Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **483,** 1993–1999 (2019)
Advance Access publication 2018 December 3

doi:10.1093/mnras/sty3278

On the existence of bright IR galaxies at $z > 2$: tension between *Herschel* and SCUBA-2 results?

Carlotta Gruppioni [1]★ and Francesca Pozzi[1,2]

[1] *Osservatorio di Astrofisica e Scienza dello Spazio – Istituto Nazionale di Astrofisica, via Gobetti 93/3, I-40129, Bologna, Italy*
[2] *Dipartimento di Fisica e Astronomia, Universitá degli Studi di Bologna, via Gobetti 93/2, I-40129, Bologna, Italy*

Accepted 2018 November 28. Received 2018 November 28; in original form 2018 July 30

ABSTRACT

Recent derivations of the galaxy star formation rate density (SFRD) obtained from sub-millimetre (sub-mm) surveys (e.g. SCUBA-2) show a tension with previous works based on *Herschel* and multi-wavelength data. Some of these works claim that the SFRD derived by pushing the *Herschel* surveys beyond $z \simeq 2$ are incorrect. However, the current sub-mm surveys obtained from SCUBA-2 data and the methods used to construct the total infrared (IR) luminosity function (LF) and the SFRD could be affected by some limitations. Here, we show how these limitations (i.e. selection bias and incompleteness effects) might affect the total IR LF, making the resulting dusty galaxy evolution of difficult interpretation. In particular, we find that the assumed spectral energy distribution (SED) plays a crucial role in the total IR LF derivation; moreover, we confirm that the long-wavelength (e.g. 850-μm) surveys can be incomplete against 'warm" SED galaxies, and that the use of a wide spectral coverage of IR wavelengths is crucial to limit the uncertainties and biases.

Key words: galaxies: evolution – galaxies: high-redshift – galaxies: luminosity function – galaxies: star formation – infrared: galaxies – submillimetre: galaxies.

1 INTRODUCTION

To directly measure the star formation (SF) activity in galaxies independently of any extinction corrections, one needs to observe in the far-infrared (far-IR) and sub-millimetre (sub-mm) domains. The amount of dust-obscured SF is directly linked to the measure of the total IR luminosity ($L_{\rm IR}$, integrated over the 8–1000 μm rest-frame range) through an empirical formula provided by Kennicutt (1998) and widely used in the literature. Therefore, how one derives the SF depends strongly on the shape of the assumed SED and on how well the far-IR bump (produced by dust, which re-emits in the IR the UV radiation from young and massive stars) is constrained. The peak and the shape of the far-IR bump provide important information about the dust temperature, the amount of dust, and the galaxy type. For this reason, it is extremely important to obtain as many photometric data as possible, in order to constrain the galaxy SEDs over a wide range of wavelengths (possibly from UV to sub-mm/mm). As a consequence, to study how the star formation rate density (SFRD) evolves with time across the Universe, it is necessary to constrain large samples of multi-wavelength galaxy SEDs over wide luminosity and redshift ranges. By counting these galaxies, per unit comoving volume, in luminosity and redshift bins, we can derive the LF, one of the most important tools to study galaxy evolution.

The extragalactic surveys performed with the *Herschel* observatory (e.g. Herschel–ATLAS, H–ATLAS, Eales et al. 2010; Herschel–GOODS, H–GOODS, Elbaz et al. 2011; PACS Evolutionary Probe, PEP, Lutz et al. 2011; Herschel Multi-tiered Extragalactic Survey, HerMES, Oliver et al. 2012) were the first ones to detect the peak of dust emission in galaxies up to high redshifts, and, due to the extensive multi-wavelength coverage in most of their fields, also to provide precise measurement of $L_{\rm IR}$ in thousands of galaxies spanning wide $L_{\rm IR}$, z ranges. The deepest *Herschel* surveys performed with PACS (observing at 70, 100 and 160 μm; Poglitsch et al. 2010) allowed us to measure $L_{\rm IR}$ and trace the SFRD evolution up to $z \sim 3$–4 (i.e. Gruppioni et al. 2013; Magnelli et al. 2013), detecting unexpectedly large numbers of very bright sources (i.e. $L_{\rm IR} > 10^{12}$ $L_{\odot}$) at $z \geq 2$). The presence of these IR-bright sources at high-z was not expected from semi-analytic models (SAMs), which indeed largely under-predict the high SFRs observed in *Herschel* galaxies at $z \sim 2$–3. Observations performed with the longer wavelength *Herschel* instrument, SPIRE (observing at 250, 350, and 500 μm; Griffin et al. 2010), detected even brighter IR galaxies and at higher z's ($\gtrsim$6; e.g. Riechers et al. 2013, Lutz 2014, Rowan-Robinson et al. 2016, Laporte et al. 2017), but with large identification uncertainties due to possible source blending. In fact, because of the large beam of the SPIRE instrument, i.e. FWHM$\simeq$18, 25, and 36 arcsec at 250, 350, and 500 μm, respectively (see Swinyard et al. 2010), source blending and mis-identification can be critical issues for sources selected at these wavelengths.

