

Towards a census of high-redshift dusty galaxies with *Herschel*

A selection of “500 μm -risers”

D. Donevski¹, V. Buat², F. Boone², C. Pappalardo^{3,4}, M. Bethermin^{1,7}, C. Schreiber⁵, F. Mazyed¹,
J. Alvarez-Marquez⁶, and S. Duivenvoorden⁸

¹ Aix-Marseille Université, CNRS, LAM, Laboratoire d’Astrophysique de Marseille, Marseille, France
e-mail: darko.donevski@lam.fr

² Université de Toulouse, UPS-OMP, IRAP, Toulouse, France

³ Centro de Astronomia e Astrofísica da Universidade de Lisboa, Observatório Astronómico de Lisboa, Tapada da Ajuda,
1349-018 Lisboa, Portugal

⁴ Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Tapada da Ajuda, 1349-018 Lisboa, Portugal

⁵ Leiden Observatory, Leiden University, 2300 RA Leiden, The Netherlands

⁶ Departamento de astrofísica, Centro de Astrobiología (CAB, CSIC-INTA), Carretera de Ajalvir, 28850 Torrejón de Ardoz,
Madrid, Spain

⁷ European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany

⁸ Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK

Received 3 September 2017 / Accepted 30 January 2018

ABSTRACT

Context. Over the last decade a large number of dusty star-forming galaxies has been discovered up to redshift $z = 2 - 3$ and recent studies have attempted to push the highly confused *Herschel* SPIRE surveys beyond that distance. To search for $z \geq 4$ galaxies they often consider the sources with fluxes rising from $250 \mu\text{m}$ to $500 \mu\text{m}$ (so-called “500 μm -risers”). *Herschel* surveys offer a unique opportunity to efficiently select a large number of these rare objects, and thus gain insight into the prodigious star-forming activity that takes place in the very distant Universe.

Aims. We aim to implement a novel method to obtain a statistical sample of 500 μm -risers and fully evaluate our selection inspecting different models of galaxy evolution.

Methods. We consider one of the largest and deepest *Herschel* surveys, the *Herschel* Virgo Cluster Survey. We develop a novel selection algorithm which links the source extraction and spectral energy distribution fitting. To fully quantify selection biases we make end-to-end simulations including clustering and lensing.

Results. We select 133 500 μm -risers over 55 deg^2 , imposing the criteria: $S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2 \text{ mJy}$ and $S_{500} > 30 \text{ mJy}$. Differential number counts are in fairly good agreement with models, displaying a better match than other existing samples. The estimated fraction of strongly lensed sources is $24^{+6}_{-5}\%$ based on models.

Conclusions. We present the faintest sample of 500 μm -risers down to $S_{250} = 13.2 \text{ mJy}$. We show that noise and strong lensing have an important impact on measured counts and redshift distribution of selected sources. We estimate the flux-corrected star formation rate density at $4 < z < 5$ with the 500 μm -risers and find it to be close to the total value measured in far-infrared. This indicates that colour selection is not a limiting effect to search for the most massive, dusty $z > 4$ sources.

Key words. galaxies: evolution – galaxies: statistics – galaxies: photometry – galaxies: star formation – galaxies: high-redshift – infrared: galaxies

1. Introduction

The abundance of dusty galaxies at high-redshifts ($z > 4$) constrains our theories about early galaxy formation, since it is generally stated that they are the progenitors of massive ellipticals seen in overdense regions of the local Universe (e.g. Eales et al. 2017; Toft et al. 2014; Simpson et al. 2014; Casey et al. 2014).

The widely accepted picture is that dusty star-forming galaxies (DSFGs) occupy the most massive halos in early epochs and lie on the most extreme tail of the galaxy stellar mass function (e.g. Ikarashi et al. 2017; Fudamoto et al. 2017; Oteo et al. 2016; Mancuso et al. 2016; Michałowski et al. 2014). These DSFGs are usually selected in the far-infrared (FIR) regime where the star formation rates can be directly measured.

Large FIR surveys, such as those conducted with the *Herschel* Space Observatory (Pilbratt et al. 2010), provide an opportunity to build a thorough census of prodigious starbursts over cosmological redshifts using a wide and blind search concept. The *Herschel* SPIRE photometer (Griffin et al. 2010) was often used for mapping large areas at wavelengths of $250 \mu\text{m}$, $350 \mu\text{m}$ and $500 \mu\text{m}$. The redshift peak of most *Herschel*-detected sources matches with the redshift where galaxies have formed most of their stars ($z \sim 2$; Pearson et al. 2013; Lapi et al. 2011; Amblard et al. 2010). Considering that rest-frame dust spectral energy distribution (SED) of a galaxy typically peaks between 70 and $100 \mu\text{m}$, colours of sources in *Herschel* SPIRE bands were used to select candidate high-redshift dusty objects. To search for $z \geq 4$ candidates, there is a particular interest to exploit the sources with red SPIRE colours,

with rising flux densities from 250 to 500 microns (so-called “500 μm -risers”). Such galaxies should lie at $z \geq 4$ unless they have dust temperatures that are notably lower than is seen in local FIR-bright equivalents (Asboth et al. 2016; Yuan et al. 2015; Dowell et al. 2014; Roseboom et al. 2012).

If the selection of 500 μm -risers is free of contaminants such as blended systems and powerful non-thermal sources (e.g. quasars), it is expected to provide us insight into very distant and dusty, star-forming galactic events. Recently, literature on dusty high- z galaxies has rapidly flourished, including a handful of serendipitously discovered 500 μm -risers (e.g. Daddi et al. 2009; Cox et al. 2011; Capak et al. 2011; Combes et al. 2012; Vieira et al. 2013; Miettinen et al. 2015; Negrello et al. 2017). However, these findings had serious shortcomings – they were limited to few objects with red SPIRE colours¹.

Being primarily focused on individual (usually strong lensing) candidates or millimeter-selected samples regardless of the galaxy colour, they poorly constrained the statistics of 500 μm -risers. To derive a larger number of potentially unlensed 500 μm -risers and analyse them in a more standardised manner, several works used map-search technique (Asboth et al. 2016; Dowell et al. 2014; Ivison et al. 2016). These led to the discovery of the most distant dusty starburst galaxies known to date: SPT0311-58 at $z = 6.902$ (Strandet et al. 2017), HFLS3 at $z = 6.34$ (Riechers et al. 2013) and G09-83808 at $z = 6.02$ (Zavala et al. 2018; Fudamoto et al. 2017). Furthermore, selecting the 500 μm -risers from lowest-resolution *Herschel* SPIRE-maps, Dowell et al. (2014) and Asboth et al. (2016) claimed significant overprediction of observed number of 500 μm -risers, one order of magnitude higher than predicted by the existing models (e.g. Hayward 2013; Béthermin et al. 2012).

Even if *Herschel* offers a direct insight into the IR emission of high- z objects, there are critical limitations, such as sensitivity of detectors and low spatial resolution. These are responsible for biases such as source confusion. The sensitivity of SPIRE allows for the direct detection of only the brightest and thus rarest objects at the tip of the luminosity function (Cowley et al. 2015; Karim et al. 2013; Oliver et al. 2010), and we therefore need large surveys to increase the statistics. The study of Dowell et al. (2014) analysed maps of three different extragalactic fields observed as part of the HerMES (*Herschel* Multi-tiered Extragalactic Survey, Oliver et al. 2010) program, while Ivison et al. (2016) and Asboth et al. (2016) probed much wider but shallower areas of the *H*-ATLAS and HeLMS (HerMES Large Mode Survey) fields, respectively (see Sect. 2.1 for details about field properties).

In this work we aim to introduce a slightly different approach to building a statistically significant sample of red, potentially $z > 4$ sources. The new selection scheme we propose is motivated by the size and the depth of the field we chose to investigate. *Herschel* Virgo Cluster Survey (HeViCS, Davies et al. 2010) is deeper than the one used in the analysis of Asboth et al. (2016) and Ivison et al. (2016), and larger than the area analysed by Dowell et al. (2014).

The paper is organised as follows: in Sect. 2 we describe the methods we used for the data analysis and our new selection criteria for 500 μm -risers. In Sects. 3 and 4 we present expected redshift/luminosity trend of selected galaxies and differential number counts. In Sect. 5 we compare our results to different

models. We perform simulations to review all selection biases and highlight the necessity of a further refinement of selection criteria in searches for $z \geq 4$ sources. The nature of 500 μm -risers and our main conclusions are outlined in Sects. 6 and 7, respectively. We assume a Planck Collaboration XIII (2016) cosmology.

2. Data analysis

2.1. HeViCS field

HeViCS is a fully sampled survey that covers a region centered at the Virgo cluster (Davies et al. 2010, 2012). It is one of the largest uniform *Herschel* surveys, and its main advantage is its sensitivity and uniformity of data. In this survey, *Herschel* observed four overlapping fields (fields V1–V4) in fast parallel-mode. The total entire survey region is 84 square degrees, where 55 square degrees are observed at unvarying depth with eight orthogonal cross scans (see Auld et al. 2013 and Pappalardo et al. 2015 for more details).

HeViCS observations reached a depth close to the confusion limit at the shortest SPIRE wavelength (250 μm). Because of the number of repeated scans, instrumental noise is significantly reduced in HeViCS maps, giving the 1σ levels of 4.9, 4.9, and 5.5 mJy at 250 μm , 350 μm , and 500 μm , respectively (Auld et al. 2013). In the overlapped, deeper regions recorded by 16 scans, these values are even smaller, namely 3.5, 3.3, and 4.0 mJy. Due to the presence of bright sources (see Sect. 2.3), global noise estimation is not a straightforward task. We excluded bright sources from the map by masking them, and after their removal the global noise was derived from the variance of the map. It reaches the 1σ values of 6.58, 7.07, and 7.68 mJy (250 μm , 350 μm , and 500 μm , respectively) for a major area covered by eight cross-scans. The extensive contributor to the overall noise measured in HeViCS maps is confusion noise, usually defined as the the variance in the sky map due to the fluctuations of unresolved sources inside the SPIRE beam. We calculate confusion using the relation $\sigma_{\text{conf}} = \sqrt{\sigma_{\text{tot}}^2 - \sigma_{\text{inst}}^2}$, where σ_{tot} is the total noise measured in the map, and σ_{inst} is the instrumental noise. Values determined for the confusion noise are 4.4, 5.2, and 5.5 mJy at the 250 μm , 350 μm , and 500 μm bands, respectively, and are almost identical to the ones presented in Auld et al. (2013). These values are also close to the confusion noise measured in HerMES maps, within twice the uncertainty of 3σ -clipping estimates from Nguyen et al. (2010) (3.8 mJy, 4.7 mJy, and 5.2 mJy at the 250 μm , 350 μm , and 500 μm bands, respectively). In this work we use only SPIRE data. In parallel, each HeViCS tile has been observed by the *Herschel* PACS instrument, but the depth of PACS data ($5\sigma_{\text{tot}} = 70$ mJy at 100 μm , Pappalardo et al. 2016) is not sufficient to directly detect our FIR-rising sources. PACS data at 100 μm and 160 μm will be added together with a deep optical-NIR map from the Next Generation Virgo Cluster Survey (NGVS, Ferrarese et al. 2012) in a following paper analysing ancillary data.

2.2. An overview of other *Herschel* fields

There are several large *Herschel* fields (e.g. with an observed area $\theta > 10$ deg²) used for detection of 500 μm -risers. A summary of their properties is shown in Table 1. All the values are taken from the literature (Oliver et al. 2012; Wang et al. 2014).

The *H*-ATLAS survey (Eales et al. 2010) is used to select red candidates by Ivison et al. (2016). *H*-ATLAS is designed to

¹ Throughout the text we adopt the following terminology regarding colours of IR-detected sources: red colours and red sources refer to 500 μm -risers, while “350 μm peakers” refer to galaxies having SED peak at 350 μm .

Table 1. Properties of different large fields mapped by *Herschel*.

Field	Mode	$N(\text{rep})$	Time [hr]	Area [deg ²]	Noise level in mJy 5σ at $250\ \mu\text{m}$
(1)	(2)	(3)	(4)	(5)	(6)
HeViCS	Parallel	8	286	55	30.5
<i>H</i> -ATLAS	Parallel	2	600	550	56.0
HerMES					
FLS	Parallel	2	17.1	6.7	25.8
Bootes NDWFS	Parallel	2	28	10.5	25.8
ELAIS-N2	Parallel	2	28	12.28	25.8
Lockman-Swire	Parallel	4	71.2	16	13.6
XMM-LSS SWIRE	Parallel	4	71.2	18.87	25.8
HeLMS	Sp.Fast	2	103.4	274	64.0

Notes. *Columns* (1) name of the field observed by *Herschel*; (2) *Herschel* observing mode; (3) total number of repeats of the observing mode in the set; (4) total time of observations; (5) field area of good pixels; (6) total noise from the literature. Noise is calculated using the relation $\sigma_{\text{tot}} = \sqrt{\sigma_{\text{conf}}^2 + \sigma_{\text{inst}}^2}$. The bottom six fields are areas with different design levels nested as a part of HerMES: total area covered in HerMES is $380\ \text{deg}^2$. Shallow HeLMS field covers the area of $274\ \text{deg}^2$, and deeper Level 1–Level 6 fields cover the total area of about $80\ \text{deg}^2$. FLS and Lockman-SWIRE have been used to probe $500\ \mu\text{m}$ -risers selection by [Dowell et al. \(2014\)](#), while [Asboth et al. \(2016\)](#) applied the same selection method in the HeLMS field.

uniformly cover $600\ \text{deg}^2$ of sky, but with its two scans, the survey did not reach the level of confusion noise at $250\ \mu\text{m}$. Studies by [Asboth et al. \(2016\)](#) and [Dowell et al. \(2014\)](#) acquired data from different fields which are part of the HerMES survey ([Oliver et al. 2012](#)). The HerMES survey observed $380\ \text{deg}^2$ of the sky. The survey has a hierarchical structure containing seven levels, ranging from very deep observations of clusters to wider fields with varying size and depth. The largest (and the shallowest) observed area is HeLMS, with its 274 square degrees. HeViCS maps consist of two or four times more scans compared to other fields listed in Table 1. This leads to a reduction of instrumental noise by a factor of $\sqrt{2}$. The global $250\ \mu\text{m}$ noise in HeViCS is smaller than in *H*-ATLAS and HeLMS. However, it is still higher than in other HerMES fields, showing that confusion is a very compelling supplier to the noise budget for point sources in the HeViCS field. We clarify that statement by repeating our analysis of regions overlapped between the tiles, which have greater coverage. We found no significant noise reduction, implying that the maps are dominated by confusion noise.

2.3. Map filtering

Prior to performing source extraction on SPIRE maps, we reduce the background contamination. The $250\ \mu\text{m}$ map of V2 field in HeViCS is strongly affected by galactic cirrus emission. This contamination peaks at around 150 – $200\ \mu\text{m}$ ([Bracco et al. 2011](#); [Valiante et al. 2016](#)), implying that it is the brightest in the shortest SPIRE bands. The main effect of cirrus emission is to increase the confusion in the maps, but the small-scale structure within it can also lead to spurious detections in the catalogues.

We follow the method applied by [Pappalardo et al. \(2015\)](#). Using the SEXtractor ([Bertin & Arnouts 1996](#)) we re-grid the $250\ \mu\text{m}$ map into meshes larger than the pixel size. The SEXtractor makes an initial pass through the pixel data, and computes an estimator for the local background in each mesh of a grid that covers the whole adopted frame. We apply repetitive iterations to estimate the mean and standard deviation of the pixel value distribution in boxes, removing outlying pixels at each iteration. An important step in this procedure is the choice of the box size, since we do not want the background

estimation to be affected by the presence of objects and random noise. The box size should generally be larger than the typical size of sources in the image, but small enough to encapsulate any background variations. We therefore fix the mesh size to 8 pixels, adopting the result from [Pappalardo et al. \(2015\)](#). The local background is clipped iteratively to reach $\pm 3\sigma$ convergence around its median. After the background subtraction, the number of sources appears uniform for regions with different cirrus emission. We use the background subtracted map as an input for the source-extraction process described in the following subsection.

2.4. Extraction of sources

Blind SPIRE source catalogues have been produced for HeViCS ([Pappalardo et al. 2015](#)). Nonetheless, when the density of sources is very high, which is the case in the highly crowded HeViCS field, blind source extraction cannot separate blended point sources in an efficient way. Additionally, it remains difficult to properly cross-match sources at different wavelengths, since central positions from blind catalogues are not well constrained. To deal with source multiplicity we chose to perform extraction of $350\ \mu\text{m}$ and $500\ \mu\text{m}$ fluxes at the exact a priori position of $250\ \mu\text{m}$ band detections, allowing much more precise identification. The potential limitation of such a method is that we might be eventually miss some $500\ \mu\text{m}$ -risers that are not included in the prior list after the first iteration of source extraction. We thus run our method iteratively and add new sources that may appear in the residuals at each iteration.

We use SUSSEXtractor ([Savage & Oliver 2007](#)), a source-finding algorithm optimised for isolated point-sources, to create the catalogue of galaxies detected in the $250\ \mu\text{m}$ map as a prior to extract the flux densities at longer wavelengths. We then implement our novel technique (see Sect. 2.6) where source deblending and single temperature modified blackbody (MBB) fitting are combined in the same procedure. In the following we explain our source-extraction pipeline (see Fig. 1 for a graphical description):

1. We run SUSSEXtractor using a fully overlapping HeViCS $250\ \mu\text{m}$ -micron map. SUSSEXtractor works on the flux-calibrated, Level 2 *Herschel* SPIRE maps. We create a point response function (PRF)-filtered image, smoothed with the

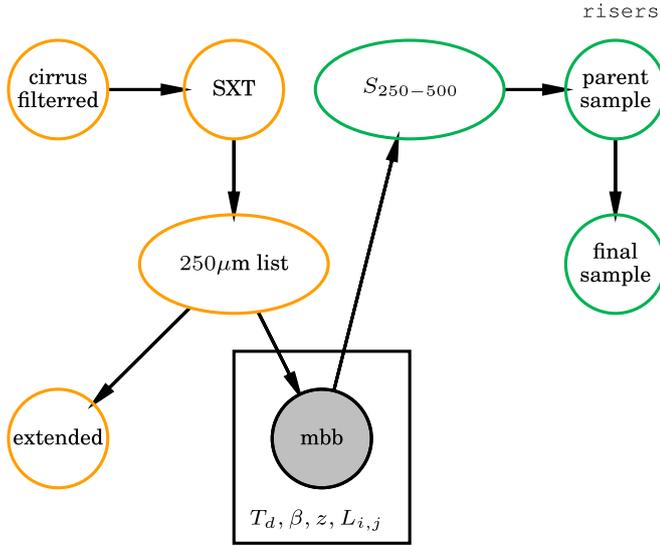


Fig. 1. Schematic representation of selection of 500 μm -risers in the HeViCS field. Coloured in orange are segments of source extraction prior to using MBB-fitter; these are (from upper left, following the arrows) subtraction of strong cirrus emission; SUSSEXtractor list of initial (threshold = 3) 250 μm detections; and assumed a priori list cleaned from extended sources. Enveloped by smaller black squares are parameters considered for the fitting procedure with MBB-fitter corresponding to all pixels in the map where we have 250 μm detections. Coloured in green are segments of our selection after performing the photometry: MBB photometry ($S_{250-500}$) at SPIRE wavelengths using 250 μm priors; parent list of 500 μm -risers (140 sources in total), not cleaned from strong synchrotron contaminants; and the final list (133 in total) of 500 μm -risers after excluding local radio sources. We apply the following selection criteria for the final sample: $S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2$ mJy, and $S_{500} > 30$ mJy.

PRF. In SUSSEXtractor, the PRF is assumed to be Gaussian by default, with the full-width-at-half-maximum (FWHM) provided by the FWHM parameter. We apply values of 17.6", 23.9", and 35.2" at 250, 350, and 500 μm , respectively. Subsequently, pixel sizes at these bands are 6", 10", and 14". The algorithm searches for a local maximum which is the highest pixel value within a distance defined by pixel region. The position assigned to the possible source is then refined by fitting a quadratic function to certain pixels in the PRF-filtered image.

2. To search for even fainter sources that are closer to the confusion limit and usually masked/hidden in highly confused fields like HeViCS, we perform an additional step looking for detections in our residual maps. Applying such an additional step, we increase the total number of sources by around 4%. Errors in the position estimated to be a source are determined as 0.6 times the ratio of the FWHM to the signal-to-noise ratio (S/N), up to a maximum of 1.0 pixel, as suggested in the literature (Iverson et al. 2007).
3. We build an initial list of 250 μm sources selecting all point sources above the threshold = 3 (Bayesian criteria in SUSSEXtractor). We also impose the flux density cut choosing values equal to or higher than 13.2 mJy, which corresponds to $S_{250} \geq 3\sigma_{\text{conf}}$. As a result of our source extraction pipeline at 250-micron maps, we listed 64 309 sources, similarly distributed among the four (V1–V4) fields.
4. List of 250 μm detections is then divided into a “no-neighbour” list of sources (250 μm sources without another detection inside the 500 μm beam) and a “close-neighbour”

list where we add all sources that have another 250 μm detection within the 500 μm beam. Prior to assigning the initial 250 μm list as an input to our modified blackbody fitting procedure, we remove potentially extended sources from the catalogue.

2.5. Extended sources

The Virgo cluster is one of the richest local clusters and we expect to have a significant number of extended sources. Objects that are extended on the SPIRE beam scale (see Wang et al. 2014 or Rigby et al. 2011) are not expected to be accurately identified with point-source extracting codes, and large galaxies may be misidentified as multiple point sources. To solve this problem, we implement the same method as used in Pappalardo et al. (2015). They define a mask using the recipe of SEXtractor, which detects a source when a fixed number of contiguous pixels is above a σ -threshold estimated from the background map. We keep the same value tested in Pappalardo et al. (2015) – 70 contiguous pixels above 1.2σ .

