Internal release of Herschel simulated data and definition of optimal technique for catalog production

Herschel and SCUBA-2

INAF – OAR for the Astrodeep project

ASTRODEEP

"Unveiling the power of the deepest images of the Universe"

THEME [SPA.2012.2.1-01]

[Exploitation of space science and exploration data]

Grant agreement for: Collaborative project

Grant agreement no: 312725

ABSTRACT

In this document we provide a description of the algorithms/software that will be implemented in the catalog production, using optimal techniques that we have developed and tested on the Herschel simulated data presented in D3.2. Deliverable Number D3.3 – Delivery date June 2015.

Prepared by: D.Elbaz, N. Bourne, T. Wang
Approved by: AEC
Date: 05/06/2015

1. Incremental recipe for catalog production in presence of confusion, application to the Herschel images

We have developed an incremental methodology to produce catalogs of sources in images where the dominant source of noise is the confusion noise, due to a large density of sources with large beam sizes. This recipe which will be used to improve the Herschel catalogs that will be released by Astrodeep allows to improve detections as well as their photometric accuracy. This recipe, described in details in Wang et al. and Shu et al., takes advantage of the following major improvements with respect to our previous catalog release :

1. the use of the covariance matrix to derive flux uncertainties in the Herschel images. This method was calibrated using realistic mock Herschel images.
2. determination of the optimum number of priors to be used for source extraction
3. improvement of the photometric accuracy on "non-clean" sources by "freezing" faint sources.
4. search for 24μm dropout sources, i.e. sources detected with Herschel-SPIRE with no 24μm prior detection (using a blind source extraction on the residual maps).
5. optimisation of the identification of distant optical counterparts to 500μm sources (as discussed in Shu et al. 2015)

In the presence of confusion, we have shown in the past that prior source extraction was the best strategy to obtain reliable flux density measurements with the following potential drawbacks :

- the definition of priors require to know a priori which positions will be most approriate to host a source in the Herschel image
- the use of priors precludes the identification of a population of unexpected galaxies which would appear in the Herschel images but be absent from the prior list.

To deal with these issues, we are using priors based on Spitzer-MIPS 24μm catalogs which are very deep, at a higher spatial resolution level and probing the dust regime as Herschel and after extracting sources with those priors, we perform another source extraction, this time blindly, in the residual map to search for new sources with no priors. However, we found that the quality of the source extraction was not independent on the number of priors that was used with a trendoff between too few priors which would lead to missing a large population of sources, and too many priors which would lead to the artificial subdivision of a single detection into several sources. In order to determine the optimal number of priors to be used, we tried all possibilities and tested them with our mock Herschel images. The interesting result that came out of this study is the finding of a optimum number of priors that can be quantified as a function of source density as measured by the number of indendent beams per source obviously depending on the beam and field size. A given image contains N[indep.beams]= (survey area)/(beam size) independent beams where (beam size)= p (FWHM/2)^2. The number of beams/source for a given number of priors, N[priors], is then : N[beam/source]= N[indep.beams]/N[priors]. This is illustrated in Figure 1 for the case of the Herschel SPIRE 350µm band. Starting with N[beam/source]= 10 (we also tried higher N values), we find that the detection limit (y-axis) remains identical for N[beam/source]≥5 and starts reacing fainter flux densities below N[beam/source]=4. The detection limit is defined here as the flux density above which at least 68% of the detected sources present a photometric accuracy better than 32%. The detection limit reaches a minimum around 1.5-2 beams/source and then rises again when too many priors are used. For a field such as GOODS, the number of priors at 250, 350 and 500 µm are 1600, 840 and 400 priors, respectively.

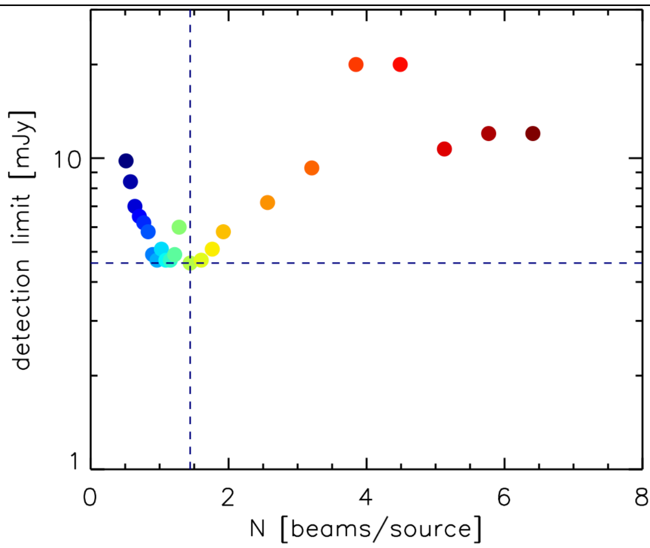

Figure 1: depth of the catalog as a function of the number of priors at 350 µm.

We use two complementary methods to compute the photometric uncertainty of individually detected sources :

- ***residual* noise, σ_{res} :** after a first run to determine 3σ detections, we launch a second run that builds a residual map (original image – 3σ detections). For each 3σ detection, we create a sub-image of 8 x FWHM on a side that we convolve with the PSF. The residual noise, σ_{res}, is the standard deviation of the pixels in that sub-image.

- ***covariance* noise, σ_{cov} :** the covariance noise is estimated from the covariance matrix that takes into account the number density of priors.

That we have tested using two different simulation methods:

- ***mock injected sources* noise, σ_{mock} :** we inject mock sources (PSFs normalized to a given flux density, S_{input}) in the real image, launch the standard source extraction technique and compare the *input* and *output* (S_{output}) flux densities of these mock sources. s_{mock} = standard devitation of (S_{output}-S_{input}).
- ***Herschel simulated image* noise, σ_{sim}:** we produce a full simulated Herschel image (see document by C.Schreiber on simulated images; gencat-report-20150505.pdf) and compare the *input* and *output* (S_{output}) flux densities of the sources in the input catalog. s_{sim} = standard devitation of (S_{output}-S_{input}).

The residual noise estimate provides in theory the most reliable noise level at the position of a given source, yet it relies on a map that was produced after subtracting 3σ detections, a process that is never perfect due to position uncertainties of the detections and the shape of the PSF that is measured on average in the map. Moreover, it may become unreliable in the presence of high densities of sources, hence we also use the covariance noise that measures the effect of high source densities.

The two methods based on simulations, σ_{sim} and σ_{mock}, are nearly redundant but with some moderate difference. In order to produce a realistic full *Herschel* image, we need to make sure to reproduce the actual noise of the real *Herschel* image, while injecting mock sources in the real image guarantees to be working with the exact same noise in the image. On the other hand, when we inject sources in the real image, these sources are affected by uncertainties linked to the presence of real sources of unknown flux densities. The full simulated *Herschel* image allows to maintain a full control of the sources of uncertainties. While both methods provide reliable flux uncertainties, the mock injected sources do not measure the flux at the position of a given detection but an average flux uncertainty for sources of a given flux density. It is therefore a good estimate of how uncertain is a given flux measurement but does not provide a local noise estimate. The full simulation provides a local noise estimate but that relies on the assumed IR SEDs that were used to build the image, hence may not reproduce perfectly well the actual environment of a given source. As a result, we have chosen to use these simulations to test σ_{sim} and σ_{mock} on both real and simulated *Herschel* images and make sure that these two estimates are reliable and we only provide these local noise estimates in our catalogs.

Point 3 in the list of improvements refers to an incremental process in which we run a second source extraction with a condition set for faint sources: these sources are forced to be fitted with a given flux density within a margin of 30% (this option initially implemented in the code FASTPHOT is being implemented in TPHOT). This process allows to avoid attributing too strong flux densities to satellite sources around the bright sources above our detection limit. The result of this second run has been compared to SCUBA-2 450μm data (described in Section 4.4) in the COSMOS field as shown in Figure 5. The result of the first run (without freezing faint sources) is shown in Figure 5-left where the filled red dots represent the 3σ sources detected in both Herschel-500μm and SCUBA2-450μm. There is a systematic offset of the Herschel measurements which overestimate the values of the flux densities by ~50%. A statistical comparison of both samples shows that with the first run alone, only 43.3% of the Herschel sources are detected at 4σ ("purity") and among the 4σ SCUBA2 detections, only 44.4% are detected by Herschel above 3σ ("completeness"). The present SCUBA2 catalog is conservatively limited to 4σ detections.

To perform the second run, we identified priors for which Herschel detections were measured below the 3σ limit and froze them to a 500μm density value that was extrapolated from their full far-IR SED from 70, 100, 160, 250 and 350μm measurements. This freezing process prevents these sources to be artificially boosted in a way that would make their SED abnormal. The flux density was frozen leaving a marging of 30% around that flux density in the PSF fitting code. The result of this second run is shown in Figure 2-right, where it appears clearly that the red dots, i.e. sources detected by both instruments, are in very good agreement and show no more systematically larger values for Herschel measurements. Thanks to this method we can now detect sources with a 66.7% completeness down to 7 mJy and a purity of 60%. This is only a conservative estimate since a peak flux analysis on the SCUBA2 mosaic shows that 50% of the Herschel-non SCUBA2 sources are actually present in the SCUBA2 map but at a flux density level typically twice lower suggesting that those Herschel sources have their flux densities boosted by neighboring sources.

