



Universiteit
Leiden
The Netherlands

Towards a census of high-redshift dusty galaxies with Herschel. A selection of "500 μm -risers"

Donevski, D.; Buat, V.; Boone, F.; Pappalardo, C.; Bethermin, M.; Schreiber, C.; ... ; Duivenvoorden, S.

Citation

Donevski, D., Buat, V., Boone, F., Pappalardo, C., Bethermin, M., Schreiber, C., ... Duivenvoorden, S. (2018). Towards a census of high-redshift dusty galaxies with Herschel. A selection of "500 μm -risers". *Astronomy And Astrophysics*, 614, A33.
doi:10.1051/0004-6361/201731888

Version: Not Applicable (or Unknown)
License: [Leiden University Non-exclusive license](#)
Downloaded from: <https://hdl.handle.net/1887/68988>

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

Unveiling High-Redshift Dusty Galaxies By *Herschel*

A selection of "500 μm -risers"

D. Donevski¹, V. Buat¹, F. Boone², C. Pappalardo^{3,4}, M. Bethermin^{1,7}, C. Schreiber⁵, F. Mazed¹, J. Alvarez-Marquez⁶, and S. Duivenvoorden⁸

¹ Aix Marseille Univ, 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 ; accepted

ABSTRACT

Context. Over the last decade a large content 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 μm to 500 μm (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 deg², imposing the criteria: $S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2$ mJy and $S_{500} > 30$ mJy. Differential number counts are in a fairly good agreement with models, displaying better match than other existing samples. Our technique allows us to recover $\sim 60\%$ of genuine 500 μm -risers with redshift distribution peaking at $z \gtrsim 4$.

Conclusions. Selecting the "500 μm -risers" that fulfil our criteria is an efficient way to pre-select $z \gtrsim 4$ galaxies from SPIRE data alone. We show that noise and weak 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 found it close to the total value measured in far-infrared. It indicates that colour selection is not a limiting effect to search for the most massive, dusty $z > 4$ sources.

Key words. Galaxies: statistics – Galaxies: evolution – Galaxies: star formation – Galaxies: high-redshift – Infrared: Galaxies: photometry: Galaxies – Submillimeter: 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 they are the progenitors of massive ellipticals seen in overdense regions of the local Universe (e.g. Casey et al. 2014).

The widely accepted picture is that dusty star-forming galaxies (DSFG) occupy most massive halos in early epochs and lie on the most extreme tail of the galaxy stellar mass function (e.g. Michałowski et al. 2014, Hayward 2013, Michałowski et al. 2012, Béthermin et al. 2012). Such massive and dusty systems 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 wide and blind concept of searches. The *Herschel* SPIRE photometer (Griffin et al. 2010) was often used for mapping large areas at wavelengths of 250 μm , 350 μm and 500 μm (PSW, PMW and PLW band subsequently). The redshift peak of most of *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 70-100 μm , colours of sources in *Herschel* SPIRE bands were used to select candidate high-redshift dusty objects. To search for $z \gtrsim 4$ candidates, there is a particular interest to exploit the sources having red SPIRE colours, with rising flux densities from 250 to 500 microns (so-called "500 μm -risers"). Such galaxies should lie at $z \gtrsim 4$ 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).

When 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 offer us insight into very distant and dusty, star-forming galactic events. Recently, a rapidly flourishing literature on dusty high- z galaxies has grown, including catalogues with handful of "500 μm -risers" (e.g. Cox et al. 2011, Riechers et al. 2013, Vieira et al. 2013, Wardlow et al. 2014, 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 standardized manner, several works used map-search technique (Asboth et al. 2016, Dowell et al. 2014, but see also Ivison et al. 2016). Both Dowell et al. (2014) and Asboth et al. (2016) used lowest-resolution *Herschel* SPIRE-maps to select "500 μm -risers". They have led to the discovery of a most distant dusty starburst galaxy (HFLS3, $z = 6.34$, Riechers et al. 2013) and claimed significant overprediction of observed number of "500 μm -risers" with those predicted by existing models (e.g. Béthermin et al. 2012).

Even if *Herschel* offers a direct insight to the IR emission of high- z objects, there are critical limitations like sensitivity of detectors and low spatial resolution. These are responsible for biases such as source confusion. The sensitivity of SPIRE instrument allows to directly detect only the brightest, thus rarest objects at the tip of luminosity function (Cowley et al. 2015, Karim et al. 2013, Oliver et al. 2010) and we therefore need large surveys to increase the statistics. Study of Dowell et al. (2014) analyzed 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 area of *H*-ATLAS and HeLMS (HerMES Large Mode Survey) field respectively (see Section 2.1 for details about field's properties).

In this work we aim to introduce a slightly different approach to build 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 organized as follows: in Section 2 we describe the methods we used for the data analysis and our new selection criteria for "500 μm -risers". In Section 3 and Section 4 we present expected redshift/luminosity trend of selected galaxies and differential number counts. In Section 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 searching for $z \gtrsim 4$ sources. The nature of "500 μm -risers" and main conclusions are outlined in Section 6 and Section 7 respectively. We assume a Planck Collaboration et al. (2016a) cosmology.

¹ Throughout the text we accept following terminology regarding source's colours: red colours and red sources refer to "500 μm -risers", while "almost red sources" refer to galaxies having SED peak around 350 μm .

2. Data Analysis

2.1. HeViCS field

HeViCS is a fully-sampled survey that covered a region centered at the Virgo cluster (Davies et al. 2010, Davies et al. 2012). It is one of the largest uniform *Herschel* surveys, and its main advantage is the sensitivity and the 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 Section 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 is 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 8 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 PSW, PMW and PLW band respectively, and 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 PSW, PMW and PLW band 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 maps 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 $\theta > 10 \text{ deg}^2$) used for detection of "500 μm -risers". 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 uniformly cover 600 deg^2 of sky, but with its two scans survey did not reach the level of confusion noise at 250 μm . Studies of Asboth et al. (2016) and Dowell et al. (2014) acquired the data from different fields which are part of HerMES survey (Oliver et al. 2012). The HerMES survey observed 380 deg^2 of the sky. The survey has a hierarchical structure containing 7 levels, ranging from very deep observations of clusters to wider fields with varying size and depth. The largest (and the shallowest)

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

Field	Mode	N(rep)	Time [hr]	Area [deg ²]	Noise level in mJy 5 σ at 250 μ m
(1)	(2)	(3)	(4)	(5)	(6)
HeViCS	Parallel	8	286	55	30.5
H-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) The 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 later six fields are areas with different design levels nested as a part of HerMES: Total area covered in HerMES is 380 deg². Shallow HeLMS field covers the area of 274 deg², and deeper Level 1- Level 6 fields cover the total area of about 80 deg². FLS and Lockman-SWIRE have been used to probe "500 μ m-risers" selection by [Dowell et al. \(2014\)](#), while [Asboth et al. \(2016\)](#) applied the same selection method in the HeLMS field.

observed area is HeLMS, with its 274 square degrees. HeViCS maps consist of two or four time more scans comparing to other fields listed in Table 1. It leads to a reduction of instrumental noise by a factor of $\sqrt{2}$. The global 250 μ 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 repeating our analysis on 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. PSW map of V2 field in HeViCS is strongly affected by galactic cirrus emission. This contamination peaks at around 150-200 μ m ([Bracco et al. 2011](#), [Valiante et al. 2016](#)), implying 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 PSW map into meshes larger than the pixel size. The SEXtractor makes an initial pass through the pixel data, and compute 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. Important step in this procedure is the choice of the box size, since we do not want the background estimation 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 source extraction process described in next subsection.

2.4. Extraction of sources

Blind SPIRE source catalogues have been produced for HeViCS ([Pappalardo et al. 2015](#)). Nonetheless, when density of sources is very high, which is the case in 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 choose to perform extraction of PMW and PLW fluxes at exact *a-priori* position of PSW detections, allowing much precise identification work. The potential limitation of such a method is that we might be eventually missing some "500 μ 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 source finding algorithm optimized for isolated point-sources, SUSSEXtractor ([Savage & Oliver \(2007\)](#)), to create the catalogue of galaxies detected in the 250 μ m map as a prior to extract the flux densities at longer wavelengths. We then implement our novel technique (see [Section 2.6](#)) where source deblending and single temperature modified blackbody (MBB) fitting are combined in the same procedure.

In following we explain our source extraction pipeline (see Fig.1 for graphical description):

1. We run SUSSExtractor with use of fully-overlapped HeViCS 250-micron map. The SUSSExtractor works on the flux-calibrated, Level 2 *Herschel* SPIRE maps. We create a point response function (PRF) filtered image, smoothed with the PRF. In SUSSExtractor PRF is assumed to be Gaussian by default, with full-width-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 PSF-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 additional step looking for detections in our residual maps. Applying such additional step, we increase total number of sources by around 4%. Errors in the position estimated to be a source are determined as $0.6 \times (\text{FWHM}/\text{signal-to-noise})$, up to a maximum of 1.0 pixels, as suggested in 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 the values equal or higher than 13.2 mJy, which corresponds to $S_{250} \gtrsim 3\sigma_{\text{conf}}$. As a result of our source extraction pipeline at 250-micron maps, we listed 64309 sources, similarly distributed among the four (V1-V4) fields.
4. List of 250-detections is then divided in "no-neighbour" list of sources (250 μm sources without another 250 detection inside the PLW beam) and "close-neighbour" list where we add all sources having another 250 detection close-by. Prior to assign initial 250 μm list as an input to our modified blackbody fitting procedure, we clean the catalogue from potentially extended sources.

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 prevail the problem, we implement the same method 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 which has been tested in Pappalardo et al. (2015) - 70 contiguous pixels above the 1.2σ . In this case, most of the sources larger than 0.7 arcmin^2 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 decrease the number of 250-detections in our parent list from 64309 to 63497.

List of sources detected at 250 μm down to 13.2 mJy, cleaned from galactic cirrus contaminants and extended sources, is fur-

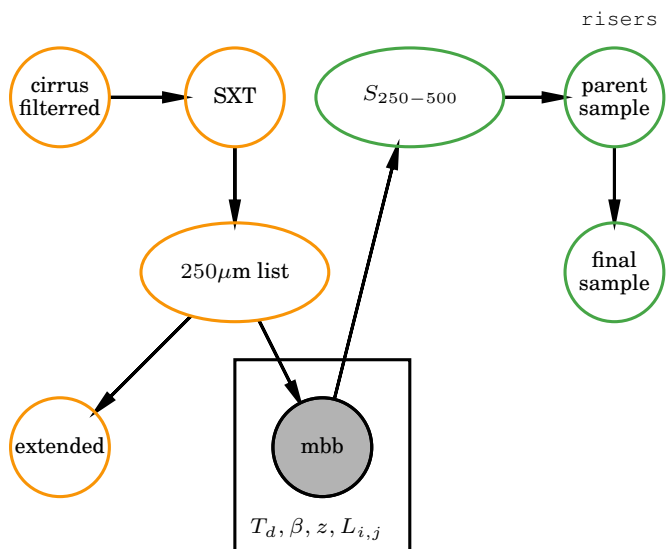


Fig. 1: Schematic representation of selection of "500 μm -risers" in the HeViCS field. Coloured in orange are segments of source extraction prior to use MBB-fitter. They are (from upper left, following the arrows): 1.) Subtraction of strong cirrus emission 2.) SUSSExtractor list of initial (threshold=3) PSW detections; 3.) Assumed *a-priori* list cleaned from extended sources. Enveloped by smaller black square are parameters considered for the fitting procedure with MBB-fitter and it corresponds to all pixels in the map where we have PSW detections; Coloured in green are segments of our selection after performing the photometry, subsequently: 1.) MBB photometry ($S_{250-500}$) at SPIRE wavelengths using PSW priors; 2.) Parent list of "500 μm -risers" (140 sources in total), not cleaned from strong synchrotron contaminants. 3.) Final list (133 in total) of "500 μm -risers" after excluding local radio-sources. We apply following selection criteria for the final sample: $S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2 \text{ mJy}$ and $S_{500} > 30 \text{ mJy}$.

ther used as *a-priori* list for our simultaneous modified blackbody fitting.