In this case, most of the sources larger than 0.7 arcmin² are rejected from the sample. Additionally, we cross-match all remaining HeViCS 250 μm detections with their nearest (within 36") counterpart in the 2MASS Extended Source Catalog (Jarrett et al. 2000). We remove any detection supposed to be a counterpart with a Kron elliptical aperture semi-major larger than 9". In total we suppressed 812 sources from the analysis, thus decreasing the number of 250 μm detections in our parent list from 64 309 to 63 497.

The list of sources detected at 250 μm down to 13.2 mJy, cleaned from galactic cirrus contaminants and extended sources, is further used as an a priori list for our simultaneous modified blackbody fitting.

2.6. Modified blackbody fitter

As described in Sect. 2.1, our maps are limited by confusion noise caused by the high density of sources relative to the resolution of the *Herschel* instrument. In cases where SPIRE images have multiple sources per beam, measured fluxes might be biased if we treat several blended sources as one. Several different approaches for source “de-blending” exist in the literature. We can separate these into three groups: Firstly, there are de-blending methods that use only positional priors as input information (e.g. Elbaz et al. 2010; Swinbank et al. 2014; Béthermin et al. 2010; Roseboom et al. 2010). The second group of de-blenders consists of methods which combine positional information with statistical techniques (Hurley et al. 2017; Safarzadeh et al. 2015); for example, Safarzadeh et al. (2015) and Hurley et al. (2017) developed the codes based on the Monte Carlo Markov chain (MCMC) sampling for prior source detections, and while Safarzadeh et al. (2015) used *Hubble* Space Telescope (HST) *H*-band sources as a positional argument, Hurley et al. (2017) developed a similar MCMC-based prior-extraction for *Spitzer* 24 μm detections. Such algorithms are efficient in obtaining Bayesian PDFs for all priors, however they might still have some limitations, the most important being the potential inability to unveil the true high- z sources in cases where they are not included in the initial list. Finally, a third group of de-blenders consists of those using SED modelling as an addition to positional arguments (MacKenzie et al. 2014; Shu et al. 2016; Merlin et al. 2016; Liu et al. 2018).

Due to the lack of *Spitzer* 24 μm data for the HeViCS field, and unsufficiently deep existing optical images, we have no opportunity to use positional priors from shorter-wavelength

surveys, and we based our analysis on SPIRE data instead. We use models to test the whole procedure of selecting 500 μm -risers in Sect. 5.

MBB-fitter is a code developed to extract sources from multiwavelength bolometric observations (Boone et al., in prep.) combining positional priors and spectral information of sources such that SEDs of fitted galaxies follow MBB shape as defined by Blain et al. (2003). Our MBB deblending approach is thus very similar to the method applied by MacKenzie et al. (2014). Although such a model can only encapsulate a segment of the general complexity of astrophysical processes in a galaxy, it can account for the SED data for a wide variety of dusty galaxies. The MBB-fitter method is described in detail and is tested on simulations in Boone et al. (in prep.). We summarise here the main points. In MBB-fitter, the map (M) of a sky region is described as a regular grid corresponding to the sky flux density distribution convolved by the point spread function (PSF) of an instrument and sampled in pixels. Thus, $M_i = M(\alpha_i, \delta_i)$ refers to the value of the map at the i th pixel with coordinates α_i, δ_i . We further assume that the sources have a morphology independent of the frequency. The flux density distributions in the three-dimensional (3D) space $s_k(\alpha, \delta, \nu)$, where ν is frequency, can therefore be decomposed into a spatial distribution term (morphology) and an SED term:

$$M_{ij} = M(\alpha_i, \delta_i, \nu_j) = \left[\sum_{k=1}^{N_{\text{sources}}} f_k(\nu_j) \tilde{s}_k(\alpha - \alpha_k, \delta - \delta_k) \otimes \text{PSF}_j \right]_i, \quad (1)$$

where PSF_j is the PSF of the map at a given frequency (ν_j), and N_{sources} is the number of sources considered for the fitting, with its reference coordinates (α_k, δ_k) . Here we assume that the thermal continuum SED of galaxies in the FIR domain can be modelled as a modified blackbody of the form (Blain et al. 2003)

$$f(\nu, T_d, \beta, \gamma) \propto \begin{cases} \nu^{3+\beta} / [\exp(h\nu/kT_d) - 1], & \text{if } \nu < \nu_w \\ \nu^{-\gamma}, & \text{if } \nu \geq \nu_w, \end{cases} \quad (2)$$

where ν_w is the lower boundary of Wien's regime, h is the Planck constant, and k is the Stefan-Boltzmann constant. Thus, Eq. (1) describes a general model for a set of maps at different wavelengths, and each source SED is defined by a set of parameters arranged into a vector of the form $P_k = [L_{\text{IR}}, T_d, z_k, \beta_k, \gamma_k]$, where z_k is the redshift of a given source, β_k is its emissivity index, γ_k is the index of the power law to substitute the Wien's regime, and L_{IR} is the IR luminosity of a source. This model is parametrized by the source SED parameters, P_k , and coordinates α_k, δ_k . The total number of parameters is therefore $N_{\text{par}} = N_s \times (N_{\text{SED}_{\text{par}}} + 2)$, where $N_{\text{SED}_{\text{par}}}$ is the number of SED parameters and N_s is the number of sources.

The introduced set of parameters reflects the global properties of the source. There are potential degeneracies, meaning that different values of parameters may give very similar SEDs. This is the most prolific challenge when using the FIR peak as a redshift indicator; dust temperature and the redshift are completely degenerate (see e.g. Pope & Chary 2010).

A defined model can be compared to a set of observed maps and its fidelity can be assessed with a figure of merit such as the chi-square, exemplified as the sum of the pixel deviations:

$$\chi(P)^2 = \sum_i^{N_{\text{pix}}} \sum_j^{N_{\text{freq}}} \frac{(M_{ij}(P) - \tilde{M}_{ij})^2}{\sigma_{ij}^2}, \quad (3)$$

where \tilde{M}_{ij} is the multiwavelength set of maps, σ_{ij} is the Gaussian noise level of the i th pixel of the map at the j th frequency, N_{pix} is the number of accounted pixels, while N_{freq} represents the number of maps, which is equal to the number of different frequencies used.

We use the Levenberg-Marquardt algorithm to find the minimum of the chi-square function in the space of parameters. The algorithm also returns an estimation of the errors on the parameters based on the covariance matrix. We note that it is not necessary to include all the pixels of maps in the chi-square computation; it can be restricted to a subset of pixels (contiguous or not), and sky regions without prior sources can be excluded.

We impose our list of 250 μm detected sources as a prior list, and use MBB-fitter to perform two-step photometry: (1) simultaneously fitting the sources affected by surrounding source-blend (12 135 sources in total), and (2) performing the photometry of sources that are more isolated without a surrounding companion inside the beam (51 362 sources in total). We set emission spectral slope and dust temperature at $\beta = 1.8$, and $T_d = 38 \pm 7$ K. These values are chosen to provide a very good description of the data and they are based on our current knowledge of dust temperatures in SPIRE-detected sources (e.g. Yuan et al. 2015; Swinbank et al. 2014; Symeonidis et al. 2013; Lapi et al. 2011, see also Schreiber et al. 2017b). An example of simultaneous MBB deblending is illustrated in Fig. 2.

2.7. Final data sample of 500 μm -risers

We select the final list of 133 500 μm -risers over the area of 55 deg^2 . Selected sources fulfil criteria accepted for the final cut: $S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2$ mJy and $S_{500} > 30$ mJy. The full catalogue is presented in Table D.1.

We performed several tests and chose to set the flux density cut in the 500 μm band at $S_{500} \approx 30$ mJy, which is related to 4σ above total noise measured in HeViCS maps. Using this value, we reach a completeness level of $\sim 80\%$ at 500 μm (see Fig. 5) avoiding larger uncertainties in colours due to lower S/Ns (see Table 6 and Sect. 5.2). A 250 μm flux cut ($S_{250} > 13.2$ mJy) corresponds to $S_{250} > 3\sigma_{\text{conf}}$ after applying an iterative 3σ clipping to remove bright sources (see Sect. 2.1). Amongst the final 500 μm -risers (133 in total) we have 11 red candidates with one or two additional 250 μm detections inside the 500 μm beam. Example two-dimensional (2D) cutouts of several 500 μm -risers from our final sample are presented in Fig. 3 (see also Appendix C). Strong radio sources that have flat spectra and very prominent FIR emission (7 objects in total) have been removed from our final list of 500 μm -risers in the HeViCS field; these objects may have colours similar to those of dusty, star-forming systems that we are interested in selecting. Contamination due to radio-bright galaxies is eliminated by cross-matching existing radio catalogues: HMQ (Flesch 2015), NVSS (Condon et al. 1998), FIRST (Helfand et al. 2015) and ALFA ALFA (Giovannelli et al. 2007). Identified radio-bright sources are classified as quasars².

In Fig. 4 we show actual flux distribution of 500 μm -risers selected in this work, along with red sources from other studies. We see that our catalogue contains, on average, more fainter objects than other existing samples. The median 500 μm flux of our sample is 38 ± 4 mJy, while median fluxes found in Dowell et al. (2014), Asboth et al. (2016), and Ivison et al. (2016) are

² For the detailed description of this category of dusty, radio sources, see Appendix A.

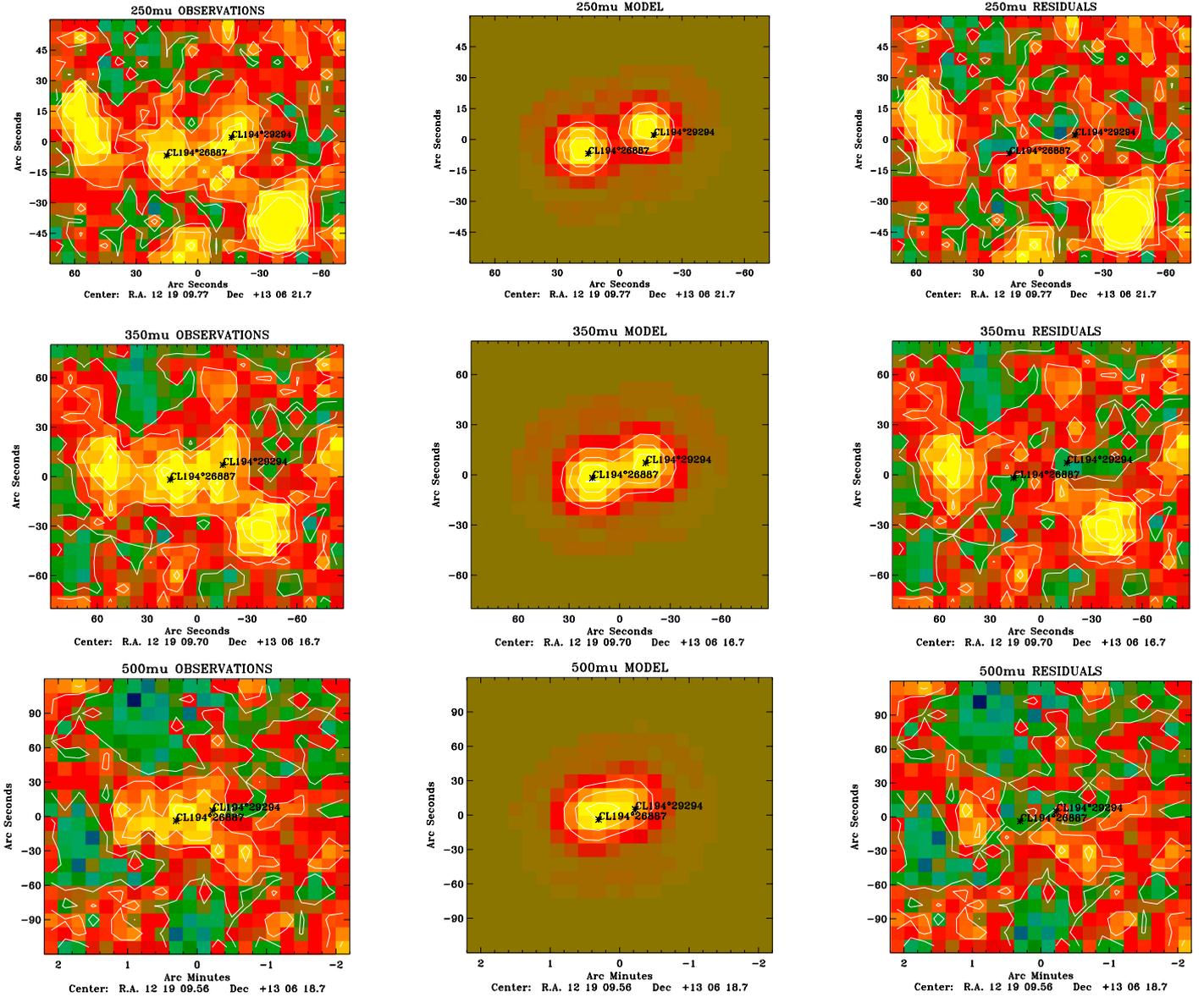


Fig. 2. Example of simultaneous prior-source deblending with MBB-fitter. Sources are taken from our “neighbour” list of 250 μm detections. Columns from left to right: 250 μm , 350 μm , and 500 μm maps of sources (first column), subtracted models (second column) and residuals (third column). We note that in this example two central sources are considered for the fitting.

45 mJy, 65 mJy, and 47 mJy, respectively. Therefore, our selection is arguably the faintest sample of 500 μm -risers available. We note that the underlying flux distribution of selected galaxies might still be affected by strong gravitational lensing, which we discuss in detail in Sect. 6.2. Along with galaxy-galaxy lensing, cluster-lens amplification might affect measured fluxes. The effect is the strongest when the deflector is located half-way between the observer and the lensed object (e.g. Kneib & Natarajan 2011). Strong lensing events are thus more numerous for clusters at $0.1 < z < 0.5$ (Johnson et al. 2014). Virgo cluster is very close along the line of sight ($z = 0.003$) and for this reason we expect negligible impact on SPIRE fluxes of 500 μm -risers. Colours of sources selected in this work, together with colours of 500 μm -risers with spectroscopic redshifts from other studies, are plotted in Fig. 6. The median observed colour of our sample is $S_{500}/S_{350} = 1.11 \pm 0.10$, whereas median colours from Dowell et al. (2014), Asboth et al. (2016), and Ivison et al. (2016) are $S_{500}/S_{350} = 1.08$, $S_{500}/S_{350} = 1.12$, and $S_{500}/S_{350} = 1.23$,

respectively. In Sects. 5.1.5 and 5.2 we illustrate and discuss the impact of the noise on observed counts and colours.

2.8. Completeness and flux accuracy

We check the quality of our data analysis by performing tests of completeness and flux accuracy. To determine these values, we use real “in-out” simulations. We calculate completeness by considering the number of injected sources with certain flux density S that are recovered in the simulated maps as real detections. Since our selection is based on a priori 250- μm positions, in this test we mostly inspect that band.

We use Monte-Carlo simulation, artificially adding Gaussian sources and projecting them at randomised positions to our real 250 μm data, before convolving with the beam. Simulated sources in 350 μm and 500 μm maps are then placed at the same positions. The FWHM values that we used to convolve are 17.6”, 23.9”, and 35.2”, that is, values we consider for our SUSSEX-tractor detections. In each HeViCS field we add sources spanning

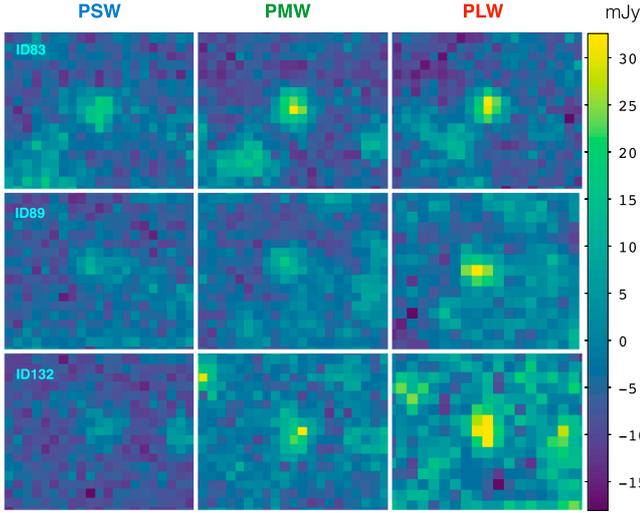


Fig. 3. 2D image cutouts as an illustration of $500\ \mu\text{m}$ -risers that fulfil our final selection criteria. Presented examples are drawn from our “no-neighbour” list, that is, sources without a clear $250\ \mu\text{m}$ detection within a radius of at least $36''$ around their centres. Colourbar shows fluxes measured in mJy.

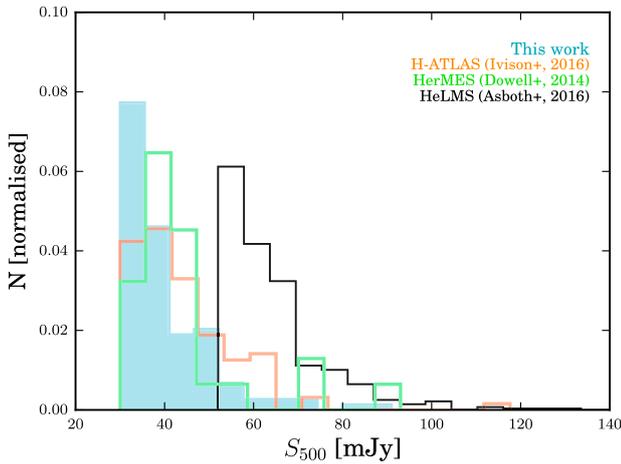


Fig. 4. Normalised distribution of $500\ \mu\text{m}$ fluxes of $500\ \mu\text{m}$ -risers. The filled blue area indicates our sample, while samples of Iverson et al. (2016), Dowell et al. (2014) and Asboth et al. (2016) are represented by orange, green and black lines, respectively.

the wide range of flux densities separated in different flux bins (first bin starts from 1–5 mJy, then 5–10 mJy, 10–20 mJy, 20–30 mJy and so on, up to 90–100 mJy). We perform the source extraction pipeline extracting the list of recovered point sources using SUSSEXtractor and searching for matched detections.

The HeViCS field is significantly crowded and we can perform mismatch with an unrelated source close to the input (x, y) coordinate. We match the input to output catalogue considering source distances smaller than half of the $250\ \mu\text{m}$ beam size ($9''$), measured from the centre of the $250\ \mu\text{m}$ detection. Quality assessment results are summarised in Table 2 and Fig. 5. The plotted completeness curves are the best-fitting logistic functions, describing the completeness as a function of input flux. We can see that positions of recovered sources do not show any significant offset to assigned inputs. More than 70% of sources with $S_{250} > 13.2\ \text{mJy}$ have an offset lower than $6''$, which is smaller than a $250\ \mu\text{m}$ pixel size. Completeness and flux accuracy are consistent over all four HeViCS fields. The top panel in Fig. 5 shows our completeness result compared to one obtained

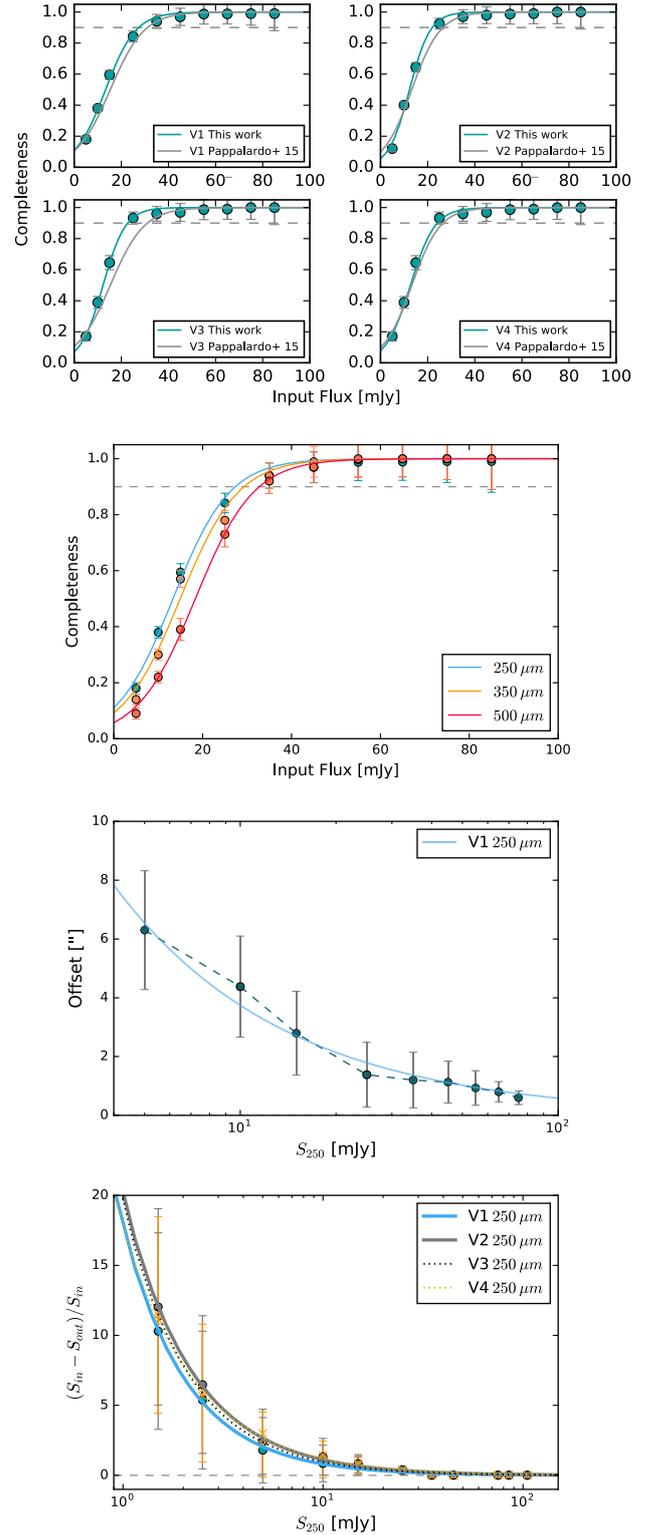


Fig. 5. For the HeViCS fields: Completeness at $250\ \mu\text{m}$ band compared to Pappalardo et al. (2015), SPIRE completeness for the V1 field, positional accuracy (average radial offset), and flux boosting (the flux difference as a function of input flux) are shown from top to bottom. Simulated (input) flux S_{in} is plotted on the x -axis for all subplots, while S_{out} refers to observed flux.

by Pappalardo et al. (2015), confirming that our results do not change notably when using different methods of flux estimation.