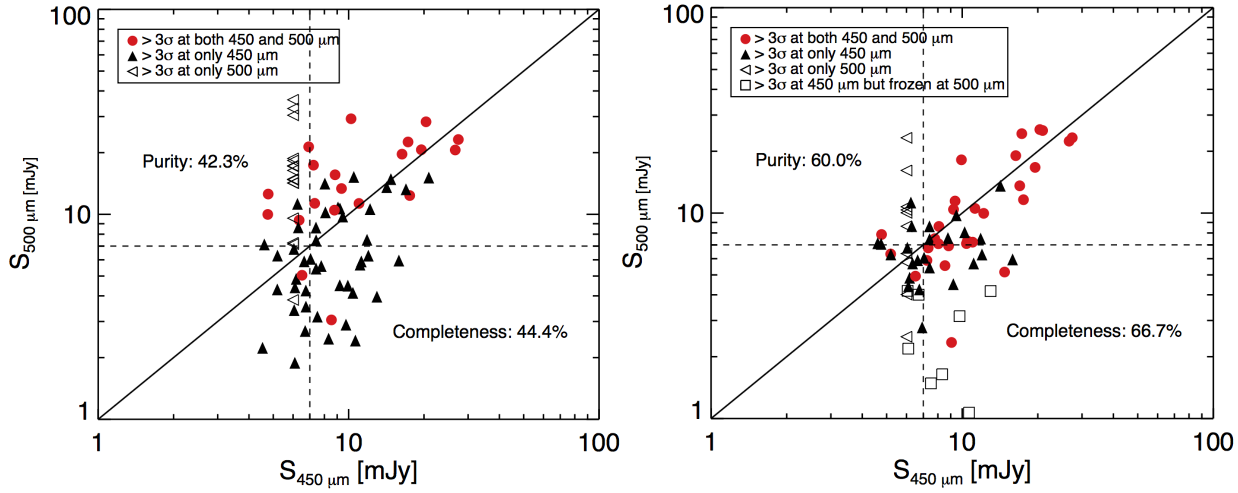

Figure 2 : $S_{500\mu m}$(Herschel) vs $S_{450\mu m}$(SCUBA2)/$S_{predicted}$ vs $S_{predicted}$ at 250μm where $S_{predicted}$ is the flux density extrapolated using the observed shorter Herschel bands (hence 70, 100 and 160μm) at the photometric redshift of the source using the correlations presented in Leiton et al. (2015).

Finally, to search for potential sources with no prior positions in the initial prior list, we apply a blind source extraction on the residual map. An example of 3 such sources is presented in Figure 3.

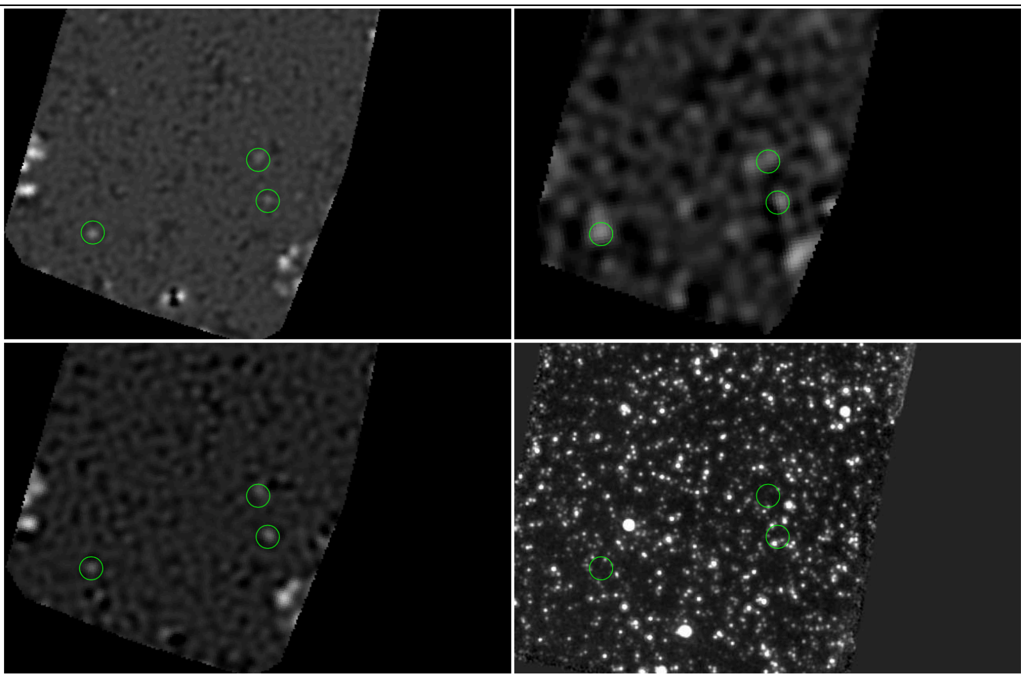

Figure 3: examples of 24μm dropout sources, i.e. sources seen in the SPIRE bands but not at 24μm. Residual images of SPIRE 250 (top left), 350 (bottom left), 500 μm (top right) and MIPS 24 μm images.

2. SCUBA-2 Tests

2.1 Introduction

The SCUBA-2 Cosmology Legacy Survey (CLS) provides sub-mm imaging over several deep survey fields, and benefits from a much better angular resolution than *Herschel* which allows for a much lower confusion noise limit. The CLS is especially useful for ASTRODEEP because it provides imaging at 7.5-arcsec resolution at 450μm in a number of ~100-arcmin^2 CANDELS pointings, constituting the deepest and highest-resolution single-dish sub-mm survey. This affords us two opportunities for testing ASTRODEEP software:

1. Owing to greatly reduced source confusion and lower confusion noise in comparison with *Herschel*, the 450μm maps provide an excellent test of the fidelity of simulated sub-mm maps produced by GENCAT.
2. The 450μm maps provide a "true" deconfused image of the sources present in the 36-arcsec-resolution *Herschel* 500μm maps, and this provides an opportunity to validate the T-PHOT deconfusion of the *Herschel* data.

A summary of these tests will be given in the following sections. In addition, we will discuss the issues with the sub-mm background estimation and subtraction that is essential for obtaining reliable results from T-PHOT.

2.2 Comparison of 450μm data and simulation

2.2.1 Residual map properties

The 450μm map of the CLS COSMOS-CANDELS field reaches an rms depth between 2.5-4.0 mJy/beam in the centre, where two DAISY pointings overlap, and 4.1-6.0 mJy/beam in the single-DAISY regions (about 1 mJy/beam after matched-filtering). The simulated map of the same field from the GENCAT/SKYMAKER packages has an rms depth of 1 mJy/beam across the full map (without filtering).

A straightforward comparison of the two maps is given by the residual between the raw, unfiltered maps, defined as $res = data - sim$. Normalising this by the instrumental noise map (of the data) *rms* allows us to define the residual signal-to-noise ratio (SNR) map as $rsn = res/rms$ (see Figure 1). Any positive source in this residual map is one that was not modelled in the simulation or was under-estimated, and a negative source is one that was over-estimated, compared with the true data.

The second panel of Figure 1 shows the result of convolving the *rsn* map with the SCUBA-2 450μm PSF (Gaussian, FWHM 7.4 arcsec) in order to maximise the SNR of point sources. Point sources were then extracted at positive and negative peaks in this map, resulting in one positive peak at SNR> 5, 66 at SNR> 3, and 54 negative peaks at SNR> 3. This suggests that the simulation has achieved a reasonably good reconstruction of the sky at 450μm, with only moderate deviations from the real measurements.

Figure 2 shows the pixel histogram of the residual map in mJy/beam, which is very close to Gaussian. The green line represents pure noise including the instrumental noise from the data and the noise in the simulation (it does not include confusion noise). The residual map histogram is slightly broader and deviates more at high fluxes as a result of differences between sources in the data and simulation.