2.6. Modified blackbody fitter

As described in Section 2.1, our maps are limited by confusion noise caused by the high density of sources relative to the resolution of the *Herschel* instrument. SPIRE images have the average number of sources per beam potentially larger than one, so measured fluxes might be biased if we treat several blended sources as one. There are different approaches for source "de-blending" which has been invoked in the literature. We can separate them in three groups: Firstly, they are "de-blending" methods that use only positional priors as an 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. 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 an algorithm is efficient in obtaining Bayesian PDFs for all

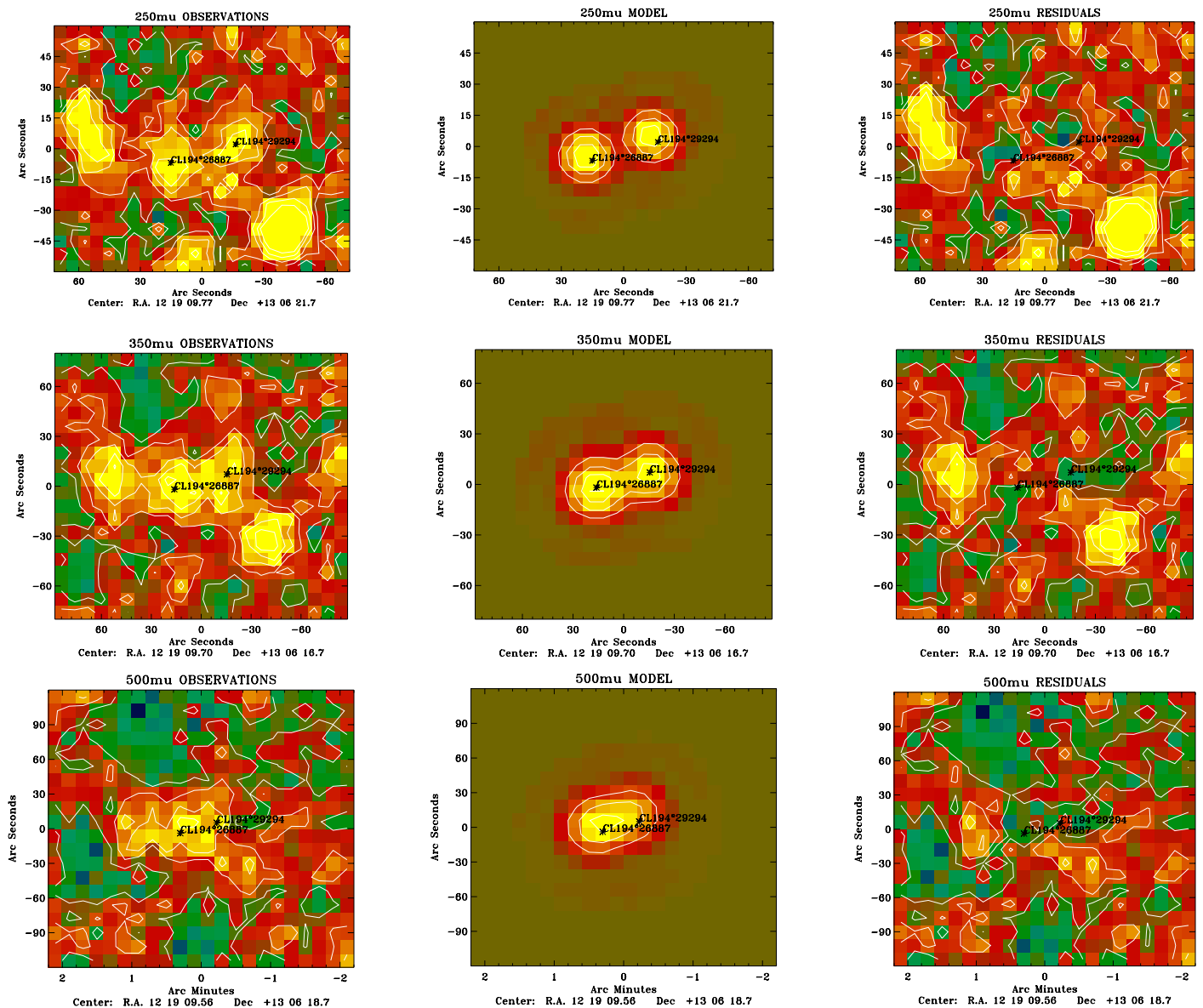


Fig. 2: Example of simultaneous prior-source deblending with MBB-fitter. Sources are taken from our "neighbour" list of 250-detections. Columns from left to right: PSW, PMW and PLW maps of sources (first column), subtracted models (second column) and residuals (third column). Note that in this example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

priors. However, they might still have some limitations, and the most important one is potential inability of unveiling the true high- z sources in the case they are not included in the initial list. Finally, third group of de-blenders consists from 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. 2017).

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 further use models to test the whole procedure of selecting "500 μm -risers" (see Section 5).

MBB-fitter is a code developed to extract sources from multiwavelength bolometric observations (Boone et al, in prep.). It combines positional priors and spectral information of sources,

such that SEDs of fitted galaxies should follow modified black-body (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 encapsulate just 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 summarize 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 3D space $s_k(\alpha, \delta, \nu)$, where ν is frequency, can

therefore be decomposed into a spatial distribution term (morphology) and a SED term:

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

where PSF_j is the PSF of the map at given frequency (ν_j), N_{sources} is a number of sources considered for the fitting, with its reference coordinates (α_k, δ_k). Here we assume that thermal continuum SED of galaxies in 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 following form $P_k = [L_{\text{IR}}, T_d, z_k, \beta_k, \gamma_k]$, where z_k is the redshift of given source, β_k 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.

Introduced set of parameters reflects 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).

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 all the pixels at maps need not to be included 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-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 (12135 sources in total), and (2) performing the photometry of sources that are more isolated without a surrounding companion inside the beam (51362 sources in total). We set emission spectral slope and dust temperature fixed 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. 2017). 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 deg². 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. We

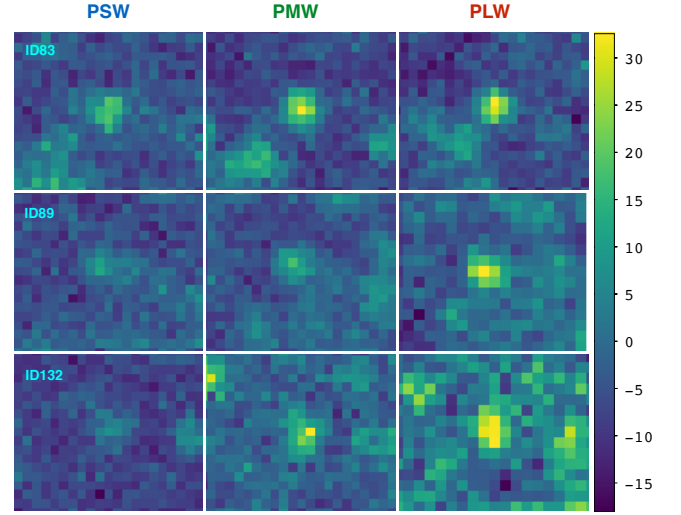


Fig. 3: 2D image cutouts as an illustration of "500 μm -risers" that fulfil our final selection criteria. Presented examples are drawn from our "no-neighbour list", meaning to be isolated, without a clear PSW detections in the radius of at least 36" around their centres. Colorbar shows fluxes measured in mJy.

performed several tests and chose to set flux density cut in PLW band at $S_{500} \approx 30$ mJy, which is related to 4σ above total noise measured in HeViCS maps. Using this value, we reach the completeness level of $\sim 80\%$ at 500 μm (see Fig.4) avoiding larger uncertainties in colours due to lower signal-to-noise ratios (see Table 5 and Section 5.2). PSW 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 Section 2.1). Amongst the final "500 μm -risers" (133 in total) we have 11 red candidates with one or two additional PSW detections inside the PLW beam. Example 2D cutouts of several "500 μm -risers" from our final sample are presented in Fig.3 (see also Appendix B and C). The final list of "500 μm -risers" in HeViCS is cleaned from strong radio sources that have flat spectra and very prominent FIR emission (7 objects in total). They may have colours similar to those of dusty, star-forming systems we are interesting to select. Contamination of our sample 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 quasars.²

2.8. Completeness and flux accuracy

We check the quality of our data analysis performing the 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 which are recovered in the simulated maps as real detections. Since our selection is based on *a priori* 250-micron positions, in this test we mostly inspect that band.

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

Table 2: The completeness fraction at SPIRE wavebands measured by injecting artificial sources into real SPIRE maps. Input fluxes are given in first column, while percentages of detected sources (average values per bin) are given in other columns.

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

We use Monte-Carlo simulation, artificially adding Gaussian sources and projecting them at randomized positions to our real PSW data, then convolving with the beam. Simulated sources in PMW and PLW maps are then placed at same positions. FWHM values we used to convolve are 17.6", 23.9" and 35.2", values we consider for our SUSSEXtractor detections. In each HeViCS field we add sources spanning 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, up to the 90-100 mJy). We perform the source extraction pipeline extracting the list of recovered point sources using SUSSEXtractor and searching for matched detections.

