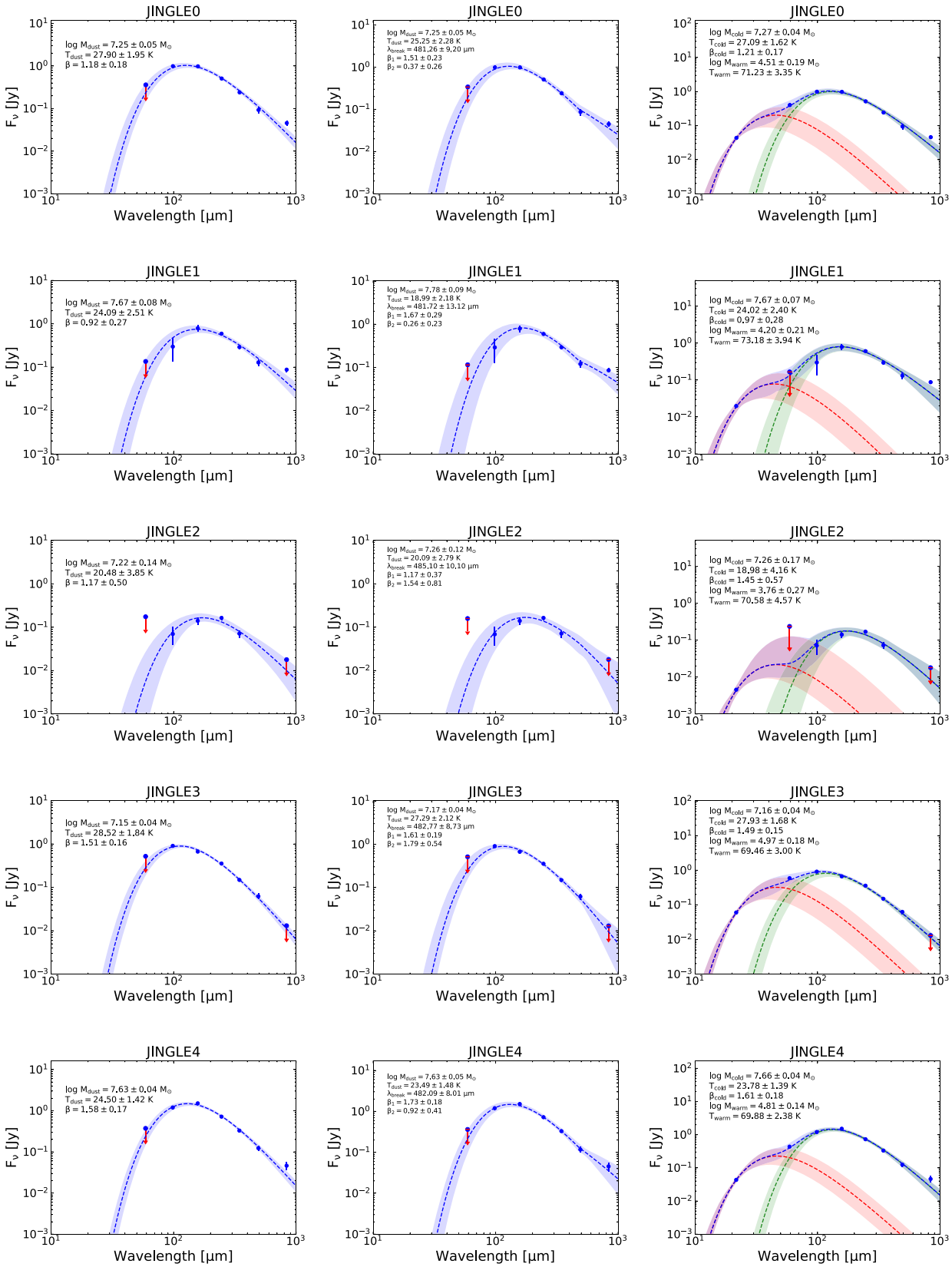

Figure D2. FIR SED of the galaxies of the JINGLE sample, fitted with the hierarchical approach using the three models: SMBB (left-hand panel), BMBB (middle panel), and TMBB (right-hand panel). The shaded regions show the lower and upper 1σ uncertainties on the SED models, defined by taking the maximum and minimum flux values of the models with likelihood values in the highest 68th percentile. Additional figures showing the entire sample of 192 JINGLE galaxies are available online.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