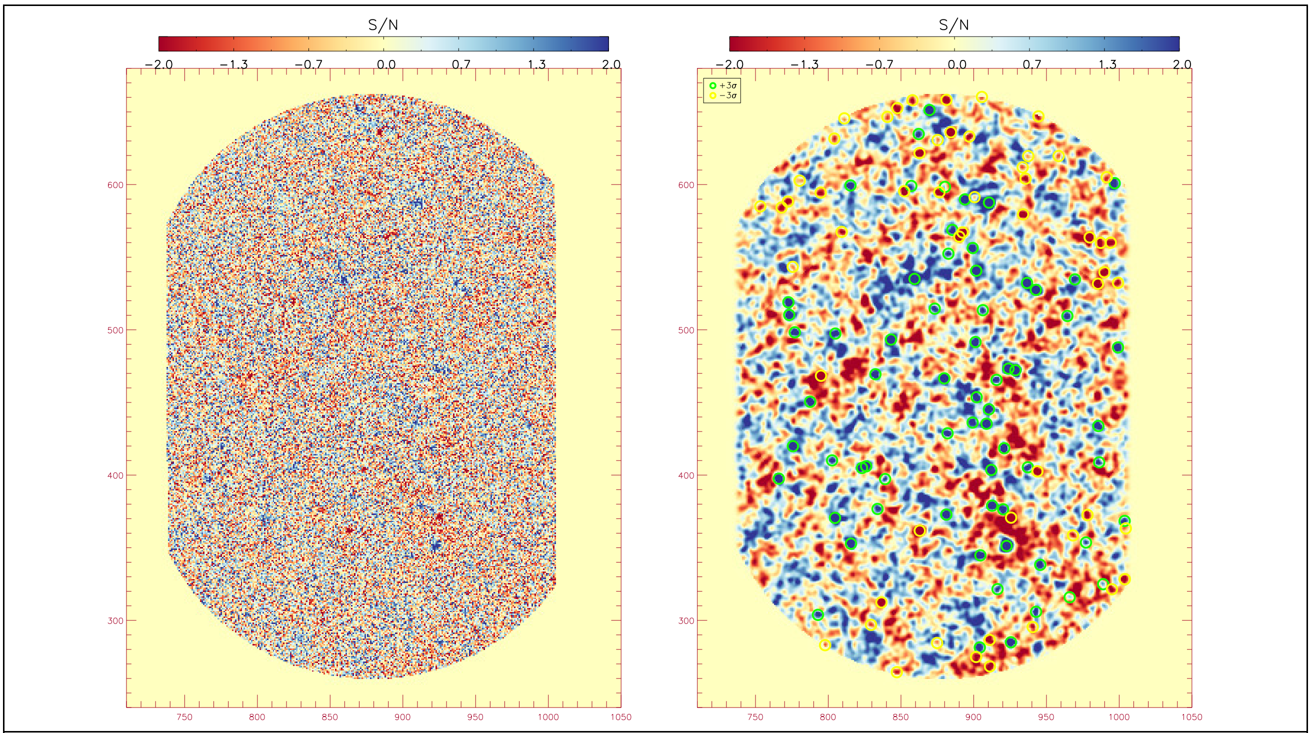

Figure 1: The residual SNR (*rsn*) map scaled between ±2 (left), and (right) after PSF-filtering and showing the locations of both positive (green) and negative (yellow) > 3σ residual ‘sources’.

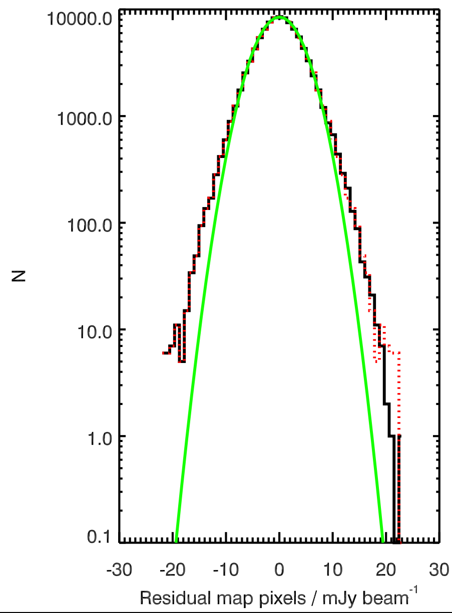

Figure 2: Histogram of pixels in the residual (*res*) map (black line), alongside a reconstructed histogram (red dotted line) produced by mirroring the negative half of the pixel histogram about the median, which demonstrates the symmetry of the residual. The green line represents the instrumental noise centred on the median value, with width σ equal to the geometric mean of the *rms* map (3.9) added in quadrature to the rms of the *sim* map (1.0).

2.2.2 Comparison of extracted sources

A more detailed comparison of the properties of the simulated and real data maps can be achieved by exploring the source population in each. Sources were extracted from each map blindly and independently, in order to compare the catalogues of bright sources. This was done within a mask that defines coverage in both maps, subtracting a background given by the median pixel value, and following a procedure similar to SUSSEXtractor (Savage & Oliver 2007) to select peaks in a PSF-filtered SNR map. Peaks were sorted in decreasing order of SNR, and for each source the precise centroid was obtained by bicubic interpolation of neighbouring pixels, before the flux was measured using a Gaussian fit at that position with width equal to the beam FWHM. The scaled Gaussian was then subtracted from the map before continuing to the next most significant source, to avoid bias of fainter sources from confusion with brighter ones. In this way we identified 97 sources at SNR> 4 in the SCUBA-2 map, and 380 in the (deeper) simulated map. Matching between these two catalogues with a search radius of 1.75 pixels (3.5 arcsec, or about 0.5×FWHM), revealed 59 matches between the two. The positions of the matched and unmatched sources are shown in Figure 3. Figure 4 shows the comparison between PSF-convolved fluxes as well as aperture fluxes for the 59 matched sources. The mean ratio of flux in the data to flux in the simulation for these sources is 1.46, and the standard deviation is 0.71.

The fluxes of sources extracted from each of the two maps can differ for various reasons, for example due to differential flux boosting from random noise peaks, or different amounts of blending between the two maps (a source might be blended with a bright neighbour in the *data* map but that neighbour could be much fainter in the *sim* map, or vice-versa). In Figure 4 we see that the fluxes of matched sources from the two maps generally follow the 1-to-1 trend but with a large amount of scatter, which can be explained by the aforementioned effects. The fact that the mean flux of matched sources is higher in the *data* map by a factor of 1.5 may be explained by flux boosting: sources are aligned randomly with respect to positive and negative peaks in the instrumental noise, but sources which fall on positive noise peaks are more likely to be extracted, leading to an overall bias towards higher fluxes of extracted sources. This bias affects the *data* sources more than the *sim* sources because of the higher instrumental noise.

It is also instructive to consider the unmatched sources, in particular those extracted from the *data* map that had no match in the *sim* map. Such cases are unlikely to result from noise fluctuations in the simulation because of the increased depth of the simulation (note that it contains three times as many sources as the *data* map). Visual inspection of the *sim* and *data* maps at the positions of the 38 unmatched *data* sources reveals that they generally result from one of the following situations:

i. The same source is detected but at a separation greater than the search radius, presumably as a result of blending altering the position in either or both maps.
ii. The source is heavily blended with a nearby neighbour, such that it is extracted as a single source in one map and as two or more in the other. This is a common problem that may be alleviated by the use of a matched filter instead of PSF-filtering.
iii. The source is present but is too faint to be extracted in the simulation. This is unlikely to result from noise as previously stated, but if the simulated flux is significantly under-estimated then this will show up as a blue peak in the residual.
iv. The source is genuinely absent from the simulation. This is difficult to distinguish from the previous option in practice, without checking for a match in the prior list, but in principle it is possible for SCUBA-2 to detect obscured, high-redshift sources that are too faint in the *Spitzer*/*HST* bands to appear in the prior list.

Situations (i) and (ii) are to be expected as a result of blending and are no cause for concern, although they demonstrate the particular difficulty in extracting meaningful source positions and photometry from confusion-limited maps. We identified 17/38 cases in categories (iii) or (iv), of which 15 appeared to be completely absent from the *sim*, indicating areas where the simulation differs significantly from the real image.

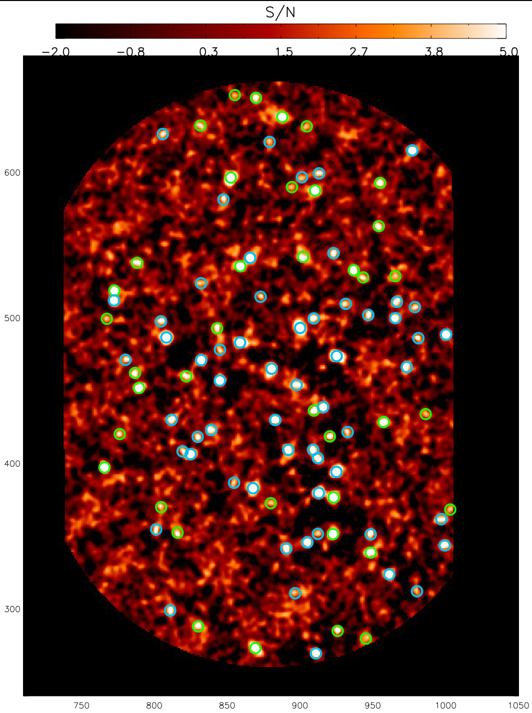

Figure 3: The PSF-filtered SNR in the SCUBA-2 map (left), and in the simulation (right). Sources extracted from each map with SNR> 4 are shown as green circles, and matches between the two lists of source positions are highlighted with a blue circle.

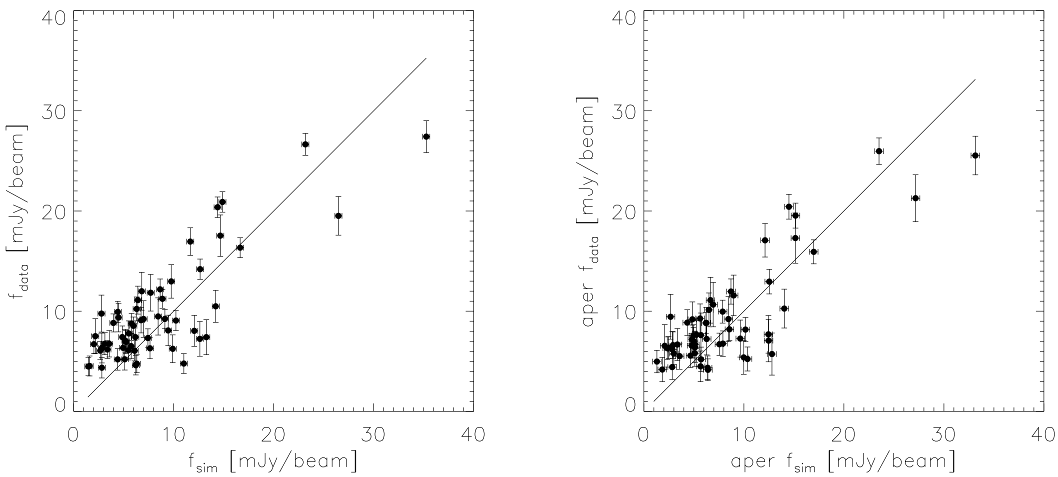

Figure 4: Each plot shows measured fluxes from the SCUBA-2 (*data*) map versus the fluxes from the simulated (*sim*) map, for the 59 matches between the two blind source-extraction catalogues. Left = PSF-filtered flux measurements; right = aperture fluxes within a radius equal to half the beam FWHM, with an aperture correction based on the PSF. The solid line is 1:1.