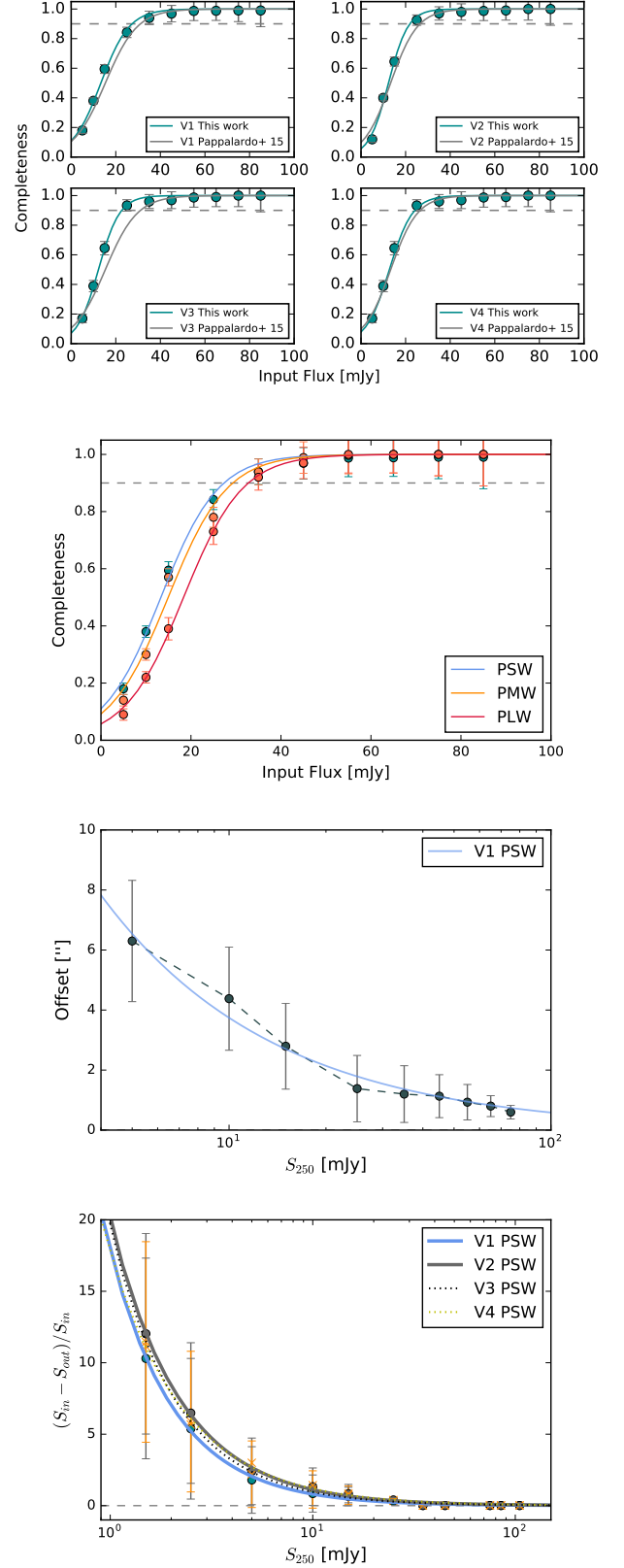


Fig. 4: For the HeViCS fields: Completeness at PSW band compared to Pappalardo et al. 2015, SPIRE completeness for the V1 field, positional accuracy (average radial offset) and PSW flux accuracy are shown from up to bottom.

The HeViCS field is significantly crowded and we can perform mismatch with an unrelated source close to the input

(x, y) coordinate. We match input to output catalogue considering source's distances smaller than half of the PSW beam size ($9''$), measured from the centre of PSW detection. Quality assessment results are summarized in Table 2 and Fig. 4. 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 $S_{250} > 13.2$ mJy have offset lower than $6''$, which is the value smaller than a PSW pixel size. Completeness and flux accuracy are consistent over all four HeViCS fields. The top panel in Fig. 4 shows our completeness result compared to one obtained by Pappalardo et al. (2015) confirming that our results do not change notably due to different methods of flux estimation.

With measured numbers in hand, we can adjust the detection cut to see variations in completeness, which is connected to function of input flux density. Our completeness at 250-micron maps shows the trend of fast decrease below the ~ 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 the saturation at the level of PSW fluxes larger than 40 mJy. For the PLW band, our completeness is above 80% at $S_{500} > 30$ mJy. This 500 μm flux we imply as the threshold for the final selection of "500 μm -risers". The bottom panel in Fig. 4 shows that fluxes of sources below $S_{250} = 7 - 10$ mJy are systematically overestimated because of "flux-boosting".

Reliability test is performed to quantify eventual spurious detections. This is important value for measuring the total number counts rather than for selection of "500 μm -risers", since we expect that eventually spurious source would be present in one waveband but not in all three SPIRE bands. However, since our prior list is based on as much as close to the confusion limit PSW detections, we make an analysis of possible false detections 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 PSW band (13.2 mJy) we found the low sample contamination of 2.7%.

3. Expected L_{IR} and z distribution of "500 μm -risers"

Without a known interferometric positions and confirmed spectroscopic redshifts of our sources, we are limited to the properties of FIR SEDs of selected "500 μm -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 should consider a partial degeneracy between $\beta - T_d$, and a perfect $T_d - z$ degeneracy (Blain et al. 2003, Pope & Chary 2010). The peak of the SED is determined by the ν/T_d term (see Eq. 2), giving 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 μm -risers" (see $T_d - z$ degeneracy illustrated in Fig. 5).

We fit a single temperature MBB model (Blain et al. 2003) to the data. We fix the power-law slope $\beta = 1.8$ and dust temperature $T_d = 38$ K, and we determine the 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. Because of aforementioned degeneracy in $(1+z)/T_d$ space, decreasing the dust temperature, e.g. from 38 K to 30 K for the fixed β , peak of the redshift distribution decreases to 3.57. Thus, we should expect that some fraction of colour-selected "500 μm -risers" reside at $z < 4$.

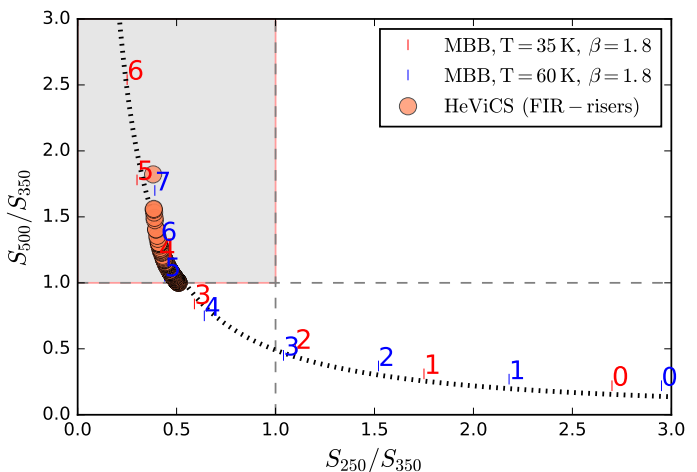


Fig. 5: SPIRE colour-colour diagram related to MBB models. SPIRE colours are plotted against possible redshifts inferred from the MBB model with fixed $\beta = 1.8$ and dust temperatures ($T_d = 35$ K (red) and $T_d = 60$ K (blue)). Combinations of redshifts are coloured in red and blue, showing the well known temperature-redshift degeneracy. Area related to "500 μm -risers" is shadowed in grey.

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

Sample of dusty galaxies at $z > 4$ are expected to be a combination of intrinsically luminous starbursts and strongly lensed systems (Béthermin et al. 2012, Casey et al. 2014). It has been shown that lack of observational constraints caused some SEDs to express too 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 comparing to their local analogues (e.g. Smith et al. 2014, Symeonidis et al. 2013). To overcome these differences we use empirical template representative for H -ATLAS galaxies (Pearson et al. 2013). The template is built for 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 it is already used in literature as a statistical framework for characterizing redshifts of sources selected from SPIRE observations (see Ivison et al. 2016). With different templates in hand we can better characterize the systematics and uncertainties of measured redshifts. Shifting these templates to fit with our FIR fluxes, we forecast redshift ranges of "500 μm -risers" in the HeViCS field. Probability distribution function (PDF) is built based on fitted χ^2 values and it allows us to derive an estimate of the median redshift. Median redshift of "500 μm -risers" is 3.87 if we apply "Cosmic Eyelash" template, 4.14 for Arp220 and 4.68 for M82. Next, we apply

H-ATLAS empirical template accepting the following best-fit parameters: $T_{\text{cold}} = 23.9$ K, $T_{\text{hot}} = 46.9$ K, and the ratio between the masses of cold and warm dust of 30.1, as same as in Pearson et al. (2013). The result is shown in Fig. 6. We determine a median redshift $\hat{z} = 4.28$, and an interquartile range of 4.08-4.75. This estimate matches 1-sigma uncertainty range of photometric redshifts calculated by Dowell et al. (2014) and Ivison et al. (2016). For red sources selected from several HerMES fields, Dowell et al. (2014) determined the photometric redshift of $\langle z \rangle = 4.9 \pm 0.9$.

To compute the IR luminosity, which is defined as the integral over the rest frame spectrum between $8 \mu\text{m}$ to $1000 \mu\text{m}$, we rely on the same approach. Our results favour the presence of very luminous systems with no regard of chosen template - in 90% of cases selected "500 μm -risers" have $L_{\text{IR}} \geq 10^{13} L_{\odot}$, with minimum and maximum values $L_{\text{IR}} = 8.1 \times 10^{12} L_{\odot}$ and $L_{\text{IR}} = 5.1 \times 10^{13} L_{\odot}$ respectively. The similar statistics is obtained if we use MBB model to fit our data. Concerning the MBB model, it has been shown that fixing β in the range of 1.5-2.0 introduces little changes in derived L_{IR} (Casey et al. 2012). Furthermore, applying the *H*-ATLAS representative template we 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 modeling of the sources with known photometric or spectroscopic redshifts, referring that "500 μm -risers" have in 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 $L_{\text{IR}} = 5.8 \times 10^{13} L_{\odot}$.

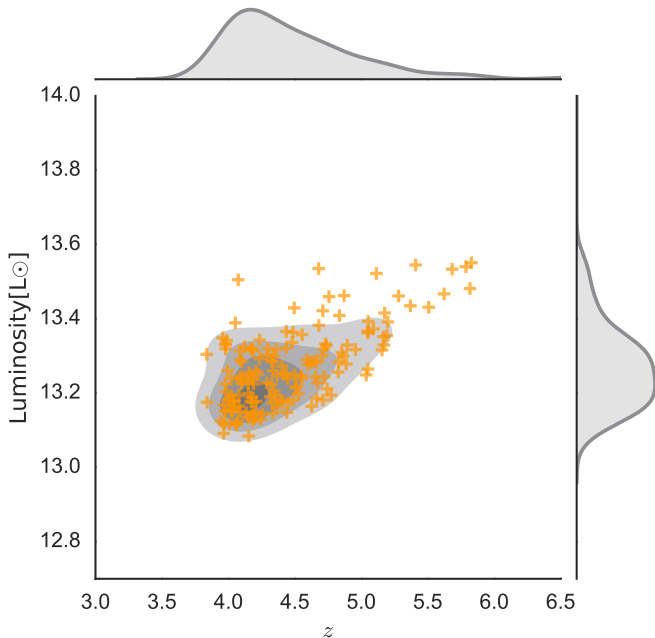


Fig. 6: Infrared luminosity of our "500 μm -risers" (y-axis) as a function of redshift. Orange crosses are values estimated applying the empirical template made from *H*-ATLAS galaxies covering the redshift range from $z = 0.5$ to $z = 4.5$ (Pearson et al. 2013). We imposed the same fitting parameters as used for *H*-ATLAS galaxies: $T_{\text{cold}} = 23.9$ K, $T_{\text{hot}} = 46.9$ K, and the cold-to-hot dust mass ratio equal to 30.1.

