

HELP: XID+, the probabilistic de-blender for *Herschel* SPIRE maps

P. D. Hurley,^{1★} S. Oliver,¹ M. Betancourt,² C. Clarke,¹ W. I. Cowley,³
 S. Duivenvoorden,¹ D. Farrah,⁴ M. Griffin,⁵ C. Lacey,³ E. Le Floc’h,⁶
 A. Papadopoulos,⁷ M. Sargent,¹ J. M. Scudder,¹ M. Vaccari,^{8,9} I. Valtchanov¹⁰
 and L. Wang^{11,12}

¹*Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Falmer, Brighton BN1 9QH, UK*

²*Department of Statistics, University of Warwick, Coventry CV4 7AL, UK*

³*Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK*

⁴*Department of Physics, Virginia Polytechnic Institute & State University, MC 0435, 910 Drillfield Drive, Blacksburg, VA 24061, USA*

⁵*Cardiff School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff CF24 3AA, UK*

⁶*Service d’Astrophysique, CEA-Saclay, Orme des Merisiers, Bat.709, F-91191 Gif-sur-Yvette, France*

⁷*European University Cyprus, 6, Diogenis Str, Engomi, PO Box 22006, 1516 Nicosia, Cyprus*

⁸*Department of Physics and Astronomy, University of the Western Cape, Robert Sobukwe Road, 7535 Bellville, Cape Town, South Africa*

⁹*Istituto di Radioastronomia, INAF, via Gobetti 101, I-40129 Bologna, Italy*

¹⁰*Herschel Science Centre, ESAC, ESA, Villafranca del Castillo, PO Box - Apdo. 50727, E-28080, Spain*

¹¹*SRON Netherlands Institute for Space Research, Landleven 12, NL-9747 AD Groningen, the Netherlands*

¹²*Kapteyn Astronomical Institute, University of Groningen, Postbus 800, NL-9700 AV Groningen, the Netherlands*

Accepted 2016 September 16. Received 2016 September 15; in original form 2016 February 1

ABSTRACT

We have developed a new prior-based source extraction tool, XID+, to carry out photometry in the *Herschel* SPIRE (Spectral and Photometric Imaging Receiver) maps at the positions of known sources. XID+ is developed using a probabilistic Bayesian framework that provides a natural framework in which to include prior information, and uses the Bayesian inference tool *Stan* to obtain the full posterior probability distribution on flux estimates. In this paper, we discuss the details of XID+ and demonstrate the basic capabilities and performance by running it on simulated SPIRE maps resembling the COSMOS field, and comparing to the current prior-based source extraction tool DESPHOT. Not only we show that XID+ performs better on metrics such as flux accuracy and flux uncertainty accuracy, but we also illustrate how obtaining the posterior probability distribution can help overcome some of the issues inherent with maximum-likelihood-based source extraction routines. We run XID+ on the COSMOS SPIRE maps from *Herschel* Multi-Tiered Extragalactic Survey using a 24- μm catalogue as a positional prior, and a uniform flux prior ranging from 0.01 to 1000 mJy. We show the marginalized SPIRE colour–colour plot and marginalized contribution to the cosmic infrared background at the SPIRE wavelengths. XID+ is a core tool arising from the *Herschel* Extragalactic Legacy Project (HELP) and we discuss how additional work within HELP providing prior information on fluxes can and will be utilized. The software is available at https://github.com/H-E-L-P/XID_plus. We also provide the data product for COSMOS. We believe this is the first time that the full posterior probability of galaxy photometry has been provided as a data product.

Key words: methods: statistical – galaxies: statistics – infrared: galaxies.

1 INTRODUCTION

Ever since the discovery of the far-infrared (IR) background by the *Cosmic Background Explorer* (COBE; Puget et al. 1996), surveys

have aimed to observe and detect the sources responsible. Most of those sources are galaxies, with the far-IR emission coming from dust.

While ground-based observatories such as SCUBA (Holland et al. 1999), and more recently SCUBA-2 (Holland et al. 2013) and ALMA can make use of IR atmospheric transmission windows to observe at the tail of the cosmic infrared background (CIB),

* E-mail: p.d.hurley@sussex.ac.uk

only space-borne facilities can observe at the peak ($\approx 140 \mu\text{m}$). The first IR space telescope, the *InfraRed Astronomical Satellite* (IRAS; Neugebauer et al. 1984), observed the whole sky in four bands centred at 12, 25, 60, and 100 μm and revealed new populations of galaxies that were optically faint but luminous in the IR (Soifer et al. 1984).

While the *Infrared Space Observatory* (ISO; Kessler et al. 1996) and the *Spitzer Space Telescope* (Werner et al. 2004) have provided deep near- and mid-IR photometry over small fields, other smaller space-borne facilities such as *AKARI* (Murakami et al. 2007) and the *Wide-field Infrared Survey Explorer* (WISE; Wright et al. 2010) have surveyed the entire sky at mid- to far-IR and near- to mid-IR wavelengths, respectively. The most recent advance in IR astronomy has been made with the ESA *Herschel* Space Observatory (Pilbratt et al. 2010). Photometry from the Photoconductor Array Camera and Spectrometer (PACS; Poglitsch et al. 2010) and Spectral and Photometric Imaging Receiver (SPIRE; Griffin et al. 2010) have given us an unprecedented view of the far-IR Universe by providing observations that measure across the peak of the far-IR background and at greater sensitivity and resolution than has been achieved previously at these wavelengths, thereby definitively setting the origin of the CIB.

With surveys such as the *Herschel* Multi-Tiered Extragalactic Survey (HerMES; Oliver et al. 2012) and the *Herschel* ATLAS survey (H-ATLAS; Eales et al. 2010), over 1000 deg^2 of the sky has been observed by the SPIRE instrument. However, due to the relatively large beam size of the SPIRE, and the galaxy density (≈ 30 per SPIRE beam for optical sources with $B < 28$), multiple galaxies can be located within the SPIRE beam. This is referred to as the problem of source confusion.

To obtain accurate photometry from the SPIRE maps, overcoming the source confusion problem is essential. One way to solve the problem is to use prior information to accurately distribute the flux in the SPIRE maps to the underlying astronomical objects. For example, if we know the location of a galaxy to a reasonable tolerance (e.g. from an optical image where resolution is better), we may expect a galaxy to be found in the SPIRE maps at the same location.

Several techniques have been developed that utilize the positions of sources detected at other wavelengths, usually 24 μm and 1.4 GHz, to disentangle the various contributions from discrete sources to the SPIRE flux in a given beam element (e.g. Roseboom et al. 2010, 2011; Chapin et al. 2011). This process is made possible by the strong correlation between the 24- μm and 1.4-GHz populations and those observed at far-IR wavelengths; >80 per cent of the cosmic IR background at SPIRE wavelengths can be accounted for by 24- μm sources with $S_{24} > 25 \mu\text{Jy}$ (e.g. Marsden et al. 2009; Pascale et al. 2009; Elbaz et al. 2010; Béthermin et al. 2012), while the strong correlation between the far-IR and radio luminosity is known to hold across a wide range in redshift and luminosity (e.g. Ivison et al. 2010). Up to the present day, most of these techniques have used a maximum-likelihood optimization approach, which suffers from two major issues. The first is that variance and covariance of source fluxes cannot be properly estimated. The second is that of overfitting when many of the input sources are intrinsically faint. The list-driven algorithm developed for HerMES (DESPHOT; Roseboom et al. 2011; Wang et al. 2014) tried to overcome this by using the non-negative weighted least absolute shrinkage and selection operator (LASSO) algorithm (Tibshirani 1996; Zou 2006; ter Braak et al. 2010), a shrinkage and selection method that introduces an additional penalty term in an attempt to reduce the number of sources needed to fit the map. However, when

multiple sources are located close-by (i.e. within the SPIRE beam), the method has been found to wrongly assign all the flux to one source.