2.2.3 Direct comparison at fixed source positions

The analysis above relies on matching blindly extracted catalogues from the two maps, which is inevitably imperfect because source extraction can be biased by blending as described above. A more direct test of the simulation can be achieved by measuring fluxes at fixed prior positions in both maps. For this comparison we use the positions of sources extracted from the *sim* map (which is deeper) and measure aperture fluxes at these positions in both maps, using an aperture radius equal to half the beam FWHM and an aperture correction based on the PSF. Note that, unlike Figure 4, this method of comparison is not subject to flux boosting in the noisier *data* map; on the other hand it ignores any sources that may be absent from the simulation. The results in Figure 5 show similar scatter as was visible in the matched catalogue fluxes. At high fluxes, the flux in the *data* is more often lower than in the *sim*. This might be expected because the positions are chosen from the peaks in the simulated map and may not be perfectly aligned with the peaks in the *data* map, or with the original prior positions, as a result of blending. At the lowest fluxes the sources are clearly not detected in the *data* map due to its higher noise level, but no bias is evident and the scatter appears to be consistent with the error bars.

2.2.4 Conclusions

The residual representing the difference between the *data* and *sim* maps is encouragingly flat and symmetrical in its pixel distribution, suggesting that the simulated sources mimic the data in a sensible way. A number of positive and negative peaks above 3 times the noise level are detected, indicating places where the simulation under- or over-predicts the flux (respectively). These peaks can be seen in Figure 1 to be clustered in some places, which may indicate spatially correlated groups of sources that the simulation fails to reproduce.
Sources extracted from the *data* and the simulation were cross-matched, resulting in 59 matches out of the 97 SNR> 4 sources extracted from the *data* map. The matched sources demonstrate correlation between their *data* and *sim* fluxes, with a large scatter, and on average the *data* flux is higher than the simulated flux by a factor 1.5. A possible explanation for this is flux-boosting or Eddington bias in the noisier *data* image.

The unmatched sources are not necessarily all absent from the simulation but may simply be offset in position due to differences in the blending between the two maps, since blending can effect the positions of extracted sources. Visual inspection of the unmatched sources indicates a possible 15 sources from the *data* map which appear to be genuinely absent (or at least very faint) in the simulation.

An independent analysis of aperture fluxes at the positions of the simulated sources in both *data* and simulated maps indicates that the fluxes of simulated sources are well correlated with their fluxes in the true data and while a slight bias is apparent this could be explained by the choice of positions.

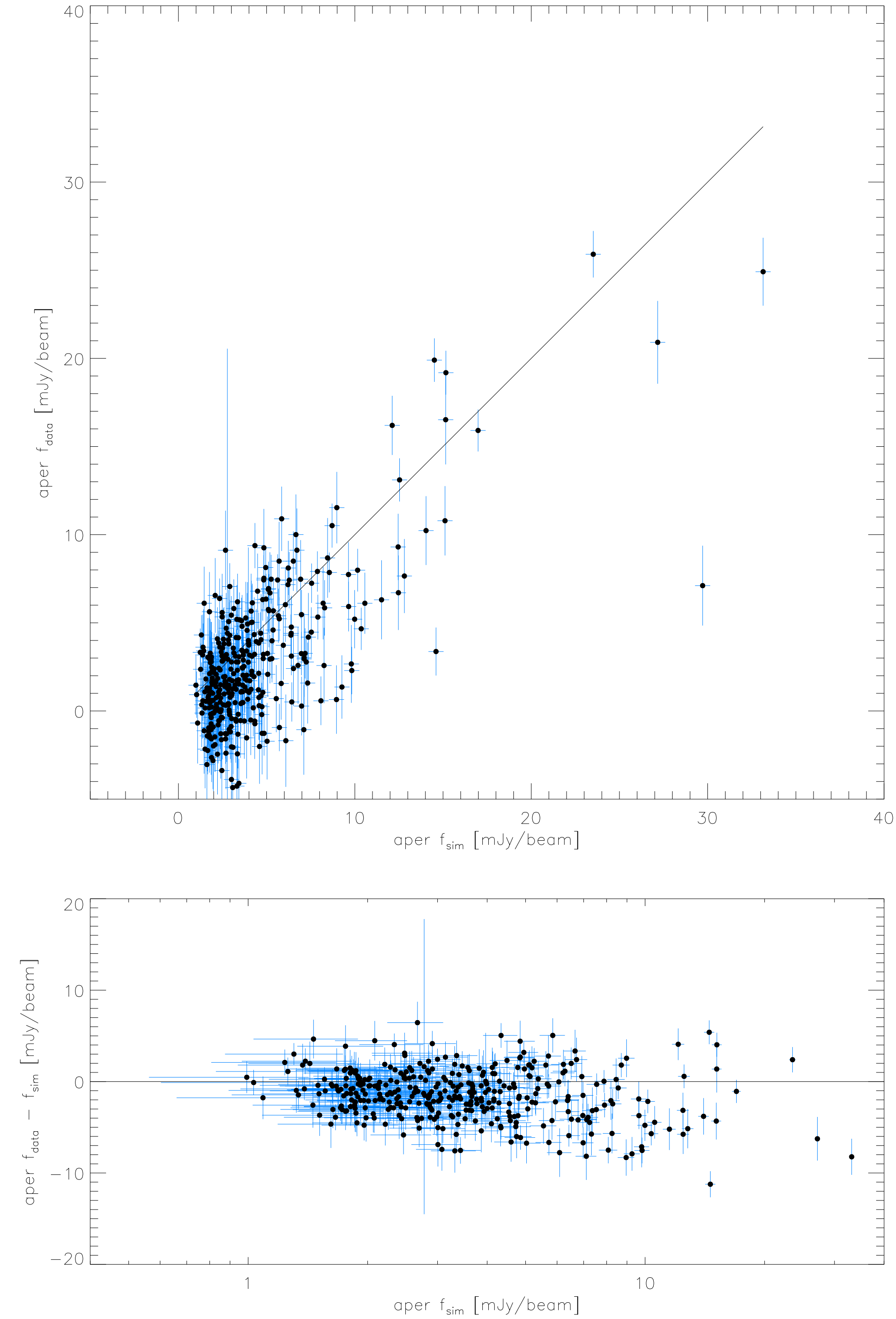

Figure 5: Top: comparison of aperture fluxes in the data and simulated maps at the positions derived from the simulated map. Bottom: difference between flux in the two maps at the same positions.

3.3 Sub-millimetre background estimation

3.3.1 The background in confused maps

An essential assumption made by T-PHOT is that the input map is background-subtracted. In the case of confusion-dominated or confusion-limited maps this can be problematic because there are no regions of empty sky background in which to measure the background, and the conceptual difference between sources and background is ambiguous. Without a suitable background subtraction, T-PHOT would be able to find a solution which is mathematically acceptable simply by balancing a large number of faint negative and positive sources, until a net zero residual is obtained (at least in parts of the map where the priors are sufficiently dense). In an attempt to more accurately reflect reality, it is possible to use the CLIP keyword to suppress negative sources, but then T-PHOT may simply fail to obtain an acceptable residual. An example of this is shown in Figure 6, which shows the T-PHOT residual/rms (*rsn*) map after fitting the *Herschel* 250μm map of the COSMOS-CANDELS field with a combined prior list from *Spitzer*/MIPS 24μm, VLA 1.4 GHz and *Spitzer*/IRAC 8μm catalogues. The raw map was not background subtracted but was scaled to have a mean pixel value of zero (a common standard for sub-mm maps). The T-PHOT residual reaches zero mean where the prior density is high, but a negative background remains in sparser regions. The flat residual in dense regions is only obtained by balancing positive and negative solutions for different sources: this is apparent when using the CLIP keyword which forces priors fitted with significant negative fluxes to be discarded, leaving a negative background over much of the map. Clearly neither of these solutions is satisfactory.

One natural option for measuring the background is to inject fake sources into a map and re-extract them, thus measuring an average offset between input and output fluxes across the map, or as a function of position within the map. The result of using this approach on the same data is shown in Figure 7, both with and without the CLIP keyword. Now the use of CLIP makes little or no difference because there are no priors fitted with significant negative fluxes. The residual in dense regions is still relatively well centred on zero, however there is significant positive background in sparser regions indicating that either there is a positive background that still needs to be removed, or that the prior list is incomplete. However, the priors from 24μm and 1.4 GHz would be expected to include most star-forming galaxies at $z<3$, while the 8μm IRAC priors should include most massive galaxies at higher redshifts that were otherwise missed.

The problem with this approach is that it only measures the background "behind" those fake sources, and the measurement includes the real sources in the map itself, most of which are faint and not simple to remove by independent means.

These are the sources that we want to model, so we do not want them to be included in the background. One would like to measure the background in the residual after these sources have been modelled and subtracted: this would give the background behind all the sources in the model. This is however a circular problem: we need an estimate of the background so that we can accurately model and subtract the sources, but we cannot measure the background until we have subtracted the sources. One cannot directly measure the background from the T-PHOT residual because the model will have been optimised to obtain a zero residual, which means it will not be optimised for the true fluxes of the sources.