APPENDIX E: TABLES

Table E1. Result parameters from the hierarchical SED fitting using the SMBB model. The parameters of the model are the dust mass ($\log M_c$), temperature (T_c), and emissivity index (β_c). The last column is the natural logarithm of the likelihood, i.e. the probability of the observed fluxes given the model parameters ($p(\boldsymbol{F}|\boldsymbol{\theta})$). This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.

JINGLE ID	SDSS name	$\log M_c$ ($M_\odot$)	T_c (K)	β_c	$\ln L$
0	J131616.82+252418.7	7.25 ± 0.05	27.92 ± 1.98	1.17 ± 0.18	11.82
1	J131453.43+270029.2	7.66 ± 0.07	24.33 ± 2.48	0.91 ± 0.26	7.28
2	J131526.03+330926.0	7.22 ± 0.14	20.69 ± 3.78	1.15 ± 0.51	11.54
3	J125606.09+274041.1	7.15 ± 0.04	28.57 ± 1.84	1.51 ± 0.16	14.33
4	J132134.91+261816.8	7.63 ± 0.04	24.50 ± 1.56	1.57 ± 0.18	12.13
5	J091728.99−003714.1	7.91 ± 0.04	24.40 ± 1.47	1.52 ± 0.16	10.40
6	J132320.14+320349.0	7.40 ± 0.11	21.51 ± 2.77	1.29 ± 0.29	16.99
7	J132051.75+312159.8	7.50 ± 0.05	23.84 ± 1.57	1.37 ± 0.18	16.89
8	J091642.17+001220.0	7.29 ± 0.06	25.76 ± 2.15	1.24 ± 0.23	12.67
9	J131547.11+315047.1	7.60 ± 0.05	23.76 ± 1.70	1.34 ± 0.19	15.55

Table E2. Result parameters from the hierarchical SED fitting using the BMBB model. The parameters are the dust mass ($\log M_c$), temperature (T_c), emissivity index before the break (β_1) and after the break (β_2), and the wavelength of the break (λ_{break}). The last column is the natural logarithm of the likelihood, i.e. the probability of the observed fluxes given the model parameters ($p(\boldsymbol{F}|\boldsymbol{\theta})$). This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.

JINGLE ID	SDSS name	$\log M_c$ ($M_\odot$)	T_c (K)	β_1	β_2	λ_{break} (μm)	$\ln L$
0	J131616.82+252418.7	7.25 ± 0.05	25.25 ± 2.28	1.51 ± 0.23	0.37 ± 0.26	481.26 ± 9.20	14.59
1	J131453.43+270029.2	7.78 ± 0.09	18.99 ± 2.18	1.67 ± 0.29	0.26 ± 0.23	481.72 ± 13.12	10.86
2	J131526.03+330926.0	7.26 ± 0.12	20.09 ± 2.79	1.17 ± 0.37	1.54 ± 0.81	485.10 ± 10.10	11.55
3	J125606.09+274041.1	7.17 ± 0.04	27.29 ± 2.12	1.61 ± 0.19	1.79 ± 0.54	482.77 ± 8.73	14.52
4	J132134.91+261816.8	7.63 ± 0.05	23.49 ± 1.48	1.73 ± 0.18	0.92 ± 0.41	482.09 ± 8.01	13.42
5	J091728.99−003714.1	7.93 ± 0.04	23.01 ± 1.14	1.71 ± 0.15	1.37 ± 0.36	483.26 ± 6.95	10.59
6	J132320.14+320349.0	7.44 ± 0.08	22.05 ± 2.16	1.08 ± 0.25	2.04 ± 0.60	485.77 ± 8.11	17.54
7	J132051.75+312159.8	7.51 ± 0.04	23.08 ± 1.25	1.46 ± 0.17	1.47 ± 0.47	483.62 ± 7.52	16.85
8	J091642.17+001220.0	7.28 ± 0.10	25.23 ± 2.14	1.34 ± 0.19	1.19 ± 0.77	483.69 ± 7.83	12.65
9	J131547.11+315047.1	7.62 ± 0.05	23.09 ± 1.33	1.44 ± 0.22	1.29 ± 0.36	483.35 ± 7.59	15.65