The solution to both of these problems is to fully explore the posterior probability distribution with Bayesian inference techniques such as Markov Chain Monte Carlo (MCMC) methods. By fully exploring the posterior, the variance and covariance between sources can be properly estimated. Also, by considering the covariance between sources (i.e. how the flux of sources affect each other), the probability of sources being very faint or bright is taken into account, removing the need for methods such as LASSO.

Up until the present day, use of MCMC techniques has been computationally unfeasible. However, advances in computational technology and algorithms such as Hamiltonian Monte Carlo (HMC) now make this sort of approach a viable alternative, as demonstrated by Safarzadeh et al. (2015), who used an MCMC-based approach to fit PACS simulated maps.

As part of the *Herschel* Extragalactic Legacy Project (HELP; Vaccari 2016, Oliver, in preparation), we have developed an alternative prior-based approach for source extraction in confusion-dominated maps. Our new method, $\text{xID}+$, is built upon a Bayesian probabilistic framework that provides a natural way in which to introduce additional prior information. By using the Bayesian inference tool (*Stan*; Stan Development Team 2015a,b) to sample the full posterior distribution, we are also able to provide more accurate flux density error estimates, whilst avoiding some of the issues associated with the maximum likelihood and LASSO fitting approach used by DESPHOT. In this paper, we show that $\text{xID}+$ outperforms DESPHOT when using just positional information. In Section 2, we discuss the algorithm, and show how the software performs on simulated SPIRE maps in Section 3. In Section 4, we apply $\text{xID}+$ on the HerMES COSMOS SPIRE maps, using a 24- μm catalogue as a prior and show the resulting marginalized SPIRE colour-colour plot and contribution to the CIB. We discuss how $\text{xID}+$ can make use of flux prior information, delivered by the HELP project (Hurley et al. in preparation) in Section 5 and make final conclusions in Section 6.

2 $\text{xID}+$ ALGORITHM

The basic goal of $\text{xID}+$ is to use the SPIRE maps to infer the likely SPIRE flux of sources we already know about. Bayesian inference is well suited to these requirements. It allows the use of prior information and provides a posterior distribution of the parameter(s) after taking into account the observed data.

We also want to provide a framework to do science directly with the maps rather than adding the additional step of first creating catalogues, which in essence is a form of lossy data compression. We therefore adopt a Bayesian probabilistic modelling approach for our $\text{xID}+$ algorithm. It aims to do the following:

- (i) map out the posterior rather than the traditional maximum likelihood point estimate, thereby providing a full account of the flux uncertainty;
- (ii) extend the use of prior information beyond just using positional information about sources.

In the following section, we describe our $\text{xID}+$ algorithm. As this algorithm builds upon knowledge gained from the original xID (a.k.a. DESPHOT) algorithm used by HerMES (Roseboom et al. 2010, 2011; Wang et al. 2014), we describe $\text{xID}+$ in the context of how it differs from DESPHOT.

2.1 Basic model

Our data (\mathbf{d}) are maps with $n_1 \times n_2 = M$ pixels. Our model assumes the maps are formed from S known sources, with flux density \mathbf{f} and a background term accounting for unknown sources. The point response function (PRF) tells us the contribution each source makes to each pixel in the map and is assumed to be a Gaussian, with full width half-maximum (FWHM) of 18.15, 25.15, and 36.3 arcsec for 250, 350, and 500 μm , respectively (Griffin et al. 2010). Our map can therefore be described as follows:

$$\mathbf{d} = \sum_{i=1}^S \mathbf{P} \mathbf{f}_i + N(0, \Sigma_{\text{instrumental}}) + N(B, \Sigma_{\text{confusion}}), \quad (1)$$

where \mathbf{d} is our model of the map, \mathbf{P} is the PRF, f_i is the flux density for source i and two independent noise terms, one for instrumental noise, the other for confusion noise that we model as Gaussian fluctuations about B , a global background.

We can rewrite the above equation in the linear form:

$$\mathbf{d} = \mathbf{A} \mathbf{f}, \quad (2)$$

where \mathbf{d} is flattened to a vector with M pixels, \mathbf{A} is a sparse $M \times S$ matrix pointing matrix. For SPIRE, the pointing matrix is calculated by taking the Gaussian PRF for each band at a 1 arcsec pixel scale, centring it on the position for each source and carrying out a nearest neighbour interpolation to establish the contribution each source makes to each pixel in the map.

As instrumental and confusion noise are independent, we can combine the two noise terms into one covariance matrix such that $\mathbf{N}_d = \Sigma_{\text{instrumental}} + \Sigma_{\text{confusion}}$. Confusion noise will be correlated across nearby pixels due to the PRF and across the three SPIRE bands. Taking these correlations into account requires the full $M \times M$ covariance matrix, which vastly increases computational time. For simplicity, we currently ignore the correlations and assume the confusion noise is constant across the map. The covariance matrix becomes a diagonal matrix, that is $N_{d,ii} = \sigma_{\text{inst},ii}^2 + \sigma_{\text{conf}}^2$.

We can now define the likelihood as the Gaussian probability function for the data given the flux densities:

$$L = p(\mathbf{d}|\mathbf{f}) \propto |\mathbf{N}_d|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{d} - \mathbf{A} \mathbf{f})^T \mathbf{N}_d^{-1} (\mathbf{d} - \mathbf{A} \mathbf{f}) \right\}. \quad (3)$$

The maximum likelihood solution to this equation can be found by setting $\chi = (\mathbf{d} - \mathbf{A} \mathbf{f})^T \mathbf{N}_d^{-1} (\mathbf{d} - \mathbf{A} \mathbf{f})$, finding the minimum and rearranging such that

$$\mathbf{f} = (\mathbf{A}^T \mathbf{N}_d^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{N}_d^{-1} \mathbf{d}. \quad (4)$$

Equation (4) can be solved directly, either by brute-force matrix inversion or via other linear methods. As discussed in Roseboom et al. (2010, 2011) and Wang et al. (2014), linear approaches ignore prior knowledge that fluxes cannot have negative flux density. They are also incapable of discriminating between real and spurious solutions, which can result in overfitting. To overcome these issues, Roseboom et al. (2011) used the non-negative weighted LASSO algorithm (Tibshirani 1996; Zou 2006; ter Braak et al. 2010).

LASSO is a shrinkage and selection method for linear regression and works by treating sources either as ‘inactive’ with flux density set to zero, or ‘active’. It switches sources on one at a time, with the order determined by reduction in chi-squared gained by turning them on. The process continues until some tolerance is reached.