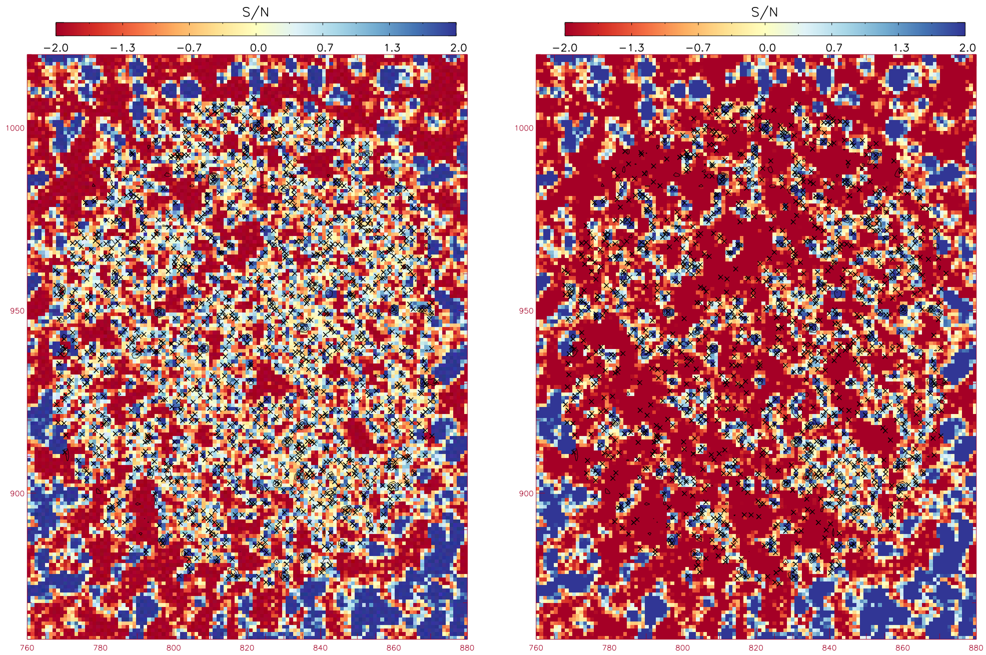

Figure 6: The T-PHOT residual after fitting part of the raw 250μm COSMOS map at the positions of 8μm+24μm+1.4GHz priors. Left: allowing all sources; right: rejecting negative sources (CLIP). Grey contours show peaks in the higher-resolution 450μm map for comparison.

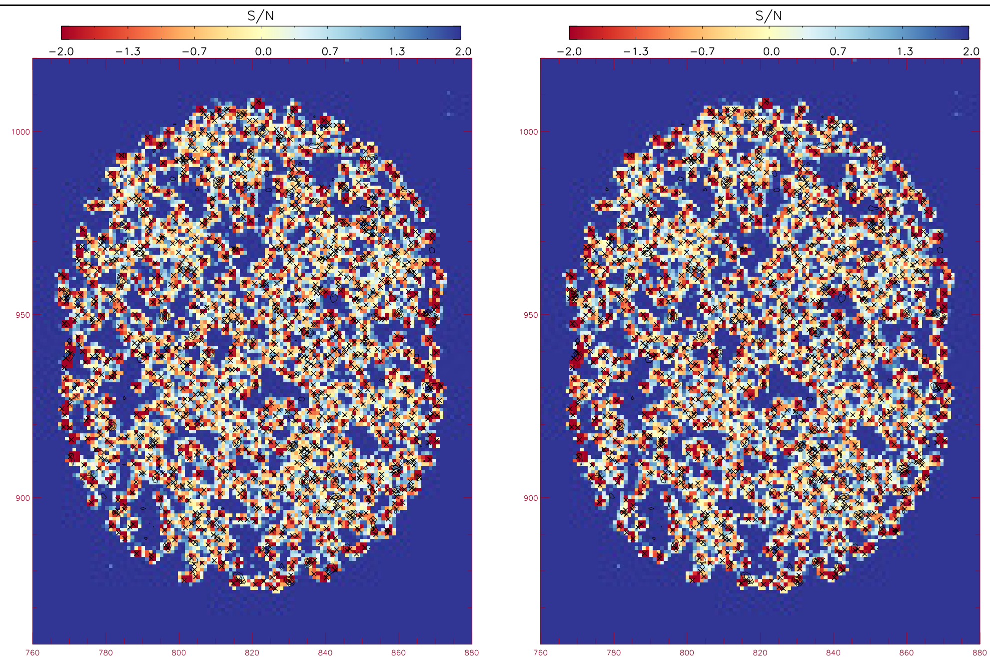

Figure 7: As Fig. 6, except after subtracting a background estimated by injecting and re-extracting fake sources.

Ideally we need an independent way to measure the background behind the sources that we want to model. This background contains any sources that are not being modelled (not in the prior list) in addition to an arbitrary sky background – arbitrary because the sub-mm maps are scaled to have a zero mean. A good first guess would be the modal value of the histogram of pixels in the raw map. This is not quite correct because this histogram is positively skewed by genuine sources, some of which we want to model. In a zero-mean map this value is negative, but in general it will be an over-estimate of the genuine background level when the map is very confused (see Figure 8).

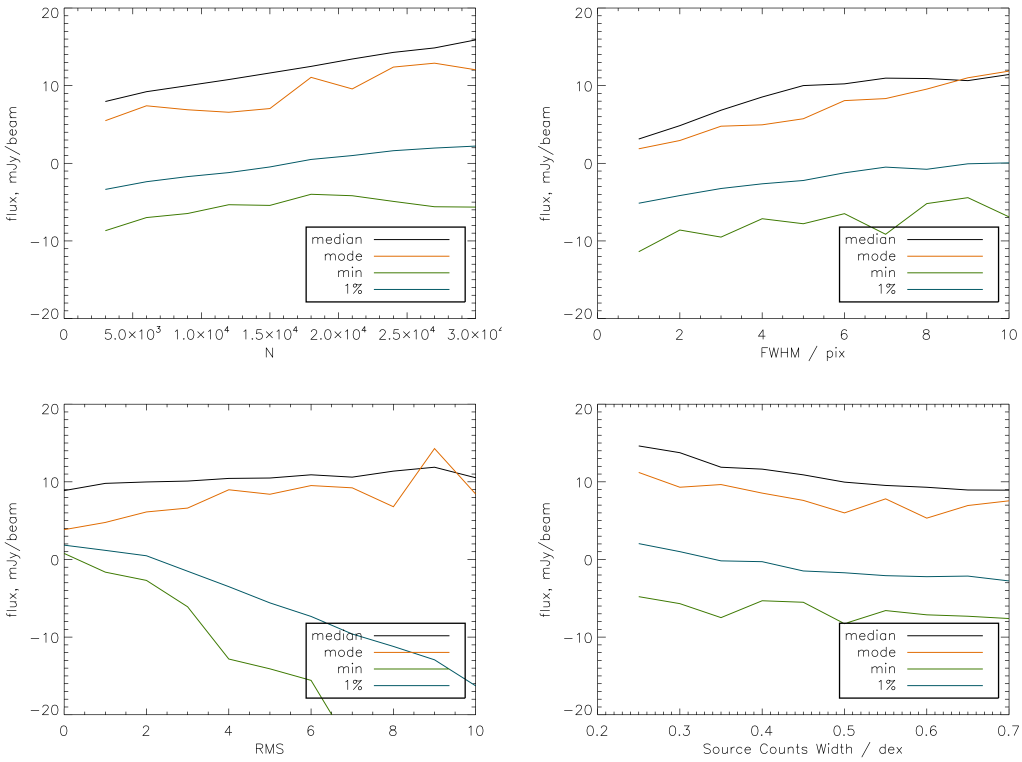

Figure 8: Confused maps were simulated with zero background and varying levels of instrumental noise, source density and source counts. The median, mode, minimum and first percentile of the pixel histogram were measured and plotted here as a function of the simulated map properties. The true background level of 0 mJy is usually between the minimum and modal values of the pixel histogram, but the relative value of these statistics depends on the number of sources in the map (top left), the beam size (top right), the rms noise level (bottom left) and the shape of the source counts (bottom right). In these simulations, the source counts were assumed to be lognormal and roughly matching the observations from Glenn et al. (2010).

3.3.2 An iterative solution

We have found the following algorithm to effectively solve the background problem:

- Estimate the mode of the raw map from the pixel histogram and use that as a first guess at the background level (this will be negative because the raw map has a mean of zero).
- Subtract this constant value from the raw map and run T-PHOT on the result, using the CLIP keyword to prevent negative sources being included in the model.
- Estimate the mode of the residual from T-PHOT: this will be slightly skewed and have approximately zero mean and slightly negative mode, because there is still some residual negative background in the map.

- Repeat from step 2, using the accumulated background value from the sum of previous estimates.

Subsequent iterations of subtracting the mode of the residual and re-running T-PHOT can be repeated until the accumulated background value asymptotes towards a stable result, the mode of the residual converges towards zero, and the chi-squared of the fit reduces towards a minimum. At this point the negative background (including all sources not modelled) has been effectively subtracted and T-PHOT is able to obtain a good fit for the remaining positive sources in the model, with a symmetrical, Gaussian residual. Examples are shown in Figures 9–14.