★ E-mail: carlotta.gruppioni@inaf.it

Because of source blending, recent works deriving the total IR LF of galaxies from a 850-μm selection (e.g. Koprowski et al. 2017 using SCUBA-2 on the James Clerk Maxwell Telescope, JCMT), claim that the high values of SFRD (or IR luminosity density, $\rho_{\rm IR}$) inferred from pushing the Herschel surveys beyond $z \simeq 2.5$ are incorrect (e.g. Koprowski et al. 2017). Indeed, these works derive a much steeper bright-end of the total IR LF than the *Herschel* ones already at $z > 1$, implying far fewer bright IR sources (and/or lower IR luminosities) and a smaller dust-obscured SFRD, at high-z. Koprowski et al. (2017) ascribes the cause of this discrepancy to the fact that the number and luminosity of $z \gtrsim 2$ sources had been severely overestimated by *Herschel* studies due to the high confusion and fraction of blended sources in SPIRE maps. Indeed, the sub-mm derived bright-end is not well constrained by data, especially at $z > 2$, and the large discrepancy is mostly in luminosity ranges where only model extrapolations are reported (without any SCUBA-2 data). The results of Koprowski et al. (2017) seem nevertheless in agreement with the Atacama Large Millimetre Array (ALMA) observations of the Hubble Ultra Deep Field by Dunlop et al. (2017) and with the results of stacking the deepest SCUBA-2 Cosmology Legacy Survey (S2CLS) images (Bourne et al. 2017), both finding a transition between obscured and unobscured SFRDs at $z \sim 3$–4, with a steep high-z decline following that shown by UV surveys (e.g. Bouwens et al. 2015). However, we must note that the area covered by the ALMA survey is too small (e.g. ∼4.5 arcmin^2; Dunlop et al. 2017) to detect the most luminous objects shaping the bright-end of the LF. Moreover, the total IR LF is strongly sensitive to the SEDs considered for deriving $L_{\rm IR}$, since integrating different SEDs over 8–1000 μm can result in $L_{\rm IR}$ differing by up to an order of magnitude, for a given 850-μm flux density.

If we cannot demonstrate without any doubt that all the bright IR sources detected by *Herschel* at $z > 1$ are not blended, we can however show with very simple arguments that the sub-mm regime (i.e. ≥850 μm) is not optimally suited for selecting the complete samples of dusty sources (at least at $z < 4$) needed to derive the total IR LFs. In fact, the 850-μm selection likely misses most of the "warmer" SED sources detected by *Herschel* and dominating the bright-end of the total IR LF (Gruppioni et al. 2013). Indeed, even before the advent of *Herschel*, radio selected high-z galaxies with far-IR luminosities comparable to those of sub-mm selected galaxies (SMGs) at similar redshifts, but not detected in the sub-mm, were found by Chapman et al. (2004) and associated to a population with hotter characteristic dust temperatures. Higher dust temperatures for these galaxies than similarly selected SMGs were then confirmed by *Herschel* observations (e.g. Chapman et al. 2010; Magnelli et al. 2010). Since even a small increase in the dust temperature implies a large decrease in observed sub-mm flux density, the shorter wavelengths of the far-IR emission peak cause the sub-mm waveband to potentially miss up to half of the most luminous, dusty galaxies at $z \sim 2$ (Chapman et al. 2004).

The flux density limits of the S2CLS survey also play a significant role once converted to IR source completeness, with the $L_{\rm IR}$ range covered by sub-mm data being very narrow around the LF knee (i.e. $L*$).

Finally, the SED choice and how well the peak of dust emission is constrained are crucial. All these factors conspire in making the total IR LFs and SFRD results from 850-μm surveys rather uncertain or incomplete.

Throughout the paper, we use a Chabrier (2003) stellar initial mass function (IMF) and adopt an ΛCDM cosmology with $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\rm m} = 0.3$, and $\Omega_\Lambda = 0.7$.

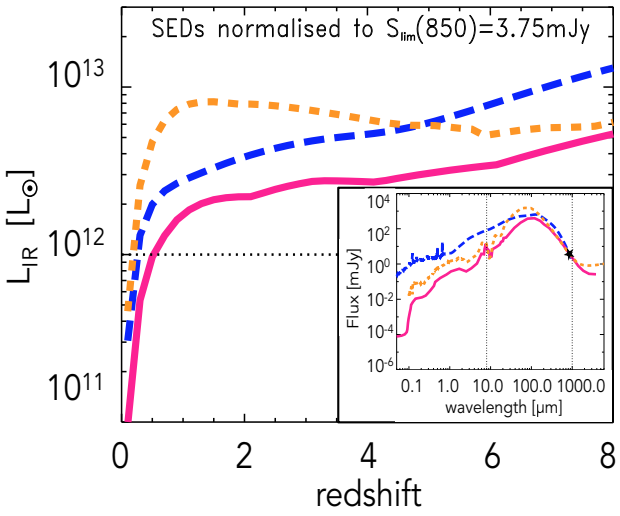

Figure 1. $L_{\rm IR}$ obtained by integrating three different SEDs (two SEDs representative of the bulk of the *Herschel* population at $z > 2$: IRAS20551, orange dashed, and a QSO, blue long-dashed, see Gruppioni et al. 2013), and the average sub-mm SED derived by Michałowski, Hjorth & Watson 2010, deep-pink solid) normalized to $S(850)$ =3.75 mJy, as a function of z. In the insert, the three SEDs are shown (normalized to the same 850-μm flux).

2 THE MAIN INGREDIENTS OF THE TOTAL IR LF AND SFRD

2.1 SED as key building block of $L_{\rm IR}$

First, we discuss and analyse the key role played by the choice of galaxy SED in the $L_{\rm IR}$ derivation. The better the far-IR bump is constrained by data, the more reliable the derived $L_{\rm IR}$ are. In fact, if the far-IR bump is not sampled by data points, one could obtain very different values of the 8–1000 μm luminosity for the same 850-μm flux, depending on the considered SEDs. To derive the total IR LFs from *Herschel* data, Gruppioni et al. (2013) fitted each object with a set of different templates and used the best-fit SED to compute the total IR LF: this provided the more realistic measure of the total IR luminosity source-by-source. On the contrary, the multi-wavelength data relative to the S2CLS surveys of Koprowski et al. (2017) were used only for estimating the photometric redshifts, while $L_{\rm IR}$ was computed from the observed 850-μm flux density by just assuming for all the sources the average sub-mm galaxy template derived by Michałowski et al. (2010), regardless of type, redshift, or luminosity.