For x_{ID+} , we want to map out the entire posterior, $p(\mathbf{f}|\mathbf{d})$, rather than find the maximum likelihood solution. This has the benefit that

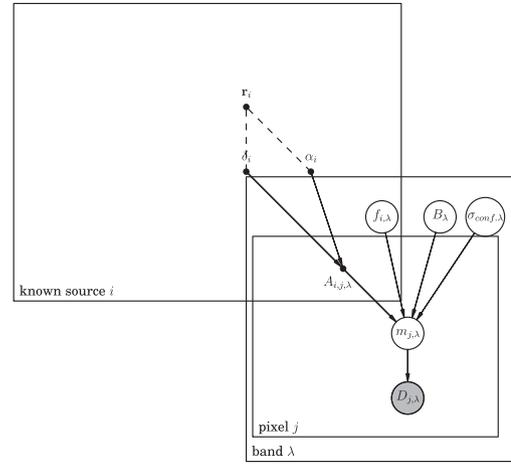


Figure 1. Our probabilistic model for x_{ID+} . Boxes represent repeated dimensions, open circles as variables, dots as deterministic (or fixed) variables. Created with DAFT (<http://daft-pgm.org/>).

it gives us more complete information on how certain we are about the predicted fluxes. The posterior can be defined as

$$p(\mathbf{f}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{f}) \times p(\mathbf{f}), \quad (5)$$

where $p(\mathbf{d}|\mathbf{f})$ is our likelihood, defined in equation (3) and $p(\mathbf{f})$ is our prior on the fluxes. For our simplest model, we use a uniform distribution for $p(\mathbf{f})$, with an upper bound of 1000 mJy and lower bound of 0.01 mJy.

In our probabilistic framework, we can illustrate our model for the map, defined in equation (2) via a probabilistic graphical model (PGM). Fig. 1 shows a plate diagram (Bishop 2006) for our PGM of the basic x_{ID+} model, where boxes indicate repeated values such as source (i), pixel (j), and band (λ). Open circles correspond to random variables and dots are deterministic (or fixed) variables, with their relative positions in the boxes indicating what indices they repeat over. For our simplest model, the positional vector of sources (\mathbf{r}_i) can be described by sky coordinates, α_i and δ_i , and are treated as deterministic (i.e. known).¹ The PRF is assumed to be a Gaussian, with FWHM of 18.15, 25.15, and 36.3 arcsec for 250, 350, and 500 μm , respectively (Griffin et al. 2010). Both these deterministic variables are used to make the pointing matrix $A_{i,j,\lambda}$, which gives the contribution source i makes to each pixel j in the map at wavelength λ . Each source has its own flux $f_{i,\lambda}$, which is a random variable. By multiplying f , A for all sources and pixels, and adding our global estimate for the background B , we can make our model for the map, m , which we can compare to the data D .

2.1.1 Stan

Now that we have our probabilistic model, we need to sample from it to obtain the posterior. We use the Bayesian inference tool, *Stan*, which is ‘a probabilistic programming language implementing full Bayesian statistical inference with MCMC sampling’. *Stan* uses the adaptive HMC No-U-Turn Sampler (NUTS) of Hoffman & Gelman (2014) to efficiently sample from the posterior. It does this

¹ In reality, positional information is often uncertain. However, these uncertainties are relatively small in comparison to the SPIRE beams. For example, the 24-micron source catalogues in Le Floc’h et al. (2009) are accurate to ≈ 2 arcsec in respect to K -band catalogues; this corresponds to a ninth of the FWHM of the SPIRE beam at 250 microns.

by using the gradient information, allowing fast traversing of high dimensional and highly correlated joint posterior distributions.

Stan has its own modelling language, in which one constructs probabilistic models. Our model for *Stan* can be found in Appendix A.

2.1.2 Estimating convergence

As with all MCMC routines, one needs to run enough chains and run them long enough to be confident that the global minimum has been found and that it has been thoroughly sampled. As default, we run four separate chains from different initial positions in parameter space. We also discard the first half of the chain as ‘warm up’ to ensure that the chains have converged to the posterior distribution. We then assess the convergence of each parameter by comparing the variation between and within chains using the diagnostics described in Gelman et al. (2013), which can be summarized as follows: Each chain is split in two and the between-chain (BC) and within-chain (WC) variance is calculated. BC and WC are then used to calculate the marginal posterior variance. This in turn can be used to estimate the potential scale reduction \hat{R} , which reduces to 1 as the number of iterations tends to infinity. An \hat{R} value > 1.2 suggests that chains require more samples. We provide \hat{R} for each parameter.

Due to the nature of MCMC, samples from MCMC routines are correlated. Inference from correlated samples is less precise than from the same number of independent draws. In order to check there are enough independent draws, we estimate the effective number of samples \hat{n}_{eff} , defined in Gelman et al. (2013). We require \hat{n}_{eff} to be 10 times the number of chains and provide the estimate for each parameter.

2.2 Map segmentation

The survey fields in HELP vary in size from 0.3 to 290 deg². Ideally, source photometry and background estimation would be done on the full image. In practice, this will be computationally unfeasible. DESPHOT segmented the map by locating islands of high signal to noise ratio (SNR) pixels enclosed by low SNR pixels.

We adopt a simpler tiling scheme that splits map data into equal area diamonds based on the Hierarchical Equal Area isoLatitude Pixelization of a sphere (HEALPIX). The resolution of the pixels are determined by the HEALPIX level, with default for XID+ set at 11, which corresponds to ≈ 1.718 arcmin. When fitting each tile, the perimeter being fitted is extended by one HEALPIX pixel with a resolution that is two levels higher (i.e. default is level 13 with a resolution of ≈ 25.77 arcmin) such that all sources that could foreseeably contribute to sources within the HEALPIX pixel of interest are taken into account. To give an example of the dimensionality, our fit to a mock simulation (described in the following section) uses a HEALPIX tile at order 9, fitting over 600 sources on average to $\approx 10\,500$ pixels at 250 microns, ≈ 5500 pixels at 350 microns, and ≈ 2500 pixels at 500 microns. That means for each tile we are fitting over 1800 parameters simultaneously.

The choice of HEALPIX pixel size affects the computational time of XID+. The required CPU time is found to scale linearly with the amount of data (i.e. number of image pixels).

2.3 Uncertainties and covariances

DESPHOT provides an estimate of the covariance matrix associated with the fluxes (N_f) from $(A^T N_f^{-1} A)^{-1}$. Due to the Cramer–Rao

inequality, this estimate is a lower limit. It also assumes the DESPHOT algorithm is linear, which is not strictly true having introduced LASSO and non-negative priors. As a result, the uncertainties are unreliable. For XID+, we have the full posterior, allowing the true variance to be properly characterized. This not only gives us a better estimate for marginalized uncertainty for each source, but it also provides the covariance information between sources (as seen in Fig. 6).

3 SIMULATIONS

In order to test and quantify the performance of XID+, we use simulated SPIRE maps of the COSMOS field, a good example of a deep map, that is, where confusion noise ($\sigma_{\text{conf.}}$) is much larger than instrumental noise ($\sigma_{\text{inst.}}$). In order to get realistic clustering, we use the mock catalogues from the latest version of the Durham semi-analytic model, GALFORM (Cowley et al. 2015; Lacey et al. 2016). The model is designed to populate Millennium-class, dark matter only, N -body simulations with a WMAP7 cosmology and minimum halo mass of $1.9 \times 10^{10} h^{-1} M_{\odot}$. The dust model is motivated by the radiative transfer code GRASIL (Silva et al. 1998) and can accurately reproduce the predictions for rest-frame wavelengths $\lambda_{\text{rest}} > 70 \mu\text{m}$. We pass this mock catalogue through the HerMES mapmaker pipeline (e.g. Levenson et al. 2010; Viero et al. 2013) to generate a mock HerMES observation, with similar noise properties to the observed COSMOS field.