With measured numbers in hand, we can adjust the detection cut to see variations in completeness, which are connected to the

Table 2. Completeness fraction at SPIRE wavebands measured by injecting artificial sources into real SPIRE maps.

Flux [mJy]	V1 [250 μ m]	V2 [250 μ m]	V3 [250 μ m]	V4 [250 μ m]
5	0.18	0.14	0.17	0.19
10	0.37	0.40	0.39	0.38
15	0.59	0.64	0.63	0.61
25	0.84	0.91	0.93	0.93
35	0.94	0.97	0.96	0.97
45	0.97	0.98	0.98	0.98
55	0.98	0.99	0.99	0.99
65	0.98	1.0	1.0	0.99
75	0.99	1.0	1.0	1.0
85	1.0	1.0	1.0	1.0
95	1.0	1.0	1.0	1.0
	[350 μ m]	[350 μ m]	[350 μ m]	[350 μ m]
5	0.14	0.19	0.12	0.11
10	0.30	0.43	0.29	0.28
15	0.57	0.64	0.53	0.55
25	0.79	0.83	0.78	0.76
35	0.92	0.94	0.93	0.93
45	0.97	0.98	0.98	0.97
55	0.98	0.99	0.99	0.98
65	0.99	1.0	0.99	0.99
75	0.99	1.0	1.0	1.0
85	1.0	1.0	1.0	1.0
95	1.0	1.0	1.0	1.0
	[500 μ m]	[500 μ m]	[500 μ m]	[500 μ m]
5	0.10	0.09	0.11	0.11
10	0.21	0.20	0.18	0.18
15	0.39	0.44	0.43	0.42
25	0.69	0.68	0.68	0.73
35	0.91	0.88	0.89	0.94
45	0.98	0.99	0.98	0.99
55	0.99	0.99	0.99	0.99
65	0.99	1.0	1.0	0.99
75	0.99	1.0	1.0	1.0
85	1.0	1.0	1.0	1.0
95	1.0	1.0	1.0	1.0

Notes. Input fluxes are given in the first column, while percentages of detected sources (average values per bin) are given in the subsequent columns.

function of input flux density. The completeness of our 250 μ m maps shows a trend of fast decrease below ~ 25 mJy, and reaches $\sim 50\%$ at ~ 13.2 mJy, the value we choose as the lowest cut for our prior S_{250} list. The fitting logistic function reaches saturation at the level of 250 μ m fluxes larger than 40 mJy. For the 500 μ m band, our completeness is above 80% at $S_{500} > 30$ mJy. We apply this 500 μ m flux as the threshold for the final selection of 500 μ m-risers. The bottom panel in Fig. 5 shows that fluxes of sources below $S_{250} = 7\text{--}10$ mJy are systematically overestimated due to “flux-boosting”.

A reliability test is performed to quantify eventual spurious detections. This is an important value for measuring the total number counts rather than for selection of 500 μ m-risers, since we expect that eventually spurious sources would be present in one waveband but not in all three SPIRE bands. However, since our prior list is based on faint 250 μ m detections close to the confusion limit, we perform an analysis of possible false detections by injecting fake sources into noise maps. We also quantify

the range of statistical outliers present in our measurements. Outlier contamination is measured as a function of output flux density. Sources are considered as contaminants if their output fluxes are more than 3σ above the value inferred for injected sources. Considering the lowest detection threshold in the 250 μ m band (13.2 mJy) we found a low sample contamination of 2.7%.

3. Expected L_{IR} and z distribution of 500 μ m-risers

Without known interferometric positions and confirmed spectroscopic redshifts of our sources, we are limited to the properties of FIR SEDs of selected 500 μ m-risers. Nevertheless, our data may be very useful in providing approximate redshift/luminosity distributions of large samples and candidates for follow-ups. To fit an MBB model to our photometric data, we consider a degeneracy between $\beta - T_d$, and $T_d - z$ (Blain et al. 2003; Pope & Chary 2010). The peak of the SED is determined by the ν/T_d term (see Eq. (2)), given that a measurement of the colours alone constrains only $(1+z)/T_d$ ratio. However, assuming reasonable priors (in our case β and T_d) it is possible to estimate a qualitative redshift distribution of 500 μ m-risers (see $T_d - z$ degeneracy illustrated in Fig. B.1).

We fit a single-temperature MBB model to the data. We fix the power-law slope $\beta = 1.8$ and dust temperature $T_d = 38$ K, and we determine a median redshift of 4.22 ± 0.49 . The choice of β is arbitrary, likely $1.5 < \beta < 2.0$ (Casey et al. 2014; MacKenzie et al. 2014; Roseboom et al. 2012), and here we chose the value which is in the middle of that range. Due to the aforementioned degeneracy in $(1+z)/T_d$ space, decreasing the dust temperature, for example from 38 K to 30 K for the fixed β , the peak of the redshift distribution decreases to $z = 3.57$ (with 33% of sources at $z > 4$). Modelled redshift tracks for different MBB parameters are provided in Appendix B (see Fig. B.1).

To further investigate the possible redshift/luminosity range, we consider a collection of SEDs of extensively studied IR-bright galaxies both from the local and higher- z Universe, namely: the compact local starburst M82, the typically interacting, local dusty system Arp220, and the “Cosmic Eyelash”, a strongly cluster-lensed dusty galaxy at $z = 2.3$ (Swinbank et al. 2010).

The sample of dusty galaxies at $z > 4$ is expected to be comprised of a combination of intrinsically luminous starbursts and strongly lensed systems (B  thermin et al. 2012; Casey et al. 2014). It has been shown that a lack of observational constraints caused some SEDs to express overly low fluxes on the long λ -side of the FIR peak (Lutz 2014; Elbaz et al. 2010). Another important point considered here is that ultra-luminous infrared galaxies (ULIRGs) at higher redshifts express a wider variance in dust temperatures compared to their local analogues (e.g. Smith et al. 2014; Symeonidis et al. 2013). To overcome these differences we use an empirical template representative of H -ATLAS galaxies (Pearson et al. 2013). The template is built from a subset of 40 H -ATLAS survey sources with accurately measured redshifts covering the range from 0.5 to 4.5. It adopts two dust components with different temperatures and has already been used in the literature as a statistical framework for characterising redshifts of sources selected from SPIRE observations (see Ivison et al. 2016). With different templates in hand we can better characterise the systematics and uncertainties of measured redshifts. Shifting these templates to fit with our FIR fluxes, we forecast redshift ranges of 500 μ m-risers in the HeViCS field. We built a probability distribution function (PDF) based on fitted χ^2 values that allows us to derive an estimate of the median redshift.

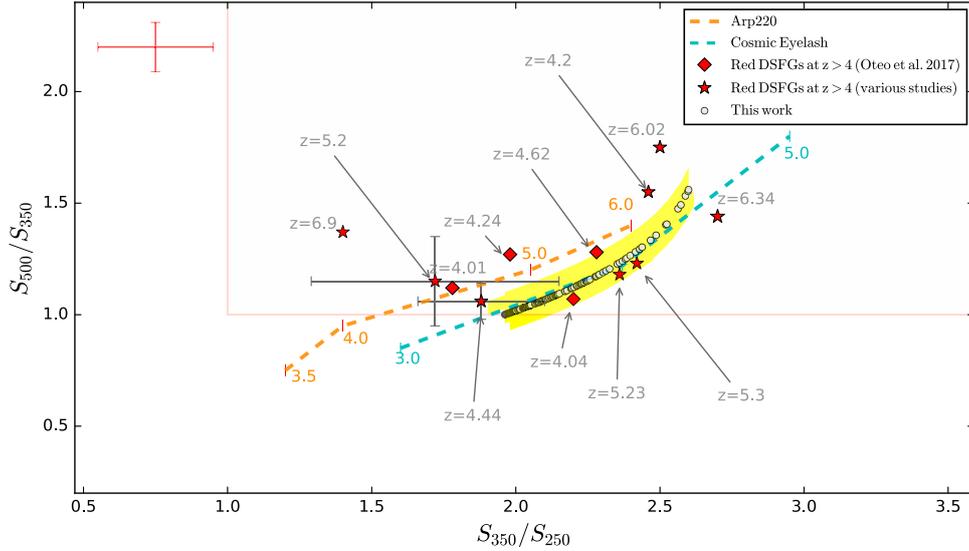


Fig. 6. SPIRE colour-colour diagram of $500\ \mu\text{m}$ -risers, overlaid with redshift tracks of Arp220 (Rangwala et al. 2011) and Cosmic Eyelash (Swinbank et al. 2010). Galaxies selected in this work are represented with circles, while yellow shaded regions describe uncertainties related to the chosen emissivity ($1.5 < \beta < 2.1$). For comparison we show $500\ \mu\text{m}$ -risers selected in different studies and with known spectroscopic redshifts. Sources marked with red stars are: $z = 6.34$ (HFLS3, Riechers et al. 2013), $z = 6.02$ (G09-8988, Zavala et al. 2018), $z = 4.44$ (FLS 5, Dowell et al. 2014), $z = 5.3$ (FLS 1, Dowell et al. 2014), $z = 5.2$ (HELMS-RED4, Asboth et al. 2016), $z = 4.04$ (GN20, Daddi et al. 2009), and $z = 4.2$ (SPT0113-46, Weiß et al. 2013). Sources marked with red diamonds are from Oteo et al. (2017b). For some of known sources we plot their colour uncertainties. Representative colour uncertainty for our sample is plotted in the upper-left corner.

With this we calculate a median redshift of $500\ \mu\text{m}$ -risers of 3.87 if we apply “Cosmic Eyelash” template, 4.14 for Arp220 and 4.68 for M82. Subsequently, we apply a *H*-ATLAS empirical template accepting the following best-fit parameters: $T_{\text{cold}} = 23.9\ \text{K}$, $T_{\text{hot}} = 46.9\ \text{K}$, and the ratio between the masses of cold and warm dust of 30.1, as in Pearson et al. (2013). The result is shown in Fig. 7. We determine a median redshift $\hat{z} = 4.28$, and an interquartile range of 4.08–4.75. This estimate matches the 1σ uncertainty range of photometric redshifts calculated by Dowell et al. (2014) and Asboth et al. (2016). To determine an average photometric redshift for red sources selected from several HerMES fields, Dowell et al. (2014) used the affine invariant method (Foreman-Mackey et al. 2013) to fit the MBB model with fixed emissivity and rest-frame wavelength peak for each source. They found $\langle z \rangle = 4.7 \pm 0.9$. The median photometric redshift estimated in Ivison et al. (2016) is somewhat lower ($\hat{z} = 3.7$), but we note that not all the sources in their catalogue of high- z candidate DSFGs are $500\ \mu\text{m}$ -risers. In Sect. 5.2 we analyse how the noise and resolution effects might impact the observed colours of DSFGs and thus the redshift distribution of selected $500\ \mu\text{m}$ -risers.

To compute the IR luminosity, which is defined as the integral over the rest frame spectrum between $8\ \mu\text{m}$ and $1000\ \mu\text{m}$, we rely on the same approach. Our results favour the presence of very luminous systems regardless of the chosen template – in 90% of cases selected $500\ \mu\text{m}$ -risers have $L_{\text{IR}} \geq 10^{13}\ L_{\odot}$, with minimum and maximum values of $L_{\text{IR}} = 8.1 \times 10^{12}\ L_{\odot}$ and $L_{\text{IR}} = 5.1 \times 10^{13}\ L_{\odot}$, respectively. Concerning the MBB model, we obtain $\hat{L}_{\text{IR}} = 2.14 \times 10^{13}\ L_{\odot}$, but it is worth noting that the model accounts for mid-IR excess using a very simplified method (see Eq. (2)). Furthermore, we apply the *H*-ATLAS representative template to determine corresponding median rest-frame IR luminosity $\hat{L}_{\text{IR}} = 1.94 \times 10^{13}\ L_{\odot}$. The result is again in consistency with Dowell et al. (2014) and Ivison et al. (2016). These studies performed SED modelling of the sources with known photometric or spectroscopic redshifts, concluding that $500\ \mu\text{m}$ -risers

have, on average, $L_{\text{IR}} \geq 10^{13}\ L_{\odot}$. For example, Ivison et al. (2016) report the median value $L_{\text{IR}} = 1.3^{+0.7}_{-0.5} \times 10^{13}\ L_{\odot}$. The corresponding luminosities in their sample range from $L_{\text{IR}} = 6.0 \times 10^{12}\ L_{\odot}$ to $L_{\text{IR}} = 5.8 \times 10^{13}\ L_{\odot}$.

In Fig. 7 we illustrate the L_{IR} -redshift trend of our $500\ \mu\text{m}$ -risers. We see that the expected redshift distribution is heavy-tailed towards the higher redshifts. This indicates that very red submillimetre galaxies are present at redshifts above the interquartile range, but could also indicate strong gravitational lensing or unresolved blending. We should consider using a much wider range of $L_{\text{IR}} - z$ values under larger uncertainties in dust temperatures. The IR luminosity emitted by an MBB source at a given T_{d} depends on the emissivity and size of the radiation area. Therefore, eventual gravitationally lensed candidates would lead to an apparent boost in dust luminosity at a given T_{d} due to a larger magnification (Greve et al. 2012). We consider and quantify the effects of strong lensing in our selection (see Sects. 5.1.5 and 6.2). We should emphasise here that the proper “training” of selected templates requires addition of higher-resolution data, and only when these are available is it possible to estimate systematic uncertainties and reject unreliable templates.

Comparison to red sources from the literature

In Fig. 6 we show how the sources from this work relate to some of the $500\ \mu\text{m}$ -risers with known redshifts selected from the literature. Redshift tracks of Arp220 and Cosmic Eyelash are added for comparison. We see that the colours of our sources agree very well with a large number of red DSFGs at $z > 4$ found in different studies. There are some exceptions, one of them being the most distant $500\ \mu\text{m}$ -riser, SPT0311-58 at $z = 6.9$ (Strandet et al. 2017). The possible reason for this might be an additional contribution from AGN, since it has been shown that SPT0311-58 experiences additional mechanical energy input, probably via AGN-driven outflows. There are also examples of sources with

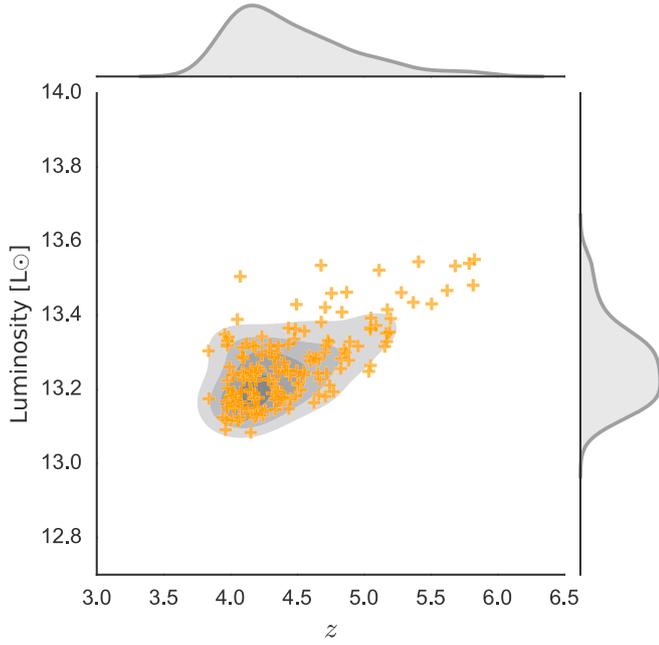


Fig. 7. Infrared luminosity of our $500\mu\text{m}$ -risers (y -axis) as a function of redshift. Orange crosses are values estimated applying the empirical template made from *H-ATLAS* galaxies covering the redshift range from $z = 0.5$ to $z = 4.5$ (Pearson et al. 2013). We imposed the same fitting parameters as used for *H-ATLAS* galaxies: $T_{\text{cold}} = 23.9$ K, $T_{\text{hot}} = 46.9$ K, and a cold-to-hot dust mass ratio equal to 30.1.

redshifts somewhat lower than what would be expected from their very red colours. These sources usually have much lower T_d compared to the median dust temperature of $500\mu\text{m}$ -risers at $z > 4$. Such an example is the $500\mu\text{m}$ -riser at $z = 4.2$ (Weiß et al. 2013), a strongly lensed red source with colours very similar to those of known $500\mu\text{m}$ -risers at $z > 6$. However this source has an estimated T_d of 30 K, compared to 50 K for all other known $500\mu\text{m}$ -risers at $z > 6$ (Zavala et al. 2018; Strandet et al. 2017; Riechers et al. 2013).

We draw attention to some interesting $500\mu\text{m}$ -risers selected in this work. According to the MBB model introduced in Sect. 2.6, it is plausible to expect a source at $z \sim 6$ if it has $T_d \geq 45$ K and satisfies colour criteria $S_{500}/S_{350} > 1.3$ and $S_{350}/S_{250} > 2.4$ (see Fig. B.1). In the literature two out of the three $500\mu\text{m}$ -risers at $z > 6$ fulfil these so-called “ultra-red” requirements (Zavala et al. 2018; Riechers et al. 2013). We have seven such sources in our sample, and we highlight them as suitable very high-redshift candidates. They are catalogued as HVS 5, HVS 6, HVS 21, HVS 35, HVS 47, HVS 75, HVS 85, and HVS 94. We also stress that five of our $500\mu\text{m}$ -risers (HVS27, HVS60, HVS 64, HVS115, and HVS 130) reside in potential overdense regions unveiled by *Planck* (Planck Collaboration 2016b). If these overdensities are not just part of random fluctuations in the galaxy number density within the *Planck* beam, this makes them candidate protoclusters of high- z DSFGs (Greenslade et al. 2018; Smolčić et al. 2017; Oteo et al. 2017a; Clements et al. 2016).

4. Differential number counts

We measure the raw $500\mu\text{m}$ differential number counts and test our selection in respect to previous studies (this section), but also models of galaxy evolution (Sect. 5). We determine our counts placing 133 $500\mu\text{m}$ -risers in seven logarithmic flux

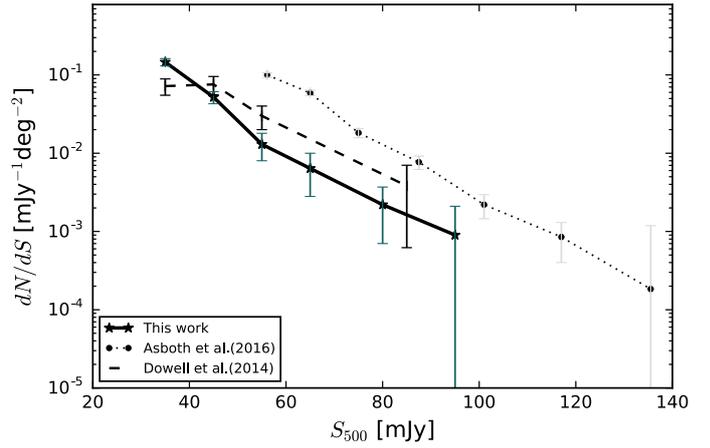


Fig. 8. Raw differential number counts of red sources from several studies. HeViCS number counts are presented with black stars, connected with a full, black line. Counts from Asboth et al. (2016) are indicated with black circles and connected with dotted lines, whilst observational findings of Dowell et al. (2014) are presented with dashed black lines. Filled points have Poisson 1σ error-bars added, and for the brightest flux bins with low statistics of sources, we plot the upper 95% confidence levels.

bins between 30 mJy and 100 mJy. The resulting plot is shown in Fig. 8. The total raw number density of $500\mu\text{m}$ -risers in 55 deg^2 of HeViCS is $N = 2.41$ sources per deg^2 , with the corresponding 1σ uncertainty of 0.34. We note that this value is uncorrected for non-red contaminants and completeness. In Sect. 5 we fully inspect and quantify biases that affect observed distribution, correcting for these effects. Furthermore, the presented differential number counts ignore eventual source multiplicity. Without interferometric data at longer wavelengths, our counts should be interpreted as the density of sources in respect to the beam resolution. Potential impact of SPIRE source multiplicity on our “FIR-riser” selection is discussed in Sect. 6.

In Table 3 we present raw differential and integral number counts. In Fig. 8 we plot our differential number counts along with values from studies of Asboth et al. (2016) and Dowell et al. (2014). They applied the so-called difference map method (*DMAP*) to select red sources. The method homogenizes beams to the size of $500\mu\text{m}$, and creates the *DMAP* where SPIRE images are all smoothed to an identical angular resolution. Prior to source identification, a weighted combination of the SPIRE maps is formed applying the relation $D = 0.920 \times M_{500} - 0.392 \times M_{250}$, where M_{500} and M_{250} are the smoothing values after matching SPIRE maps to the $500\mu\text{m}$ resolution. To select $500\mu\text{m}$ -risers, they performed a red colour map-based search, starting from lowest-resolution $500\mu\text{m}$ maps.