In Fig. 1, we show the variation of $L_{\rm IR}$ with z, obtained by integrating three different SEDs (normalized to the same 850-μm flux density, corresponding to the limiting flux of the SCUBA-2 Ultra Deep Survey, UDS, by Koprowski et al. 2017: $S(850)$ =3.75 mJy): the average sub-mm SED by Michałowski et al. (2010) considered for all the SCUBA-2 sources (deep-pink solid), IRAS 20551 (orange dashed)m, and a typical QSO SED (blue long-dashed). The latter templates are those that best reproduce the SEDs of the bulk of *Herschel* sources responsible for the observed bright-end of the total IR LF at $z > 2$ (Gruppioni et al. 2013). The three templates are shown in the insert at the bottom right corner of the figure. It is clear that by integrating the three templates over the 8–1000 μm range, one would get very different results, even starting from the same 850-μm flux: it is therefore crucial to have data in the rest

Table 1. Total IR LF ratio (single SED versus SED library).

$<log_{10}(L_{IR}/L_{\odot})>$	$\Phi_{michalowskiSED}/\Phi_{gruppioniSEDs}$	
	$0.0 < z < 0.3$	$2.5 < z < 3.0$
8.2	4.3	–
8.7	2.1	–
9.2	0.5	–
9.7	0.4	–
10.2	0.2	–
10.7	0.2	–
11.2	0.02	–
11.7	0.0	–
12.2	0.0	4.6
12.7	–	0.9
13.2	–	0.3
13.7	–	0.1

mid-/far-IR range to better constrain the SEDs and L_{IR}. Up to at least $z \simeq 8$, the two best-fit *Herschel* SEDs provide much higher values of L_{IR} than the Michałowski et al. (2010) template. It is therefore not surprising that the SCUBA-2-based total IR LF has such a steep bright-end (see figs. 7 and 8 of Koprowski et al. 2017): L_{IR} can be severely underestimated (up to a factor of 5 at $z \simeq 1$–2) causing the luminous objects to fall in the wrong (e.g. lower) luminosity bin. In order to quantify the effect of the SED, we have recomputed the total IR LF with L_{IR} derived from the Michałowski et al. (2010) template (normalized to the measured 160-μm flux) for all our PEP sources: the resulting LFs show a different shape in any z-bins (i.e. steeper than the Gruppioni et al. 2013, with lower or absent data points in the brighter L-bins and higher in the faint L-bin). The effect is more evident at lower z, while at higher redshifts the difference is less pronounced, due to the fact that the 850-μm band with increasing z samples rest-frame wavelengths closer and closer to the SED peak, although the steepening of the LF is still significant. In Table 1, we report the ratio between the two IR LFs calculated with the Michałowski et al. (2010) single SED and with the template library used by Gruppioni et al. (2013), in two redshift bins (i.e. at $0.0 < z < 0.3$ and $2.5 < z < 3.0$).

2.2 Sample selection and completeness

Another factor that might conspire to depress the bright-end of the Koprowski et al. (2017) LF is the IR source incompleteness due to the sub-mm selection. As an example, in Fig. 2, we show the 850-μm flux expected for the three template SEDs considered above, for a source detected at 160 μm at the limiting flux of the PEP GOODS-S survey ($S(160) = 2.4$ mJy). A source with the Michałowski et al. (2010) SED (deep-pink), selected at the limits of *Herschel*–PEP, will not be detectable in the S2CLS UDS at $z < 3$, while it will not be detectable up to $z \simeq 3.5$–4 if it had a QSO- or IRAS 20551-like SED. Since the IR sources detected by *Herschel* show a wide diversity of SEDs, fitted by different templates, and the templates that best reproduce the observed SEDs of most of the $z \gtrsim 2$ *Herschel* galaxies making up the controversial bright-end of the IR LF are significantly 'warmer" than the Michałowski et al. (2010) one (see, e.g. figs. 1 and 18 of Gruppioni et al. 2013), we conclude that an 850-μm survey (unless extremely deep, e.g. to <0.1 mJy) is, by definition, unable to detect the majority of the *Herschel* sources at redshifts lower than $z \simeq 3$–4. As an empirical probe of the above assertion, in Fig. 3 we show the expected 850-μm flux density for all the *Herschel*–PEP 160-μm sources in the COSMOS (*top*) and GOODS-S (*bottom*) fields contributing to the total IR LF and

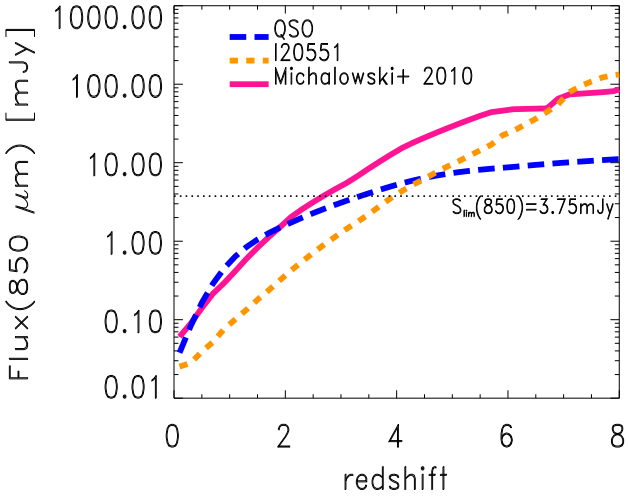

Figure 2. Expected 850-μm flux for a *Herschel* source detected at the limiting flux of the deepest PEP survey (GOODS-S, $S(160\ \mu m) = 2.4$ mJy), based on the three different SEDs described in fig. 1. The horizontal dotted line shows the limiting flux of the 850-μm SCUBA-2 UDS survey.