A mock 100- μm input catalogue, similar to what would be expected of a PACS catalogue, is generated by taking the mock catalogue and making a cut at a flux limit of 50 μJy (similar to that used for a 24- μm input catalogue), giving a total of 64 719 sources over 3.4 deg². We use this as our prior input catalogue for both XID+ and DESPHOT. In order to compare performance, we look at three measures: precision, flux accuracy, and flux uncertainty accuracy. For XID+, we only consider sources whose output median flux is above 1 mJy. Likewise, with DESPHOT, we only consider sources that have a maximum likelihood flux of greater than 1 mJy.

3.1 Flux precision

Precision is a measure of how well the flux is believed to be constrained. For our posterior sample, this relates to the spread of the sample and so we use the interquartile range (75th–25th percentile; IQR) as our measure of precision. Fig. 2 shows the 16th, 50th, and 84th percentile (i.e. median and median $\pm \sigma$) of the IQR for six bins in true flux. IQR is normalized as a function of input flux for both XID+ and DESPHOT. As one would expect, IQR/ S_{True} decreases as a function of input flux, indicating a higher precision is achieved for the brighter sources. While 250 and 350 μm outputs achieve a similar level of precision, the outputs for 500 μm do not reach the same level of precision. In comparison to DESPHOT, XID+ is marginally less precise for all three bands, though as we show later, DESPHOT’s smaller precision comes at a price of severely underestimating uncertainty.

3.2 Flux accuracy

Flux accuracy is a measure of how far away the estimated flux is from the truth. We use the difference between the median flux estimate from our posterior and the true flux from the simulation, normalized by true flux, as our estimate of flux accuracy. Fig. 3 shows how flux accuracy changes as a function of input flux for all three bands. As before, we show the 16th, 50th, and 84th percentile

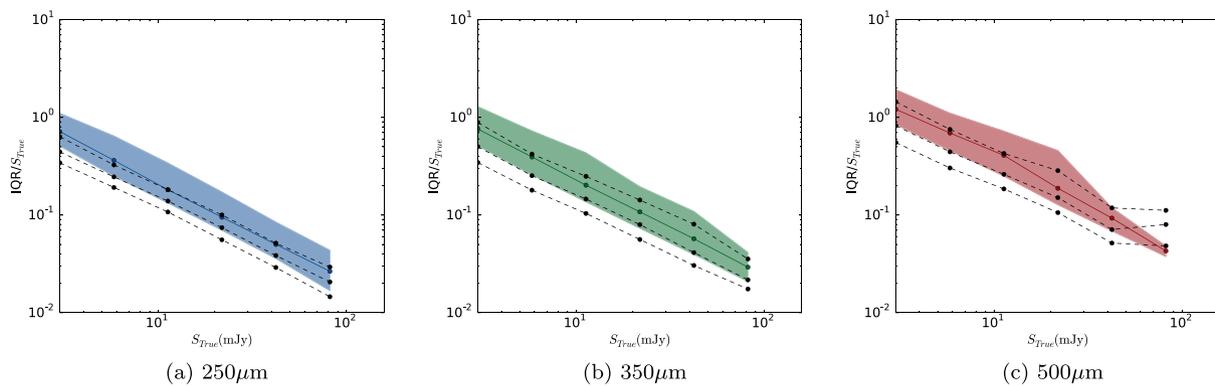


Figure 2. The 16th, 50th, and 84th percentiles for inverse precision, or IQR of *XID+* (coloured line and shaded region) and *DESPHOT* (black dashed lines), as a function of true flux, for the 250- μm (blue), 350- μm (green), and 500- μm (red) SPIRE bands.

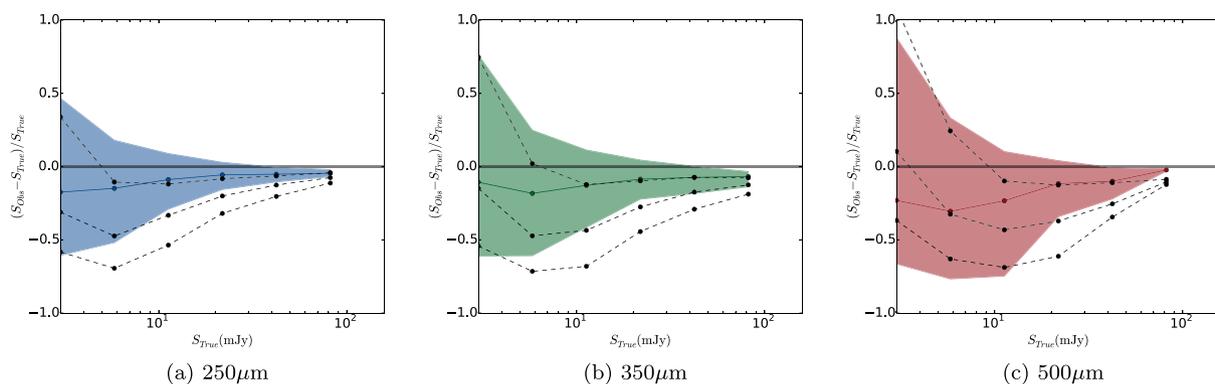


Figure 3. The 16th, 50th, and 84th percentiles for flux accuracy of *XID+* (coloured line and shaded region) and *DESPHOT* (black dashed lines), as a function of true flux, for the 250- μm (blue), 350- μm (green), and 500- μm (red) SPIRE bands. A horizontal thick line is shown at zero for clarity.

for six bins in true flux. For 250 μm , *XID+* reaches an offset smaller than 10 per cent by ≈ 5 mJy, whereas *DESPHOT* underestimates the flux for all but the very brightest sources. For 350 and 500 μm , the offset from the truth is less than 10 per cent by ≈ 10 mJy, whereas *DESPHOT* continues to underestimate for all but the very brightest sources. There remains a slight offset of ≈ 5 per cent for all fluxes. This is likely due to our inability to model the correlated component of the confusion noise.

3.3 Flux uncertainty accuracy

Estimated flux values should be within 1σ of the true value 68.27 per cent of the time and within 2σ 95.45 per cent of the time. We can quantify how many sigma away the true value is from the median, in terms of a *Z* score. A *Z* score of 1 corresponding to being 1σ above the median. Fig. 4 shows the flux uncertainty accuracy (or *Z* score) as a function of input flux. For *DESPHOT*, uncertainties are assumed to have a normal distribution, truncated at zero. With *XID+*, we have the full posterior and do not have to make an assumption on the shape of the uncertainty distribution. If we assume the posterior has good frequentist coverage, then we can calculate uncertainty accuracy by taking the percentile at which the true flux value falls within the posterior, and convert the percentile to a corresponding sigma level.