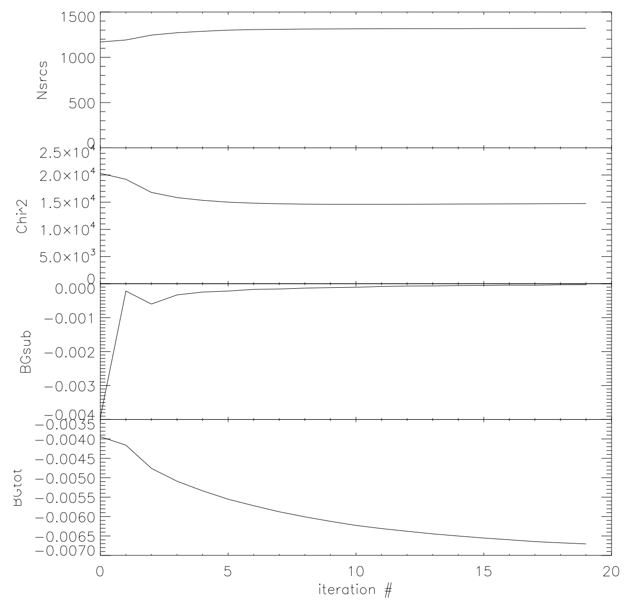

Figure 9: Convergence towards a background solution for the 250μm map. For each iteration, the plot shows the number of sources in the T-PHOT fit, chi-squared of the fit, the mode of the residual (the background to be subtracted in the next iteration), and the accumulated total background level from all previous iterations.

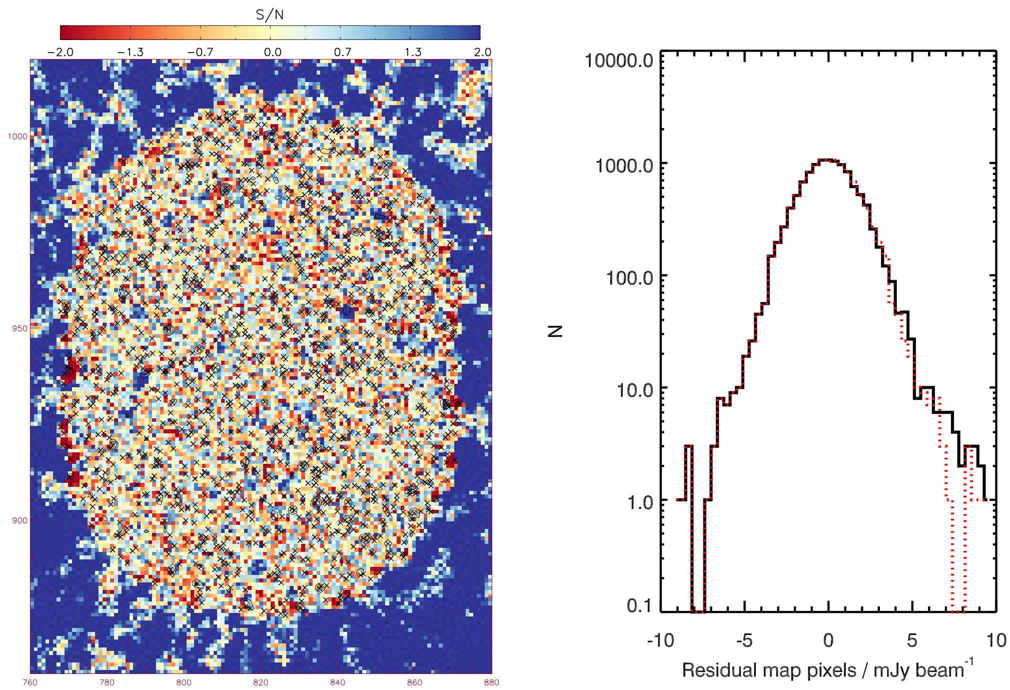

Figure 10: The T-PHOT residual at 250μm with CLIP=true, after iterating towards a background solution (iteration #19). Also shown is the histogram of pixels within the modelled region.

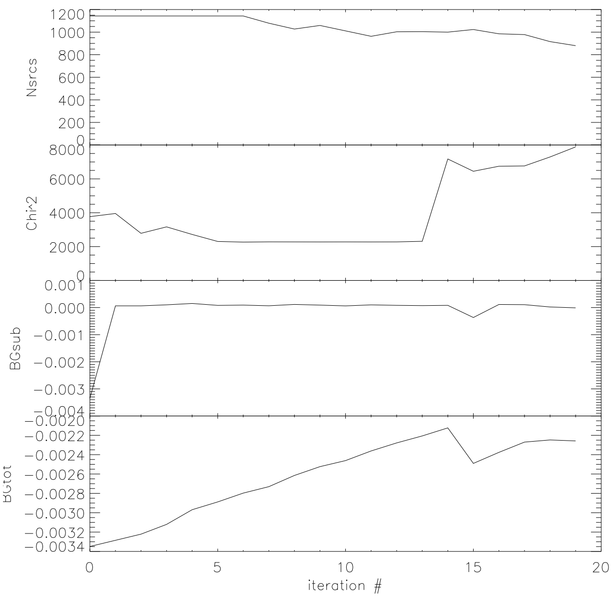

Figure 11: Convergence towards a background solution for the 500μm map (as in previous figure).

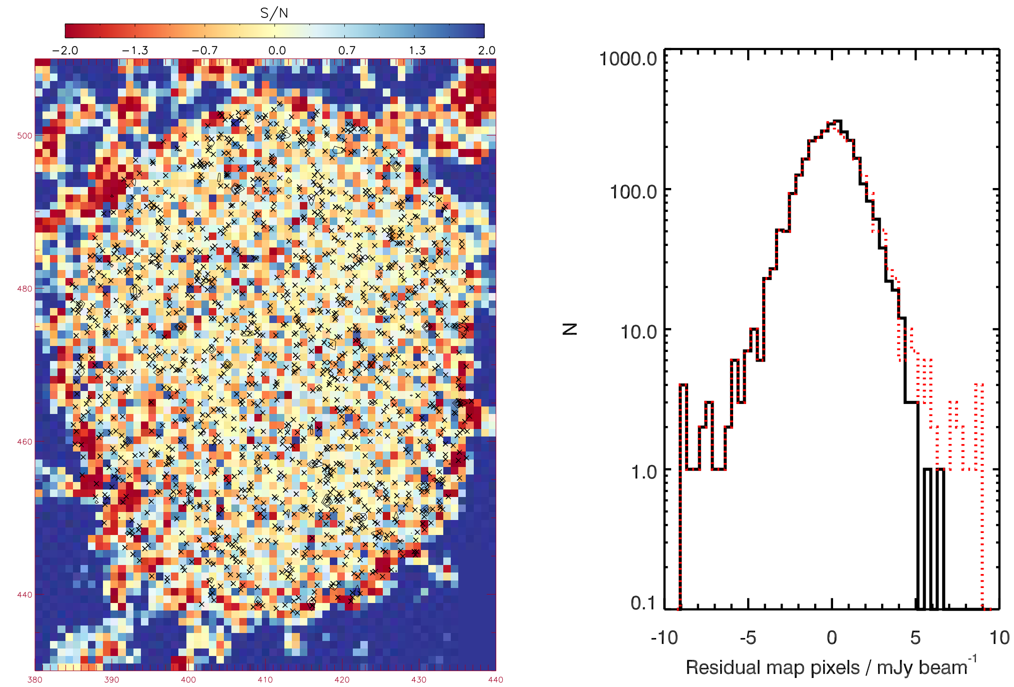

Figure 12: The T-PHOT residual at 500μm with CLIP=true, after iterating towards a background solution (iteration #6). Also shown is the histogram of pixels within the modelled region.

This method suffers from two particular problems. Firstly, it is very time-consuming since every iteration involves a full run of T-PHOT. Convergence usually occurs within 5–10 iterations, depending on the prior density relative to the map resolution (for example, convergence is quicker in the 850μm map shown in Figure 13 than in the slightly lower-resolution 250μm map in Figure 9). Moreover, by its nature, the background value depends not only on the the map but also on the prior catalogue used, since for a sparser catalogue, the background would contain more positive flux from sources not being modelled. One must therefore re-run these iterative fits every time the prior catalogue is changed.

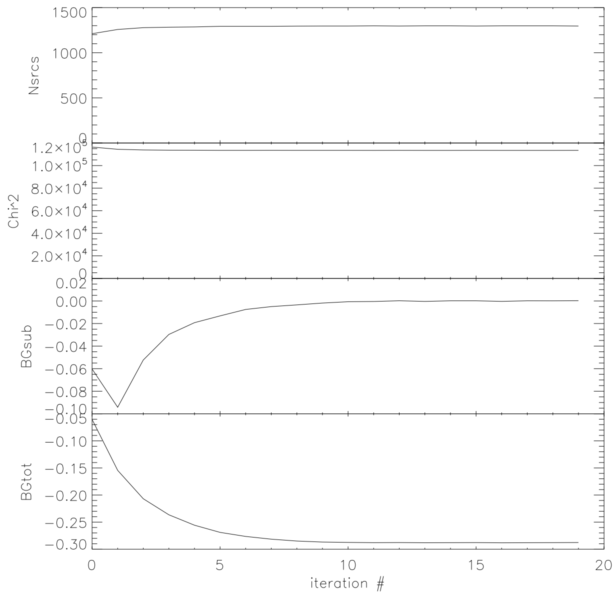

Figure 13: Convergence towards a background solution for the 850μm map (as in previous figure).