In Fig. 6 we illustrate the L_{IR} -redshift trend of our "500 μm -risers". We see that expected redshift distribution is not fully uniform, expressing the heavy-tail towards the higher redshifts. It indicates that very red submm galaxies are present at redshifts above the interquartile range, but it could also be due to strong gravitational lensing or unresolved blending. We should account with the possibility to have a much wider range of $L_{\text{IR}} - z$ values under larger uncertainties in dust temperatures. The IR luminosity emitted by a MBB source at a given T_d depends on the emissivity and size of 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 effects of strong lensing in our selection (see Section 5.1.5 and Section 6.2). We should emphasize here that the proper "training" of selected templates requires addition of higher-resolution data, and just in that case we can estimate systematic uncertainties and reject unreliable templates.

4. Differential number counts

We measure the raw 500 μm differential number counts and test our selection in respect to previous studies (this section), but also models of galaxy evolution (Section 5). We determine our counts placing 133 "500 μm -risers" in seven logarithmic flux bins between 30 mJy and 100 mJy. The resulting plot is shown in Fig. 7. The total raw number density of "500 μm -risers" in 55 deg² of HeViCS is $N = 2.41$ sources per deg², with the corresponding 1σ uncertainty of 0.34. Note that this value is uncorrected for non-red contaminants and completeness. In Section 5.2 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 Section 6.

In Fig. 7 we plot our differential number counts along with values from studies of Asboth et al. (2016) and Dowell et al. (2014). They applied so-called difference map method (*DMAP*) to select red sources. The method homogenizes a beams to the size of PLW, and creates the *DMAP* where SPIRE images are all smoothed to an identical angular resolution. Prior to source identification, 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 PLW resolution. To select "500 μm -risers", they performed red colour map-based search, starting from lowest resolution PLW maps.

From Fig. 7 it can be seen that differential number counts notably differ, especially in the lowest two flux bins where our measurements are in consistency 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 HeViCS curve, where a steeper decrease is present. For bins starting from $S_{500} > 50$ mJy, our counts are below the values measured by Dowell et al. (2014) and Asboth et al. (2016). Depending on a chosen flux bin discrepancy factor ranges from 3 to 9.

Apart from samples mentioned above, it is possible to find other "500 μm -risers" in the literature, but they are not selected in 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

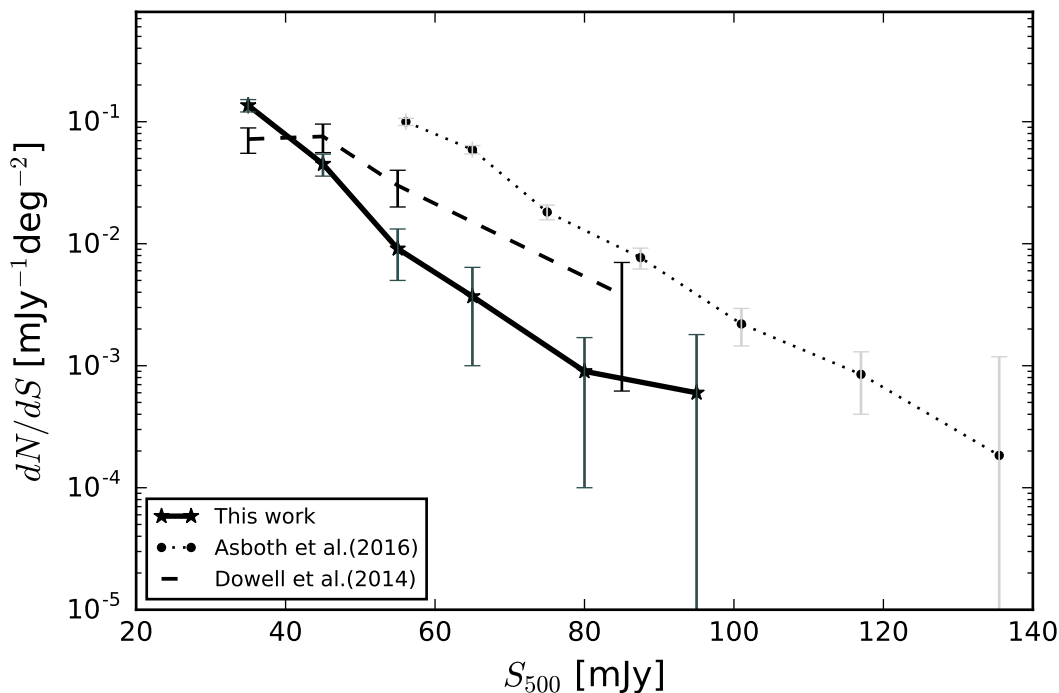


Fig. 7: Raw differential number counts of red sources from several studies. HeViCS number counts are presented with black stars, connected with a full, black line. Counts from [Asboth et al. \(2016\)](#) are presented with black circles and connected with dotted line, whilst observational findings of [Dowell et al. \(2014\)](#) are presented with dashed black line. 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.

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 predicts 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. \(2016a\)](#) 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. Summary of models and their main ingredients is given in Table 3. Common for 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 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 main sequence, expressing very high specific star formation rates ($\text{sSFR} = \text{SFR}/M_\star$). Shape of the SEDs is controlled by the galaxy type (MS or SB) and mean intensity of the radiation field $\langle U \rangle$, which couples with the temperature of dust. Differences between the models are described briefly in following subsections - the most important are scatter from the main sequence, parameters chosen to fit stellar masses, and redshift evolution of a dust temperature. Models presented here have also slightly different description of mergers.

5.1.1. Béthermin et al. (2012) model

The B12 model describes the stellar mass function (SMF) with a Schechter function with redshift invariant characteristic mass parameter (see eq. 4 in [Béthermin et al. 2012](#), but also [Peng et al. 2010](#)). IR SED template are based on [Draine & Li \(2007\)](#)

Table 3: Comparison of models used in our analysis.

	Models		
	B12 Béthermin et al. (2012)	B17 Bethérmin et al. (2017)	S16 Schreiber et al. (2016a)
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

(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. Later value is redshift invariant in Béthermin et al. (2012) model, and evolves with redshift in other two presented models. For redshifts $z > 4$, model of Bethérmin et al. (2017) assumes single Schechter function fixing the $\Phi_1^*=0$, while model of Schreiber et al. (2016a) 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) sSFR increases with redshift up to $z = 2.5$ and then flattens. This trend is independent of chosen range of stellar masses. In later two models the evolution is different, see Eq. 6 in Bethérmin et al. (2017); (3) Modelled width of the main-sequence (as log-normal scatter).

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 submm counts ($\text{PLW} \gtrsim 100$ mJy). The AGN contribution is statistically associated based on results from Aird et al. (2012) and Mullaney et al. (2011).

5.1.2. Bethérmin 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 submm number counts due to the source blending (Karim et al. 2013, Simpson et al. 2015). The model accounts for 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 $\langle U \rangle$. Evolution of dust temperature of a main sequence SF 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 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. (2016a) 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 ~ 5 . The distribution of stellar masses is described by double power-law Schechter fit, with parameters evolving with redshift. These parameters are chosen accordingly 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 Galilano et al. (2011). Redshift evolution of a dust temperature is modelled as

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

The "starburstiness" term R_{SB} is used to quantify the SFR offset between MS and SB galaxies ($\text{R}_{\text{SB}} = \text{SFR}/\text{SFR}_{\text{MS}}$). Dependence of sSFR of 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 summarize, in S16 both sSFR and T_d evolve continuously with redshift, which is not the case

for other two models (see Table 3). 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 previous section. Our goal is to compare the number counts of observed "500 μm -risers" to the ones predicted by models.³ 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 (B  thermin 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, B  thermin et al. 2017). The size of simulated area is equal to the size of HeLMS field, thus perfectly suited for a comparison of our selection technique to one performed by Asboth et al. (2016). With 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) on modelled sources. Left panel in Fig. 8 shows observed counts of HeViCS-selected "500 μm -risers" plotted against the model predictions. We see that observed counts are in a fairly good agreement with models, while corresponding power-law slope show no significant difference from the slopes anticipated by models. Observed values are steep at the fainter-end and flatter at the bright-end ($S_{500} > 80 - 90$ mJy) which is due to flux magnification by gravitational lensing. Our HeViCS differential number counts have the perfect agreement with B12, but we have to note there are recent evidences (see B  thermin et al. 2017 for more details) that the simulated catalogue based on the B12 model overpredicts the number counts at PLW fluxes below 50 mJy by a factor of 2-3. Predictions of B17 and S16 are consistent with the 1σ error bars of observed counts in the higher flux bins (60 mJy < S_{500} < 100 mJy). In lower flux bins (30 mJy < S_{500} < 60 mJy) discrepancy factors are between 1 and 6 depending on a chosen bin. In Section 5.1.5 we discuss in more details potential reasons of a present difference. 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.

The number of sources satisfying our selection in B17 and S16 are 0.45 deg^{-2} and 0.40 deg^{-2} respectively - interestingly, contribution of weakly lensed and unlensed populations ($\mu < 2$) ranges from 49% (based on B17) to 100% (based on S16). It confirms that HeViCS selection of "500 μm -risers" is not biased towards the strong magnifications.⁴ Both B17 and S16 agree that all galaxies in the lowest flux bin (30-40 mJy) that match our "500 μm -riser" criteria are unlensed. It is very important indication, consistent with that selecting the galaxies by their red colours down to $S_{500} \simeq 30$ mJy and $S_{250} \simeq 13.2$ mJy is the way to detect the population of unlensed "500 μm -risers". These are potentially some of the most intrinsically bright and most massive dusty starburst galaxies.

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

⁴ In B17 star formation rate (SFR) is limited to a maximum 1000 $\text{M}_{\odot}\text{yr}^{-1}$ which implies that all red sources above the certain PLW flux cut ($S_{500} > 60$ mJy) must be strongly lensed.

5.1.5. Effects of noise on measured counts

The results described in Section 5.1.4 neglect the effect of noise on the number counts of "500 μm -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 (Section 2.1). The comparison between observations and models with simulated Gaussian noise are illustrated in the right panel in Fig. 8. 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. 8).

Considering the addition of noise, the number of simulated sources that appear as "500 μm -risers" in B17 jumps to 473, resulting to 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 - the 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 model even in the faintest PLW flux regime. We find that noise often tends to increase the number of gen-

Table 4: Comparison of modelled number counts before and after adding the Gaussian noise. Counts are calculated by imposing our criteria to select "500 μm -risers".

	Predicted density of 500 μm -risers ⁽¹⁾	Faint 500 μm -risers ⁽²⁾ ($S_{500} < 30$ mJy)
B17	0.45 deg^{-2} (51% lensed)	-
B17 + noise	1.73 deg^{-2} (26% lensed)	52 %
S16	0.4 deg^{-2}	-
S16 +noise	1.54 deg^{-2}	67 %

(1) Modelled density of 500 μm -risers before and after the addition of random Gaussian noise. Criteria used to select "500 μm -risers" are: $S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2$ mJy and $S_{500} > 30$ 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 μm -risers" selection criteria after the addition of modelled Gaussian noise.

uinely red sources that are slightly below our PLW flux cut. These galaxies have modelled flux densities $S_{500} < 30$ mJy, but they can pass our final "500 μm -risers" selection after the addition of a Gaussian noise. As shown in Table 4 contribution of these sources to the modelled "500 μm -risers" vary between 52-67%.