For the 250- μm band, and sources $\lesssim 25$ mJy, *XID+* produces a *Z* score distribution that is slightly above that expected if uncertainties are correctly estimated (i.e. distribution is centred around zero, with width ≈ 1). Above 25 mJy, the median *Z* score increases, indicating

that flux uncertainties are being underestimated. In comparison, for all fluxes, the uncertainty distribution from *DESPHOT* are above 1 and increases to over 2 for the brighter sources. There are also a large number of sources with a *Z* score greater than 3 (as seen by the higher density in bins at a *Z* score of 3). This indicates that *DESPHOT* is a poor estimator with the majority of sources in *DESPHOT* lying more than 1σ away from their true flux. The flux uncertainty accuracies for 350 and 500 μm show a similar behaviour, though not as severe.

3.4 Convergence

As described in Section 2.1.2, we provide \hat{R} as an estimate of convergence and \hat{n}_{eff} as a measure of independence within the sample. Fig. 5 show the histogram for \hat{R} and \hat{n}_{eff} for the three bands, with our thresholds for the statistics shown by dotted lines. In our fit to the simulated SPIRE maps, we use four chains, each with 1500 iterations (half of which are discarded as warm up). This leads to over 99.99 per cent of the sources having an \hat{R} and \hat{n}_{eff} within the threshold for all three bands, indicating our solution is well converged. In cases where convergence has not been reached, the number of iterations can be increased.

3.5 Correlated sources

For sources that are close together (i.e. within FWHM of the PRF), the uncertainty on the flux estimates can be correlated. One of the advantages of obtaining the full posterior is that we get a proper

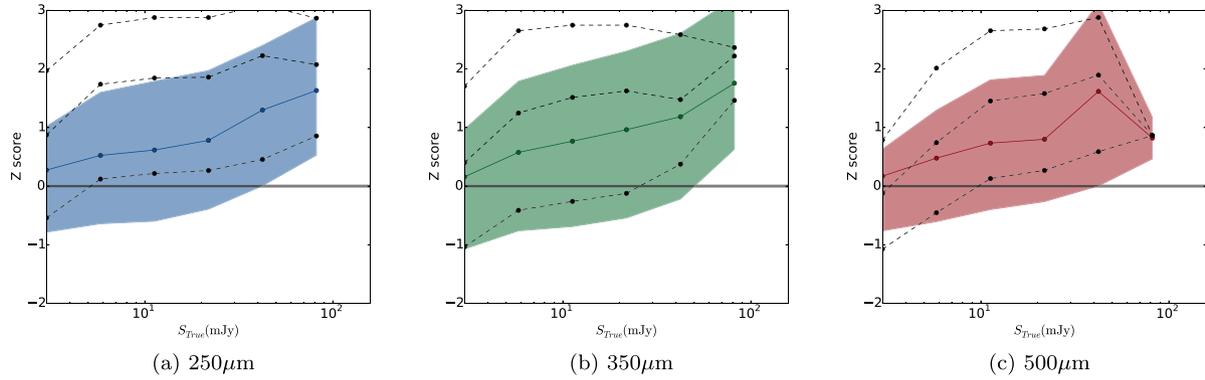


Figure 4. The 16th, 50th, and 84th percentiles for Z score, or flux density error of *XID+* (coloured line and shaded region) and *DESPHOT* (black dashed lines), as a function of true flux, for the 250- μm (blue), 350- μm (green) and 500- μm (red) SPIRE bands. A horizontal thick line is shown at zero for clarity.

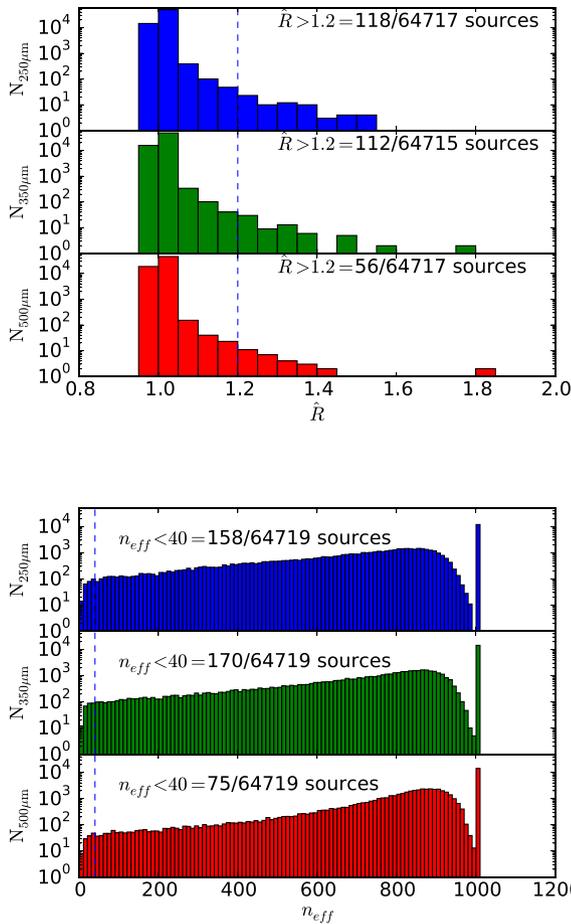


Figure 5. \hat{R} and \hat{n}_{eff} values for all sources fitted in the simulation. The majority of sources have converged and have enough effective samples.

estimate of uncertainty and its correlation. This is particularly apparent when comparing flux estimates with *DESPHOT*, which, by using the LASSO algorithm forces one source to have all the flux and the other nearby source to zero. Fig. 6 shows an example of two sources that are 2 arcsec apart. The 250- μm flux estimate from both *XID+* and *DESPHOT* are shown in green and blue, respectively. The posterior provided by *XID+*, fully captures the correlated uncertainty, whereas the ‘winner takes all’ approach from *DESPHOT* clearly fails to estimate the true flux for both sources.

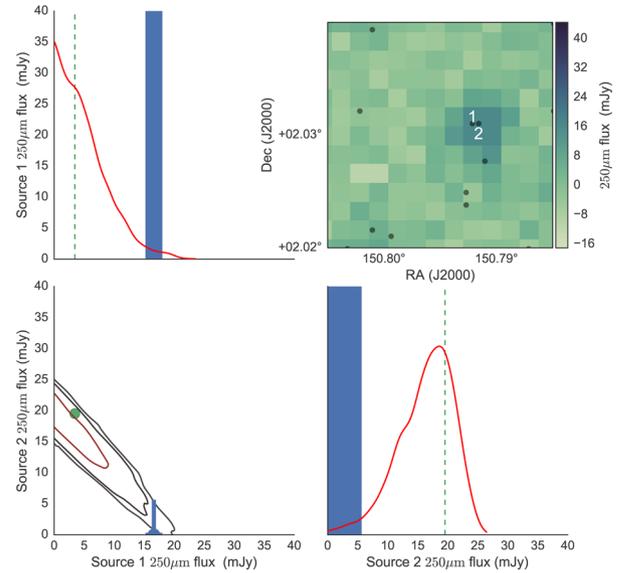


Figure 6. Joint and marginalized posterior plot of two correlated sources that are 2 arcsec apart, with 1σ , 2σ and 3σ contours overlotted. The true flux is shown by the green circle and green dashed lines. *DESPHOT* (blue error bar and vertical filled region spanning 1σ) assigns all the flux to one source which is actually the fainter of the two, whereas with *XID+* we get the full uncertainty information from the posterior.