From Fig. 8 it can be seen that differential number counts notably differ, especially in the lowest two flux bins where our measurements are consistent with Dowell et al. (2014). The slope obtained by Dowell et al. (2014) in the first two flux bins is almost flat, which is different from the HeViCS curve where a steeper decrease is present. While our number counts show fairly good agreement with those measured by Dowell et al. (2014) especially in lower flux bins, they are well below the values measured by Asboth et al. (2016). Depending on the chosen flux bin, the discrepancy factor ranges from 3 to 10. We note that discrepancy between our number counts and those from Asboth et al. (2016) are due to several effects. Asboth et al. (2016) included possible blends in their final catalogue under the assumption that at

Table 3. Differential and integral number counts.

Flux (500 μm)		Diff. counts	Int. counts ¹
S_{\min}	S_{\max}	dN/dS	$N(>S_{\min})$
[mJy]	[mJy]	$[\times 10^{-4} \text{ mJy}^{-1} \text{ deg}^{-2}]$	$[\text{deg}^{-2}]$
30	40	1527 ± 164	2.41 ± 0.22
40	50	581 ± 90	0.91 ± 0.14
50	60	145 ± 51	0.29 ± 0.09
60	70	72 ± 36	0.14 ± 0.05
70	90	27 ± 15	0.07 ± 0.03
90	110	9 ± 5	0.02 ± 0.01

Notes. ⁽¹⁾In order to calculate the cumulative (integral) number counts we have summed over all the 500 μm -risers above the specified flux density (S_{\min}). Here we present raw differential and integrated number counts.

least one component of a blend is a red DSFG. This introduces a non-negligible number of contaminants, which is the effect we analyse in Sect. 5.2. The effect of blending on our number counts is reduced since we apply a new technique of source extraction based on positional and SED priors. Additionally, Béthermin et al. (2017) simulated the same process of source extraction and selection as in Asboth et al. (2016) and found the number of observed 500 μm -risers to be higher by a factor of eight compared to the number of genuine (modelled) 500 μm -risers. They concluded that a combination of noise, resolution effects, and the steepness of flux density distributions produces numerous red artefacts that match 500 μm -risers criteria. In Sects. 5.1.5 and 5.2 we present a detailed analysis of the effect of noise on our selection.

Apart from the samples mentioned above, it is possible to find other 500 μm -risers in the literature, but they are not selected in a uniform manner (e.g. Casey et al. 2012; Miettinen et al. 2015; Negrello et al. 2017). 500 μm -risers are serendipitously detected in relatively wide and shallow surveys that use wavelengths longer than 500 μm for initial detection (e.g. South Pole Telescope, see Vieira et al. 2013). These samples contain some of brightest 500 μm -risers, all of them significantly amplified by gravitational lensing ($S_{500} > 100$ mJy).

5. Comparison to models of galaxy evolution

In this section we compare our results with expectations from the most recent models of galaxy evolution. To evaluate our selection method, we perform simulations by generating realistic SPIRE maps from catalogues of simulated galaxies.

5.1. Models

In the literature there is a number of methods aiming to predict the evolution of number counts at different IR wavelengths. In this work we consider galaxy evolution models based on multi-band surveys. In general such models rely on the combination of observed SED templates of galaxies and luminosity functions. It has been shown (Dowell et al. 2014) that some galaxy evolution models, mostly from “pre-*Herschel* era”, underpredict either the total number of high- z , red sources (Le Borgne et al. 2009), or the number of red sources which lie at $z > 4$ (Béthermin et al. 2011). There are some exceptions (e.g. Franceschini et al. 2010), that anticipate larger number of red, high- z galaxies, but predict at the same time large amount of very low- z ($z < 2$) red objects which seems unphysical. Additionally, we expect that

“almost red” sources (described as sources having a peak in their FIR SEDs at wavelengths between 350 and 500 μm) contaminate the sample of “500 μm -risers”, making observational artefacts caused by the noise. It is then very important to have models that predict significant number of such sources, to check the systematics in the detection rate of these contaminants. In this work we consider phenomenological models of Béthermin et al. (2012), Schreiber et al. (2016) and Béthermin et al. (2017), hereafter denoted as B12, S16, and B17 respectively. These models were built on *Herschel* data and accurately match the total IR number counts. A summary of the models and their main ingredients is given in Table 4. Common to all models is their usage of stellar mass function (SMF) as a starting point from which properties of galaxies are generated. They share the same general description of star-forming galaxies, with star formation rate (SFR) assigned based on the dichotomy model of Sargent et al. (2012). It decomposes the bolometric FIR-luminosity function with main sequence (MS) and starburst (SB) galaxies. The MS galaxies are described as secularly evolving galaxies tightly relating stellar mass and SFR, while SB galaxies show an offset from the MS, expressing very high specific SFRs ($s\text{SFR} = \text{SFR}/M_{\star}$). The shape of the SEDs is controlled by the galaxy type (MS or SB) and the mean intensity of the radiation field $\langle U \rangle$, which couples with the temperature of dust. Differences between the models are described briefly in the following subsections - the most important are scatter from the MS, parameters chosen to fit stellar masses, and redshift evolution of a dust temperature. Models presented here also have a slightly different description of mergers.

5.1.1. Béthermin et al. (2012) model

The B12 model describes the SMF with a Schechter function with a redshift invariant characteristic mass parameter (see Eq. (4) in Béthermin et al. 2012, but also Peng et al. 2010). Here, the infrared SED template is based on Draine & Li (2007) models, with the mean radiation field $\langle U \rangle$ evolving with redshift up to $z = 2.0$ (Magdis et al. 2012). The dispersion of the MS log-normal distribution is $\sigma_{\text{MS}} = 0.15$ dex, following Sargent et al. (2012) and Salmi et al. (2012). The model takes into account strong lensing effects reckoning the magnification rate PDF from Hezaveh & Holder (2011). These lensed sources contribute $\sim 20\%$ to the bright-end submillimetre counts ($\text{PLW} \geq 100$ mJy). The AGN contribution is statistically associated based on results from Aird et al. (2012) and Mullaney et al. (2011).

5.1.2. Béthermin et al. (2017) model

The B17 model is based on similar prescriptions as B12, but implies several improvements in order to match recent interferometric results that disclosed notable overestimation of submillimetre number counts due to source blending (Karim et al. 2013; Simpson et al. 2015). The model includes detailed consideration of clustering and resolution effects. To describe the physical clustering, an abundance matching procedure is used to populate the dark-matter halos of a light cone constructed from the Bolshoi-Planck simulation (Rodríguez-Puebla et al. 2016). The MS scatter is updated ($\sigma_{\text{MS}} = 0.3$ dex) in order to match the measurements of Ilbert et al. (2015) and Schreiber et al. (2015). The model uses a new parametric form to fit the redshift evolution of the radiation field $\langle U \rangle$. The evolution of the dust temperature of MS star-forming galaxies stops at $z = 4$ (instead at $z = 2.0$ as in B12) and remains constant at higher values. Contribution of AGNs and strong lensing effects are modelled as in B12. The

Table 4. Comparison of models used in our analysis.

	Models		
	B12 Béthermin et al. (2012)	B17 Béthermin et al. (2017)	S16 Schreiber et al. (2016)
Formalism ¹	2SFM	2SFM	2SFM
sSFR ²	Evolves up to $z = 2.5$	Evolves up to $z = 4$	Evolves continuously
Dispersion (σ_{MS}) ³	0.15 dex	0.3 dex	0.3 dex
Strong lensing	Yes	Yes	No
Passive galaxies	Yes	Yes	Yes
Evolution of T_{d}	up to $z = 2$	up to $z = 4$	Continuous
AGN contribution	Yes	Yes	No

Notes. ⁽¹⁾All the models are based on two SF mode formalism. Stellar mass function (SMF) is described by a double Schechter function: $\phi(M_*)d(M_*) = \exp(-\frac{M_*}{M^*})[\Phi_1^*(\frac{M_*}{M^*}) + \Phi_2^*(\frac{M_*}{M^*})]$, where M^* is the characteristic mass of the Schechter break. The M^* is the redshift invariant in the Béthermin et al. (2012) model, and evolves with redshift in the other two presented models. For redshifts $z > 4$, the model of Béthermin et al. (2017) assumes a single Schechter function fixing Φ_1^* at zero, while the model of Schreiber et al. (2016) adopts double Schechter fitting to results of Grazian et al. (2015) for $4.5 < z < 7$. ⁽²⁾The specific star-formation rate, defined as $\text{sSFR} = \text{SFR}/M_*$. In Béthermin et al. (2012) the sSFR increases with redshift up to $z = 2.5$ and then flattens. This trend is independent of the chosen range of stellar masses. In the latter two models the evolution is different, see Eq. (6) in Béthermin et al. (2017). ⁽³⁾Modelled width of the main sequence (as log-normal scatter).

weak lensing regime is modelled with magnifications that are randomly drawn from a Gaussian distribution. Their width and mean values are derived based on a cosmological simulation of Hilbert et al. (2007).

5.1.3. Schreiber et al. (2016) model

Similarly to B12 and B17, the S16 model is based on the MS-SB dichotomy. MS galaxies are modelled based on a stellar mass and redshift from Schreiber et al. (2015) and Pannella et al. (2015). The scatter of the main sequence $\sigma_{\text{MS}} = 0.3$ dex. Randomly-selected galaxies (5%) are placed in the SB mode by enhancing their SFR by a factor of approximately five. The distribution of stellar masses is described by a double power-law Schechter fit, with parameters evolving with redshift. These parameters are chosen according to observational data from (Schreiber et al. 2015; see also Grazian et al. 2015). A new set of template SEDs is used to model the dust emission of star-forming galaxies. These SEDs are based on the physically motivated dust model of Galliano et al. (2011). Redshift evolution of a dust temperature is modelled as

$$T_{\text{d}}[\text{K}] = \begin{cases} 4.65 \times (z - 2) + 31, & \text{for MS} \\ T_{\text{d}}^{\text{MS}} + 6.6 \times \log_{10}(R_{\text{SB}}), & \text{for SB} \end{cases} \quad (4)$$

The “starburstiness” term R_{SB} is used to quantify the SFR offset between MS and SB galaxies ($R_{\text{SB}} = \text{SFR}/\text{SFR}_{\text{MS}}$). Dependence of sSFR on the stellar mass shows that sSFR is constant at lower masses, and drops at the highest masses, $M > 10^{10.5} M_*$ (e.g. Schreiber et al. 2015; Whitaker et al. 2015; Magnelli et al. 2014). This trend is similar to the one used in B17, but different from the fit used in B12. To summarise, in S16 both sSFR and T_{d} evolve continuously with redshift, which is not the case for the other two models (see Table 4). Effects of strong lensing and AGNs are not included in the model.

5.1.4. Mock catalogues

We set mock catalogues based on models described in the previous section. Our goal is to compare the number counts of

observed 500 μm -risers to the ones predicted by models³. The catalogues based on B12, B17, and S16 cover sufficiently large areas to offer accurate inspection of our observational criteria in the HeViCS field. Namely, for the B12 model we used the catalogue which is a result of 500 deg^2 simulations (Béthermin et al. 2012). We also used the catalogue created by the B17 model covering the area of 274 deg^2 (Simulated Infrared Dusty Extragalactic Sky, SIDES, Béthermin et al. 2017). The size of the simulated area is equal to the size of the HeLMS field, thus perfectly suited for a comparison of our selection technique to the one used by Asboth et al. (2016). With the S16 model, we generated a mock catalogue covering the size of the HeViCS field (55 deg^2).

We further apply our 500 μm -risers selection criteria ($S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2$ mJy and $S_{500} > 30$ mJy) to modelled sources. In Fig. 9 we see that observed counts are in a fairly good agreement with models, while corresponding power-law slopes show no significant difference from the slopes anticipated by models. Observed values are steep at the fainter end and flatter at the brighter end ($S_{500} > 80\text{--}90$ mJy) which is due to flux magnification by gravitational lensing. Our HeViCS differential number counts show a perfect agreement with B12, but we must note some recent evidence (see Béthermin et al. 2017 for more details) suggesting that the simulated catalogue based on the B12 model overpredicts the number counts at 500 μm fluxes below 50 mJy by a factor of two to three. The number of sources satisfying our selection in B17 and S16 is 0.45 deg^{-2} and 0.40 deg^{-2} , respectively. Predictions of B17 and S16 are consistent with the 1σ error bars of observed counts in the higher flux bins ($60 \text{ mJy} < S_{500} < 100 \text{ mJy}$). In lower flux bins ($30 \text{ mJy} < S_{500} < 60 \text{ mJy}$) discrepancy factors are between 1 and 6 depending on the chosen bin. In Sect. 5.1.5 we discuss in more detail potential reasons for these differences. However, it is clear that we have a much better agreement with empirical models than previous studies (Asboth et al. 2016) who claim observational discrepancy of an order of magnitude.

³ In the following, we make the distinction between intrinsic and observed quantities. The former ones are the true, modelled properties of galaxies, free of measurement errors and systematics. The latter are measured values, affected by biases.

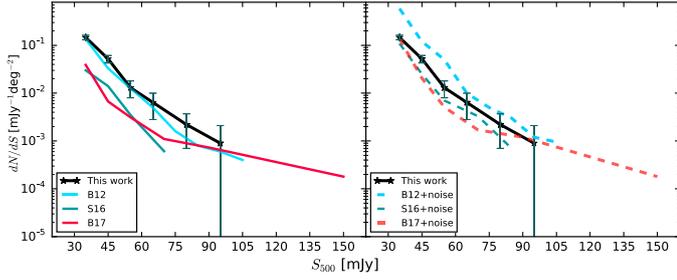


Fig. 9. Observed differential number counts of $500\ \mu\text{m}$ -risers in the HeViCS field (black line) compared to models. Expected values from models are overplotted as coloured lines. *Left*: comparison of observed and modelled counts. Models are represented by full, coloured lines: cyan for B12 (Béthermin et al. 2012), red for B17 (Béthermin et al. 2017), and green for S16 (Schreiber et al. 2016). The effect of simulated noise is ignored. *Right*: comparison between observed and modelled counts if the effect of noise is simulated. The effect of confusion and instrumental noise is simulated by adding a random Gaussian noise to the modelled fluxes. Differential counts are then represented by dashed, coloured lines.

Table 5. Comparison of modelled number counts before and after adding the Gaussian noise.

	Predicted density of $500\ \mu\text{m}$ -risers ¹	Faint $500\ \mu\text{m}$ -risers ² ($S_{500} < 30\ \text{mJy}$)
B17	$0.45\ \text{deg}^{-2}$ (51% lensed)	–
B17 + noise	$1.73\ \text{deg}^{-2}$ (24% lensed)	52%
S16	$0.4\ \text{deg}^{-2}$	–
S16 + noise	$1.54\ \text{deg}^{-2}$	67%

Notes. Counts are calculated by imposing our criteria to select $500\ \mu\text{m}$ -risers. ⁽¹⁾Modelled density of $500\ \mu\text{m}$ -risers before and after the addition of random Gaussian noise. Criteria used to select $500\ \mu\text{m}$ -risers are: $S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2\ \text{mJy}$ and $S_{500} > 30\ \text{mJy}$. Numbers reported in brackets refer to the percentage of strongly lensed galaxies; ⁽²⁾Contribution of intrinsically red but faint sources ($S_{500} < 30\ \text{mJy}$) to the population of modelled sources that match our $500\ \mu\text{m}$ -risers selection criteria after the addition of modelled Gaussian noise.

5.1.5. Effects of noise on measured counts

The results described in Sect. 5.1.4 neglect the effect of noise on the number counts of $500\ \mu\text{m}$ -risers. Therefore, we simulate the effect of both confusion and instrumental noise by adding a random Gaussian noise drawn from the values measured in the HeViCS field (Sect. 2.1). The comparison between observations and models with simulated Gaussian noise is illustrated in the right panel of Fig. 9. The significant increase of counts in the lowest two flux bins is clearly seen, while their slopes do not express significant change compared to intrinsic ones (left panel in Fig. 9).

Considering the addition of noise, the number of simulated sources that appear as $500\ \mu\text{m}$ -risers in B17 jumps to 473, resulting in an increase of number counts from 0.45 per deg^2 to 1.73 per deg^2 . We obtain a very similar result with S16 – a number density increases from 0.4 per deg^2 to 1.54 per deg^2 . As a consequence, observed HeViCS values match predictions of the B17 and S16 models even in the faintest $500\ \mu\text{m}$ flux regime⁴.

⁴ In B17, SFR is limited to a maximum $1000\ M_{\odot}\ \text{yr}^{-1}$.

We find that noise often tends to increase the number of genuinely red sources that are slightly below our $500\ \mu\text{m}$ flux cut. These galaxies have modelled flux densities $S_{500} < 30\ \text{mJy}$, but they can pass our final $500\ \mu\text{m}$ -risers selection after the addition of a Gaussian noise. As shown in Table 5 contribution of these sources to the modelled $500\ \mu\text{m}$ -risers varies between 52 and 67%. In Sect. 4, we mentioned that for the wide and shallow HeLMS field, Béthermin et al. (2017) found an increase of observed number counts by a factor of eight compared to their modelled values. This strong increase is caused by the combination of noise and clustering, and it is two times higher than the noise-driven increase of density of detected $500\ \mu\text{m}$ -risers in the HeViCS field (see Table 6). We thus conclude that the effect of noise is more prominent for shallower fields, while for deeper fields (e.g. HeViCS and deep HerMES fields analysed in Dowell et al. 2014) the noise has only a mild impact on measured number counts.

The effect of noise also changes the relative contribution of modelled lensed and unlensed $500\ \mu\text{m}$ -risers. The percentage of weakly lensed and non-lensed $500\ \mu\text{m}$ -risers in B17 increases from 49% (modelled values) to 76% (modelled values+Gaussian noise). Closer inspection reveals that weak lensing ($1 < \mu < 2$) is a contributor to the flux budget for a non-negligible number of sources (166 out of 473, or 35%). Fainter, weakly lensed red sources can pass our final $500\ \mu\text{m}$ -risers criteria more often if they are on a positive fluctuation of the magnification. This has an important role in producing observed $500\ \mu\text{m}$ -risers without the support of strong lensing.

5.2. Simulated maps

Mock mapping simulations are needed to explore the exact nature of selected sources and possible biases of our *Herschel* SPIRE selection. These are mostly induced by the limited angular resolution combined with the noise effects. Having used two different models, we can uncover systematic uncertainty surrounding the purity of detected sources. This is a crucial step if we want to investigate our photometric pipeline of $500\ \mu\text{m}$ -risers, since it is necessary to count the numbers of sources we potentially miss and to determine fraction of contaminants.

We generated a set of simulated SPIRE maps with the use of Empirical Galaxy Generator Code (EGG, Schreiber et al. 2017a). We filled our mock maps with sources drawn from S16 and B17 mock catalogues. We converted mock catalogues into maps, assigning theoretical *Herschel* PSF and the same noise properties as measured in our real HeViCS images. We chose to simulate clustering. The mock galaxies from the catalogues are then placed on the sky at coordinates with a fixed angular two-point correlation function to implement this effect. Positions of galaxies in S16 are clustered by default using of Soneira & Peebles (1978) algorithm, which produces a two-point correlation function with a power-law shape. To assign clustered coordinates for mock maps created from B17, we follow prescriptions from a simulation described in Béthermin et al. (2017), which adopts a sophisticated clustering model with the sub-halo abundance matching procedure used to populate the dark-matter halos of a light cone constructed from the Bolshoi-Planck simulation (Rodríguez-Puebla et al. 2016). With this we ensure that simulated maps contain a realistic background of clustered sources that can contribute to the confusion noise.

Due to computational dependencies, we generated maps covering the area of one of the four HeViCS fields ($16\ \text{deg}^2$ each). As

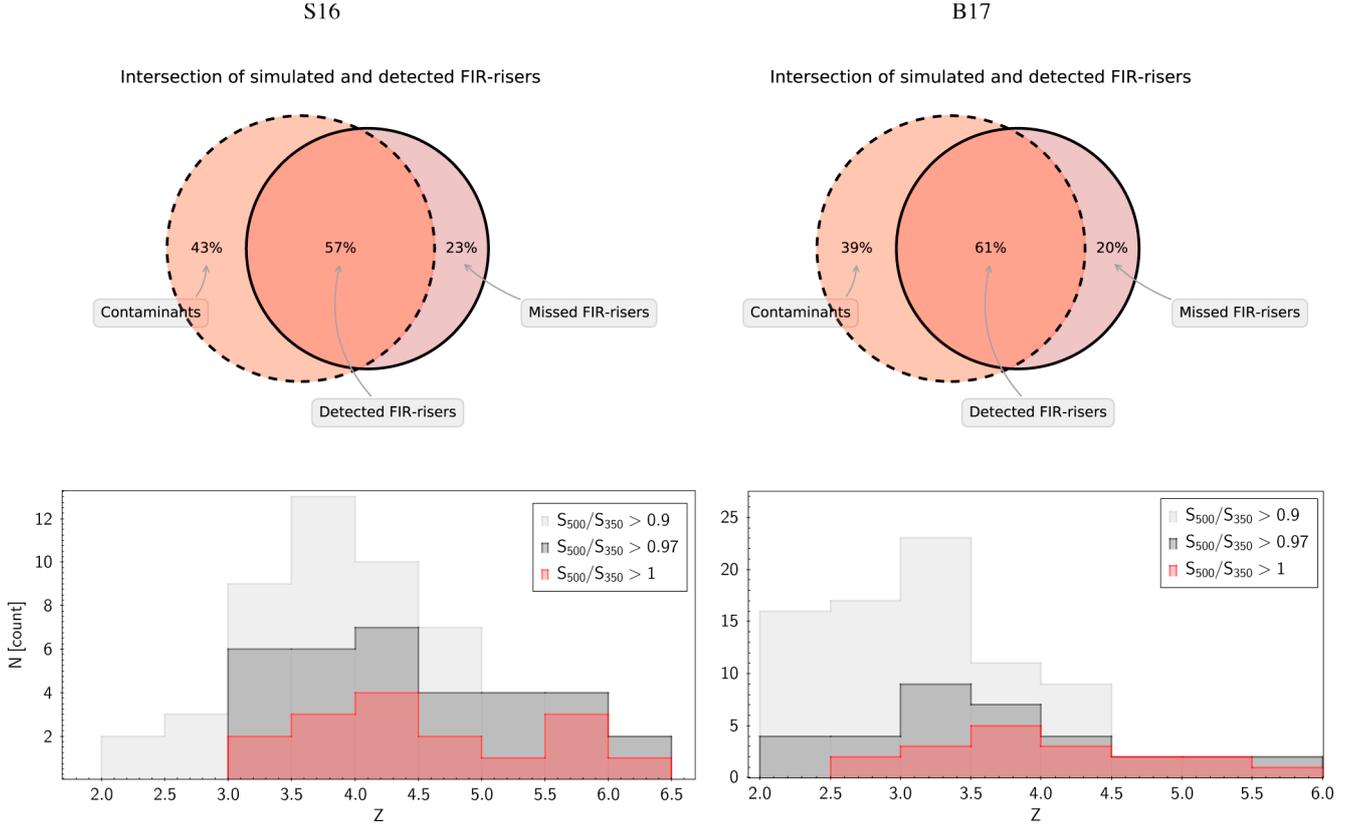


Fig. 10. *Upper panels:* galaxies detected in mock maps. “FIR-riser” criteria are imposed as for real HeViCS maps ($S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2$ mJy, $S_{500} > 30$ mJy). Intersected area coloured in dark orange depicts recovered red sources. Light orange area represents detected contaminants. Violet area represents missed sources. Those genuinely red galaxies are present in our catalogue, but not as $500\ \mu\text{m}$ -risers. *Lower panels:* redshift distribution of modelled galaxies. Different S_{500}/S_{350} colour cuts are imposed. These cuts are related to statistical properties of the confusion noise, which is the greatest contributor to colour uncertainties (see also Table 6). Left and right panels refer to S16 and B17, respectively.

shown in Sect. 2.4, our source extraction lead to a homogeneous distribution of sources over all HeViCS fields, so we can easily extrapolate our results for the whole observed area of $55\ \text{deg}^2$. We use the methods described above and perform the same source-detection pipeline as we did in our raw HeViCS maps. We also check positional and flux accuracies in the same way we did for real HeViCS images.