The effect of noise also changes the relative contribution of modelled lensed and unlensed "500 μm -risers". The percentage of weakly lensed and non-lensed "risers" in B17 increases from 49% (modelled values) to 71% (modelled values+Gaussian noise). A closer inspection suggests 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 μm -risers" criteria more often if they are on a positive fluctuations of the magnification. This has an important role in producing observed "500 μm -risers" without a support of strong lensing.

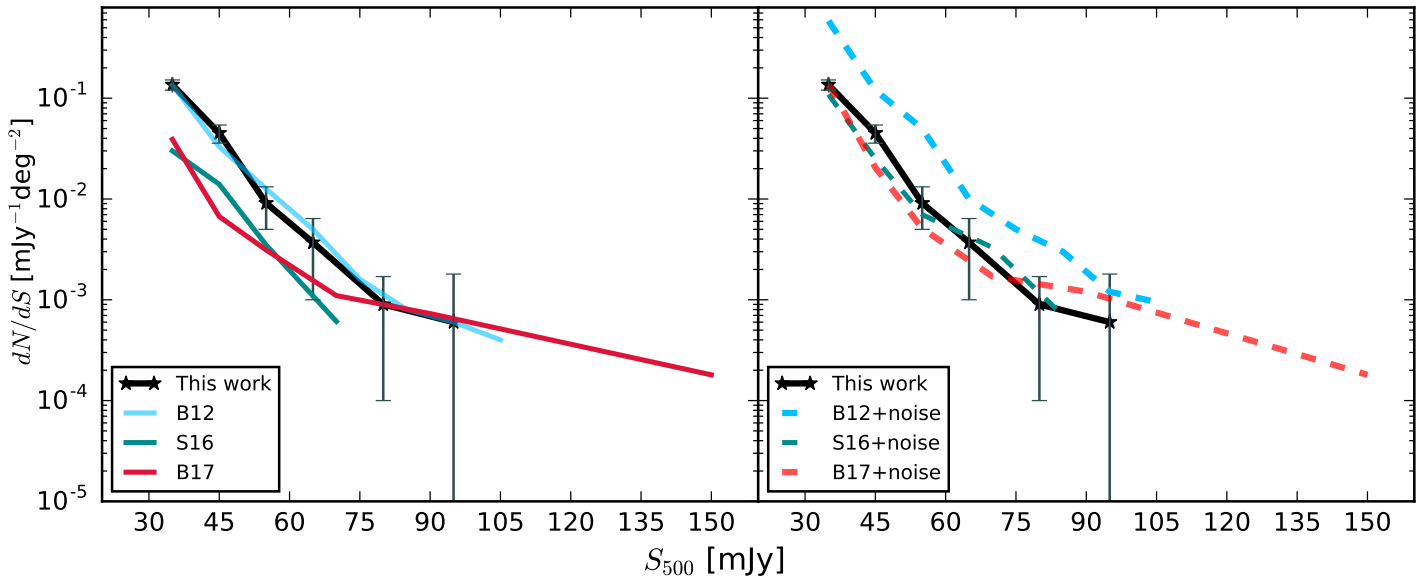


Fig. 8: Observed differential number counts in the HeViCS field (black line) confronted to models. Expected values from models are overplotted as coloured lines. *Left:* Comparison between observed and modelled counts. Models are represented by full, coloured lines: cyan for B12 (B  thermin et al. 2012), red for B17 (Beth  rmin et al. 2017) and green for S16 (Schreiber et al. 2016a). 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.

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 about the purity of detected sources. This is a crucial step if we want to investigate our photometric pipeline of "500 μm -risers", since it is necessary to count with the numbers of sources we potentially miss and to determine fraction of contaminants.

We generate a set of simulated SPIRE maps with use of Empirical Galaxy Generator Code (EGG, Schreiber et al. 2016b). We filled our mock maps with the sources that has been drawn from S16 and B17 mock catalogues. We convert 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 with use 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 simulation described in B  thermin et al. (2017). It adopts sophisticated clustering model with the subhalo abundance matching procedure used to populate the dark-matter halos of a light cone constructed from the Bolshoi-Planck simulation (Rodr  guez-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 deg² each). As shown in Section 2.4, our source extraction lead to an homogeneous distribution of sources over all HeViCS fields, so we can easily extrapolate our results for the whole observed area of

55 deg². 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 summarized and presented in Fig. 9. We detected all modelled "500 μm -risers" via our detection procedure, but not all of them are characterised with their intrinsic colours. We thus disclose three groups of sources described below.

1. Recovered "500 μm -risers"

Dark orange region in Fig. 9 represents the fraction of recovered "500 μm -risers". These are intrinsically red sources selected as "500 μm -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 μm -risers" is around 60% for both models (57% for S16 and 61% for B17).

As expected from the mock catalogue analysis (Section 5.1.4), population of recovered risers is a mix of both lensed and unlensed galaxies. We confirm that our criteria to select "500 μm -risers" unveil significant group of dusty and potentially very high- z sources that are not biased toward higher magnifications.

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

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

where σ_{PMW} and σ_{PLW} are 1-sigma total noise ratios measured for appropriate bands in the HeViCS field (see Section 2.1). The differences between intrinsic and observed

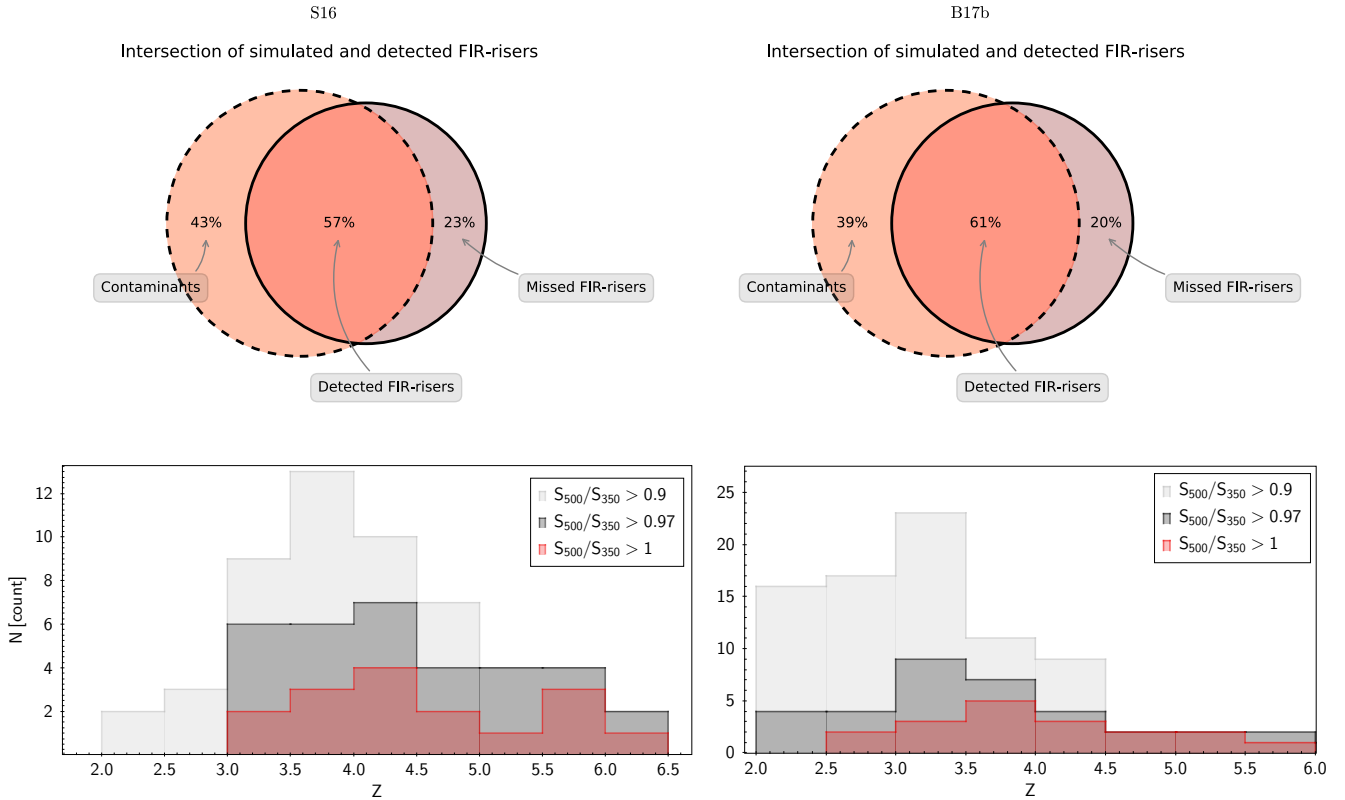


Fig. 9: Left and right panel attribute to S16 and B17 respectively. *Upper panels:* Galaxies detected in mock maps. "FIR-riser" criteria are imposed as for real HeViCS maps ($S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2$ mJy, $S_{500} > 30$ mJy). Intersected area coloured in dark orange depicts recovered 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 μ m-risers". *Lower panels:* Redshift distribution of modelled galaxies. Different S_{500}/S_{350} colour cuts are imposed. Those cuts are related to statistical properties of the confusion noise, which are the most responsible for colour uncertainties. (see also Table 5).

colours of recovered 500 μ m-risers are presented in Table 5. We see that the largest difference is for the lowest flux bin, where we also expect largest flux uncertainties due to the lower S/N ratio. Redshift distribution of recovered "500 μ m-risers" in Fig.9 reflects differences between models. Following S16, all red galaxies are at $z > 3$, and most of them (70%) 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

Light-orange region in Fig.9 represents a fraction of contaminants. Sources are detected as "500 μ 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 one are genuinely red, but fainter galaxies, with modelled PLW 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 μ 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 between $0.95 < S_{500}/S_{350} < 1.0$. 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 PLW excess is caused if clustered neighbours are at $z = 1.3 \pm 0.1$. These sources are massive galaxies with $M_* > 10^{9.8} M_\odot$. It confirms the conclusion from Section 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 μ m-risers"

With our extraction procedure we detected all modelled "500 μ m-risers" from simulated maps. However there is a fraction of sources for which we failed to properly characterize their colours with MBB-fitter. Violet region in Fig.9 represents sources having $S_{500}/S_{350} < 1$ or $S_{500} < 30$ mJy. In this way we missed to select around 20% of genuine "500 μ m-risers". According to models, missed "500 μ m-risers"

"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 PSW beam. Therefore, missed sources passes both PSW and PLW flux selection cuts but their observed colours are non-red. We quantify that the difference between observed and modelled colours of missed "risers" is shifted bluewards by a factor of ~ 0.03 , thus giving negative Δ . Interestingly, not just observed, but the modelled Δ of missed sources exhibits negligible change due to noise effects. It tends 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 (see Section 5.2.1).