3.6 Defining detections

For sources close to or below the noise level of the map, the data will be unable to constrain our model. The flux posterior for these sources will be a subtle combination of the flux prior and an upper limit imposed by the noise level of the map. This results in a non-Gaussian posterior flux distribution for those sources that cannot be constrained. To illustrate how the shape of the posterior flux distribution changes, Fig. 7 shows the $\frac{84\text{th}-50\text{th}}{50\text{th}-16\text{th}}$ percentiles of the flux posterior distribution as a function of the 50th percentile for the three SPIRE bands. For 250 and 350 μm , uncertainties become Gaussian around 4 mJy, while for 500 μm , it occurs around 6 mJy. In principle, the full posterior distribution could be used for all objects, including those not fully constrained by the data. However, Fig. 7 provides a convenient way in which to define the level at which there are robust detections. We note that these limits will change depending on prior list and the maps being fitted.

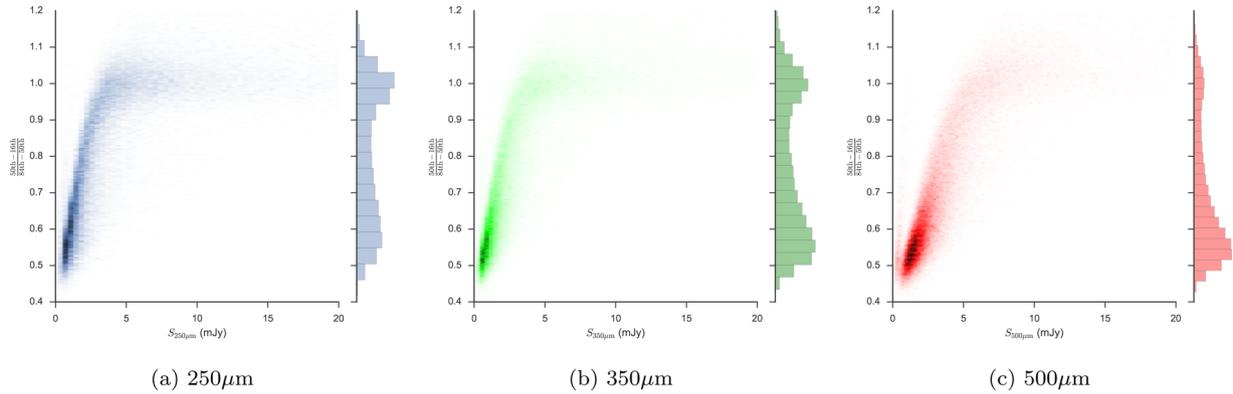


Figure 7. The $\frac{84\text{th}-50\text{th}}{50\text{th}-16\text{th}}$ percentiles of the flux posterior distribution as a function of the 50th percentile for the three SPIRE bands. The posterior is approximately Gaussian when the ratio is around 1. For 250 and 350 μm , this occurs at around 4 mJy, while for 500 μm , it occurs around 6 mJy.

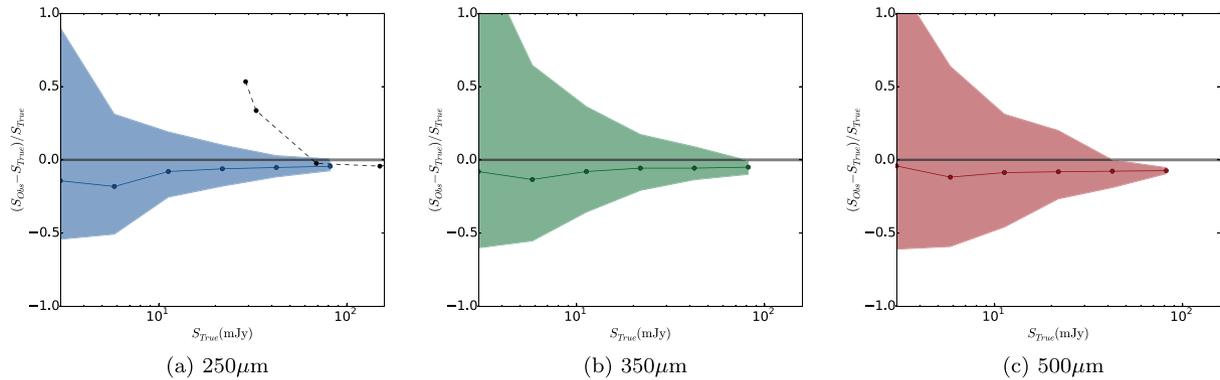


Figure 8. The 16th, 50th, and 84th percentiles for flux accuracy of *xid+* (coloured line and shaded region) and the median flux accuracy for LAMBDA R (black dashed line), as a function of true flux, for the 250- μm (blue), 350- μm (green) and 500- μm (red) SPIRE bands. A horizontal thick line is shown at zero for clarity.

3.7 Comparison to stacking

To further demonstrate the performance of *xid+*, we compare the average *xid+* fluxes for objects grouped into six stellar mass and six redshift bins, against those from stacking. Stacking takes the position of known sources and for each source, cuts out a thumbnail from a map centred on the source position. By averaging the thumbnail images, one reduces the noise and produces an image of the average galaxy and an average flux estimate for the list of known sources used to make the stack. Fig. B1 in Appendix B shows the 250, 350, and 500 μm stacked, average true flux from the simulation, and average *xid+* fluxes for 36 different groups of objects. For objects above the detection limit defined in Section 3.6, the average flux of the groups from *xid+* (in red) is able to recover the true average flux (in green), all be it with the same minor underestimate (≈ 5 per cent) seen in the flux accuracy plots. In comparison, the stacked flux (in blue) performs at a similar level. As expected, the performance for both methods deteriorate when the number of objects in the stack is low. This stacking comparison shows that for groups of objects above the detection limit, *xid+* reaches a comparable performance to stacking, but with the distinct advantage in having the flux for every objects, compared to stacking which provides an average flux, assuming the underlying flux distribution is Gaussian.

3.8 Performance with a shallower prior list

So far, we have shown the performance of *xid+* with the type of prior lists that are available for the *HerMES* fields (i.e. both relatively

deep and selected at a similar wavelength to the SPIRE bands). We now demonstrate the expected performance of *xid+* for H-ATLAS fields, where the prior list is likely to be derived from shallow, optical ancillary data.

We use the same simulated maps as before, but now our prior list only use sources with an *r*-band magnitude < 19.8 or a 250- μm flux > 15.48 mJy. With this cut, our prior list contains 5536 sources. Not only does this cut mimic the type of ancillary data available in the H-ATLAS fields (i.e. shallow optical data and additional bright SPIRE sources detected via blind source detection), but is the same cut used to test the Lambda Adaptive Multi-band Deblending Algorithm in *R* (LAMBDA R; Wright et al. 2016).

Fig. 8 shows the flux accuracy of *xid+*, using the shallower prior list. Although the dispersion in accuracy is worse at the lower fluxes, the median accuracy remains close to zero for all fluxes. For the 250 μm , we compare our flux accuracy with that obtained by LAMBDA R. While the flux accuracy of *xid+* is relatively constant from 10 mJy, the flux accuracy of LAMBDA R only reaches comparable accuracy levels at fluxes greater than 70 mJy, illustrating *xid+* can reach far lower flux levels than LAMBDA R. We note both precision and flux uncertainty accuracy are similar to that obtained in Sections 3.1 and 3.3.