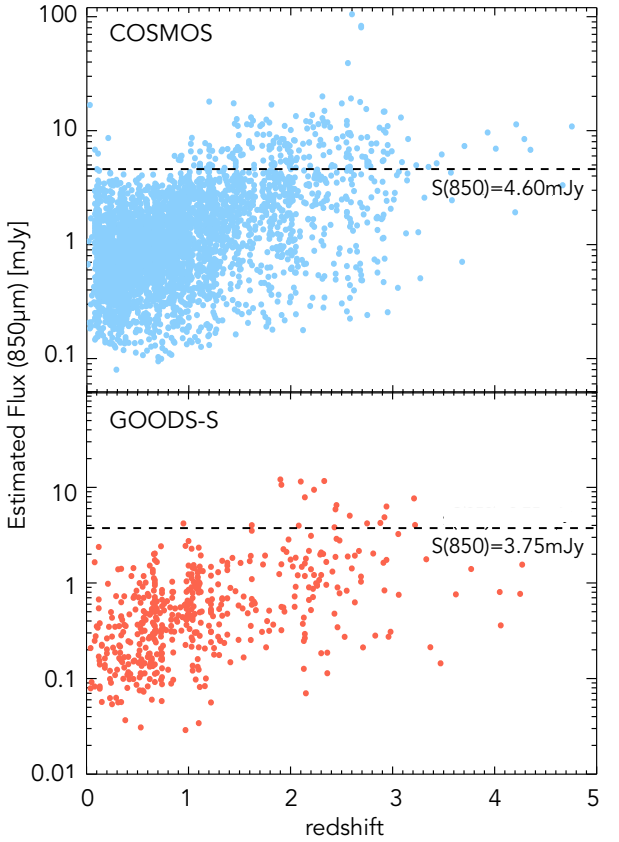

Figure 3. Flux at 850 μm expected for the *Herschel*–PEP 160-μm sources detected in the COSMOS (*top*, light-blue filled circles) and GOODS-S (*bottom*, orange-red circles) fields (based on their best-fit SEDs). The horizontal dotted lines show the 850-μm limits of the S2CLS surveys in the COSMOS and UDS fields.

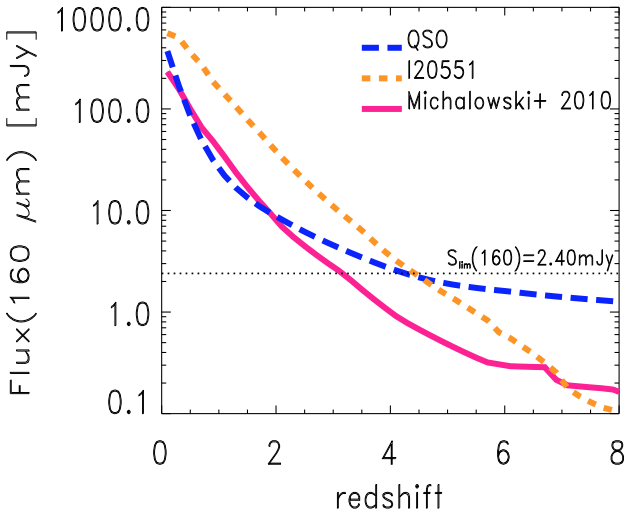

Figure 4. Expected 160-μm flux of a SCUBA-2 source with $S(850\ \mu m) = 3.75$ mJy, based on the SEDs of Fig. 2. The horizontal dotted line shows the limiting flux of the deepest 160-μm PEP survey in the GOODS-S.

SFRD (Gruppioni et al. 2013, 2015): each flux density has been obtained by interpolating the best-fit template SED (reproducing data from UV to far-IR) for each source, and has to be considered only indicative (extrapolation from the longer wavelength sampled by *Herschel* to 850-μm can be very uncertain, being in the steep Rayleigh–Jeans domain). Most of the *Herschel*–PEP sources are expected to be undetectable in the SCUBA-2 surveys of Koprowski et al. (2017), and indeed from a cross-match between the PEP and the S2CLS catalogues we find that only ∼5 per cent of the PEP 160-μm sources considered for the LF in the COSMOS field have a counterpart in the S2CLS catalogue (this percentage reduces to 2 (1) per cent if we limit to $z < 2$ (1)). We therefore cannot consider the total IR LF and SFRD derived from the S2CLS surveys as a complete derivation, since a population of bright and 'warm" IR sources significantly contributing to these quantities is likely missed by them.

The reverse case of the expected detectability of an S2CLS source in the PEP survey (e.g. in the GOODS-S field) is plotted in Fig. 4: depending on the SEDs, the SCUBA-2 sources are expected to be detected above the PEP flux density limit at $z \lesssim 3$ ($\lesssim$4 for an IRAS 20551 or QSO template). We therefore expect the deepest PEP 160-μm survey to be almost complete with respect to the faintest S2CLS sources, at least up to $z \simeq 3$–4. Indeed, from a simple catalogue–catalogue cross-match within the same fields and in the overlapping areas, we found that ∼85 per cent of the S2CLS sources detected at $S/N > 4\sigma$ have a PEP 160-μm counterpart in the GOODS-N field and ∼80 per cent in COSMOS (note that the PEP and S2CLS surveys in these two fields are shallower than those considered in Fig. 4, that reach the deepest PEP and S2CLS limits, but are not in the same area; moreover, no redshift information was given for the S2CLS sources, so we are not able to test the percentage of counterparts as a function of z).

2.3 Multiple sources

The main reasons ascribed by Koprowski et al. (2017) to the large discrepancy between their and *Herschel* LFs are 'problems in source identification and redshift estimation arising from the large-beam long-wavelength SPIRE data, as well as potential blending issues".

Table 2. Total IR LF ratio ($f_{SPIRE} \times 0.5$ versus original).