Results from our simulations are summarised and presented in Fig. 10. We detected all modelled $500\ \mu\text{m}$ -risers via our detection procedure, but not all of them are characterised using their intrinsic colours. We thus disclose three groups of sources described below.

1. *Recovered $500\ \mu\text{m}$ -risers.* The dark orange region in Fig. 9 indicates the fraction of recovered $500\ \mu\text{m}$ -risers. These are intrinsically red sources selected as $500\ \mu\text{m}$ -risers when we applied the same selection criteria as for the real HeViCS maps ($S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2$ mJy and $S_{500} > 30$ mJy). The percentage of recovered $500\ \mu\text{m}$ -risers is around 60% for both models (57% for S16 and 61% for B17).

As expected from the mock catalogue analysis (Sect. 5.1.4), the population of recovered risers is a mix of both lensed and unlensed galaxies. We confirm that our criteria to select $500\ \mu\text{m}$ -risers unveil a significant group of dusty and potentially very high- z sources that are not biased towards higher magnifications.

As introduced in Sect. 5.1.5 noise effects might produce colours somewhat redder than original ones. We quantify this as a “reddening”, namely:

$$\Delta = \left(\frac{S_{500} + \sigma_{\text{PLW}}}{S_{350} + \sigma_{\text{PMW}}} \right) - \frac{S_{500}}{S_{350}}, \quad (5)$$

where σ_{PLWm} and σ_{PMW} are 1σ total noise ratios measured for appropriate bands in the HeViCS field (see Sect. 2.1). The differences between intrinsic and observed colours of recovered $500\ \mu\text{m}$ -risers are presented in Table 6. We see that the largest difference is for the lowest flux bin, where we also expect the largest flux uncertainties due to the lower S/N. The redshift distribution of recovered $500\ \mu\text{m}$ -risers in Fig. 10 reflects differences between models. Following S16, all red galaxies are at $z > 3$, and most of them (70%) are at $z > 4$. The mean redshift value is $\langle z \rangle = 4.56 \pm 0.94$. The B17 model has a comparable high redshift tail, but this model peaks at lower redshift, causing the mean value ($\langle z \rangle = 3.89 \pm 0.9$) to be lower than in S16.

2. *Contaminants.* The light-orange region in Fig. 10 indicates a fraction of contaminants. Sources are detected as $500\ \mu\text{m}$ -risers, but having assigned their modelled colours and fluxes, they should not pass our selection criteria. The total percentage of contaminants is 39% for B17 and 43% for S16. They consist of two different populations of sources: The first ones are genuinely red, but are fainter galaxies, with modelled $500\ \mu\text{m}$ fluxes not bright enough for the final catalogue inclusion ($S_{500} < 30$ mJy). Nonetheless, their observed fluxes and colours are sufficiently high to pass “ $500\ \mu\text{m}$ -risers” criteria. The relative contribution of these contaminants is 60% for B17 and 72% for S16. The second population of contaminants are non-red sources with narrowly ranged intrinsic colours in the range $0.95 < S_{500}/S_{350} <$

Table 6. Intrinsic vs. observed colours of simulated 500 μm -risers.

Flux bin	S_{500}/S_{350} [B17 _{cat}]	S_{500}/S_{350} [B17 _{map}]	Δ
30–40 mJy	1.03 ± 0.09	1.07 ± 0.1	0.04 (0.012)
40–50 mJy	1.10 ± 0.05	1.12 ± 0.06	0.02 (0.009)
>50 mJy	1.15 ± 0.11	1.16 ± 0.18	0.01 (0.008)

Notes. Catalogue and maps are based on B17. Column 2: catalogued colours from B17. Noise is not added to initial flux values; Column 3: measured colours in simulated maps; Column 4: colour difference between observed and simulated flux values. It is quantified as “reddening” ($\Delta = \left(\frac{S_{500} + \sigma_{PLW}}{S_{350} + \sigma_{PMW}}\right) - \frac{S_{500}}{S_{350}}$). Values in brackets show the difference between catalogued fluxes before and after the addition of modelled Gaussian noise.

1.0. They are weakly lensed or unlensed “almost red” sources. We also find that non-red contaminants are accompanied with sources clustered at $1 < z < 2$. The largest 500 μm excess is caused if clustered neighbours are at $z = 1.3 \pm 0.1$. These sources are massive galaxies with $M_* > 10^{9.8} M_\odot$. This confirms the conclusion from Sect. 5.1.4 that combined with weak lensing, confusion which arises from clustering and noise has a tendency to increase the number of observed red sources. Having exploited both models, we see that all detected contaminants are high- z dusty galaxies, with just one source at $z < 3$. Redshift distribution of contaminants peaks at 3.48 for B17 ($z_{\min} = 2.2$ and $z_{\max} = 5.11$), and 4.17 for S16 ($z_{\min} = 3.28$ and $z_{\max} = 5.86$). Additionally, according to both models, contaminants are galaxies with warm dust temperatures ($40 \text{ K} < T_d < 50 \text{ K}$).

3. *Missed 500 μm -risers.* With our extraction procedure we detected all modelled 500 μm -risers from simulated maps. However there is a fraction of sources for which we failed to properly characterize their colours with MBB-fitter. The violet region in Fig. 10 indicates sources with $S_{500}/S_{350} < 1$ or $S_{500} < 30 \text{ mJy}$. In this way we failed to select around 20% of the genuine 500 μm -risers. According to models, missed 500 μm -risers have colours and fluxes very close to the lowest selection threshold. In our simulated maps, they often immerse in complex blends, having a nearby, unresolved blue source at $0.5 < z < 1.0$ inside the 250 μm beam. Therefore, missed sources pass both 250 μm and 500 μm flux selection cuts but their observed colours are non-red. We find that the difference between observed and modelled colours of missed 500 μm -risers is shifted bluewards by a factor of approximately 0.03, thus giving negative Δ . Interestingly, not just the observed, but also the modelled Δ of missed sources exhibits negligible change due to noise effects; they tend to make their colours marginally redder or even bluer compared to the average value we measure for 500 μm -risers ($\langle \Delta \rangle = 0.01$ for the whole population, while for missed sources $\langle \Delta \rangle < 0.003$ having consulted both models). The final colour-criterion is thus highly sensitive to multiple effects including the strong clustering of blue sources. The redshift distribution of missed 500 μm -risers shows that 50% of them are weakly lensed galaxies at $z > 4$. It becomes vital to account for the missed population of red sources while searching for $z > 4$ dusty galaxies.

Modifying criteria to select galaxies at $z > 4$?

The percentage of simulated red sources recovered by our detection pipeline is not as high as the values presented in Asboth et al. (2016) and Dowell et al. (2014). These studies reported

Table 7. Relation of different colour cuts to the number of observed/missed red sources and their redshift distributions.

Models	Colour-cuts					
	$S_{500/350} > 0.9$		$S_{500/350} > 0.97$		$S_{500/350} > 1.0$	
	m	$z > 4$	m	$z > 4$	m	$z > 4$
S16	3%	43%	18%	62%	23%	72%
B17	0%	23%	15%	33%	20%	47%

Notes. Sources observed in simulated SPIRE maps are divided into three columns according to different colour limits ($S_{500/350} > 0.9$, $S_{500/350} > 0.97$, $S_{500/350} > 1.0$). For each of these columns m gives a percentage of missed genuine 500 μm -risers, while column $z > 4$ gives the percentage of $z > 4$ sources contained in selected populations.

purity of almost 90%. Yet, there is an important and significant difference between our simulations. While we created mock images assuming the colours of sources that have been drawn directly from the simulated catalogues, Asboth et al. (2016) and Dowell et al. (2014) omitted modelled red sources from their maps, injecting 500 μm -risers with SPIRE colours fixed to the median value measured in their raw maps. Such an approach may lead to a bias, since we found (see Sect. 5.1) that most unlensed 500 μm -risers have S_{500}/S_{350} very close to one. Additionally, S16 and B17 predict a large number of almost red FIR sources (e.g. $0.9 < S_{500}/S_{350} < 1.0$ and $S_{250} > 13.2 \text{ mJy}$). The modelled number density of such defined sources is between 8 and 10 per deg^2 . They are included in our initial “prior” catalogue, and due to effects of noise, clustering and/or lensing, some of them might be observed “redwards”, thus being responsible for lower purity index.

Considering the colour differences (parameter Δ in Table 6), it appears clear that distinguishing between the population with $S_{500}/S_{350} = 0.97$ and $S_{500}/S_{350} = 1$ is very troublesome in crowded environments, even for fields with significantly reduced instrumental noise. As expected, colour uncertainties increase towards the lower fluxes. We conclude that some flexibility on a colour threshold is needed to account for noise and environmental effects. Here we test several colour cuts. The goal of such a test is twofold: to find the colour threshold that increases the number of true (modelled) 500 μm -risers, and to introduce at the same time low contamination of $2 < z < 4$ sources. We quantify the contribution of these lower- z objects for each colour limit. The result of the analysis are summarised in Table 7 and histograms in the lower panel of Fig. 10. Having imposed $S_{500}/S_{350} > 0.9$ we anticipate that almost all 500 μm -risers will be found independently of the chosen model. Nevertheless, contribution of $2 < z < 4$ sources is significant, and varies from 57% (S16) to 77% (B17). Redshift distribution peaks at $z = 3.25$ for B17 and $z = 3.92$ for S16. If we increase the colour cut requirement to $S_{500}/S_{350} > 0.97$, the percentage of detected simulated intrinsic 500 μm -risers ranges between 82 and 85%. Contamination is heavily reduced, and just a few galaxies are at $z < 3$ (e.g. in S16 all the galaxies that satisfy this colour cut are at $z > 3$). Furthermore, redshift peak shifts to higher values, $z = 3.68$ for B17 and $z = 4.29$ for S16.

Following these results, we suggest that a cut at $S_{500}/S_{350} > 0.97$ might be imposed to search for $z \gtrsim 4$ galaxies in the future. It is a good compromise that accounts for several effects which decrease the percentage of recovered red galaxies. At the same time, it works well against the lower- z contaminants, especially those at $z < 3$, limiting their contribution to a maximum of

15%. Recent follow-up observations of red sources selected in the *H*-ATLAS field (Iverson et al. 2016) also support our reasoning for modifying 500 μm -risers colour criteria. They report that just 30% of sources selected as 500 μm -risers are lying at $z > 4$, claiming that such a low fraction is perhaps due to significant number of spurious-red contaminants in their catalogue. To account for a high level of noise in the *H*-ATLAS field, authors agreed that some refined selection technique is needed. In this paper we keep our final selection based on the traditional threshold of $S_{500}/S_{350} > 1$ for ease of comparison with existing studies.

6. What are 500 μm -risers?

6.1. Problem of multiplicity

The potential problem in performing refined selection criteria for 500 μm -risers is our limited knowledge of source multiplicity. Different rates of multiplicities are claimed by studies found in the literature (e.g. Cowley et al. 2015; Simpson et al. 2015). Their conclusions are tightly related to selection criteria and instruments used to measure the flux (single dish or interferometer). Some observations suggest that multiplicity fraction may be dependent on the observed FIR flux, where is more liable for brighter sources to have multiple components (Karim et al. 2013; Bussmann et al. 2015). This assumption includes strongly lensed sources as well, since they are known to be very bright in SPIRE bands. A more recent SPIRE multiplicity study was done by Scudder et al. (2016). They used an improved version of the XID+ code (Hurley et al. 2017) to search for many potential contributing sources per 250 μm detection. They considered dusty galaxies covering a wide range of photometric redshifts and 250 μm flux densities. For 250 μm fluxes between 30 mJy and 45 mJy they found that a combination of the two brightest components accounts for approximately 90% of the flux in the 250 μm source. Additionally, the brightest component contributes more than 60% of the total flux.

Since we work with SPIRE data only, our multiplicity analysis is based on simulations. We perform photometry on simulated SPIRE maps and compare the flux emitted from the brightest galaxy to the total single-dish flux ratio. For each source detected as a 500 μm -riser in our mock maps, we assign the brightest component in the radius no larger than half of the 250 μm beam ($9''$). We then measure the ratio between the flux of this brightest galaxy from mock catalogues and the measured single-dish flux in our simulated maps. We place computed values in several flux bins. Our results are presented in Fig. 11. For galaxies detected as 500 μm -risers, we found that 72 to 85% of observed 250 μm flux is emitted by the brightest galaxy (78% on average).

Overall, we see almost insignificant change for S16 and B17, with the trend expressing a slight increase towards higher 250 μm fluxes. Repeating the same analysis for the 500 μm beam we find the similar trend with a somewhat lower brightest-galaxy fraction (64% on average) due to stronger resolution effects. Even from such a simple multiplicity analysis, it seems plausible to expect that selection of red sources from prior 250 μm detections does not experience strong resolution effects. The situation is somewhat different for missed red sources – the contribution of brightest components is lower (47–65%) with a tendency to decrease towards larger 250 μm fluxes. This is due to a presence of “blue” blends close to the position of 500 μm -risers ($r < 9''$).

Scudder et al. (2016) found the ratio of the fraction of the brightest 250 μm components to the total flux measured in

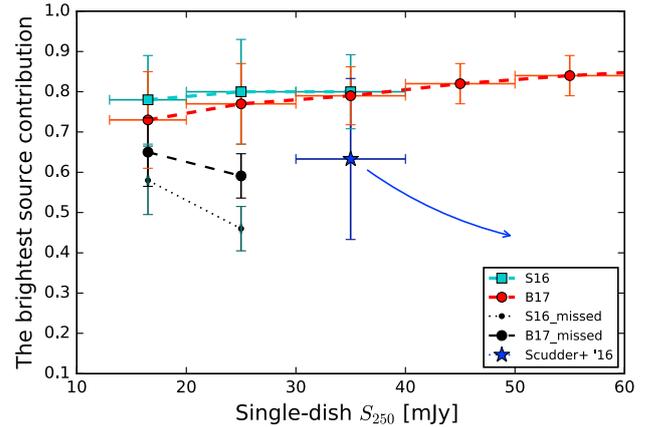


Fig. 11. Multiplicity of red sources detected in mock maps. Average fraction of the 250 μm flux density emitted by the brightest galaxy in the *Herschel* beam is plotted as a function of total, single-dish 250 μm flux measured in our simulated maps. Horizontal error bars indicate the width of the chosen flux bin, while vertical error bars show the standard deviation of the distribution. The blue star refers to the study of Scudder et al. (2016), and the blue arrow indicates the observed trend they found towards brighter fluxes.

simulated maps to be higher than one. They used deep multi-wavelength data and applied a different definition of the flux density fraction. Additionally, their analysis considered sources regardless of SPIRE colours making a direct comparison difficult here. Our multiplicity predictions appear to be in good agreement with the measurements of Greenslade et al. (in prep.). They performed a Submillimeter Array (SMA) follow up program for 36 500 μm -risers selected from various fields and found a multiplicity rate of 33% with brightest components contributing 50–75% of the SMA flux. To shed new light on this issue, the contamination of sources at intermediate redshifts ($1 < z < 2$) seen in deep NGVS and PACS images will be discussed in a following paper.

6.2. Lensing and clustering

Considering the increase with redshift of the optical depth to lensing and the magnification bias, we account for the possibility that some of our 500 μm -risers are strongly lensed (B ethermin et al. 2017; Negrello et al. 2017; Cai et al. 2013). Without interferometric data allowing us to identify the true lensing fraction of our 500 μm -risers, we cannot precisely correct for the lensed population in our final sample. We thus investigate their relative contribution by making the comparison with recent ALMA/NOEMA follow-up results of 500 μm -risers selected in other fields (*H*-ATLAS, HerMES and HeLMS, Oteo et al. 2017b, see also Fudamoto et al. 2017), and using predictions from B17. Our findings are summarised in Fig. 12 and Table 8.

The sample used for high-resolution ALMA follow-up in Oteo et al. (2017b) contains 44 red DSFGs. They were chosen from complete samples of red DSFGs presented in Iverson et al. (2016), Asboth et al. (2016), and Dowell et al. (2014). The sources span the wide range of 500 μm fluxes, from 30 mJy to 162 mJy. We note here that the sample is highly incomplete, accounting for only the reddest sources whose colours are consistent with high photometric redshifts, estimated to be $z \sim 4$ –6. Observed total fraction of strongly lensed galaxies in Oteo et al. (2017b) is 40% (18 out of 44). However, most of the lensed sources are in the bright flux regime of red DSFGs

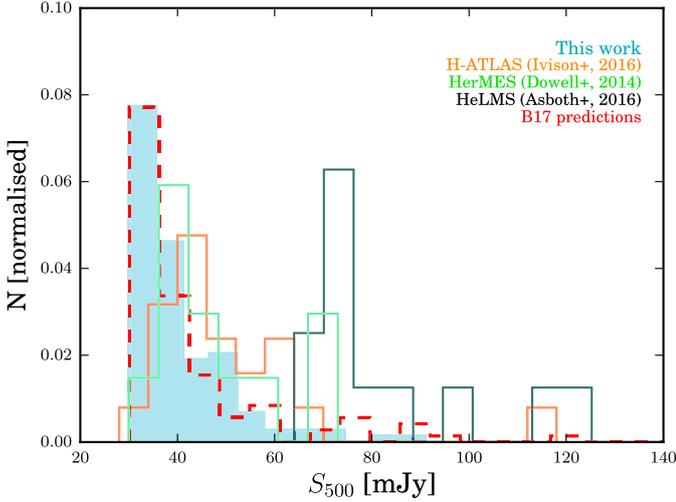


Fig. 12. Flux distribution of $500\mu\text{m}$ -risers from this work (coloured blue area) compared to $500\mu\text{m}$ -risers considered for a submillimetre interferometric follow-up in different studies. Orange, black, and green stepped lines represent sub-samples selected from Ivison et al. (2016), Asboth et al. (2016), and Dowell et al. (2014), respectively. Flux distribution of red sources from the B17 that are modelled applying the same selection criteria presented in this work is represented by the dashed red line.

Table 8. A relative contribution of strongly lensed sources to $500\mu\text{m}$ -risers selected in different studies (1).

	Median S_{500} [mJy]	Lensed [observed]	Lensed [modelled (B17)]
<i>H</i> -ATLAS	45 ± 14	$23^{+13}_{-11}\%$ (5/21)	$28^{+8}_{-9}\%$
HeLMS	78 ± 19	$69^{+21}_{-25}\%$ (9/13)	$75^{+10}_{-13}\%$
HerMES	47 ± 9	$40^{+10}_{-8}\%$ (4/10)	$36^{+9}_{-7}\%$
This work	38 ± 4	–	$24^{+6}_{-5}\%$

Notes. Flux distribution of $500\mu\text{m}$ -risers from this work compared to 44 red DSFGs presented in Oteo et al. (2017b). These are *H*-ATLAS, HerMES, and HeLMS sub-samples observed with ALMA and NOEMA at $870\mu\text{m}$, 1 mm, and 3 mm (Oteo et al. 2017b; Fudamoto et al. 2017). The second column shows median $500\mu\text{m}$ fluxes along with the median $500\mu\text{m}$ flux from this work. The relative contribution of observed, strongly lensed sources in various selections is shown in the third column. The fourth column shows the relative contribution of modelled, strongly lensed sources from B17. A relative contribution of strongly lensed sources is modelled applying the same selection cuts of $500\mu\text{m}$ -risers for tabulated extragalactic fields.