4 COSMOS FIELD

Having satisfactorily demonstrated the performance on simulations, we have run *xid+* on the *HerMES* COSMOS SPIRE maps from the

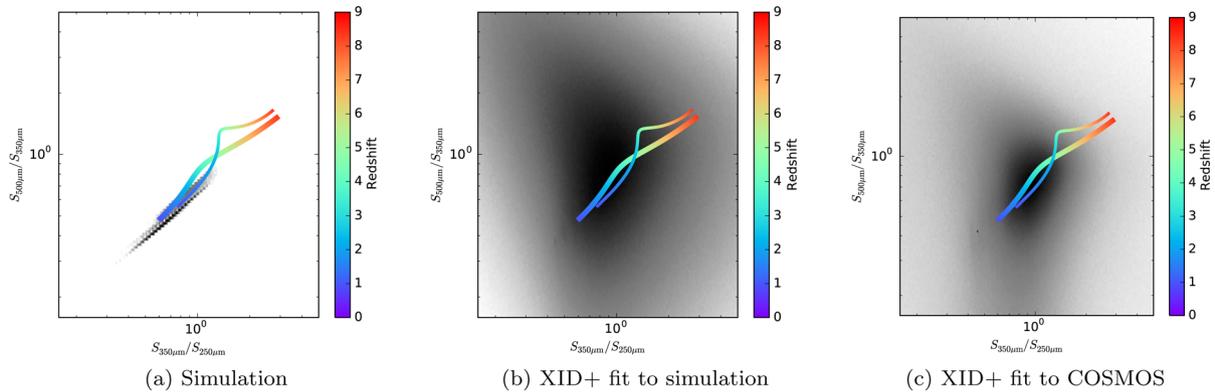


Figure 9. The marginalized SPIRE colour–colour probability density (in black) for the sources used in the simulation (left), `xID+` fit to simulated sources (centre), and `xID+` fit to MIPS 24- μ m sources in the COSMOS field (right). Overplotted are the redshifted spectral energy distributions for a red star-forming galaxy (thick line) and blue star-forming galaxy (thin line) as defined by Berta et al. (2013).

Table 1. Contribution to the CIRB at the SPIRE wavelengths from various measures, including the absolute measurement made by FIRAS.

	CIRB 250 μ m ($\text{nWm}^{-2}\text{sr}^{-1}$)	CIRB 350 μ m ($\text{nWm}^{-2}\text{sr}^{-1}$)	CIRB 500 μ m ($\text{nWm}^{-2}\text{sr}^{-1}$)
FIRAS (Lagache & Puget 2000)	11.8 ± 2.9	6.4 ± 1.6	2.7 ± 0.7
SPIRE-resolved sources (Oliver et al. 2010)	1.73 ± 0.33 (15 per cent)	0.63 ± 0.18 (10 per cent)	0.15 ± 0.07 (6 per cent)
<code>xID+</code> with 24- μ m sources	5.573 ± 0.003 (47 per cent)	2.805 ± 0.002 (45 per cent)	1.24 ± 0.002 (46 per cent)
Stacking of 24- μ m sources (B��thermin et al. 2012)	$7.40^{+1.42}_{-1.43}$ (73 per cent)	$4.50^{+0.90}_{-0.90}$ (63 per cent)	$1.54^{+0.30}_{-0.30}$ (55 per cent)

2nd Data release. As a prior, we take the MIPS 24- μ m catalogue (Le Flocc’h et al. 2009), which covers an area of 2.265 deg² and includes 52 092 sources with a 24- μ m flux density above 60 μ Jy, which corresponds to a signal-to-noise cut of 3.

Fig. 9 shows the marginalized probability density of our MIPS 24- μ m prior catalogue in SPIRE colour–colour space and is constructed by combining the 1500 samples from the posterior, for all 52 092 sources. The redshift tracks for a red and blue star-forming galaxy spectral energy distribution template, empirically derived from Herschel sources (Berta et al. 2013), are overplotted and run through the highest density region at redshifts of around 2 to 3. As a comparison, we show the true colour distribution for our mock simulation used in Section 3 and the corresponding distribution from the `xID+` fit. The probability distribution from the fit to the simulation is far wider than the truth. This is not surprising as, at present, the three SPIRE bands are being fit independently, and so any correlation in colour is being dispersed by confusion noise and non-detections. Interestingly, the fit to the real COSMOS data shows a more constrained probability distribution. Whether this is a consequence of slightly different selection, or less variation in spectral energy distribution shape will require further investigation.

Table 1 shows the total contribution to the CIB at 250, 350, and 500 μ m from our MIPS 24- μ m prior catalogue, alongside the Far Infrared Absolute Spectrophotometer (FIRAS) absolute measurements from Lagache & Puget (2000), the contribution to the CIRB from SPIRE-resolved sources (Oliver et al. 2010) and the contribution from stacking MIPS 24- μ m sources (B  thermin et al. 2012). By going to a depth of 60 μ Jy in the 24- μ m catalogue, we can explain 47, 44, and 46 per cent of the nominal measured values at 250, 350, and 500 μ m (Lagache & Puget 2000). This compares favourably to the 15, 10, and 6 per cent that SPIRE resolves at the (40 beams)^{−1} depth (Oliver et al. 2010). We resolve less of the CIRB than achieved by stacking MIPS 24 μ m, with the missing contribution belonging to those sources that `xID+` cannot constrain.

However, unlike stacking, we know how that flux is distributed amongst sources. The remaining CIRB that is not associated with micrometre sources, will be coming from other sources not detected at 24 μ m.

Our final data product consists of a catalogue, summarizing the SPIRE fluxes via the 16th, 50th, and 84th percentile (i.e. median and median $\pm\sigma$); the median background; and the convergence statistics. We also make available on request, the 3000 samples from the posterior probability distribution, each of which can be thought of as a possible catalogue in probability space.

5 DISCUSSION

We have run `xID+` on simulated maps of the COSMOS field and compared it to `DESPHOT` using three main metrics: flux accuracy, precision, and uncertainty accuracy. On accuracy, `xID+` performs significantly better in all three bands, and although appearing marginally less precise, the loss of precision relates to more realistic estimates for flux uncertainties.

The higher performance gained by `xID+` comes from fully capturing the posterior probability distribution on flux estimates. By exploring the posterior, we get a proper handle on uncertainties and no longer have to employ penalisation techniques such as LASSO, which are known to behave erratically.

By using a probabilistic approach, we have a framework where we can introduce prior information on the source fluxes in a transparent manner. For this basic version, we use a simple uniform prior, with bounds at 10^{−2} and 10³ mJy. However, as can be seen in Fig. 6, where sources are correlated, if we have prior knowledge on the flux of one of the sources, it can help us determine a more precise flux for the other.

As demonstrated by Safarzadeh et al. (2015), these priors could come from fitting spectral energy distribution models to multiwavelength ancillary data. Another alternative is to use machine-learning

algorithms to ‘learn’ the expected flux from the statistical population. As part of HELP, the testing and benchmarking of suitable methods for deriving SPIRE flux priors will be presented in Hurley et al. (in preparation).