$<\log_{10}(L_{IR}/L_{\odot})>$	$\Phi_{michalowskiSED}/\Phi_{gruppioniSEDs}$	
	$0.0 < z < 0.3$	$2.5 < z < 3.0$
8.7	2.4	–
9.2	1.1	–
9.7	1.1	–
10.2	1.0	–
10.7	0.7	–
11.2	1.0	–
11.7	1.0	–
12.2	0.9	2.0
12.7	–	0.9
13.2	–	0.6
13.7	–	0.4

Source blending (or confusion) means that more than one astronomical source may be present within the beam, hence some sources can have their fluxes boosted. We have tested the effect of possible blending of SPIRE sources in our calculations by recomputing L_{IR} and the total IR LF by halving the 250, 350, and 500-μm fluxes (i.e. in the extreme assumption that all the SPIRE fluxes come from the blend of two sources) and re-fitting the SEDs. The comparison between the new total IR LFs and the original ones indicates a slight steepening when reducing the SPIRE fluxes, though not as extreme as found by using the Michałowski et al. (2010) template for all the sources (see Section 2). In Table 2, we report the ratios between the two IR LFs in two redshift bins (i.e. at $0.0 < z < 0.3$ and $2.5 < z < 3.0$). Therefore, we can conclude that, even if all the SPIRE fluxes would result from the blend of two sources (e.g. a factor of 2 higher than the real ones), the error on the calculation of the total IR LF would be smaller by constraining the SEDs with all the other available data, than by using an average template for all the sources neglecting the observed data. However, these ratios are much lower than the differences claimed by Koprowski et al. (2017) at high L_{IR} (i.e. factors >100 at $z > 1.2$–1.3), that would imply that the SPIRE bright fluxes were likely due to the blend of more than two sources (though without a 24-μm counterpart): these fainter sources artificially boosting the bright end of the *Herschel* LF by such large amounts, if resolved by SCUBA-2, should then produce a steeper faint-end in the SCUBA-2 total IR LF (i.e. steeper than found by just halving the SPIRE fluxes), to compensate the much lower bright-end. At odds with this expectation, Koprowski et al. (2017) find a faint-end of the total IR LF in fairly good agreement with Gruppioni et al. (2013) and significantly shallower than the Magnelli et al. (2013) one.

Recent works observing sub-mm sources (e.g. SPIRE, APEX, or SCUBA-2) with ALMA, report high fractions of multiple ALMA counterparts, especially for the brightest SPIRE 500-μm sources (e.g. Hodge et al. 2013; Bussmann et al. 2015), though Bussmann et al. (2015) find that the ALMA counterparts of the *Herschel* targets in some cases are located so close to each other that the most plausible hypothesis seems to be interactions and mergers. Other works observing similarly bright SPIRE 500-μm sources with SCUBA-2 seem to find very little evidence of source confusion (e.g. Bakx et al. 2018), while Hill et al. (2018) estimate the probability that a 10 mJy single-dish sub-mm source resolves into two or more galaxies to be <15 per cent. However, the *Herschel* targets so far observed with ALMA are sources selected at 500-μm in SPIRE images (i.e. at longer wavelength, with much larger FWHM than PACS), while the *Herschel* catalogues considered for deriving the total IR LF are selected at PACS wavelengths (i.e. 160 μm), from maps with

FWHM smaller than or similar to the SCUBA-2 ones at 850 μm. In particular, the works of Gruppioni et al. (2013) is based on blind 160-μm PACS catalogues, while that of Magnelli et al. (2013) on a PACS catalogue at 160 μm obtained with 24-μm prior positions. PACS is diffraction-limited and the photometre PSF at 160 μm is ≃11 arcsec (see Poglitsch et al. 2010). The *Herschel*–PACS confusion at 100 and 160 μm is 0.15 and 0.68 mJy, respectively (e.g. Magnelli et al. 2013), while the estimated SCUBA-2 confusion at 850 μm is ∼2 mJy (e.g. Chen et al. 2013a), therefore both PEP and SC2LS surveys are not expected to be confused at the reached fluxes. Instead, the SPIRE confusion estimated by Nguyen et al. (2010) is ∼5.8, 6.3, 6.8 mJy at 250, 350, and 500 μm, respectively, with HerMES reaching fluxes close to (or fainter than) these values. However, the HerMES SPIRE fluxes used to construct the SEDs of the 160-μm sources have been extracted by starting from *Spitzer* MIPS 24-μm source positions (Roseboom et al. 2010; Oliver et al. 2012). Although the SPIRE instrument has FWHM of 18, 25, and 36 arcsec at 250, 350, and 500 μm, respectively, the positions of sources detected at shorter wavelengths, i.e. 24 μm, have been utilized in order to disentangle the various contributions from discrete sources to the SPIRE flux (Roseboom et al. 2010; Elbaz et al. 2011). Thus, SPIRE fluxes have been de-blended using 24-μm priors, assuming that the positions of all sources contributing significantly to the SPIRE map are known (i.e. previously detected at 24 μm), and that only the SPIRE flux density of each of these sources is unknown. This significantly reduces the confusion (e.g. by a factor of ≃20–30 per cent, Roseboom et al. 2010), although it might produce incomplete catalogues (i.e. missing SPIRE sources undetected in the prior band).

2.4 The total IR LF

A large discrepancy between the total IR LF derived from JCMT/SCUBA-2 and *Herschel* data has been ascribed to $L_{\rm IR}$ overestimated by *Herschel* works due to large *Herschel* beam size and source blending. However, as shown in the previous sections, the 8–1000-μm luminosity derivation is a delicate task, depending on several factors that, if not taken into account correctly, may lead to severe under-/over-estimation of the proper value of $L_{\rm IR}$. We believe that in the SCUBA-2 results these factors have not been considered in the appropriate way for a fair comparison.