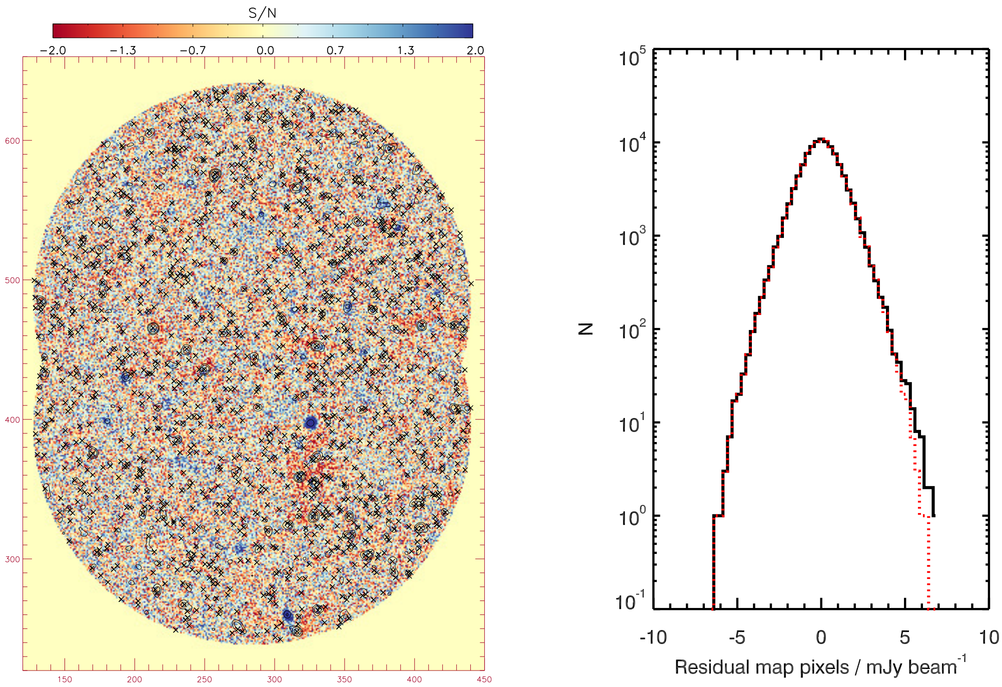

Figure 14: The T-PHOT residual at 850μm with CLIP=true, after iterating towards a background solution (iteration #11). Also shown is the histogram of pixels within the modelled region.

The second problem is that the iterations do not always converge neatly and the solution can either diverge indefinitely or “jump” to a very different background level, with a different number of sources and similar or larger chi-squared. This is particularly a problem in the lowest-resolution 500μm images, as in the example Figure 11. The reason for this is that successive iterations are not obtaining a more precise estimate of the background, but are simply reaching alternative (perhaps equally valid) solutions with a different number of sources in the map. This can happen because the density of sources in the prior catalogue is too large in comparison to the beam size and the solutions are degenerate. This can be avoided by careful choice of priors.

3.3.3 Conclusion

We have identified a reliable technique for accurately estimating and subtracting the background level in confusion-limited/confusion-dominated maps, in order to remove bias from T-PHOT flux measurements. The ideal situation would be to include the background as an additional free parameter in the T-PHOT fitting but currently this is not possible, hence the iteration method provides the next-best alternative.

3.4 Comparison of 450μm and 500μm T-PHOT results

3.4.1 Validating T-PHOT deconfusion

Here we describe a test of the T-PHOT deconfusion of 500μm imaging from *Herschel*. The basis of this test is that the SCUBA-2 450μm map provides a “true” deconfused image of the 500μm sources which are highly confused in the *Herschel* map at 36-arcsec resolution. The 450μm resolution of 7.5 arcsec is close to that of MIPS 24μm, which is commonly used as a prior for deconfusing *Herschel* images.

The COSMOS-CANDELS field was used for this comparison because this field has the deepest 450μm data to date. For a prior list we use the latest version of the CLS 450μm catalogue extracted from the same catalogue. It is reasonable to assume that a catalogue extracted from the 450μm data is a good prior for the 500μm map because the same sources will be present in both maps due to the similarity of the photometric bandpasses, and a denser prior list would over-parameterize the model leading to degeneracies in the fitting (i.e. the model would have too many parameters relative to the number of degrees of freedom in the data).

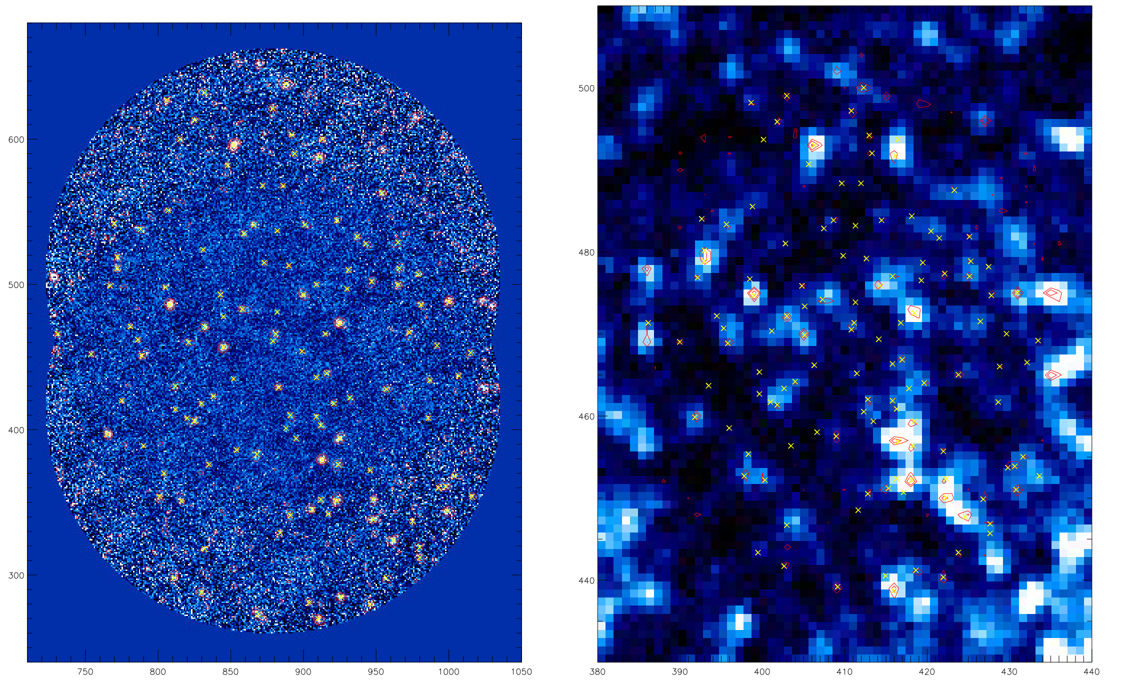

Figure 15: The 450µm SCUBA-2 map (left) and the 500µm *Herschel* map (right) of the same field in COSMOS-CANDELS. Red contours represent the 450µm map and yellow crosses show the positions of 450µm sources.

3.4.2 Results

Figure 15 shows the SCUBA-2 and *Herschel* maps side-by-side, with the prior positions marked. T-PHOT was run with these inputs following the iteration procedure described in the previous section to subtract the background level from each map. Residuals of the fits in Figure 16 contain some positive peaks and negative regions, indicating that the priors are incomplete and some sources remain that are not modelled, leading to some residual negative background. The histograms of pixels in the residuals shown in Figure 17) peak close to zero, indicating good background subtraction, although the peak is slightly negative in the 500µm histogram indicating a small negative background. The 450µm histogram is close to pure noise, although with a small excess of positive flux at □ 10 mJy/beam, while there is a significant amount of positive flux remaining in the 500µm residual.

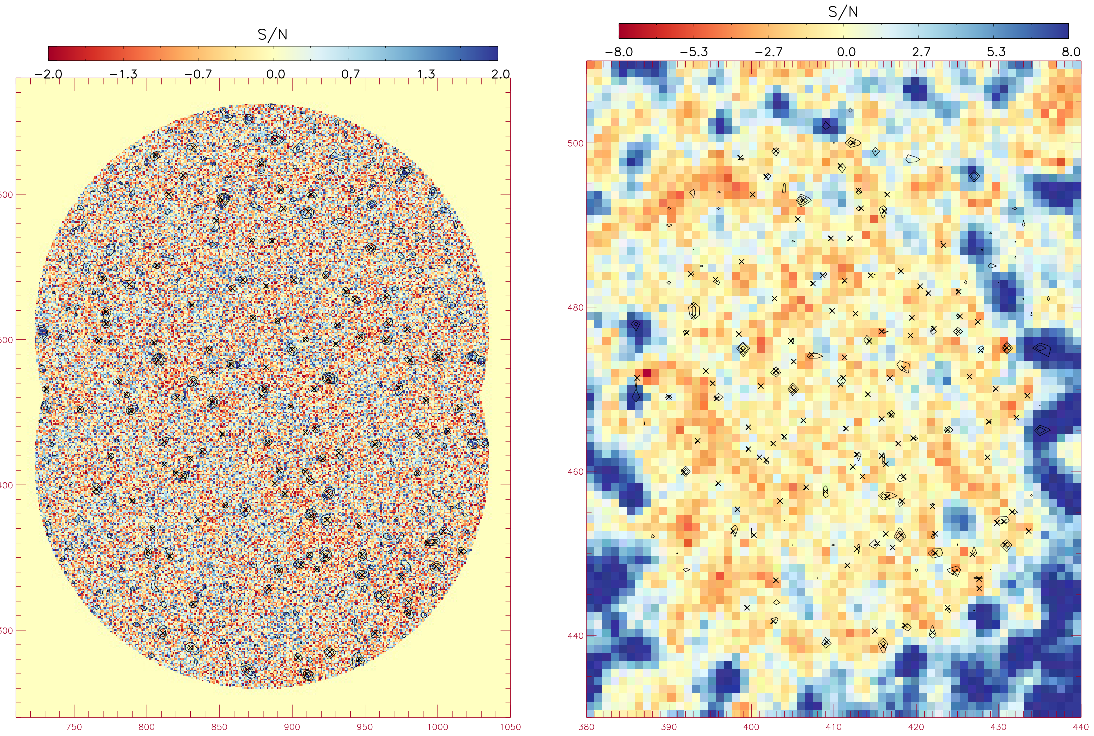

Figure 16: The T-PHOT residuals of the 450μm SCUBA-2 map (left) and the 500μm *Herschel* map (right). Black crosses show the positions of the priors (450μm sources), and grey contours follow the original 450μm map.