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

Table E3. Result parameters from the hierarchical SED fitting using the TMBB model. The parameters are the dust mass ($\log M_c$), temperature (T_c), and emissivity index (β_c) of the cold dust component, and the dust mass ($\log M_w$), and temperature (T_w) of the warm dust component. The emissivity index of the warm component has been fixed to $\beta_w = 1.5$. The last column is the natural logarithm of the likelihood, i.e. the probability of the observed fluxes given the model parameters ($p(\boldsymbol{F}|\boldsymbol{\theta})$). This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.

JINGLE ID	SDSS name	$\log M_c$ ($M_\odot$)	T_c (K)	β_c	$\log M_w$ ($M_\odot$)	T_w (K)	$\ln L$
0	J131616.82+252418.7	7.27 ± 0.04	27.09 ± 1.62	1.21 ± 0.17	4.51 ± 0.19	71.23 ± 3.35	18.37
1	J131453.43+270029.2	7.67 ± 0.07	24.02 ± 2.40	0.97 ± 0.28	4.20 ± 0.21	73.18 ± 3.94	11.78
2	J131526.03+330926.0	7.26 ± 0.17	18.98 ± 4.16	1.45 ± 0.57	3.76 ± 0.27	70.58 ± 4.57	18.19
3	J125606.09+274041.1	7.16 ± 0.04	27.93 ± 1.68	1.49 ± 0.15	4.97 ± 0.18	69.46 ± 3.00	20.09
4	J132134.91+261816.8	7.66 ± 0.04	23.78 ± 1.39	1.61 ± 0.18	4.81 ± 0.14	69.88 ± 2.38	18.57
5	J091728.99−003714.1	7.93 ± 0.04	23.71 ± 1.30	1.55 ± 0.15	4.97 ± 0.14	70.26 ± 2.39	15.79
6	J132320.14+320349.0	7.44 ± 0.12	20.53 ± 2.85	1.40 ± 0.33	4.00 ± 0.23	70.68 ± 3.66	22.82
7	J132051.75+312159.8	7.54 ± 0.05	22.68 ± 1.40	1.45 ± 0.18	4.32 ± 0.16	70.96 ± 2.82	22.50
8	J091642.17+001220.0	7.27 ± 0.05	26.87 ± 2.29	1.14 ± 0.26	4.51 ± 0.19	72.35 ± 3.40	17.26
9	J131547.11+315047.1	7.64 ± 0.05	22.74 ± 1.49	1.41 ± 0.18	4.40 ± 0.16	71.10 ± 2.79	21.20

Table E4. Results of the analysis of the correlation between dust emissivity index β and combinations of other galaxy properties. The table shows the coefficients a_j of the best polynomial expression $\beta_{\rm model}(x_1, \ldots, x_k) = \sum_{j=1}^{k} a_j \log(x_j) + b$, to estimate β using combinations of two or three galaxy properties.