We note that the claimed inconsistencies between the far-IR and the sub-mm-derived LFs come from very few sub-mm luminosity bins, likely close to the LF knee (e.g. $L*$). In fact, if we compare fig. 3 and figs. 7 and 8 of Koprowski et al. (2017), we note that the best-fit Schechter function is obtained at 250 μm in four large z bins from 0.5 to 4.5, with ≥5 luminosity bins around $L*$, then extrapolated to total IR luminosity and reported in all the redshift bins where the *Herschel* LF was derived (i.e. 5 z-bins between 0.4 and 2.3 when comparing with Magnelli et al. 2013 and 10 z-bins between 0.1 and 4.2 when comparing with Gruppioni et al. 2013). S2CLS data are not shown in figs. 7 and 8 of Koprowski et al. (2017), but only the Schechter function, which by-definition has a steeper bright-end than the modified-Schechter (commonly used to reproduce IR data; see Saunders et al. 1990). The functional shape nevertheless doesn't seem to play a crucial role in the total IR LFs discrepancy, as we have verified by fitting the rest-frame 250-μm LF reported by Koprowski et al. (2017) with a modified Schechter function, then converting it to total IR LF through the $L_{250\mu{\rm m}}/L_{\rm IR}$ ratio given by the Michałowski et al. (2010) template, and finally interpolating in the parametre-space to match the same redshift bins of Gruppioni et al. (2013) (in order to follow exactly the same procedure described in the SCUBA-2 work). The results of this test are illustrated in Fig. 5: the deep-pink solid line shows the best-fit modified-Schechter function and the dark-green dashed line the Koprowski et al. (2017) Schechter curve (both converted from 250-μm, as described above). We find that the two functions are not significantly different in the plotted range, in any case not enough to explain the discrepancy with data, although both curves obtained from our test appear closer to the *Herschel*–PEP data points (black dots) than the one reported in fig. 8 of Koprowski et al. (2017). The bright-end of the total IR LFs at $z > 1$ is not well reproduced by neither the Schechter nor the modified-Schechter function, although a much better agreement with the data is obtained by considering the $L_{250\mu{\rm m}}/L_{\rm IR}$ ratio for a spiral galaxy template, and for the IRAS 20551 template, respectively at $z <$ and >2 (see, e.g. the orange dot–dashed and the blue dot–dot-dot–dashed lines in Fig. 5, showing the modified-Schechter and the Schechter function, respectively, obtained with these ratios). These latter templates are more similar to the average SEDs of the populations that dominate the *Herschel*–PEP LFs in the two redshift intervals (although the calculation reported here is rougher than a proper one performed on an object-by-object basis). From this test, we can therefore conclude that the considered template SEDs, combined with incompleteness issues, are likely the principal players in the observed discrepancies.

2.5 The 850-μm source counts

The comparison between the 850-μm source counts obtained from the S2CLS and from other literature surveys, and the ones estimated from the *Herschel* total IR LFs by Koprowski et al. (2017) also showed a significant inconsistency (see figure 9 of Koprowski et al. 2017). In particular, the *Herschel*-derived 850-μm counts seemed to predict an order of magnitude more sources than observed at bright 850-μm flux densities (e.g. >10 mJy). This severe overprediction of the 850-μm counts produced by the *Herschel* IR LFs led the authors to the conclusion that the high values reported from these studies most likely reflect problems in source identification and redshift estimation arising from the large-beam long-wavelength SPIRE data, as well as potential blending issues. However, to obtain this result Koprowski et al. (2017) used a single SED for all the *Herschel* sources and a unique evolution for the whole LF (while different evolutionary paths had been found for different populations; Gruppioni et al. 2013). Here we demonstrate how the simple use of different ingredients (e.g. the best-fit SED for each source and different evolutions for the different IR populations) provides results based on the *Herschel* IR LF much closer to the observations, though without any fine tuning aiming at reproducing the observed 850-μm source counts and redshift distributions. In Fig. 6, we show the 850-μm integral source counts obtained by integrating the *Herschel* LFs and the evolutions found for each SED-class of *Herschel* galaxies by Gruppioni et al. (2013), then converting to sub-mm wavelength using the best-fit SEDs for each source. Note that the *Herschel* LF is well constrained by data only to $z \sim 3$, with an upper limit derivation at $3.0 < z \leq 4.2$. At higher redshifts, we need to extrapolate the evolutions: the results shown in the plot have been obtained by assuming a negative density evolution for all the populations (e.g. $\propto (1+z)^{-3}$) at $z > 3$. The bright-end of the 850-μm counts is slightly overestimated if we compare to the Geach et al. (2017) and Simpson et al. (2015) works, but is very close to the Casey et al. (2013) and Chen et al. (2013b) points. In any case, it is very far (and much closer to the data) from what Koprowski et al. (2017) claim to be the prediction obtained by integrating the Grup-