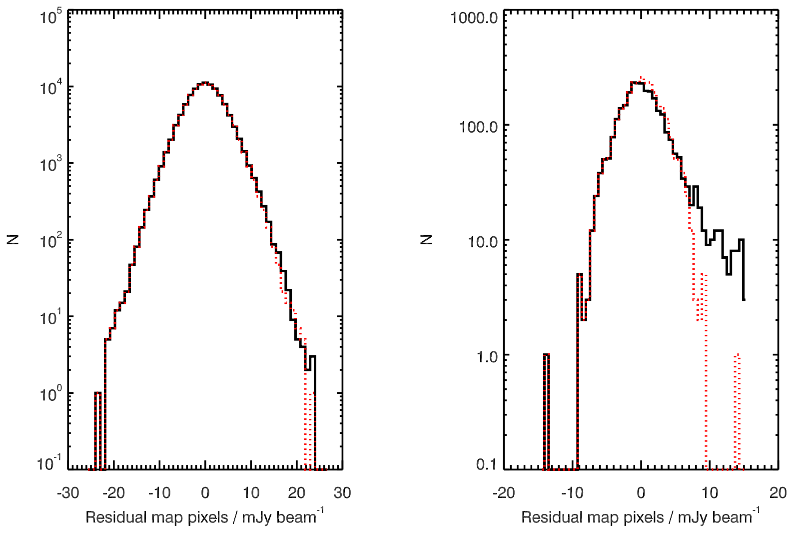

Figure 17: Histograms of the t-phot residuals from 450μm and 500μm maps.

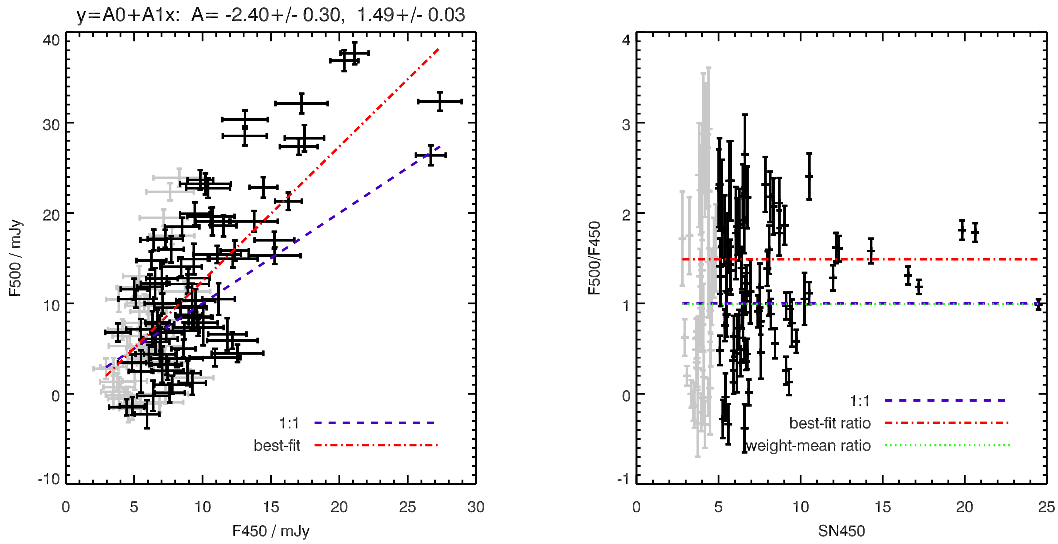

Figure 18: Comparison of flux measurements from T-PHOT fits to the 450 and 500μm maps (left), and 500/450μm flux ratio versus 450μm SNR (right). In the left-hand panel, the dashed violet line indicates equal fluxes at 450 and 500μm, and the red dot-dashed line indicates the best linear fit given by the parameters listed above the plot. In the right-hand panel, the purple line indicates flux ratio of 1.0, the red dot-dashed line indicates the flux ratio of 1.5 from the linear fit in the left-hand panel, and the green dotted line indicates the error-weighted mean flux ratio, which is 1.0.

Figure 18 shows a comparison between fluxes measured at the same prior positions in the two maps. The 450μm fluxes are treated as the "truth" and are in agreement with the CLS catalogue from which the priors were selected. In comparison, the 500μm fluxes from T-PHOT are well correlated with these, and show no bias towards negative values at low fluxes indicating that the background subtraction was adequate. The 500μm fluxes appear to be typically higher than 450μm fluxes, by an average ratio of 1.5, although when taking into account the errors from the covariance matrix, the error-weighted mean flux ratio is 1.0.

In general the errors on 500μm fluxes appear to be under-estimated by T-PHOT, since the scatter in flux ratios is high compared to the error bars. This is likely due to inadequate specification of the priors: the covariance matrix includes only the statistical uncertainty in the model within the confines of the priors, but not additional uncertainties due to confusion with background sources. This might be remedied by retrospectively adding a systematic error to the results representing additional confusion noise, and/or by specifying a more complete prior list.

The effects of using a denser prior catalogue can be seen in Figure 12, which shows the residual from fitting priors at combined positions from 8μm, 24μm and 1.4GHz catalogues. This residual is much flatter than that in Figure 16, although the noise in the residual appears to have been suppressed in regions where the prior density is highest, indicating that there were too many free parameters in the model (too many priors). This is reflected in the very large error bars on fluxes obtained from this model, shown in Figure 19, which result from degeneracies in the model. Clearly these results are useless for constraining the 500μm fluxes.

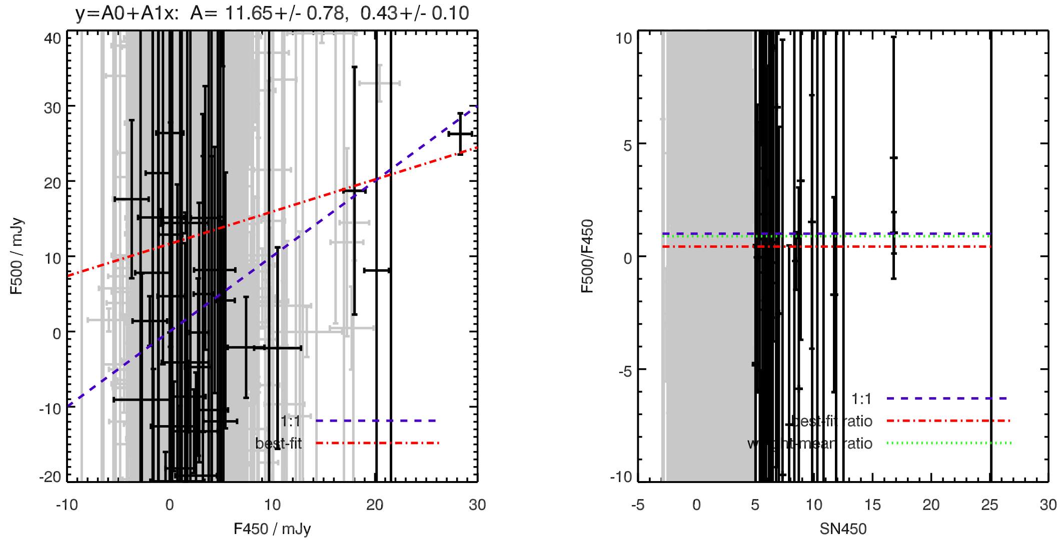

Figure 19: As Figure 18, but replacing the 450μm priors with the 8μm+24μm+1.4GHz prior list (see residual in Figure 12).

3.4.3 Conclusions

Fluxes obtained from T-PHOT in the 500μm image show considerable scatter when compared with the 450μm catalogue, even for the brightest sources. This could result from an inadequate prior list based on the 450μm catalogue, which does not include relatively faint but numerous sources that are important for accurately modelling the map. Although such sources are subtracted as part of the overall background, this does not account for their correlations with the brighter sources in the model. The error bars from T-PHOT include only the uncertainties due to confusion between sources in the model, and should be supplemented with an estimate of the additional uncertainty from residual confusion and instrumental noise.

A denser prior list combining 8μm, 24μm and 1.4GHz catalogues led to the opposite problem, where the errors on fluxes are too large for the measurements to be useful, since the model is degenerate. These results underline the importance of optimising the prior list to avoid under- or over-parameterizing the model. In the case of the 500μm map, an optimal prior list might for example combine 24μm and 1.4GHz sources, which will include most star-forming galaxies at $z < 3$, with either 450μm or 850μm sources, which will include any bright higher-redshift sources missed at the other wavelengths. Further work would be needed to test whether this improves the fidelity of T-PHOT photometry.

References

Glenn, J. et al. (2010) MNRAS, 409, 109
Savage, R. S., Oliver, S. (2007) ApJ, 661, 1339