	Emissivity index β								
No. of param.	$\log M_*$ ($M_\odot$)	log SFR ($M_\odot$ yr^{-1})	log area (kpc^2)	12 + log(O/H)	$\log M_{\rm dust}$ ($M_\odot$)	$\log M_{\rm H\,I}$ ($M_\odot$)	Intercept	BIC (1)	R (2)
2 param.	0.22 ± 0.03	-0.01 ± 0.03					-0.44 ± 0.25	226.29	0.53
	0.42 ± 0.02		-0.37 ± 0.03				-1.97 ± 0.18	53.19	0.64
	0.08 ± 0.02			0.73 ± 0.10			-5.34 ± 0.66	164.7	0.61
	0.49 ± 0.04				-0.31 ± 0.04		-0.79 ± 0.15	172.34	0.55
	0.39 ± 0.02					-0.3 ± 0.02	0.64 ± 0.15	68.67	0.63
		0.32 ± 0.02	-0.22 ± 0.03				2.02 ± 0.04	228.43	0.47
		0.04 ± 0.02		0.88 ± 0.08			-5.89 ± 0.7	172.41	0.61
		0.15 ± 0.04			0.04 ± 0.03		1.43 ± 0.26	299.18	0.45
		0.44 ± 0.03				-0.35 ± 0.03	5.01 ± 0.27	138.35	0.54
			-0.21 ± 0.02	1.39 ± 0.07			-10.03 ± 0.63	91.25	0.62
			-0.43 ± 0.03		0.45 ± 0.03		-1.15 ± 0.17	151.77	0.55
			0.05 ± 0.03			-0.01 ± 0.03	1.65 ± 0.22	435.98	0.21
				1.07 ± 0.09	-0.03 ± 0.02		-7.35 ± 0.65	173.7	0.61
				1.19 ± 0.07		-0.13 ± 0.02	-7.42 ± 0.55	131.88	0.62
					0.48 ± 0.03	-0.43 ± 0.03	2.05 ± 0.16	117.7	0.58
3 param.	0.37 ± 0.03	0.08 ± 0.03	-0.38 ± 0.03				-1.46 ± 0.26	51.04	0.65
	0.09 ± 0.03	-0.01 ± 0.03		0.73 ± 0.09			-5.46 ± 0.74	169.97	0.61
	0.55 ± 0.04	0.26 ± 0.04			-0.59 ± 0.06		0.78 ± 0.27	132.11	0.56
	0.28 ± 0.02	0.23 ± 0.03				-0.41 ± 0.03	2.76 ± 0.32	21.40	0.66
	0.28 ± 0.03		-0.38 ± 0.03	0.8 ± 0.09			-7.48 ± 0.64	−14.37	0.70
	0.41 ± 0.04		-0.37 ± 0.03		0.01 ± 0.05		-1.97 ± 0.19	58.63	0.64
	0.46 ± 0.02		-0.25 ± 0.03			-0.18 ± 0.03	-0.85 ± 0.26	18.08	0.67
	0.39 ± 0.04			0.85 ± 0.09	-0.37 ± 0.04		-6.67 ± 0.68	96.94	0.64
	0.28 ± 0.03			0.54 ± 0.1		-0.28 ± 0.02	-3.14 ± 0.67	41.97	0.67
	0.32 ± 0.04				0.11 ± 0.06	-0.34 ± 0.03	0.92 ± 0.22	70.94	0.63
		0.19 ± 0.02	-0.32 ± 0.03	1.15 ± 0.08			-7.77 ± 0.69	33.69	0.66
		0.04 ± 0.04	-0.42 ± 0.03		0.41 ± 0.05		-0.89 ± 0.35	156.46	0.55
		0.46 ± 0.02	-0.06 ± 0.03			-0.32 ± 0.03	4.77 ± 0.28	140.08	0.54
		0.23 ± 0.04		1.17 ± 0.09	-0.25 ± 0.04		-6.49 ± 0.65	142.45	0.62
		0.29 ± 0.03		0.8 ± 0.08		-0.33 ± 0.03	-2.14 ± 0.72	38.58	0.66
		0.23 ± 0.04			0.3 ± 0.04	-0.46 ± 0.03	3.7 ± 0.31	85.63	0.60
			-0.40 ± 0.03	1.00 ± 0.09	0.25 ± 0.03		-8.25 ± 0.7	38.36	0.65
			-0.21 ± 0.03	1.38 ± 0.08		0.00 ± 0.03	-10.02 ± 0.67	96.77	0.62
			-0.31 ± 0.03		0.63 ± 0.03	-0.35 ± 0.03	0.56 ± 0.22	42.80	0.63
				0.63 ± 0.10	0.29 ± 0.04	-0.34 ± 0.03	-2.89 ± 0.78	84.11	0.63