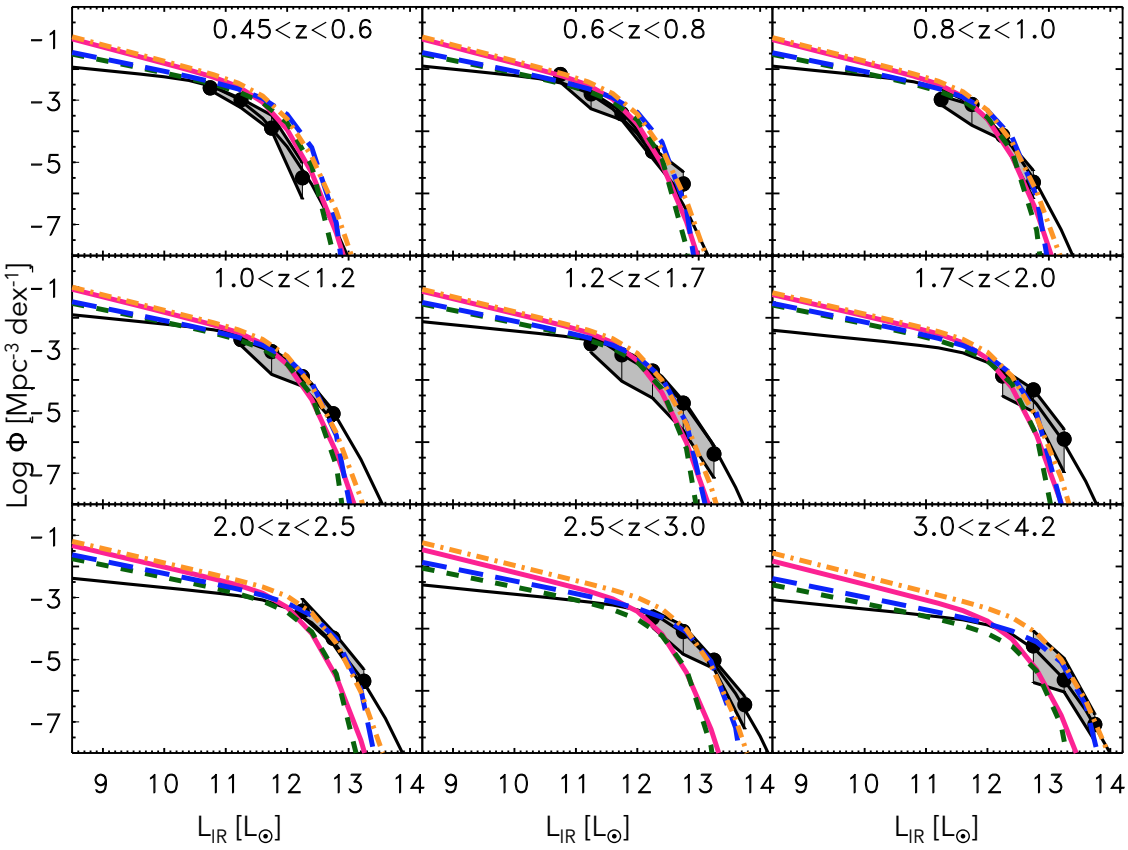

Figure 5. Comparison between the *Herschel*–PEP total IR LF by Gruppioni et al. (2013), shown by the black filled circles within the grey filled area, and different functional forms best-fitting the rest-frame 250-μm LF from SCUBA-2 data derived by Koprowski et al. (2017). The dark-green dashed line shows the original Schechter function converted to total IR LF by using the $L_{250\mu m}/L_{IR}$ ratio given by the Michałowski et al. (2010) template, then interpolated in z to match the same redshift bins. The deep-pink dot–dot–dot–dashed line shows the result of the same procedure for a modified-Schechter function (e.g. Saunders et al. 1990). The blue long–dashed and the orange dot–dashed lines show the Schechter and modified-Schecter functions, respectively, obtained by considering the $L_{250\mu m}/L_{IR}$ ratio given by a spiral galaxy template at $z < 2$ and by the IRAS 20551 template at $z > 2$.

pioni et al. (2013) evolving LFs. At fluxes <10 mJy, the agreement between our *Herschel* predictions and the sub-mm observations is very good.

3 CONCLUSIONS

In order to understand the cause of the large discrepancies observed in the total IR LFs derived from *Herschel* and SCUBA-2 survey data, we have run through all the steps that lead to the total IR LF calculation, by analysing in detail all the assumptions that can be made and the incompleteness than can affect the selection at different wavelengths (i.e. 160 μm or 850 μm). We have shown how the choice of different ingredients can lead to estimates of L_{IR} and, consequently, of the total IR LF, that can significantly disagree with each other.

The main conclusions of this work can be summarized as follows:

(i) Since a wide diversity of SEDs is found in far-IR selected samples of galaxies, and since different SED-types are found to dominate the IR populations at different redshifts, the use of a single template for a whole survey is probably too restrictive: the more data are considered for constraining the galaxy SEDs (especially around the far-IR bump), the more precisely L_{IR} is derived. If the luminosities are inaccurate, objects are counted in the wrong luminosity bin, biasing the resulting total IR LF.

(ii) If the sample completeness is not accurately estimated and corrected, the resulting source number density can be wrong. In particular, the current SCUBA-2 surveys appear to be incomplete (mainly at z <3) against galaxies with 'warm" SEDs, which are indeed found to be the major contributors to the bright-end of the *Herschel* IR LF and of the SFRD at $z \gtrsim 2$.

(iii) The functional form considered to fit the LF data (i.e. Schechter or modified-Schechter function) does not seem to play a major role in the total IR LFs discrepancies.

(iv) By integrating the *Herschel* total IR LFs, considering the best-fit SED for each source and different evolutions for each galaxy population (as derived by Gruppioni et al. 2013), we have obtained an estimate of the 850-μm source counts in agreement with observations (and different from what has been obtained by Koprowski et al. (2017) by integrating the same IR LFs).

We therefore conclude that the observed differences have to be ascribed principally to the use of a single SED template in the calculations, and to a considerable incompleteness against a significant population of IR galaxies detected in far-IR surveys, but not in sub-mm ones.