Table 5: Intrinsic vs. observed colours of simulated "500 μm -risers". Catalogue and maps are based on B17.

Flux bin	$S_{500}/S_{350}^{(1)}$ [B17 _{cat}]	$S_{500}/S_{350}^{(2)}$ [B17 _{map}]	$\Delta^{(3)}$
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)

(1) Catalogued colours from B17. Noise is not added to initial flux values; (2) Measured colours in simulated maps.; (3) 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.

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

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 purity of almost 90 per cent. 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 "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 Section 5.1) that most of unlensed "500 μm -risers" have S_{500}/S_{350} very close to one. Additionally, S16 and B17 predict large number of "almost red" FIR sources (e.g. $0.9 < S_{500}/S_{350} < 1.0$ and $S_{250} > 13.2$ mJy). The modelled number density of such defined sources is between 8-10 per deg². 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 5), 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

Table 6: Relation of different colour cuts to the number of observed/missed red sources and their redshift distributions. Sources observed in simulated SPIRE maps are divided in three columns according to different colour limits ($S_{500}/S_{350} > 0.9$, $S_{500}/S_{350} > 0.97$, $S_{500}/S_{350} > 1.0$). In each of three cells left column (m) substitute to percentage of missed, genuine "500 μm -risers", while the right column represents the percentage of $z > 4$ sources contained in selected populations.

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

of such a test is twofold: to find the colour threshold that increases the number of true (modelled) "500 μm -risers", but introducing in the same time low contamination of $2 < z < 4$ sources. We quantify what is the contribution of these lower- z objects for each colour limit. The result of the analysis is summarized in Table 6 and histograms in the lower panel of Fig. 9. Having imposed $S_{500}/S_{350} > 0.9$ we anticipate that almost all "500 μm -risers" will be found independently of chosen models. Nevertheless, contribution of $2 < z < 4$ sources is significant, and vary 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 $S_{500}/S_{350} > 0.97$, percentage of detected simulated intrinsic "500 μm -risers" ranges between the 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. In the same time, it works well against the lower- z contaminants, especially those at $z < 3$, limiting their contribution to a maximum 15%. Recent follow-up observations of red sources selected in H-ATLAS field (Ivov et al. 2016) also support our quest for modifying "500 μm -risers" colour criteria. They report just 30% of sources selected as "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 traditional threshold $S_{500}/S_{350} > 1$ due to easier comparison to 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 instrument 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, Busmann et al. 2015). This assumption includes strongly lensed sources as well, since they are known to be very bright in SPIRE bands. More recent SPIRE multiplicity study is done by Scudder et al. (2016). They used improved version of XID+ code (Hurley et al. 2017) to search for many potential contributing sources per 250 μm detection. They have considered dusty galaxies covering a wide range of photometric redshifts and 250 μm flux densities. For PSW fluxes between 30 mJy and 45 mJy they found that combination of the two brightest components accounts for approximately 90% of the flux in the 250 μm source. Additionally, the brightest component contributes with more than 60% to 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 "500 μm -riser" in our mock maps, we assign the brightest component in the radius no larger than the half of PSW 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. 10. For galaxies detected as "500 μm -risers", we found that 72% to 85% of observed PSW flux is emitted by the brightest galaxy (78% in average).

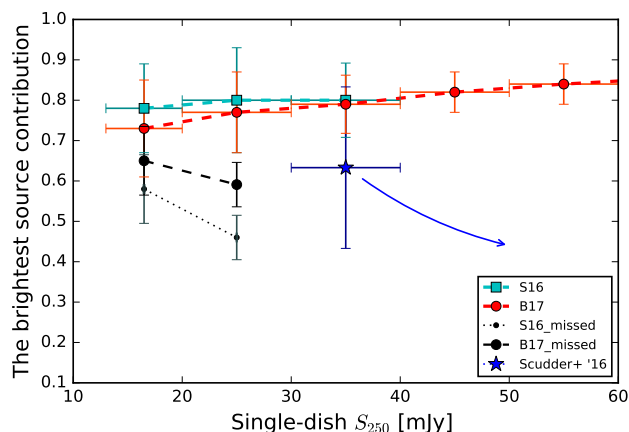


Fig. 10: Multiplicity of red sources detected in mock maps. Average fraction of the PSW flux density emitted by the brightest galaxy in the *Herschel* beam is plotted as a function of total, single-dish 250 μm flux measured in our simulated maps. Horizontal error bars indicate the width of the chosen flux bin, while vertical error bars show the standard deviation of the distribution. Blue star refers to the study of Scudder et al. (2016), and blue arrow indicates observed trend they found towards brighter fluxes.

Overall, we see almost insignificant change for S16 and B17, with the trend expressing slight increase towards higher PSW fluxes. Repeating the same analysis for PLW beam we find the similar trend with somewhat lower brightest galaxy fraction (64 % in 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-micron detections does not experience strong resolution effects. Situation is somewhat different for missed red sources - the contribution of brightest components is lower (47-65%) with the tendency of decreasing towards larger PSW fluxes. This is due to a presence of "blue" blends close to position of "500 μm -risers" ($r < 9''$).

Fraction of the brightest PSW component to the total flux measured in simulated maps is higher than one found in Scudder et al. (2016). They use deep multiwavelength data and apply different definition of the flux density fraction. Additionally, their analysis consider sources regardless of SPIRE colours making it difficult for a direct comparison here. Our multiplicity predictions appear to be in a good agreement with the measurements of Greenslade et al. (in prep.). They performed Submillimeter Array (SMA) follow up program for 36 of "500 μm -risers" selected from various fields and found the multiplicity rate of 33% with brightest components contributing 50 – 75% to SMA flux. To shed a new light on this issue, the contamination of sources at intermediate redshifts ($1 < z < 2$) seen in deep NGVS and PACS images will be discussing in our following paper.

6.2. Lensing and clustering

Simulations presented in Section 5.2 have some limitations meaning that they are affected by cosmic variance beyond simple Poissonian, thus containing under- or overdensities at some redshifts. On top of that, pure probabilistic treatment is used to model the lensing, lacking the physical connection to the overdensity of lower- z systems. Here we examine impact of clustering, lensing and environments considering archival data and studies found in literature.

Strongly lensed galaxies are expected to dominate the very bright end of number counts of selected "500 μm -risers". In literature limiting number of strongly lensed "500 μm -risers" has been serendipitously found by different selection methods (Negrello et al. 2017, Spilker et al. 2016, Nayyeri et al. 2016, Cañameras et al. 2015, Wardlow et al. 2014, Negrello et al. 2007). Our HeViCS sample of "500 μm -risers" has the number density of 0.018 deg^{-2} of very bright ($S_{500} \sim 100$ mJy), potentially strongly lensed sources. This is in consistency with the study of Negrello et al. (2017) who catalogued 7 strongly lensed "500 μm -risers" over 600 deg^{-2} in the *H*-ATLAS field (0.012 deg^{-2}). To further cross-check the statistical effect of lensing on the selection of "500 μm -risers", we look at the density of WISE (Wide-field Infrared Survey Explorer, Wright et al. 2010) sources around our "500 μm -risers. Here we assume that denser foreground concentrations lead to higher lensing amplifications. We find that 50% of HeViCS selected "500 μm -risers" have at least one 5σ detected WISE source (W1 band, 5σ) inside the PSW beam. We probe several apertures ranging from 3.6" to 36" (in step of 5")⁵, and measure no significant overdensity of WISE detections near the positions of our 133 targets. It might indicate

⁵ While the WISE astrometric uncertainty is negligible, we assume the astrometric uncertainty is due to PSW S/N ratio given by $\sigma_{\text{pos}} = 0.6 \times \frac{\text{FWHM}}{S/N}$.

that our final sample of "500 μm -risers" in HeViCS is marginally influenced by strongly lensed sources.⁶

Apart from strong lensing, clustering might be responsible in producing observed excess in number counts of "500 μm -risers". Several studies of clustering of submm-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 been shown they are hosted in moderately strongly clustered massive halos. *Herschel* detected star-forming galaxies are in average more strongly clustered at higher-redshifts ($z \sim 2$) than nearby objects (B  thermin et al. 2014, Ono et al. 2014). Since the beam size in *Herschel* is a direct function of the wavelength, we expect largest SPIRE beam (PLW) to be more affected than other two beams.

In addition, various studies have found the strong environmental impact on colours of dusty, high- z galaxies (Umehata et al. 2017, Smol     et al. 2017, Kato et al. 2016, Dannerbauer et al. 2017, Clements et al. 2014, Hodge et al. 2013). It is suggested that protocluster galaxies might be the dominant population at very high redshifts, since their expected contribution to the cosmic star-formation rate density (CSFRD) rises (Chiang et al. 2017). Furthermore, Smol     et al. (2017) found a significantly higher presence of DSFGs occupying overdense environments at $z > 3$ than at lower redshifts. To make a brief statistical check of environmental influence on our "500 μm -risers" selection, we examine *Planck* High- z Candidate Catalogue (PHZ, Planck Collaboration et al. 2016b). Given the large beam size (FWHM=4.5') and the high confusion noise (> 100 mJy at 857 GHz), *Planck*-identified sources are likely highly clustered objects. This assumption is confirmed by several follow-up observations (e.g. MacKenzie et al. 2017, Clements et al. 2016). They have revealed that majority of PHZ sources appear as overdensities of DSFGs, having peaks in their SEDs around 350 μm . We cross-match PHZ catalogue with our list of "500 μm -risers" and find that 8 "500 μm -risers" are hosted by PHZ objects. For 5 of 8 cross-matched sources we find they have, in average, more "almost red" galaxies ($S_{350} > S_{500}$) inside the *Planck* beam than randomly chosen 133 galaxies from the field. We run our test 1000 times to confirm the significance. If these overdensities are not just part of random fluctuations in the galaxy number density within the *Planck* beam, it makes them suitable candidate protoclusters of high- z DSFGs. More detailed (proto)clustering analysis is out of scope of a present paper, and we consider it for our following study.

To improve our understanding of how lensing and clustering might shape "500 μm -risers" 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 confusion, environmental effects and refined treatment of a lens modelling.

6.3. Star-formation rate density

The bright individually-detected "500 μm -risers" play an important role in understanding the evolution of massive systems. Here we test if such a selection is suitable to measure the total star formation rate density (SFRD). We make a rough estimate of the SFRD assuming the statistical properties of selected "500 μm -risers". We determine SFRD at $4 < z < 5$ applying the equation

(Vincoletto 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. Bottom number (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):

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

We use redshift distribution from models (e.g. S16, see Section 5.) and adopt the S16 luminosity distribution of "500 μm -risers" selected with our criteria. We further apply the result from our simulations to correct for the number of missed and contaminant sources. 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 $H_0 = 67.8$ km/s/Mpc (Planck Collaboration et al. 2016a), we find that $\text{SFRD} = 2.3 \times 10^{-4} M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$, for "500 μm -risers" at $4 < z < 5$. This result is shown with a filled, red star in Fig.11. 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 constrains with SPIRE data only. We additionally cross-check this result and make the same estimation of SFRD but adopting results from observed $L_{\text{IR}} - z$ distribution (see Section 3.). Assigning the median IR luminosity ($1.94 \times 10^{13} L_\odot$), we find $\text{SFRD} = 3.3 \times 10^{-4} M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$. Our estimation is also consistent with recent observational findings from Duivenvoorden et al. (MNRAS, submitted). Note also that offset between the study of Dowell et al. (2014) and our own may be explained by the fact that estimation by Dowell et al. (2014) assumed substantially higher IR luminosities and purity ratios.