($S_{500} > 52$ mJy)⁵. To make the sample suitable for comparison with our sources and model predictions, we split it into three subsamples according to methods used for their selections: 21 sources from Ivison et al. (2016), 13 sources from Asboth et al. (2016), and 10 sources from Dowell et al. (2014). In Fig. 12 we show $500\mu\text{m}$ flux distribution of our sample alongside the sub-samples drawn from Oteo et al. (2017b). We see that

⁵ Here we distinguish between fainter ($S_{500} < 52$ mJy), and brighter $500\mu\text{m}$ -risers ($S_{500} > 52$ mJy). The chosen separation is based on model predictions (B17), where highly magnified sources are contributing more than 50% to the population of red DSFGs at fluxes higher than $S_{500} > 52$ mJy.

$500\mu\text{m}$ -risers selected in this work have average fluxes fainter than other presented galaxies. We show the observed number of strongly lensed objects in *H*-ATLAS, HerMES, and HeLMS in the second column of Table 8. The sub-sample of $500\mu\text{m}$ -risers in HeLMS contains the largest fraction of lensed sources (75%), while the lowest observed fraction is for *H*-ATLAS (23%).

We further use B17 and consider the same selection criteria as initial studies (Ivison et al. 2016; Asboth et al. 2016; Dowell et al. 2014). For each sub-sample we examine the range of fluxes identical to what is presented in Oteo et al. (2017b) and simulate the effect noise as explained in Sect. 5.1.5. From Table 8 we see a very good match between model predictions and observations. We estimate the predicted contribution of strongly lensed sources to our sample of $500\mu\text{m}$ -risers. Applying our selection criteria we find that B17 predicts $24^{+6}_{-5}\%$ fraction of strongly lensed $500\mu\text{m}$ -risers in our sample. More precisely, 17% of the strongly magnified galaxies are expected to lie at a fainter flux regime ($S_{500} < 52$ mJy) where most of our sources reside (119 out of 133, or 89%). The median redshift of modelled, strongly lensed sources is $\hat{z} = 4.2 \pm 0.4$, whilst median magnification differs by $\hat{\mu} = 5.3 \pm 3$ and $\hat{\mu} = 8.1 \pm 5$ for the fainter and brighter flux regime, respectively. This implies that strong lensing can affect the observed luminosity function of our $500\mu\text{m}$ -risers and lowers the fraction of intrinsically bright sources. After correcting for lensing, we expect to have 24% of galaxies with luminosities falling in the range $1.3 \times 10^{12} L_{\odot} < L_{\text{IR}} < 10^{13} L_{\odot}$ (classified as ULIRGs), and 76% of sources with $L_{\text{IR}} > 10^{13} L_{\odot}$, classified as hyperluminous IR galaxies (HyLIRGs).

Apart from strong lensing, clustering might be responsible for producing observed excess in number counts of $500\mu\text{m}$ -risers. Several studies of the clustering of submillimetre bright galaxies at $1 < z < 3$ (e.g. $M_{*} > 10^{11} M_{\odot}$, and $L_{\text{IR}} > 10^{12} L_{\odot}$, see Farrah et al. 2006; Wilkinson et al. 2017) have shown that they are hosted in moderately strongly clustered massive halos. *Herschel* detected that star-forming galaxies are on average more strongly clustered at higher redshifts ($z \sim 2$) than nearby objects (Béthermin et al. 2014; Ono et al. 2014). Since the beam size in *Herschel* is a direct function of the wavelength, we expect the largest SPIRE beam ($500\mu\text{m}$) to be more affected than the other two beams. Béthermin et al. (2017) have studied in detail the influence of clustering on FIR continuum observations. They found that the confusion which arises from clustering increases the number of observed $500\mu\text{m}$ -risers. The number of such contaminants causes the peak of intrinsic redshift distribution of $500\mu\text{m}$ -risers to be somewhat lower than what would be expected from observed SPIRE colours⁶. Therefore, future spectroscopic confirmations of full samples of selected $500\mu\text{m}$ -risers are needed to quantify and compare the exact fraction of sources at $z > 4$.

Simulations presented in Sect. 5.2 have some limitations in that they are affected by cosmic variance beyond simple Poissonian, and thus contain under- or overdensities at some redshifts. Furthermore, a purely probabilistic treatment is used to model the lensing, lacking the physical connection to the overdensity of lower- z systems. To improve our understanding of how lensing and clustering might shape $500\mu\text{m}$ -riser selection, different solutions can be considered. Along with deeper, interferometric observations, more complex cosmological simulations are needed. They should include larger halo volumes, high-resolution to unveil fainter galaxies responsible for

⁶ Ivison et al. (2016) and Duivenvoorden et al. (2018) found that more than half of DSFGs from their samples have $z_{\text{phot}} < 4$. Photometric redshifts are estimated from SPIRE and SCUBA-2/LABOCA data.

confusion, environmental effects, and refined treatment of a lens modelling.

6.3. Star-formation rate density

Bright, individually detected $500\mu\text{m}$ -risers are important for understanding the evolution of massive systems. Here we test if such a selection is suitable to measure the total SFR density (SFRD). We make a rough estimate of the SFRD assuming the statistical properties of selected $500\mu\text{m}$ -risers. We determine SFRD at $4 < z < 5$ applying the equation (Vincioletto et al. 2012; Hogg 1999):

$$\rho_*(z) = \frac{\sum \text{SFR}_{\text{IR}}}{\frac{4\pi}{3} \int_{z=4}^{z=5} \frac{c/H_0}{\sqrt{\Omega_M(1+z)^3 + \Omega_\Lambda}} dz}, \quad (6)$$

where $\rho_*(z)$ is SFRD, defined as a total sum of SFRs per co-moving volume, while SFR_{IR} is determined from IR luminosity. The denominator in Eq. (6) is the co-moving volume contained within $4 < z < 5$, and IR luminosity is converted to SFR applying the standard conversion formula from Kennicutt (1998):

$$\text{SFR}_{\text{IR}} (M_\odot \text{yr}^{-1}) = 1.71 \times 10^{10} L_{\text{IR}} (L_\odot). \quad (7)$$

We apply our selection criteria to model (B17). Since B17 provides a magnification factor for each source, we further use the modelled, lensing-corrected luminosities and redshift distribution of $500\mu\text{m}$ -risers. We then correct for the number of missed and contaminant sources applying the result from our simulations (Fig. 11, Sect. 5.2). We place IR luminosities in a wide bin by the co-moving volume per deg^2 . Scaling with the observed area and applying cosmological parameters $\Omega_M = 0.307$, $\Omega_\Lambda = 0.693$, and $H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-2}$ (Planck Collaboration XIII 2016), we find that $\text{SFRD} = 1.99 \times 10^{-4} M_\odot \text{yr}^{-1} \text{ Mpc}^{-3}$ for $500\mu\text{m}$ -risers at $4 < z < 5$. This result is shown with a filled red star in Fig. 13. Large uncertainties ($\pm 1.4 \times 10^{-4}$) determined by Monte Carlo bootstrapping are due to the Poisson noise and the fact that we make constraints with SPIRE data only. We additionally cross-check this result and make the same estimation of SFRD but adopting median IR luminosity ($1.94 \times 10^{13} L_\odot$) from observed $L_{\text{IR}} - z$ distribution (see Sect. 3).

Based on the expected fraction of strongly amplified objects found in Sect. 6.2 (24%), we randomly chose sources from our catalogue to correct for lensing, namely 17% of sources at $S_{500} < 52 \text{ mJy}$, and 83% at $S_{500} > 52 \text{ mJy}$. These percentages correspond to predictions from B17. We then adopt a median magnification from B17 that corresponds to a chosen flux-regime ($\mu = 5.3$ for fainter, and $\mu = 8.1$ for brighter “ $500\mu\text{m}$ -risers”), and produce the observed, lensing-corrected luminosity distribution. We find $\text{SFRD} = 2.87 \times 10^{-4} M_\odot \text{yr}^{-1} \text{ Mpc}^{-3}$. Our model-based SFRD estimation is consistent with recent observational findings from Duivenvoorden et al. (2018). We note also that offset between the study of Dowell et al. (2014) and our own may be explained by the fact that the estimation by Dowell et al. (2014) assumed substantially higher IR luminosities and purity ratios.

Furthermore, we can use the shape of the luminosity function to correct for the fainter red sources. We extrapolate the contribution of $500\mu\text{m}$ -risers to SFRD – firstly we remove the $500\mu\text{m}$ flux limit (thus accounting for red sources fainter than $S_{500} < 30 \text{ mJy}$), and then we remove the lower $250\mu\text{m}$ flux cut, thus accounting for all $500\mu\text{m}$ -risers below the *Herschel* sensitivity limit. Extrapolated contributions are depicted by empty red triangles in Fig. 13).

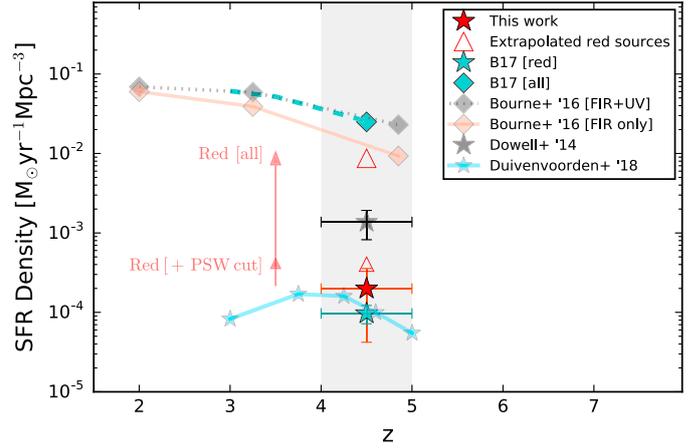


Fig. 13. SFRD as a function of redshift. The filled red star indicates our measurement. The estimated SFRDs of extrapolated $500\mu\text{m}$ -risers are shown with empty red-edged triangles: smaller when $250\mu\text{m}$ cut is imposed ($S_{250} > 13.2 \text{ mJy}$), and larger if all flux cuts are removed. Flux corrections are annotated with arrows. We show for comparison results from other $500\mu\text{m}$ -riser selections: B17 (cyan), Dowell et al. (2014; black) and Duivenvoorden et al. (2018; blue). Horizontal bars reflect the bin size. FIR and total FIR+UV observed SFRD (Bourne et al. 2017), as well as total simulated SFRD (B17), are marked with diamonds.

Here we assume the luminosity distribution from B17, but we should mention that we obtain a very similar result (higher by a factor of 1.2) by applying the luminosity function from S16. As expected, the largest portion of sources responsible for a stellar budget at $z > 4$ is missed due to sensitivity limitations of *Herschel* instruments and not to colour selection.

The contribution of the emission of DSFGs to the total IR luminosity functions up to $z = 5$ is partially investigated (e.g. Koprowski et al. 2017; Bourne et al. 2016; Rowan-Robinson et al. 2016; Madau & Dickinson 2014; Gruppioni et al. 2013; Burgarella et al. 2013). However, there is no consensus on whether or not UV estimates have notably underestimated the contribution of dust-enshrouded star formation in DSFGs because our knowledge at $z > 3$ is severely incomplete, mostly due to optical obscuration. Attempting to extend the cosmic SFRD up to $z = 6$, Rowan-Robinson et al. (2016) exploited existing $500\mu\text{m}$ *Herschel* selections of $500\mu\text{m}$ -risers. They report no decline in dust-obscured SFRDs in the range $z = 3-6$, and very high SFRD values at $4 < z < 5$, an order of magnitude higher than our estimation. However, our result is in much better agreement with studies of Bourne et al. (2016; marked with diamonds in Fig. 13), who used very deep millimetre imaging to reduce the effects of confusion on cosmic SFRD measurements.

7. Conclusions

We have performed a systematic search of dusty background galaxies on a 55 deg^2 area. We introduce an innovative method for selecting $500\mu\text{m}$ -risers ($S_{500} > S_{350} > S_{250}$). In order to estimate the $250-500\mu\text{m}$ flux of blended sources, we use the multiwavelength MBB-fitter code. We assign positions of galaxies detected in the $250\mu\text{m}$ map as a prior to extracting the flux densities at longer wavelengths, iteratively fitting their SEDs as modified blackbodies. We select 133 $500\mu\text{m}$ -risers that fulfil the following criteria: $S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2 \text{ mJy}$ and $S_{500} > 30 \text{ mJy}$. We reject all strong synchrotron sources by cross-matching available radio catalogues.

We use statistical properties of selected galaxies and estimate a median redshift value of $\hat{z} = 4.28$ and a corresponding median rest-frame IR luminosity of $\hat{L}_{\text{IR}} = 1.94 \times 10^{13} L_{\odot}$.

The total raw number density of selected sources is $2.41 \pm 0.34 \text{ deg}^{-2}$. To evaluate our results, we retrieve mock catalogues based on models of Béthermin et al. (2012, 2017) and Schreiber et al. (2016). Differential number counts are fairly consistent with model predictions. Namely, towards the brighter end ($S_{500} > 70 \text{ mJy}$), HeViCS counts have both slope and values identical to those anticipated by models. At the lowest flux end, discrepancy is the effect of noise and lensing. We find that noise tends to increase number counts of $500 \mu\text{m}$ -risers. It may be high enough to fully match with observations within 1σ uncertainty. We confirm that selecting $500 \mu\text{m}$ -risers applying our selection criteria is a direct way to detect a significant population of unlensed, dusty, high- z galaxies.

We build simulated maps. Clustering and lensing are modelled in order to fully resolve effects that produce colour uncertainties. We propose a modified colour criterion ($S_{500}/S_{350} > 0.97$) that should be probed in future selections of potentially $z > 4$ dusty sources. Motivation to apply this criterion is twofold: it would account for colour uncertainties that arise from effects of noise, clustering, and lensing, and it increases the sample of candidate $z > 4$ galaxies. Anticipated contribution of $2 < z < 3$ contaminants satisfying this colour cut is small, and reaches the maximum of 13%.

We inspect the effect of multiplicity working with simulated SPIRE data. The brightest galaxy inside the beam contributes on average 75% and 64% to the total single-dish flux measured at $250 \mu\text{m}$ and $500 \mu\text{m}$, respectively.

We use the model of Béthermin et al. (2017) and find that $24^{+6}_{-5}\%$ of $500 \mu\text{m}$ -risers selected with our criteria would be strongly lensed. After correcting measured luminosities for lensing, we expect 24% of sources to have luminosities of $10^{12} L_{\odot} < L_{\text{IR}} < 10^{13} L_{\odot}$ (classified as ULIRGs) and 76% of sources to have $L_{\text{IR}} > 10^{13} L_{\odot}$ (classified as HyLIRGs).

We use the statistical properties of our galaxies to determine the role of $500 \mu\text{m}$ -risers in the SFRD for $4 < z < 5$. We show that after correcting for fainter sources, the projected contribution of $500 \mu\text{m}$ -risers is just a factor of 2 below the total, dust-corrected SFRD.

Acknowledgements. We are thankful to the anonymous referee for very constructive comments and points which significantly improved this paper. We would like to acknowledge David Corre, Denis Burgarella, Alessandro Boselli, Tom Bakx and Eric Jullo for useful discussions. The project leading to this publication has received funding from Excellence Initiative of Aix-Marseille University – A*MIDEX, a French “Investissements d’Avenir” programme (ANR-11-IDEX-0001-02) and from the OCEVU Labex (ANR-11-LABX-0060), French government program managed by the ANR. C.P. acknowledges support from the Science and Technology Foundation (FCT, Portugal) through the Postdoctoral Fellowship SFRH/BPD/90559/2012, PEst-OE/FIS/UI2751/2014, PTDC/FIS-AST/2194/2012, and through the support to the IA activity via the UID/FIS/04434/2013 fund. HeViCS is a *Herschel* open time key program (PI Davies, Davies et al. 2010). All the data are publicly available through *Herschel* Science Archive <http://archives.esac.esa.int/hsa/whsa/#home>. *Herschel* is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA. SPIRE has been developed by a consortium of institutes led by Cardiff University (UK) and including Univ. Lethbridge (Canada); NAOC (China); CEA, LAM (France); IFSI, Univ. Padua (Italy); IAC (Spain); Stockholm Observatory (Sweden); Imperial College London, RAL, UCL-MSSL, UKATC, Univ. Sussex (UK); and Caltech, JPL, NHSC, Univ. Colorado (USA). This development has been supported by national funding agencies: CSA (Canada); NAOC (China); CEA, CNES, CNRS (France); ASI (Italy); MCINN (Spain); SNSB (Sweden); STFC (UK); and NASA (USA). This publication makes use of public data products from the Two Micron All Sky Survey (2MASS), which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center California Institute of Technology, funded by the National

Aeronautics and Space Administration and the National Science Foundation. We express our gratitude to Daft <https://github.com/dfm/daft>, a probabilistic graphical model available under the MIT License.

References

- Aird, J., Coil, A. L., Moustakas, J., et al. 2012, *ApJ*, 746, 90
 Amblard, A., Cooray, A., Serra, P., et al. 2010, *A&A*, 518, L9
 Asboth, V., Conley, A., Sayers, J., et al. 2016, *MNRAS*, 462, 1989
 Auld, R., Bianchi, S., Smith, M. W. L., et al. 2013, *MNRAS*, 428, 1880
 Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
 Béthermin, M., Dole, H., Cousin, M., & Bavouzet, N. 2010, *A&A*, 516, A43
 Béthermin, M., Dole, H., Lagache, G., Le Borgne, D., & Penin, A. 2011, *A&A*, 529, A4
 Béthermin, M., Daddi, E., Magdis, G., et al. 2012, *ApJ*, 757, L23
 Béthermin, M., Kilbinger, M., Daddi, E., et al. 2014, *A&A*, 567, A103
 Béthermin, M., Wu, H.-Y., Lagache, G., et al. 2017, *A&A*, 607, A89
 Blain, A. W., Barnard, V. E., & Chapman, S. C. 2003, *MNRAS*, 338, 733
 Bourne, N., Dunne, L., Maddox, S. J., et al. 2016, *MNRAS*, 462, 1714
 Bourne, N., Dunlop, J. S., Merlin, E., et al. 2017, *MNRAS*, 467, 1360
 Bracco, A., Cooray, A., Veneziani, M., et al. 2011, *MNRAS*, 412, 1151
 Burgarella, D., Buat, V., Gruppioni, C., et al. 2013, *A&A*, 554, A70
 Bussmann, R. S., Riechers, D., Fialkov, A., et al. 2015, *ApJ*, 812, 43
 Cai, Z.-Y., Lapi, A., Xia, J.-Q., et al. 2013, *ApJ*, 768, 21
 Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, *Nature*, 470, 233
 Casey, C. M., Berta, S., Béthermin, M., et al. 2012, *ApJ*, 761, 140
 Casey, C. M., Narayanan, D., & Cooray, A. 2014, *Phys. Rep.*, 541, 45
 Clements, D. L., Braglia, F., Petitpas, G., et al. 2016, *MNRAS*, 461, 1719
 Combes, F., Rex, M., Rawle, T. D., et al. 2012, *A&A*, 538, L4
 Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, *AJ*, 115, 1693
 Cowley, W. I., Lacey, C. G., Baugh, C. M., & Cole, S. 2015, *MNRAS*, 446, 1784
 Cox, P., Krips, M., Neri, R., et al. 2011, *ApJ*, 740, 63
 da Cunha, E., Groves, B., Walter, F., et al. 2013, *ApJ*, 766, 13
 Daddi, E., Dannerbauer, H., Stern, D., et al. 2009, *ApJ*, 694, 1517
 Davies, J. I., Baes, M., Bendo, G. J., et al. 2010, *A&A*, 518, L48
 Davies, J. I., Bianchi, S., Cortese, L., et al. 2012, *MNRAS*, 419, 3505
 Delhaize, J., Smolcic, V., Delvecchio, I., et al. 2017, *A&A*, 602, A4
 Dowell, C. D., Conley, A., Glenn, J., et al. 2014, *ApJ*, 780, 75
 Draine, B. T., & Li, A. 2007, *ApJ*, 657, 810
 Duivenvoorden, S., Oliver, S., Scudler, J. M., et al. 2018, *MNRAS*, 477, 1099
 Eales, S. A., Smith, M. W. L., Wilson, C. D., et al. 2010, *A&A*, 518, L62
 Eales, S., de Vis, P., Smith, M. W. L., et al. 2017, *MNRAS*, 465, 3125
 Elbaz, D., Hwang, H. S., Magnelli, B., et al. 2010, *A&A*, 518, L29
 Farrah, D., Lonsdale, C. J., Borys, C., et al. 2006, *ApJ*, 641, L17
 Ferrarese, L., Côté, P., Cuillandre, J.-C., et al. 2012, *ApJS*, 200, 4
 Flesch, E. W. 2015, *PASA*, 32, e010
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
 Franceschini, A., Rodighiero, G., Vaccari, M., et al. 2010, *A&A*, 517, A74
 Fudamoto, Y., Ivison, R. J., Oteo, I., et al. 2017, *MNRAS*, 472, 2028
 Galliano, F., Hony, S., Bernard, J.-P., et al. 2011, *A&A*, 536, A88
 Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2007, *AJ*, 133, 2569
 Grazian, A., Fontana, A., Santini, P., et al. 2015, *A&A*, 575, A96
 Greenslade, J., Clements, D. L., Cheng, T., et al. 2018, *MNRAS*, 476, 3336
 Greve, T. R., Vieira, J. D., Weiß, A., et al. 2012, *ApJ*, 756, 101
 Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, *A&A*, 518, L3
 Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, *MNRAS*, 432, 23
 Hayward, C. C. 2013, *MNRAS*, 432, L85
 Helfand, D. J., White, R. L., & Becker, R. H. 2015, *ApJ*, 801, 26
 Hezaveh, Y. D., & Holder, G. P. 2011, *ApJ*, 734, 52
 Hilbert, S., White, S. D. M., Hartlap, J., & Schneider, P. 2007, *MNRAS*, 382, 121
 Hogg, D. W. 1999, ArXiv e-prints [arXiv: astro-ph/9905116]
 Hurlley, P. D., Oliver, S., Betancourt, M., et al. 2017, *MNRAS*, 464, 885
 Ikarashi, S., Ivison, R. J., Caputi, K. I., et al. 2017, *ApJ*, 835, 286
 Ilbert, O., Arnouts, S., Le Floc’h, E., et al. 2015, *A&A*, 579, A2
 Ivison, R. J., Greve, T. R., Dunlop, J. S., et al. 2007, *MNRAS*, 380, 199
 Ivison, R. J., Lewis, A. J. R., Weiss, A., et al. 2016, *ApJ*, 832, 78
 Jarrett, T.-H., Chester, T., Cutri, R., et al. 2000, *AJ*, 120, 298
 Johnson, T. L., Sharon, K., Bayliss, M. B., et al. 2014, *ApJ*, 797, 48
 Karim, A., Swinbank, A. M., Hodge, J. A., et al. 2013, *MNRAS*, 432, 2
 Kennicutt, Jr., R. C. 1998, *ARA&A*, 36, 189
 Kneib, J.-P., & Natarajan, P. 2011, *A&ARv*, 19, 47
 Koprowski, M. P., Dunlop, J. S., Michałowski, M. J., et al. 2017, *MNRAS*, 471, 4155
 Lapi, A., González-Nuevo, J., Fan, L., et al. 2011, *ApJ*, 742, 24