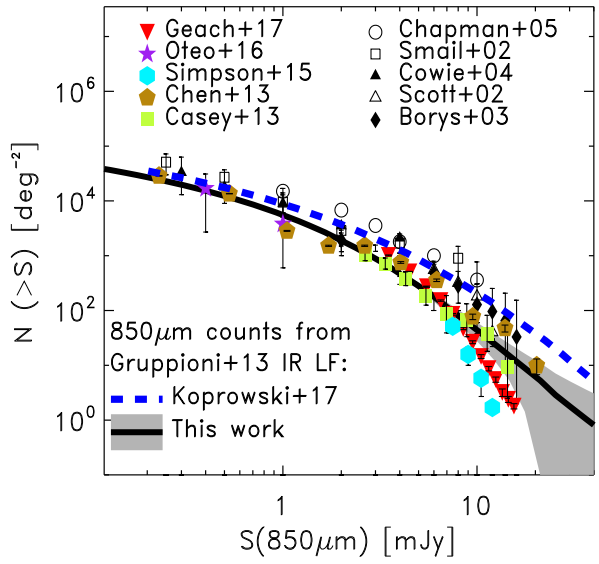

Figure 6. Integral source counts at 850 μm from SCUBA, SCUBA-2, and ALMA data from the literature are shown as different symbols (as explained in the legend) and compared with the number counts predicted from the *Herschel* total IR LF (Gruppioni et al. 2013), as derived in this work, shown by the black solid line within grey uncertainty area, and by Koprowski et al. (2017), shown by the blue dashed line.

Finally, we must note that the *Herschel* PACS LFs (i.e. Gruppioni et al. 2013, Magnelli et al. 2013) are in perfect agreement with totally independent derivations, either from *Spitzer*-24 μm (Rodighiero et al. 2010), SPIRE (Lapi et al. 2011; Marchetti et al. 2016; Rowan-Robinson et al. 2016), VLA-3 GHz (Novak et al. 2017), and –35 GHz data (Riechers et al. 2018). The latter work, reporting the measurement of the CO luminosity function at $z \sim 2$–3 and ∼5–7, as part of the CO Luminosity Density at High Redshift (COLDz) survey, shows a very good agreement with the empirical predictions by Vallini et al. (2016), which are indeed based on the Gruppioni et al. (2013) *Herschel* IR LF. In fact, the Vallini et al. (2016) CO LF seem to be the only ones reproducing the excess of bright CO sources compared to the semi-analytical predictions observed at high-z. Moreover, the new method to super-deblend *Herschel* images based on priors at other wavelenghts presented by Liu et al. (2018), provides values of SFRD very similar to those previously derived by Gruppioni et al. (2013) and Magnelli et al. (2013).

ACKNOWLEDGEMENTS

We thank an anonymous referee and the scientific editor for helpful comments that led to significant improvements in the paper. A particular thank goes to M. Negrello and G. De Zotti for fruitful discussions and for urging us to work on this subject. We acknowledge funding from the INAF PRIN-SKA 2017 program ESKAPE 1.05.01.88.04.

REFERENCES

Bakx T. J. L. C. et al., 2018, MNRAS, 473, 1751
Bourne N. et al., 2017, MNRAS, 467, 1360
Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015, ApJ, 811, 140
Bussmann R. S. et al., 2015, ApJ, 812, 43
Casey C. M. et al., 2013, MNRAS, 436, 1919
Chabrier G., 2003, PASP, 115, 763
Chapman S. C. et al., 2010, MNRAS, 409, L13
Chapman S. C., Smail I., Blain A. W., Ivison R. J., 2004, ApJ, 614, 671
Chen C.-C., Cowie L. L., Barger A. J., Casey C. M., Lee N., Sanders D. B., Wang W.-H., Williams J. P., 2014, ApJ, 784, 9
Chen C.-C., Cowie L. L., Barger A. J., Casey C. M., Lee N., Sanders D. B., Wang W.-H., Williams J. P., 2014, ApJ, 789, 12
Dunlop J. S. et al., 2017, MNRAS, 466, 861
Eales S. et al., 2010, PASP, 122, 499
Elbaz D. et al., 2011, A&A, 533, A119
Geach J. E. et al., 2017, MNRAS, 465, 1789
Griffin M. J. et al., 2010, A&A, 518, L3
Gruppioni C. et al., 2013, MNRAS, 432, 23
Gruppioni C. et al., 2015, MNRAS, 451, 3419
Hill R. et al., 2018, MNRAS, 477, 2042
Hodge J. et al., 2013, ApJ, 768, 91
Kennicutt R. C., Jr., 1998b, ARA&A, 36, 189
Koprowski M. P., Dunlop J. S., Michałowski M. J., Coppin K. E. K., Geach J. E., McLure R. J., Scott D., van der Werf P. P., 2017, MNRAS, 471, 4155
Lapi A. et al., 2011, ApJ, 742, 24
Laporte N. et al., 2017, ApJ, 837, L21
Liu D. et al., 2018, ApJ, 853, 172
Lutz D. et al., 2011, A&A, 532, A90
Lutz D., 2014, ARA&A, 52, 373
Magnelli B. et al., 2010, A&A, 518, L28
Magnelli B. et al., 2013, A&A, 553, A132
Marchetti L. et al., 2016, MNRAS, 456, 1999
Michałowski M., Hjorth J., Watson D., 2010, A&A, 514, A67
Nguyen H. T. et al., 2010, A&A, 518, L5
Novak M. et al., 2017, A&A, 602, A5
Oliver S. J. et al., 2012, MNRAS, 424, 1614
Poglitsch A. et al., 2010, A&A, 518, L2
Riechers D. A. et al., 2013, Nature, 496, 329
Riechers D. A. et al., 2018, preprint (arXiv:1808.04371)
Rodighiero G. et al., 2010, A&A, 515, A8
Roseboom I. G. et al., 2010, MNRAS, 409, 48
Rowan-Robinson M. et al., 2016, MNRAS, 461, 1100
Saunders W., Rowan-Robinson M., Lawrence A., Efstathiou G., Kaiser N., Ellis R. S., Frenk C. S., 1990, MNRAS, 242, 318
Simpson J. M. et al., 2015, ApJ, 807, 128
Swinyard B. M. et al., 2010, A&A, 518, L4
Vallini L., Gruppioni C., Pozzi F., Vignali C., Zamorani G., 2016, MNRAS, 456, L40

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