Furthermore, we can use the shape of luminosity function to correct for the fainter red sources. We extrapolate the contribution of "500 μm -risers" to SFRD - firstly we remove PLW flux limit (thus accounting for a red sources fainter than $S_{500} < 30$ mJy), and then we remove lower PSW flux cut, thus accounting for all "500 μm -risers" below the *Herschel* sensitivity limit. Extrapolated contributions are depicted by empty red triangles in Fig.11).

Here we assume the luminosity distribution from S16, but we should mention that very similar result (lower by a factor of 1.4) we obtain applying the luminosity function from B17. As expected, the largest portion of sources responsible for a stellar budget at $z > 4$ are missed due to sensitivity limitations of *Herschel* instruments and not because of colour selection.

The contribution of the emission of dusty star-forming galaxies (DSFGs) to the total infrared (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 about whether or not ultraviolet (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 PLW *Herschel* selections of "500 μm -risers". They report no decline in dust-obscured SFRD between $z = 3 - 6$, and very high SFRD value at $4 < z < 5$ order of magnitude higher than our estimation. However, our result is in a much better agreement with studies of Bourne et al.

⁶ To clarify this statement, high-resolution submm follow-ups are needed. For the sample from Ivison et al. (2016), recent ALMA follow-ups (Fudamoto et al. 2017) shown that the fraction of strongly lensed "risers" is higher than expected.

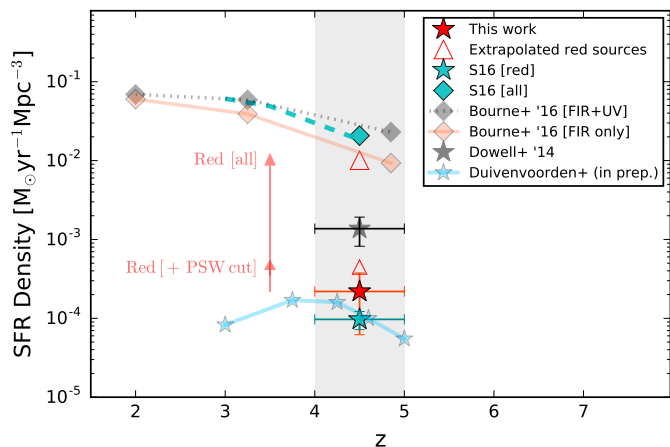


Fig. 11: SFRD as a function of redshift. Filled red star represents our measurement. The estimated SFRDs of extrapolated "500 μ m-risers" are shown with empty red-edged triangles: smaller when PSW cut is imposed ($S_{250} > 13.2$ mJy), and larger triangle if all flux cuts are removed. Flux corrections are annotated by arrows. We show for comparison results from other "500 μ m-risers" selections: S16 (cyan, [Dowell et al. \(2014\)](#)) (black) and Duivenvoorden et al. (MNRAS, submitted) (blue). Horizontal bars reflect the bin size. FIR and total FIR+UV observed SFRD ([Bourne et al. 2017](#)), as well as total simulated SFRD (S16), are marked with diamonds.

(2016) (marked with diamonds in Fig.11) and [Koprowski et al. \(2017\)](#) 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 onto 55 deg² area. We introduce an innovative method to select "500 μ m-risers" ($S_{500} > S_{350} > S_{250}$). In order to estimate the 250-500 μ m flux of blended sources, we use the multi-wavelength MBB-fitter code. We assign positions of galaxies detected in the 250 μ m map as a prior to extract the flux densities at longer wavelengths, iteratively fitting their SEDs as modified blackbodies.

- We select 133 "500 μ m-risers" which fulfil following criteria: $S_{500} > S_{350} > S_{250}$, $S_{250} > 13.2$ mJy and $S_{500} > 30$ 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 $\hat{z} = 4.28$ and the corresponding median rest-frame IR luminosity $\hat{L}_{\text{IR}} = 1.94 \times 10^{13} L_{\odot}$. All our "500 μ m-risers" are thus classified as high- z "hyper-luminous" dusty galaxies (so-called "HyLIRGs").
- The total raw number density of selected sources is 2.41 ± 0.34 deg⁻². To evaluate our results, we retrieve mock catalogues based on models of [Béthermin et al. \(2012\)](#), [Béthermin et al. \(2017\)](#) and [Schreiber et al. \(2016a\)](#). Differential number counts are fairly consistent with model predictions. Namely, towards the brighter end ($S_{500} > 70$ mJy), HeViCS counts have both slope and values identical to ones anticipated by models. At the lowest flux end, discrepancy is effect of noise and weak lensing. We find that noise tends to increase number counts of "500 μ m-risers". It may be high

enough to fully match with observations within 1-sigma uncertainty. We confirm that selecting "500 μ m-risers" applying our selection criteria is a direct way to detect 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 an modified colour criterion ($S_{500}/S_{350} > 0.97$) that should be probed in the future selections of potentially $z > 4$ dusty sources. Motivation to apply this criterion is twofold: (1) it would account for colour uncertainties that arise from effects of noise, clustering and weak lensing; (2) it increases the sample of candidate $z > 4$ galaxies. Anticipated contribution of $2 < z < 3$ contaminants satisfying this colour cut is small, and reach the maximum of 13%.
- We inspect the effect of multiplicity working with simulated SPIRE data. The brightest galaxy inside the beam contributes in average 75% and 64% to the total single-dish flux measured at 250 μ m and 500 μ m respectively.
- We use statistical properties of our galaxies to determine the role of "500 μ m-risers" to the SFRD for $4 < z < 5$. We show that after correcting for fainter sources, projected contribution of "500 μ m-risers" is factor of 2-3 below the total, dust-corrected SFRD.

Acknowledgements. We would like to acknowledge David Corre, Denis Burgarella, Alessandro Boselli and Tom Bakx 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). d'Avenir" 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, *Monthly Notices of the Royal Astronomical Society*, 462, 1889
 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., Daddi, E., Magdis, G., et al. 2012, *ApJ*, 757, L23
 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., Kilbinger, M., Daddi, E., et al. 2014, *A&A*, 567, A103
 Béthermin, M., Wu, H.-Y., Lagache, G., et al. 2017, *ArXiv e-prints*
 Blain, A. W., Barnard, V. E., & Chapman, S. C. 2003, *MNRAS*, 338, 733
 Bourne, N., Dunlop, J. S., Merlin, E., et al. 2017, *MNRAS*, 467, 1360

- Bourne, N., Dunne, L., Maddox, S. J., et al. 2016, *MNRAS*, 462, 1714
- 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
- Cañameras, R., Nesvadba, N. P. H., Guery, D., et al. 2015, *A&A*, 581, A105
- 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
- Chiang, Y.-K., Overzier, R. A., Gebhardt, K., & Henriques, B. 2017, *ApJ*, 844, L23
- Clements, D. L., Braglia, F., Petitpas, G., et al. 2016, *MNRAS*, 461, 1719
- Clements, D. L., Braglia, F. G., Hyde, A. K., et al. 2014, *MNRAS*, 439, 1193
- 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
- Dannerbauer, H., Lehnert, M. D., Emonts, B. H. C., et al. 2017, *ArXiv e-prints*
- 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, *ArXiv e-prints*
- Dowell, C. D., Conley, A., Glenn, J., et al. 2014, *The Astrophysical Journal*, 780, 75
- Draine, B. T. & Li, A. 2007, *ApJ*, 657, 810
- Eales, S. A., Smith, M. W. L., Wilson, C. D., et al. 2010, *A&A*, 518, L62
- 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
- Franceschini, A., Rodighiero, G., Vaccari, M., et al. 2010, *A&A*, 517, A74
- Fudamoto, Y., Ivison, R. J., Oteo, I., et al. 2017, *ArXiv e-prints*
- Galliano, F., Hony, S., Bernard, J.-P., et al. 2011, *A&A*, 536, A88
- Giovannelli, 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
- 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
- Hodge, J. A., Carilli, C. L., Walter, F., Daddi, E., & Riechers, D. 2013, *ApJ*, 776, 22
- Hogg, D. W. 1999, *ArXiv Astrophysics e-prints*
- Hurley, P. D., Oliver, S., Betancourt, M., et al. 2017, *MNRAS*, 464, 885
- 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
- Karim, A., Swinbank, A. M., Hodge, J. A., et al. 2013, *MNRAS*, 432, 2
- Kato, Y., Matsuda, Y., Smail, I., et al. 2016, *MNRAS*, 460, 3861
- Kennicutt, Jr., R. C. 1998, *ARA&A*, 36, 189
- Koprowski, M. P., Dunlop, J. S., Michałowski, M. J., et al. 2017, *ArXiv e-prints*
- 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. 2017, *ArXiv e-prints*
- Lutz, D. 2014, *ARA&A*, 52, 373
- Ma, Z. & Yan, H. 2015, *ApJ*, 811, 58
- MacKenzie, T. P., Scott, D., Bianconi, M., et al. 2017, *MNRAS*, 468, 4006
- 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
- Merlin, E., Bourne, N., Castellano, M., et al. 2016, *A&A*, 595, A97
- Michałowski, M. J., Dunlop, J. S., Cirasuolo, M., et al. 2012, *A&A*, 541, A85
- 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
- Nayyeri, H., Keele, M., Cooray, A., et al. 2016, *ApJ*, 823, 17
- Negrello, M., Amber, S., Amvrosiadis, A., et al. 2017, *MNRAS*, 465, 3558
- Negrello, M., Perrotta, F., González-Nuevo, J., et al. 2007, *MNRAS*, 377, 1557
- 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., Bock, J., Altieri, B., et al. 2012, *MNRAS*, 424, 1614
- Oliver, S. J., Wang, L., Smith, A. J., et al. 2010, *A&A*, 518, L21
- Ono, Y., Ouchi, M., Kuroi, Y., & Momose, R. 2014, *ApJ*, 795, 5
- 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, Ade, P. A. R., Aghanim, N., et al. 2016a, *A&A*, 594, A13
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, *A&A*, 596, A100
- Pope, A. & Chary, R.-R. 2010, *ApJ*, 715, L171
- 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., Ivison, R. J., Greve, T. R., et al. 2012, *MNRAS*, 419, 2758
- Roseboom, I. G., Oliver, S. J., Kunz, M., et al. 2010, *MNRAS*, 409, 48
- 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., Elbaz, D., Pannella, M., et al. 2016a, *A&A*, 589, A35
- Schreiber, C., Elbaz, D., Pannella, M., et al. 2016b, *ArXiv e-prints*
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, *A&A*, 575, A74
- Schreiber, C., Pannella, M., Leiton, R., et al. 2017, *A&A*, 599, A134
- Scudder, J. M., Oliver, S., Hurley, P. D., et al. 2016, *Monthly Notices of the Royal Astronomical Society*, 460, 1119
- Shu, X. W., Elbaz, D., Bourne, N., et al. 2016, *ApJS*, 222, 4
- 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
- Spilker, J. S., Marrone, D. P., Aravena, M., et al. 2016, *ApJ*, 826, 112
- Swinbank, A. M., Simpson, J. M., Smail, I., et al. 2014, *MNRAS*, 438, 1267
- Swinbank, A. M., Smail, I., Longmore, S., et al. 2010, *Nature*, 464, 733
- Symeonidis, M., Vaccari, M., Berta, S., et al. 2013, *MNRAS*, 431, 2317
- Umehata, H., Tamura, Y., Kohno, K., et al. 2017, *ApJ*, 835, 98
- 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
- Wardlow, J. L., Malhotra, S., Zheng, Z., et al. 2014, *ApJ*, 787, 9
- 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
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868
- 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