Note. (1) BIC (Schwarz 1978), calculated as BIC $= -2 \cdot \ln(L) + q \cdot \ln(m)$ where L is the likelihood (i.e. the probability of the data given the parameter $p(\boldsymbol{F}|\boldsymbol{\theta})$), q is the number of free parameters of the model, and m is the number of data points (wavebands). (2) Pearson correlation coefficient.

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019

Table E5. Same as Table E4, but for the correlation between dust temperature T and galaxy properties.

				Dust temperature T					
No. of param.	log M_* ($M_\odot$)	log SFR ($M_\odot$ yr^{-1})	log area (kpc^2)	12 + log(O/H)	log M_{dust} ($M_\odot$)	log M_{H1} ($M_\odot$)	Intercept	BIC (1)	R (2)
2 param.	-2.14 ± 0.20	2.50 ± 0.22					44.24 ± 1.93	914.52	0.54
	0.10 ± 0.18		-0.59 ± 0.24				22.62 ± 1.53	1046.44	0.24
	0.24 ± 0.17			-2.86 ± 0.77			45.26 ± 5.44	1038.07	0.18
	-1.07 ± 0.34				0.89 ± 0.35		26.50 ± 1.20	1045.98	0.14
	-0.59 ± 0.17					0.53 ± 0.19	23.62 ± 1.27	1045.69	0.17
		1.93 ± 0.18	-2.17 ± 0.22				26.21 ± 0.33	931.29	0.53
		1.50 ± 0.15		-5.74 ± 0.61			72.73 ± 5.45	948.15	0.41
		4.19 ± 0.29			-3.73 ± 0.30		51.88 ± 2.20	849.60	0.68
		1.78 ± 0.22				-1.52 ± 0.23	37.27 ± 2.12	985.74	0.36
			-0.21 ± 0.18	-1.69 ± 0.59			37.74 ± 5.05	1038.7	0.22
			-1.19 ± 0.30		0.65 ± 0.22		19.53 ± 1.40	1038.33	0.25
			-1.38 ± 0.24			1.00 ± 0.22	15.42 ± 1.78	1026.52	0.29
				-3.47 ± 0.72	0.45 ± 0.18		49.50 ± 5.74	1032.98	0.17
				-2.65 ± 0.60		0.33 ± 0.15	42.65 ± 4.65	1035.02	0.18
					-0.54 ± 0.22	0.54 ± 0.25	21.84 ± 1.24	1051.14	0.14
3 param.	-1.56 ± 0.23	2.83 ± 0.22	-1.40 ± 0.25				40.58 ± 2.08	887.83	0.61
	-1.65 ± 0.23	2.52 ± 0.21		-2.89 ± 0.74			64.51 ± 5.56	905.66	0.55
	0.19 ± 0.37	4.24 ± 0.29			-3.96 ± 0.48		51.72 ± 2.11	854.86	0.68
	-1.92 ± 0.21	3.04 ± 0.25				-0.97 ± 0.24	51.26 ± 2.56	903.19	0.58
	0.64 ± 0.22		-0.65 ± 0.24	-2.98 ± 0.72			43.22 ± 5.51	1036.39	0.25
	-1.24 ± 0.34		-1.3 ± 0.29		1.9 ± 0.41		22.31 ± 1.50	1030.77	0.26
	-0.20 ± 0.19		-1.25 ± 0.27			1.07 ± 0.24	16.49 ± 1.95	1031.0	0.29
	-0.64 ± 0.36			-3.11 ± 0.79	1.02 ± 0.33		48.37 ± 5.39	1035.13	0.18
	-0.04 ± 0.24			-2.57 ± 0.75		0.35 ± 0.21	42.16 ± 5.75	1040.49	0.19
	-0.94 ± 0.36				0.53 ± 0.46	0.32 ± 0.28	25.04 ± 1.74	1050.03	0.16
		2.43 ± 0.19	-1.85 ± 0.22	-4.66 ± 0.63			66.19 ± 5.50	885.12	0.61
		4.09 ± 0.29	-0.50 ± 0.29		-3.32 ± 0.38		49.48 ± 2.51	851.99	0.69
		2.15 ± 0.22	-1.94 ± 0.26			-0.47 ± 0.27	30.33 ± 2.38	933.86	0.54
		4.06 ± 0.29		-1.85 ± 0.75	-3.31 ± 0.31		64.70 ± 5.44	848.76	0.68
		2.85 ± 0.23		-6.20 ± 0.62		-1.71 ± 0.22	92.99 ± 6.09	897.60	0.57
		4.19 ± 0.31			-3.70 ± 0.32	-0.06 ± 0.27	52.12 ± 2.55	855.07	0.68
			-1.42 ± 0.29	-4.09 ± 0.71	1.49 ± 0.26		48.84 ± 4.95	1013.67	0.30
			-1.10 ± 0.27	-1.74 ± 0.60		1.01 ± 0.22	30.02 ± 5.51	1023.57	0.29
			-1.46 ± 0.29		0.13 ± 0.26	0.93 ± 0.26	15.14 ± 1.93	1031.79	0.29
				-3.47 ± 0.81	0.46 ± 0.34	-0.01 ± 0.29	49.36 ± 6.84	1038.49	0.17