- Le Borgne, D., Elbaz, D., Ocvirk, P., & Pichon, C. 2009, *A&A*, 504, 727
- Liu, D., Daddi, E., Dickinson, M., et al. 2018, *ApJ*, 853, 172
- Lutz, D. 2014, *ARA&A*, 52, 373
- Ma, Z., & Yan, H. 2015, *ApJ*, 811, 58
- MacKenzie, T. P., Scott, D., Smail, I., et al. 2014, *MNRAS*, 445, 201
- Madau, P., & Dickinson, M. 2014, *ARA&A*, 52, 415
- Magdis, G. E., Daddi, E., Béthermin, M., et al. 2012, *ApJ*, 760, 6
- Magnelli, B., Lutz, D., Saintonge, A., et al. 2014, *A&A*, 561, A86
- Mancuso, C., Lapi, A., Shi, J., et al. 2016, *ApJ*, 823, 128
- Merlin, E., Bourne, N., Castellano, M., et al. 2016, *A&A*, 595, A97
- Michałowski, M. J., Hayward, C. C., Dunlop, J. S., et al. 2014, *A&A*, 571, A75
- Miettinen, O., Novak, M., Smolčić, V., et al. 2015, *A&A*, 584, A32
- Mullaney, J. R., Alexander, D. M., Goulding, A. D., & Hickox, R. C. 2011, *MNRAS*, 414, 1082
- Negrello, M., Amber, S., Amvrosiadis, A., et al. 2017, *MNRAS*, 465, 3558
- Nguyen, H. T., Schulz, B., Levenson, L., et al. 2010, *A&A*, 518, L5
- Norris, R. P., Hopkins, A. M., Afonso, J., et al. 2011, *PASA*, 28, 215
- Oliver, S. J., Wang, L., Smith, A. J., et al. 2010, *A&A*, 518, L21
- Oliver, S. J., Bock, J., Altieri, B., et al. 2012, *MNRAS*, 424, 1614
- Ono, Y., Ouchi, M., Kurono, Y., & Momose, R. 2014, *ApJ*, 795, 5
- Oteo, I., Ivison, R. J., Dunne, L., et al. 2016, *ApJ*, 827, 34
- Oteo, I., Ivison, R. J., Dunne, L., et al. 2017a, *ApJ*, 856, 72
- Oteo, I., Ivison, R. J., Negrello, M., et al. 2017b, *ApJ*, submitted [arXiv: 1709.04191]
- Pannella, M., Elbaz, D., Daddi, E., et al. 2015, *ApJ*, 807, 141
- Pappalardo, C., Bendo, G. J., Bianchi, S., et al. 2015, *A&A*, 573, A129
- Pappalardo, C., Bizzocchi, L., Fritz, J., et al. 2016, *A&A*, 589, A11
- Pearson, E. A., Eales, S., Dunne, L., et al. 2013, *MNRAS*, 435, 2753
- Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, *ApJ*, 721, 193
- Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, *A&A*, 518, L1
- Planck Collaboration XIII. 2016, *A&A*, 594, A13
- Planck Collaboration Int. XXXIX. 2016, *A&A*, 596, A100
- Pope, A., & Chary, R.-R. 2010, *ApJ*, 715, L171
- Rangwala, N., Maloney, P. R., Glenn, J., et al. 2011, *ApJ*, 743, 94
- Riechers, D. A., Bradford, C. M., Clements, D. L., et al. 2013, *Nature*, 496, 329
- Rigby, E. E., Maddox, S. J., Dunne, L., et al. 2011, *MNRAS*, 415, 2336
- Rodríguez-Puebla, A., Behroozi, P., Primack, J., et al. 2016, *MNRAS*, 462, 893
- Roseboom, I. G., Oliver, S. J., Kunz, M., et al. 2010, *MNRAS*, 409, 48
- Roseboom, I. G., Ivison, R. J., Greve, T. R., et al. 2012, *MNRAS*, 419, 2758
- Rowan-Robinson, M., Oliver, S., Wang, L., et al. 2016, *MNRAS*, 461, 1100
- Safarzadeh, M., Ferguson, H. C., Lu, Y., Inami, H., & Somerville, R. S. 2015, *ApJ*, 798, 91
- Salmi, F., Daddi, E., Elbaz, D., et al. 2012, *ApJ*, 754, L14
- Sargent, M. T., Béthermin, M., Daddi, E., & Elbaz, D. 2012, *ApJ*, 747, L31
- Savage, R. S., & Oliver, S. 2007, *ApJ*, 661, 1339
- Schober, J., Schleicher, D. R. G., & Klessen, R. S. 2017, *MNRAS*, 468, 946
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, *A&A*, 575, A74
- Schreiber, C., Elbaz, D., Pannella, M., et al. 2016, *A&A*, 589, A35
- Schreiber, C., Elbaz, D., Pannella, M., et al. 2017a, *A&A*, 602, A96
- Schreiber, C., Pannella, M., Leiton, R., et al. 2017b, *A&A*, 599, A134
- Scudder, J. M., Oliver, S., Hurley, P. D., et al. 2016, *MNRAS*, 460, 1119
- Shu, X. W., Elbaz, D., Bourne, N., et al. 2016, *ApJS*, 222, 4
- Simpson, J. M., Swinbank, A. M., Smail, I., et al. 2014, *ApJ*, 788, 125
- Simpson, J. M., Smail, I., Swinbank, A. M., et al. 2015, *ApJ*, 807, 128
- Smith, D. J. B., Jarvis, M. J., Hardcastle, M. J., et al. 2014, *MNRAS*, 445, 2232
- Smolčić, V., Miettinen, O., Tomičić, N., et al. 2017, *A&A*, 597, A4
- Soneira, R. M., & Peebles, P. J. E. 1978, *AJ*, 83, 845
- Strandet, M. L., Weiss, A., De Breuck, C., et al. 2017, *ApJ*, 842, L15
- Swinbank, A. M., Smail, I., Longmore, S., et al. 2010, *Nature*, 464, 733
- Swinbank, A. M., Simpson, J. M., Smail, I., et al. 2014, *MNRAS*, 438, 1267
- Symeonidis, M., Vaccari, M., Berta, S., et al. 2013, *MNRAS*, 431, 2317
- Toft, S., Smolčić, V., Magnelli, B., et al. 2014, *ApJ*, 782, 68
- Valiante, E., Smith, M. W. L., Eales, S., et al. 2016, *MNRAS*, 462, 3146
- Vieira, J. D., Marrone, D. P., Chapman, S. C., et al. 2013, *Nature*, 495, 344
- Vincoletto, L., Matteucci, F., Calura, F., Silva, L., & Granato, G. 2012, *MNRAS*, 421, 3116
- Wang, L., Viero, M., Clarke, C., et al. 2014, *MNRAS*, 444, 2870
- Weiß, A., De Breuck, C., Marrone, D. P., et al. 2013, *ApJ*, 767, 88
- Whitaker, K. E., Franx, M., Bezanson, R., et al. 2015, *ApJ*, 811, L12
- Wilkinson, A., Almaini, O., Chen, C.-C., et al. 2017, *MNRAS*, 464, 1380
- Yuan, F., Buat, V., Burgarella, D., et al. 2015, *A&A*, 582, A90
- Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, *ApJ*, 554, 803
- Zavala, J. A., Montaña, A., Hughes, D. H., et al. 2018, *Nat. Astron.*, 2, 56

Appendix A: Dusty radio sources

As mentioned in Sect. 2.5, our final list of 500 μm -risers has been cleared of strong radio sources and quasars (QSOs, 7 objects in total). We measure significant FIR emission close to position of these sources, and here we briefly investigate its nature. Radio emission from quasars may be due to star formation in the quasar host galaxy, to a jet launched by the supermassive black hole, or to relativistic particles accelerated in a wide-angle radiatively driven outflow. Our MBB fitting approach realtively efficiently “protects” star-forming selection against the flat (in terms of SED) and dusty synchrotron sources. When MBB-fitter photometry of such flat sources is performed, most of the χ^2 values lie outside the range we accept as good for our final selection ($\chi^2 < 3.84$, see Appendix B). Between the seven objects removed from the final list, three of them are defined as radio-loud quasars (QSO) or blazars, with catalogued L -band 1.4 GHz radio emission larger than 1 mJy (5.04 mJy, 799.97 mJy and 6.2 mJy respectively), and their redshifts span the range from 0.6 to 0.965. An example is the brightest radio source catalogued as HVS387 in our catalogue or J122222.5+041316 in the FIRST database. It has IR excess and very strong UV and gamma-ray emission. The source has a spectroscopic redshift of $z = 0.964$ and according to the classical PSF-fitting photometry performed with, for example, SUSSEXtractor or DA0phot it could pass our red source criteria, but it cannot be fitted accurately using the MBB-fitter. Since they may have colours similar to those of high- z DSFGs, heterogeneously selected 500 μm -risers that exist in the literature thus should be considered with caution, since they may be polluted by synchrotron-dominated sources whose IR energy is not star-formation related. Another possibility is that some of those QSOs act as a foreground lens, amplifying the signal from the higher-redshift, dusty system.

Another four QSO-like sources are found by cross-matching with the Half Million Quasar catalogue (HMQ, Flesch 2015). They are optically bright, high- z QSO-like sources close to the position of our SPIRE detection ($r < 3''$). They have high photometric redshifts - on average $z = 2.74$, with one source lying at $z > 4$. To fully inspect the relation of these objects to the population that we aimed to select in the current study, in our following paper we consider optical and PACS data, together with archival WISE data. Here we can determine the expected radio emission of our final group of 133 500 μm -risers implementing FIR-radio correlation:

$$q_{\text{IR}} = \log\left(\frac{F_{\text{FIR}}}{3.75 \times 10^{12} \text{ W m}^{-2}}\right) - \log\left(\frac{S_{1.4}}{\text{W m}^{-2} \text{ Hz}^{-1}}\right), \quad (\text{A.1})$$

where $S_{1.4}$ is the continuum radio emission flux at 1.4 GHz per frequency such that $S_{1.4} \propto \nu^{-\alpha}$ and α is the radio spectral index, positive in the vast majority of sources. F_{FIR} is the rest-frame FIR dust emission flux. There is considerable evidence that local star-forming galaxies express a tight FIR-radio correlation. For example, Yun et al. (2001) analysed the sample of 1800 IRAS galaxies and measured this value to be $q_{\text{IR}} = 2.37 \pm 0.01$ with a dispersion of 0.25 dex. From this we estimate that for our $z \sim 4$ candidates 1400 MHz flux densities are in the range between 25 μJy and 90 μJy . Assuming the faintest level of existing radio surveys covering Virgo (~ 0.75 –1 mJy at 1.4 GHz), we can reveal just local cluster members and radio bright quasars, but not dusty background, star-forming galaxies.

However, FIR-radio correlation cannot be implemented accurately in order to estimate expected radio flux of optically selected quasars which have radio flux below the FIRST (NVSS)

limit. As we add QSO contribution to both radio and IR, correlation could break down. This could be particularly due to the origin of the FIR/submm emission which includes a small contribution from the AGN torus, being predominantly linked to (kpc-scale) heated dust by the AGN. Recent studies have enlarged the sample of *Herschel*-detected optically selected QSOs (e.g. Ma & Yan 2015). They, however, claimed a statistically large group of objects for which SEDs are due to the cold-dust components within the host galaxies, consistent with being heated by active star formation. Lacking the L_{IR} correlation with the black hole masses or the X-ray QSO luminosities of the quasars, this could support a scenario where their thermal SED is not AGN-driven. Thus, our subsample of QSO-like objects could be a very interesting case study for probing the connection between AGN activity and host galaxy star-formation. Existing large sky radio catalogues are not deep enough to distinguish between very dusty radio-dominated galaxies at higher redshifts and star-forming systems. This issue will be addressed with deep radio observations performed over the area covering Virgo; Evolutionary Map of the Universe (EMU survey, Norris et al. 2011) will make a deep ($10''$, rms = 10 $\mu\text{Jy beam}^{-1}$) continuum maps.

It is clear that in addition to FIR/submm data, radio data are of significant importance to further uncover the true nature of selected dusty sources. Since photometric redshifts calculated using the SPIRE data alone are highly uncertain, the addition of radio data (via FIR-radio correlation) might be expected to improve the photometric redshift accuracy, for example to $\Delta z/(1+z) = 0.15$ (Roseboom et al. 2012). From this, it is vital to understand the true nature of the FIR to radio correlation at redshifts higher than 2, since some recent results have claimed much larger scatter and even break at redshifts larger than 4 (e.g. Miettinen et al. 2015; Delhaize et al. 2017, see also Schober et al. 2017). Thus the redshift independent FIR-radio correlation cannot be used straightforwardly to break temperature-redshift degeneracy as suggested in most previous papers found in the literature (e.g. Pope & Chary 2010). Approaching existing and future interferometers (JVLA, SKA), we can put constraints on the photometric redshifts of our red sources, and can also uncover the sizes of their star-forming regions and investigate how they compare to other tracers such as FIR or CO.

Appendix B: Colour-redshift diagram related to the MBB model

In the FIR regime, SED peak is provoked by thermal emission from dust of different temperatures, and it can be used in extragalactic surveys such as *Herschel* to identify objects in a specific redshift range. To test our theory that selecting 500 μm -risers is the way to select $z \gtrsim 4$ candidates, we run a Monte Carlo simulation to sample all possible colours as a function of redshift. We test which range of redshifts we should expect assuming the MBB model with the fixed various ranges of parameters. We confirm that if sources are not sufficiently cold, they reside at $z \gtrsim 4$. Thus, we are particularly interested in unveiling MBB-fitted sources with $S_{500}/S_{350} > 1$ and $S_{500}/S_{250} > 2$. Plotted in Fig. B.1 are some such examples, where we applied different combinations of fixed T_{d} and β . At redshifts $z > 6$ the heating of dust by the cosmic microwave background (CMB) could be more significant and we take that into account in our modelling of observed SED following da Cunha et al. (2013).

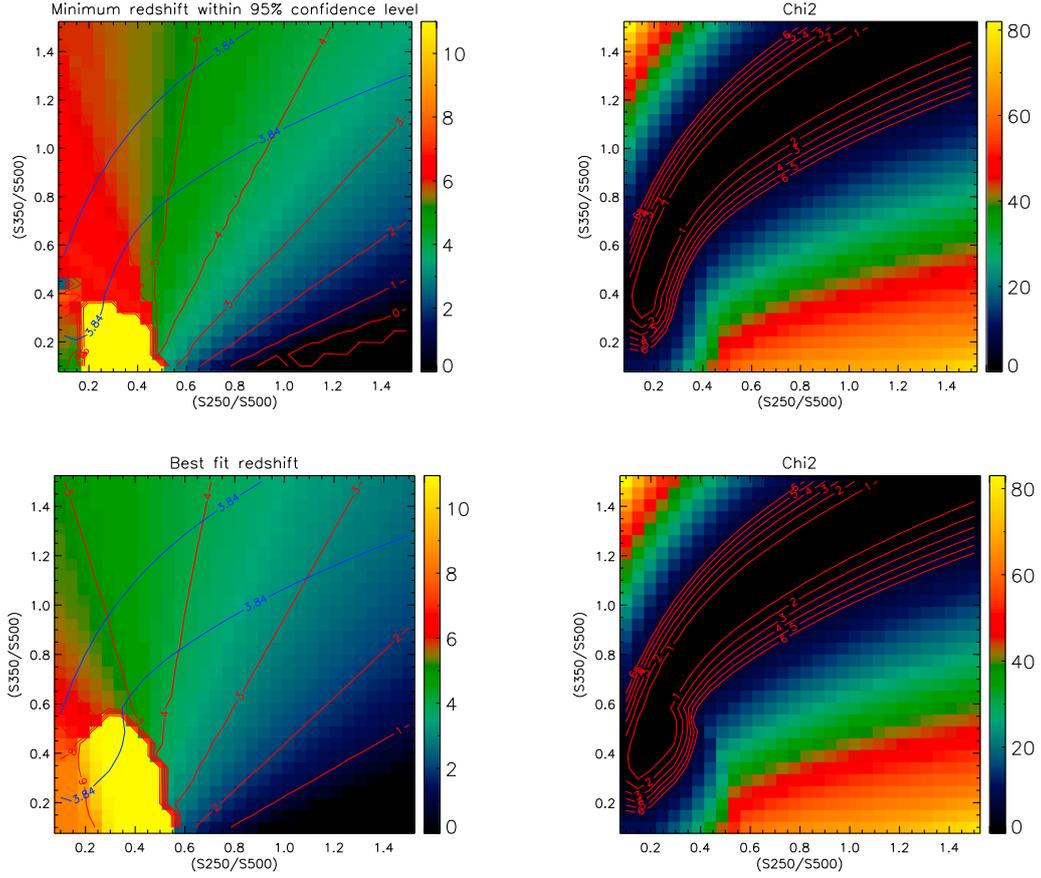


Fig. B.1. Criteria for selecting $500\ \mu\text{m}$ -risers. Plotted is a SPIRE colour diagram related to the MBB model. The plot is a result of a Monte Carlo simulation in which we start with a single-temperature modified blackbody with the flux $S_{500} = 30\ \text{mJy}$. The mock SEDs are generated to account for different fixed T_d and β illustrating temperature-redshift degeneracy. Plotted against all possible colours are simulated sources with $T_d = 45\ \text{K}$ and $\beta = 1.8$ (*upper panel*) and $T_d = 38\ \text{K}$ and $\beta = 1.8$ (*lower panel*). *Left*: colour-redshift plot. Red lines represent redshift tracks rising from right to left. The coloured background indicates the average redshift in these colour-colour spaces. Blue lines define the 95% confidence limit region. A correction factor due to modelled effect of CMB heating on colours is applied. However, our simulations show that we do not expect a major contribution from that effect at sources lying below $z \sim 7$; *right*: different chi-square fitting values reflected on SPIRE colour-colour ratios. For our final $500\ \mu\text{m}$ -risers selection we kept only those sources whose chi-square resides inside the 95% confidence limit region ($\chi^2 < 3.84$).

Appendix C: An example of faint and deblended $500\ \mu\text{m}$ -risers

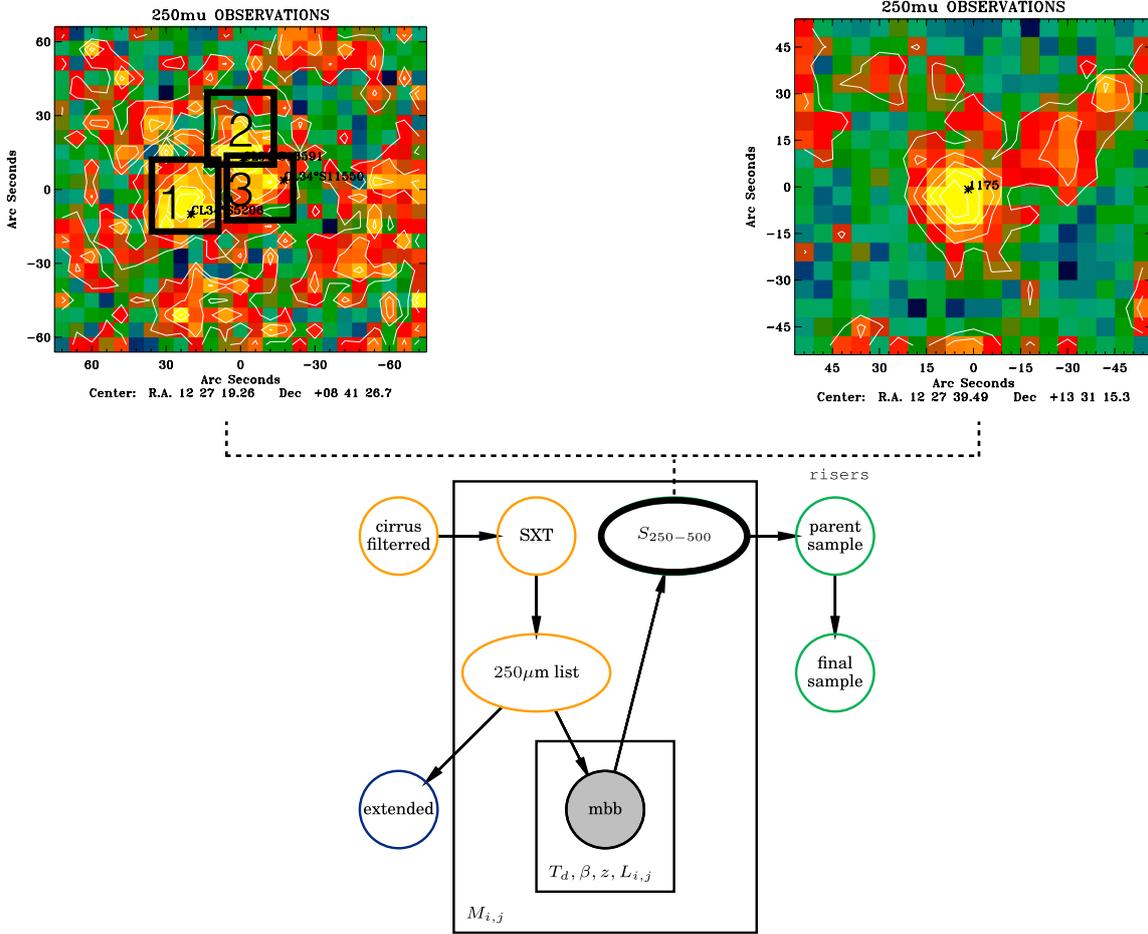


Fig. C.1. Example 2D cutouts of confused (*left*) and isolated point source detections (*right*) from our $250\ \mu\text{m}$ catalogue. The first group of sources are fitted simultaneously, while the second is fitted as a single source.