Appendix A: Dusty radio sources

As mentioned in Section 2.5, our final list of "500 μm -risers" is cleaned from 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 "protect" quiet well 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 χ^2 values lie outside the range we accept as good to our final selection ($\chi^2 < 3.84$, see Appendix B). Between 7 objects removed from the final list, 3 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. Source has a spectroscopic redshift of $z = 0.964$ and according to the classical PSF-fitting photometry performed with e.g. 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 4 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 - in average $z = 2.74$, with one sources lying at $z > 4$. To fully inspect relation of these objects to the population we aim 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 expecting radio emission of our final group of 133 "500 μm -risers" implementing FIR-radio correlation:

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

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 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) analyzed 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 to 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 might break down. It 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 sample of *Herschel* detected optically selected QSOs (e.g. Ma & Yan 2015). They, however, claimed 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, it could support scenario that their thermal SED is not AGN-driven. Thus, our subsample of QSO-like objects might be 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 ($10''$, $rms=10 \mu\text{Jy/beam}$) continuum maps.

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

Appendix B: Color-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 approach that selecting "500 μm -risers" is the way to select $z \geq 4$ candidates, we run 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 $z \geq 4$. Thus, we are particularly interested to unveil MBB-fitted sources having $S_{500}/S_{350} > 1$ and $S_{500}/S_{250} > 2$. Plotted in Fig. B.1 are 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).

Appendix C: An example of faint and deblended "500 μm -risers"

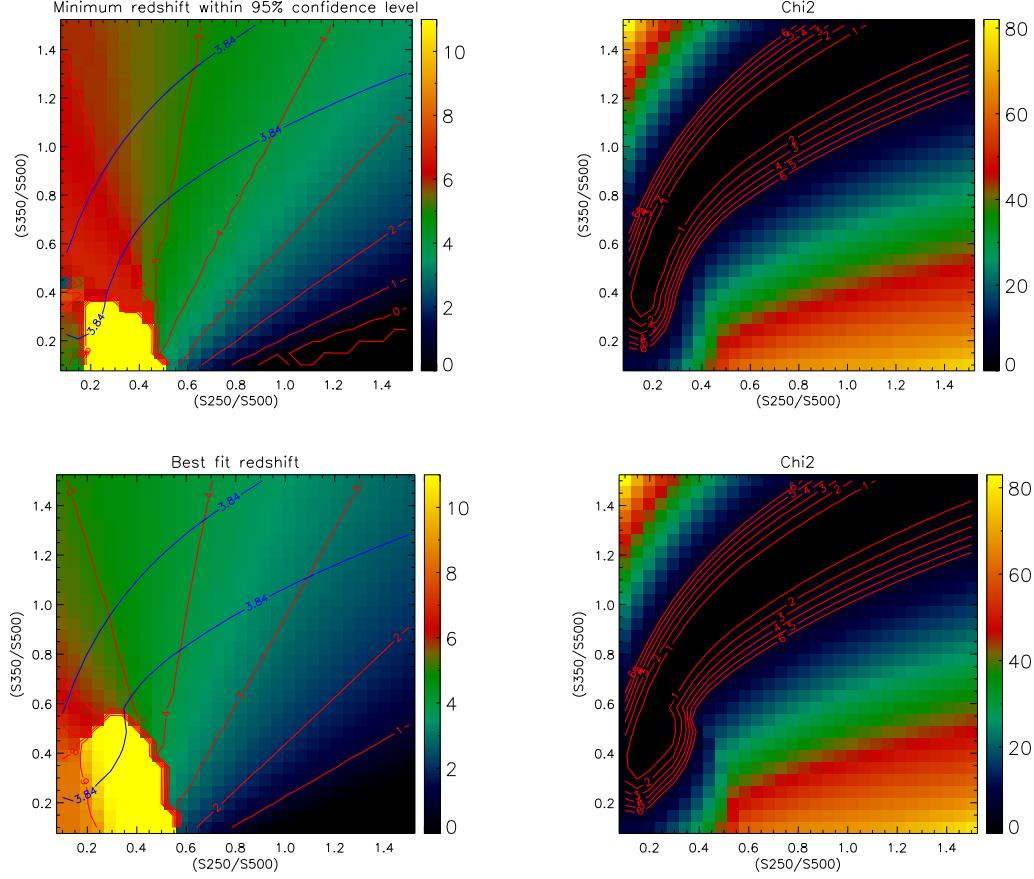


Fig. B.1: Criteria for selecting "500 μ m-risers". Plotted is SPIRE colour diagram related to the modified blackbody (MBB) model. The plot is a result of a Monte Carlo simulation in which we start with a single-temperature modified blackbody with the flux $S_{500}=30$ mJy. The mock SEDs are generated to account different fixed T_d and β illustrating temperature-redshift degeneracy. Plotted against all possible colours are simulated sources with $T_d=45$ K and $\beta = 1.8$ (upper panel) and $T_d=38$ 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. Correction factor due to modelled effect of CMB heating on colours is applied. However, our simulations shown that we do not expect 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 μ m-risers" selection we kept just those sources whose chi-square reside inside 95% confidence limit region ($\chi^2 < 3.84$).

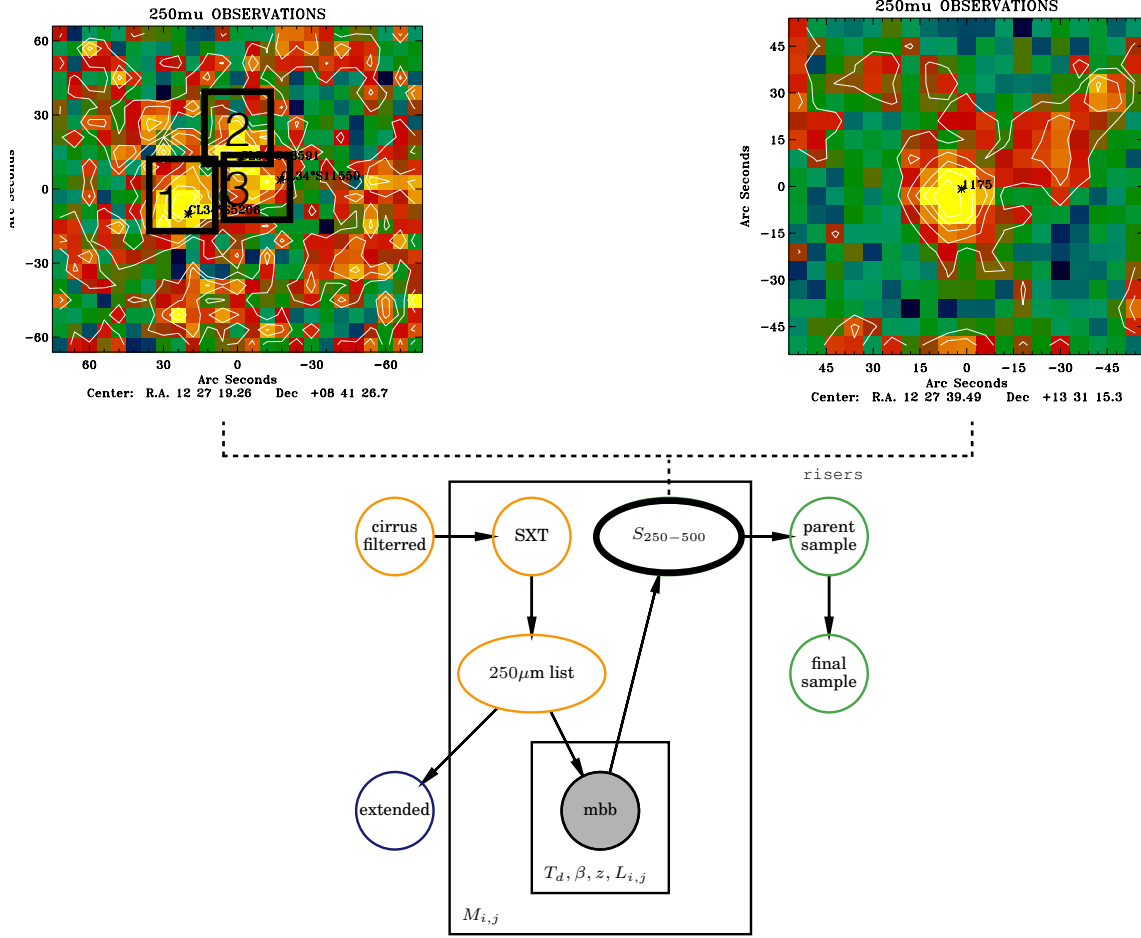


Fig. C.1: Example 2D cutouts of confused (left) and isolated point source detections (right) from our PSW catalogue. The first group of sources are fitted simultaneously, while second one is fitted as a single source.

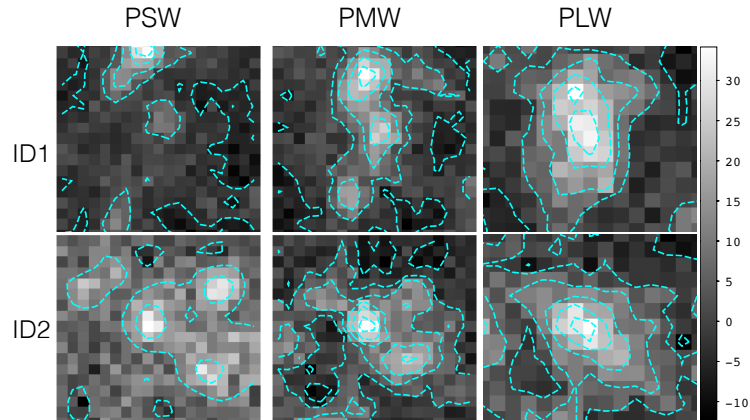


Fig. C.2: Example 60''x 60'' cutouts of "500 μm-risers" from our final sample. Both sources are detected in the second iteration of our extraction procedure, when we add sources that appeared bright in our residual maps. Upper panel: PSW, PMW and PLW cutout of an multisource case where two sources are de-blended for their PMW and PLW emission. Lower panel: An isolated case, without a prominent PMW or PLW blend in respect to the beam.