Note. (1) BIC (Schwarz 1978), calculated as BIC $= -2 \cdot \ln(L) + q \cdot \ln(m)$ where L is the likelihood (i.e. the probability of the data given the parameter $p(\boldsymbol{F}|\boldsymbol{\theta})$), q is the number of free parameters of the model, and m is the number of data points (wavebands). (2) Pearson correlation coefficient.

[1]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*
[2]*Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium*
[3]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[4]*School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff CF24 3AA, UK*
[5]*Department of Physics and Astronomy, McMaster University, Hamilton, ON L8S 4M1, Canada*
[6]*Centre for Astrophysics Research, University of Hertfordshire, College Lane AL10 9AB, UK*
[7]*Sub-department of Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[8]*Blackett Laboratory, Physics Department, Imperial College, London SW7 2AZ, UK*
[9]*School of Physical Sciences and Computing, Jeremiah Horrocks Institute, University of Central Lancashire, Preston, Lancashire PR1 2HE, UK*
[10]*Korea Astronomy and Space Science Institute, 776 Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Republic of Korea*
[11]*Institute of Astronomy and Astrophysics, Academia Sinica, PO Box 23-141, Taipei 10617, Taiwan*
[12]*Faculty of Physics, Astronomical Observatory Institute, Adam Mickiewicz University, ul. Sloneczna 36, PL-60-286 Poznań, Poland*
[13]*Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK*
[14]*Shanghai Astronomical Observatory, 80 Nandan Road, Xuhui District, Shanghai 200030, China*
[15]*European Southern Observatory, Alonso de Córdova 3107, Vitacura, Casilla 19001, Santiago, Chile*
[16]*Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, China*
[17]*Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Nanjing 210008, China*

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

Downloaded from https://academic.oup.com/mnras/article-abstract/489/3/4389/5552673 by guest on 16 October 2019