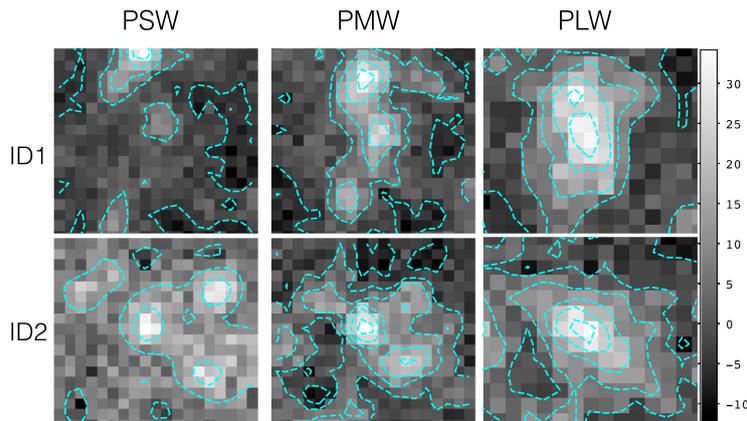


Fig. C.2. Example $60'' \times 60''$ cutouts of $500\ \mu\text{m}$ -risers from our final sample. Both sources are detected in the second iteration of our extraction procedure, when we add sources that appeared bright in our residual maps. *Upper panel*: $250\ \mu\text{m}$, $350\ \mu\text{m}$ and $500\ \mu\text{m}$ cutout of a multisource case where two sources are de-blended for their $350\ \mu\text{m}$ and $500\ \mu\text{m}$ emission. *Lower panel*: an isolated case, without a prominent $350\ \mu\text{m}$ or $500\ \mu\text{m}$ blend with respect to the beam.

Appendix D: Additional table

Table D.1. Selected sources.

ID (HVS)	Source name	RA [J2000] [h m s]	Dec [J2000] [° ' '']	S_{250} [mJy]	S_{250}^{err} [mJy]	S_{350} [mJy]	S_{350}^{err} [mJy]	S_{500} [mJy]	S_{500}^{err} [mJy]
1	HVS J120715+1453.9	12:07:15.07	+14:53:17.9	17.81	1.75	35.07	3.46	35.94	3.54
2	HVS J120724+1430.0	12:07:24.76	+14:30:14.0	20.96	1.87	40.52	3.61	40.9	3.64
3	HVS J120909+1433.7	12:09:09.18	+14:33:26.7	14.03	1.67	32.46	3.87	38.84	4.63
4	HVS J120936+1518.0	12:09:36.01	+15:18:36.0	17.95	1.69	37.63	3.55	40.07	3.78
5	HVS J121140+1416.8	12:11:40.83	+14:16:10.8	9.86	2.23	25.62	5.86	39.8	5.68
6	HVS J121540+1307.9	12:15:40.18	+13:07:36.9	11.82	1.61	28.72	3.91	37.17	5.06
7	HVS J121847+1537.3	12:18:47.43	+15:37:39.3	18.22	1.73	36.82	3.49	38.61	3.67
8	HVS J122017+1535.9	12:20:17.55	+15:35:14.9	18.14	1.86	34.69	3.56	34.69	3.56
9	HVS J122217+1534.7	12:22:17.80	+15:34:11.7	14.62	1.74	29.46	3.52	30.18	3.61
10	HVS J123037+1207.1	12:30:37.15	+12:07:15.1	16.34	1.75	33.94	3.64	35.82	3.84
11	HVS J122654+1214.6	12:26:54.33	+12:14:22.6	15.88	1.79	31.50	3.55	31.82	3.59
12	HVS J122445+0930.7	12:24:45.12	+09:30:14.7	13.34	1.65	29.88	3.70	34.92	4.32
13	HVS J123304+1309.0	12:33:04.61	+13:09:16.0	15.45	1.77	31.86	3.65	34.05	3.90
14	HVS J123027+1237.7	12:30:27.61	+12:37:39.7	13.59	2.04	28.34	3.63	30.01	3.85
15	HVS J122438+0953.5	12:24:38.28	+09:53:41.6	17.27	1.73	35.37	3.54	37.56	3.76
16	HVS J122646+1152.1	12:26:46.27	+11:52:35.0	17.76	2.09	35.35	3.49	36.57	3.61
17	HVS J122416+1208.1	12:24:16.90	+12:08:16.1	23.64	1.79	46.08	3.49	46.82	3.56
18	HVS J122938+1355.7	12:29:38.92	+13:55:49.7	21.81	1.81	42.91	3.56	43.95	3.65
19	HVS J121814+1529.1	12:18:14.42	+15:29:36.1	15.73	1.57	36.9	3.68	45.05	4.49
20	HVS J122358+0725.1	12:23:58.59	+07:25:34.1	13.43	1.67	32.19	3.92	41.36	5.04
21	HVS J121740+1613.0	12:17:40.07	+16:13:26.0	9.3	1.72	24.17	4.48	37.71	6.99
22	HVS J121723+0418.0	12:17:23.84	+04:18:58.0	14.86	1.75	30.38	3.55	31.27	3.65
23	HVS J122816+0824.3	12:28:16.15	+08:24:02.3	13.41	1.58	31.65	3.74	39.0	4.61
24	HVS J123138+1020.0	12:31:38.77	+10:20:59.0	11.99	1.58	28.29	3.73	35.44	4.67
25	HVS J123027+1237.3	12:30:27.76	+12:37:37.3	14.32	2.14	31.91	3.56	36.42	3.70
26	HVS J123232+0701.9	12:32:32.99	+07:01:32.9	18.56	2.05	37.31	3.51	38.07	3.58
27	HVS J123435+0839.9	12:34:35.03	+08:39:47.9	15.65	1.73	31.67	3.49	33.25	3.68
28	HVS J121613+0253.1	12:16:13.80	+02:53:11.1	36.75	1.78	70.84	3.44	71.31	3.46
29	HVS J121744+0721.5	12:17:44.72	+07:21:27.5	21.99	1.85	46.42	3.46	49.75	3.39
30	HVS J121409+0516.7	12:14:09.70	+05:16:37.7	15.53	1.73	31.68	3.52	32.85	3.65
31	HVS J121801+0348.4	12:18:01.13	+03:48:10.4	14.03	2.11	31.22	3.73	35.51	4.25
32	HVS J121016+0527.1	12:10:16.25	+05:27:22.1	15.28	2.16	31.37	3.62	32.72	3.77
33	HVS J122526+0414.7	12:25:26.88	+04:14:52.7	13.68	2.15	29.71	3.62	32.83	3.98
34	HVS J122521+0407.8	12:25:23.23	+04:07:18.8	12.45	2.38	28.51	3.84	33.64	4.53
35	HVS J122821+0407.4	12:28:21.13	+04:07:32.4	12.81	1.60	32.32	4.04	45.43	5.67
36	HVS J123917+1112.4	12:39:17.85	+11:12:42.4	13.24	2.21	28.33	3.58	30.82	3.91
37	HVS J122854+1328.6	12:28:54.43	+13:28:05.6	12.72	1.65	28.21	3.67	32.57	4.23
38	HVS J122649+0502.2	12:26:49.90	+05:02:01.2	13.34	1.65	29.03	3.62	32.78	4.11
39	HVS J121755+0721.2	12:17:55.74	+07:21:00.2	12.69	1.63	29.85	3.87	36.63	4.73
40	HVS J122505+0735.6	12:25:05.61	+07:35:44.6	24.77	1.77	50.85	3.25	53.04	3.07
41	HVS J121654+1604.8	12:16:54.97	+16:04:14.8	14.05	1.73	29.65	3.65	31.78	3.91
42	HVS J122446+0540.0	12:24:46.52	+05:40:18.0	14.22	1.79	29.09	3.67	30.25	3.81
43	HVS J122031+1551.5	12:20:31.49	+15:51:26.5	16.36	1.65	37.25	3.72	43.61	4.35
44	HVS J123419+1054.2	12:34:19.79	+10:54:00.2	17.46	1.57	40.73	3.66	49.37	4.48
45	HVS J122810+0512.6	12:28:10.78	+05:12:02.6	15.66	1.73	31.94	3.58	33.11	3.66
46	HVS J123200+0821.9	12:32:00.58	+08:21:37.9	18.42	1.86	41.18	3.65	47.19	3.67
47	HVS J121443+0344.0	12:14:43.75	+03:44:29.0	10.08	1.63	25.98	4.19	38.98	6.3
48	HVS J123220+1214.0	12:32:20.88	+12:14:37.0	13.46	1.67	28.0	3.53	30.18	3.80
49	HVS J121851+0301.8	12:18:51.09	+03:01:53.8	16.38	1.69	34.8	3.59	37.58	3.87
50	HVS J122000+0422.0	12:20:00.48	+04:22:01.0	18.28	1.61	40.2	3.54	45.09	3.97
51	HVS J122744+0948.1	12:27:44.28	+09:48:01.1	12.12	1.59	27.86	3.66	33.04	4.34
52	HVS J122623+0621.1	12:26:23.79	+06:21:26.1	16.81	1.55	40.81	3.78	52.72	4.88

Notes. Column descriptions: *column 1*: catalogue numbers (HVS); *column 2*: full source name (source names are based on their positions for the J2000). *columns 3–4*: coordinates of sources (RA, Dec); *column 4–9*: SPIRE fluxes (S_{250} , S_{350} , S_{500}) and accompanied errors in the flux estimation (S^{err}). Flux errors are estimated as explained in Sect. 2.6. We note, however, that errors are strongly correlated, since we performed “forced” MBB photometry at SPIRE wavelengths.

Table D.1. continued

ID (HVS)	Source name	RA [J2000] [h m s]	Dec [J2000] [° ' '']	S_{250} [mJy]	S_{250}^{err} [mJy]	S_{350} [mJy]	S_{350}^{err} [mJy]	S_{500} [mJy]	S_{500}^{err} [mJy]
53	HVS J122939+1355.5	12:29:39.00	+13:55:49.5	18.62	1.73	39.84	3.53	44.21	3.64
54	HVS J122000+0524.9	12:20:00.60	+05:24:12.9	17.27	1.77	34.4	3.54	34.89	3.59
55	HVS J122124+0911.4	12:21:24.48	+09:11:46.4	18.33	1.65	37.64	3.40	39.28	3.55
56	HVS J120934+1526.0	12:09:34.04	+15:26:47.0	29.55	1.69	60.62	3.48	63.18	3.62
57	HVS J122728+0939.0	12:27:28.12	+09:39:00.0	16.81	1.80	35.48	3.51	38.03	3.48
58	HVS J123348+0703.5	12:33:48.69	+07:03:55.5	17.86	1.69	37.32	3.54	39.62	3.76
59	HVS J122248+1635.6	12:22:48.00	+16:35:19.6	18.42	1.68	41.2	3.64	47.15	4.17
60	HVS J121346+0438.2	12:13:46.22	+04:38:36.2	19.86	1.79	41.28	3.67	43.59	3.88
61	HVS J122101+1449.9	12:21:01.77	+14:49:40.9	22.0	1.64	47.46	3.46	52.08	3.79
62	HVS J122026+0731.4	12:20:26.24	+07:31:28.4	21.89	2.18	43.96	3.37	44.9	2.89
63	HVS J122157+0526.6	12:21:57.97	+05:26:35.6	18.95	1.73	38.51	3.52	39.79	3.64
64	HVS J121349+0438.1	12:13:49.72	+04:38:28.1	31.41	1.58	73.99	3.69	90.97	4.54
65	HVS J123044+0950.7	12:30:44.45	+09:50:49.7	14.14	2.09	28.58	3.51	29.99	3.69
66	HVS J122547+1529.7	12:25:47.42	+15:29:32.7	21.98	2.13	45.11	3.53	47.04	3.68
67	HVS J122142+1611.4	12:21:42.97	+16:11:50.4	21.81	2.10	42.79	3.53	42.81	3.53
68	HVS J123312+0713.5	12:33:12.72	+07:13:27.5	15.85	2.27	30.32	3.45	30.34	3.46
69	HVS J122634+1400.0	12:26:34.07	+14:00:10.0	18.18	2.11	38.23	3.6	40.85	3.84
70	HVS J121742+0351.1	12:17:42.83	+03:51:28.1	11.77	2.19	26.34	3.69	30.18	4.23
71	HVS J122148+1443.8	12:21:48.69	+14:43:52.8	17.55	2.15	34.65	3.27	34.86	3.29
72	HVS J122415+0958.8	12:24:15.19	+09:58:41.8	22.87	2.23	46.32	3.50	47.68	3.60
73	HVS J121815+0307.5	12:18:15.98	+03:07:43.5	14.73	1.75	31.08	3.39	33.31	3.35
74	HVS J122854+1328.3	12:28:54.46	+13:28:05.3	15.08	1.54	33.27	3.41	37.45	3.84
75	HVS J121900+1550.6	12:19:00.96	+15:50:19.6	14.52	1.56	36.32	3.90	49.9	5.36
76	HVS J121949+1610.5	12:19:49.92	+16:10:09.5	17.93	1.81	37.84	3.42	40.56	3.41
77	HVS J121827+1544.7	12:18:27.36	+15:44:15.7	16.02	1.87	32.89	3.49	34.31	3.43
78	HVS J121911+1618.4	12:19:11.52	+16:18:09.4	13.02	1.67	28.65	3.67	32.79	4.21
79	HVS J123032+1014.2	12:30:32.38	+10:14:07.2	23.69	2.06	47.57	3.38	48.59	3.02
80	HVS J123000+0415.5	12:30:00.59	+04:15:21.5	19.6	1.61	40.25	3.31	42.0	3.47
81	HVS J122024+1434.2	12:20:24.48	+14:34:23.2	12.5	1.64	28.23	3.72	33.34	4.39
82	HVS J122610+1027.9	12:26:10.56	+10:27:02.9	20.16	1.78	42.07	3.62	45.49	3.92
83	HVS J123346+1007.0	12:33:46.32	+10:07:03.0	13.78	2.17	28.29	3.60	30.1	3.84
84	HVS J121604+1246.7	12:16:04.08	+12:46:21.7	13.34	1.75	28.85	3.68	32.37	4.14
85	HVS J123126+1249.7	12:31:26.16	+12:49:43.7	8.93	1.76	23.24	4.61	36.93	7.34
86	HVS J123120+1412.2	12:31:20.88	+14:12:38.2	13.72	1.68	28.88	3.55	30.87	3.81
87	HVS J122627+0336.6	12:26:27.53	+03:36:31.6	16.27	1.79	32.58	3.58	33.9	3.72
88	HVS J123155+1420.9	12:31:55.20	+14:20:17.9	13.41	1.71	28.45	3.63	30.66	3.91
89	HVS J121037+0517.2	12:10:37.19	+05:17:22.2	15.93	1.86	32.7	3.40	34.11	3.28
90	HVS J122212+1141.7	12:22:12.24	+11:41:33.7	15.94	1.86	32.72	3.41	34.13	3.35
91	HVS J122527+1153.0	12:25:27.12	+11:53:28.0	13.64	1.72	28.83	3.63	31.58	3.98
92	HVS J122449+0951.2	12:24:49.90	+09:51:12.2	15.92	1.71	32.89	3.53	34.51	3.74
93	HVS J122727+0525.6	12:27:27.06	+05:25:29.6	23.19	2.13	46.57	3.40	47.57	2.97
94	HVS J121601+1231.4	12:16:01.18	+12:31:11.4	14.78	2.13	36.54	3.84	48.96	5.18
95	HVS J122845+0929.8	12:28:45.36	+09:29:34.8	14.58	1.73	29.13	3.46	30.23	3.59
96	HVS J123007+1338.1	12:30:07.92	+13:38:45.1	21.29	2.05	42.75	3.30	43.67	2.90
97	HVS J122508+0756.7	12:25:08.64	+07:56:02.7	11.9	1.63	26.49	3.63	30.72	4.21
98	HVS J121843+0328.8	12:18:43.68	+03:28:10.8	18.27	1.81	35.72	3.54	36.38	3.61
99	HVS J121713+1202.9	12:17:13.96	+12:02:42.9	18.93	2.09	38.0	3.42	38.82	3.57
100	HVS J122519+0518.1	12:25:19.92	+05:18:40.1	16.7	1.78	31.94	3.42	31.96	3.62
101	HVS J122709+0856.1	12:27:09.42	+08:56:26.1	17.52	1.68	38.42	3.69	42.96	4.13
102	HVS J122156+1527.0	12:21:56.03	+15:27:57.0	18.44	2.02	37.86	3.41	39.49	3.77

Table D.1. continued

ID (HVS)	Source name	RA [J2000] [h m s]	Dec [J2000] [° ' "]	S_{250} [mJy]	S_{250}^{err} [mJy]	S_{350} [mJy]	S_{350}^{err} [mJy]	S_{500} [mJy]	S_{500}^{err} [mJy]
103	HVS J123003+1318.3	12:30:03.80	+13:18:04.3	22.1	1.77	46.63	3.47	49.98	3.69
104	HVS J122619+0443.7	12:26:19.44	+04:43:37.7	14.24	1.82	30.05	3.47	32.21	3.77
105	HVS J121735+0413.8	12:17:35.04	+04:13:42.8	15.58	1.98	31.98	3.52	33.36	4.37
106	HVS J121451+0412.5	12:14:51.58	+04:12:44.5	11.74	1.67	26.36	3.75	30.31	4.32
107	HVS J121617+0408.0	12:16:17.76	+04:08:21.0	13.08	1.62	29.6	3.66	35.0	4.33
108	HVS J123427+0714.0	12:34:27.98	+07:14:54.0	25.39	2.03	50.99	3.31	52.08	3.93
109	HVS J121543+0619.5	12:15:43.92	+06:19:12.5	16.35	1.75	33.65	3.61	35.9	3.85
110	HVS J123007+1332.1	12:30:07.93	+13:32:07.1	18.17	1.92	37.3	3.42	38.91	3.54
111	HVS J122414+1014.6	12:24:14.05	+10:14:36.6	21.98	2.02	45.12	3.38	47.06	3.78
112	HVS J121304+0519.0	12:13:04.30	+05:19:28.0	11.33	1.66	25.84	3.74	30.3	4.42
113	HVS J122517+1121.3	12:25:17.75	+11:21:56.3	13.98	2.53	30.82	3.46	34.68	3.87
114	HVS J122100+1306.9	12:21:00.68	+13:06:51.9	30.94	2.13	62.14	3.38	63.47	3.95
115	HVS J122709+0540.2	12:27:09.60	+05:40:00.2	20.09	2.86	38.74	3.52	39.02	3.57
116	HVS J123213+0740.8	12:32:13.44	+07:40:30.8	12.36	2.68	27.12	3.70	30.92	4.21
117	HVS J122109+0946.9	12:21:09.49	+09:46:33.9	30.53	2.12	61.3	3.30	62.62	3.83
118	HVS J121904+1338.9	12:19:04.56	+13:38:53.9	25.03	1.96	51.39	3.44	53.6	3.21
119	HVS J123619+0909.5	12:36:19.04	+09:09:44.5	26.06	2.13	52.33	3.31	53.45	3.84
120	HVS J121818+0506.4	12:18:18.40	+05:06:48.4	35.85	2.05	71.98	3.38	73.53	3.04
121	HVS J122611+1518.8	12:26:11.82	+15:18:33.8	24.69	2.02	50.7	3.41	52.88	3.23
122	HVS J123032+0641.8	12:30:32.44	+06:41:54.8	31.33	2.04	62.91	3.29	64.26	3.93
123	HVS J121110+1220.3	12:11:10.88	+12:20:35.3	24.74	2.06	49.68	3.30	50.75	3.90
124	HVS J123113+1220.1	12:31:13.14	+12:20:43.1	16.38	2.14	33.62	3.44	35.07	3.29
125	HVS J122646+1049.3	12:26:46.02	+10:49:06.3	16.22	1.66	34.18	3.50	36.57	3.75
126	HVS J121900+1516.5	12:19:00.34	+15:16:49.5	21.75	1.92	44.65	3.41	46.57	3.23
127	HVS J122806+0928.7	12:28:06.04	+09:28:11.7	39.83	2.03	79.97	3.32	81.69	3.97
128	HVS J123000+0803.6	12:30:00.33	+08:03:26.6	21.04	1.99	43.19	3.41	45.05	3.15
129	HVS J121609+1601.2	12:16:09.60	+16:01:59.2	14.11	1.64	31.25	3.64	36.04	4.19
130	HVS J122241+0708.4	12:22:41.71	+07:08:26.4	21.1	1.71	44.93	3.65	49.44	4.03
131	HVS J122335+1108.0	12:23:35.52	+11:08:56.0	12.35	1.63	26.86	3.61	30.33	4.08
132	HVS J123304+1326.6	12:33:04.80	+13:26:27.6	14.75	1.68	32.51	3.70	37.26	4.25
133	HVS J123317+1235.2	12:33:17.76	+12:35:31.2	14.29	1.77	28.74	3.58	30.03	3.74