More generally, the probabilistic model used in XID+ can easily be expanded, allowing distributions such as the flux distribution (or number counts) to be modelled explicitly. In principle, and with additional information such as redshifts, the probabilistic model could become detailed enough to simultaneously fit luminosity functions and the location of locus of the star formation rate M^* relation at different redshift and we will explore these expansions in future papers.

6 CONCLUSIONS

In this paper, we have introduced the prior-based source detection software, XID+. By using the Bayesian inference tool *Stan*, we are able to fully sample the posterior probability distribution, which in turn gives a better understanding of the uncertainty associated with the source flux.

Having run XID+ on simulated maps, we have shown this is extremely advantageous for maps that are confusion limited, such as the *Herschel* observations that are part of *HerMES*. In comparison to the current maximum likelihood based software *DESPHOT*, XID+ performs far better in all three main metrics: flux accuracy, precision, and uncertainty accuracy.

We have run XID+ on the *HerMES* COSMOS SPIRE maps from the 2nd Data release, using the MIPS 24- μm catalogue (Le Floch et al. 2009) as our prior. Using the full posterior, we have created a marginalized SPIRE colour–colour plot, illustrating the probability distribution of our MIPS 24- μm catalogue in SPIRE colour–colour space. We have also shown that the MIPS 24- μm sources contribute 47, 44, and 46 per cent to the CIB at 250, 350, and 500 μm . We provide the catalogue as part of the fourth *HerMES* data release² and posterior probability distribution samples as a data product as part of HELP.³ As far as we are aware, this is the first time the full posterior probability distribution is made available as a data product for list-driven photometry.

ACKNOWLEDGEMENTS

We thank the anonymous referee for suggestions that greatly enhanced this work. The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013/ under grant agreement number 607254. This publication reflects only the author’s view and the European Union is not responsible for any use that may be made of the information contained therein.

SO acknowledges support from the Science and Technology Facilities Council (grant number ST/L000652/1) and MB is supported under EPSRC grant EP/J016934/1. HCSS/HSpot/HIPE are joint developments by the *Herschel* Science Ground Segment Consortium, consisting of ESA, NASA *Herschel* Science Center, and HIFI, PACS and SPIRE consortia.

SPIRE has been developed by a consortium of institutes led by Cardiff University (UK) and including University of Lethbridge (Canada); NAOC (China); CEA, LAM (France); IFSI, University of Padua (Italy); IAC (Spain); Stockholm Observatory (Sweden); Imperial College London, RAL, UCL-MSSL, UKATC, University

of Sussex (UK); and Caltech, JPL, NHSC, University of 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, UKSA (UK); and NASA (USA). This research has made use of data from *HerMES* project (<http://hermes.sussex.ac.uk/>). *HerMES* is a *Herschel* Key Programme utilizing Guaranteed Time from the SPIRE instrument team, ESAC scientists and a mission scientist. The *HerMES* data were accessed through the *Herschel* Database in Marseille (HeDaM – <http://hedam.lam.fr>) operated by CeSAM and hosted by the Laboratoire d’Astrophysique de Marseille.

This work is based in part on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. Special thanks goes to Yannick Roehly for organization of code, and Louise Winters for management within HELP. This paper make use of *MATPLOTLIB* (Hunter 2007) and *ASTROPY*, a community-developed core PYTHON package for Astronomy (Astropy Collaboration et al. 2013).

REFERENCES

- Astropy Collaboration et al., 2013, *A&A*, 558, A33
 Berta S. et al., 2013, *A&A*, 551, A100
 Béthermin M. et al., 2012, *A&A*, 542, A58
 Bishop C. M., 2006, *Pattern Recognition and Machine Learning*. Springer Science+ Business Media, New York
 Chapin E. L. et al., 2011, *MNRAS*, 411, 505
 Cowley W. L., Lacey C. G., Baugh C. M., Cole S., 2015, *MNRAS*, 446, 1784
 Eales S. et al., 2010, *PASP*, 122, 499
 Elbaz D. et al., 2010, *A&A*, 518, L29
 Gelman A., Carlin J., Stern H., Dunson D., Vehtari A., Rubin D., 2013, *Bayesian Data Analysis*, 3rd edn. Chapman & Hall/CRC Texts in Statistical Science, Boca Raton, FL
 Griffin M. J. et al., 2010, *A&A*, 518, L3
 Hoffman M. D., Gelman A., 2014, *J. Mach. Learn. Res.*, 15, 1593
 Holland W. S. et al., 1999, *MNRAS*, 303, 659
 Holland W. S. et al., 2013, *MNRAS*, 430, 2513
 Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
 Ivison R. J. et al., 2010, *A&A*, 518, L31
 Kessler M. F. et al., 1996, *A&A*, 315, L27
 Lacey C. G. et al., 2016, *MNRAS*, 462, 3854
 Lagache G., Puget J. L., 2000, *A&A*, 355, 17
 Le Floch E. et al., 2009, *ApJ*, 703, 222
 Levenson L. et al., 2010, *MNRAS*, 409, 83
 Marsden G. et al., 2009, *ApJ*, 707, 1729
 Murakami H. et al., 2007, *PASJ*, 59, 369
 Neugebauer G. et al., 1984, *ApJ*, 278, L1
 Oliver S. J. et al., 2010, *A&A*, 518, L21
 Oliver S. J. et al., 2012, *MNRAS*, 424, 1614
 Pascale E. et al., 2009, *ApJ*, 707, 1740
 Pilbratt G. L. et al., 2010, *A&A*, 518, L1
 Poglitsch A. et al., 2010, *A&A*, 518, L2
 Puget J.-L., Abergel A., Bernard J.-P., Boulanger F., Burton W. B., Desert F.-X., Hartmann D., 1996, *A&A*, 308, L5
 Roseboom I. G. et al., 2010, *arXiv*
 Roseboom I. G. et al., 2011, *MNRAS*, 419, 2758
 Safarzadeh M., Ferguson H. C., Lu Y., Inami H., Somerville R. S., 2015, *ApJ*, 798, 91
 Silva L., Granato G. L., Bressan A., Danese L., 1998, *ApJ*, 509, 103
 Soifer B. T. et al., 1984, *ApJ*, 278, L71
 Stan Development Team 2015a, *PyStan: The Python Interface to Stan*, Version 2.7.0
 Stan Development Team 2015b, *Stan: A C++ Library for Probability and Sampling*, Version 2.7.0
 ter Braak C. J., Boer M. P., Totir L. R., Winkler C. R., Smith O. S., Bink M. C., 2010, *Genetics*, 185, 1045

² <http://hedam.lam.fr/HerMES/index/dr4>

³ Data products will be made publicly available at the end of HELP in December 2017; however, early access can be granted on request.

Tibshirani R., 1996, J. R. Stat. Soc. B, 58, 267

Vaccari M., 2016, Proc. SALT Sci. Conf. 2015, Stellenbosch Institute of Advanced Study, South Africa. Available at: <http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=250>

Viero M. P. et al., 2013, ApJ, 772, 77

Wang L. et al., 2014, MNRAS, 444, 2870

Werner M. W. et al., 2004, ApJS, 154, 1

Wright E. L. et al., 2010, AJ, 140, 1868

Wright A. H. et al., 2016, MNRAS, 460, 765

Zou H., 2006, J. Am. Stat. Assoc., 101, 1418

APPENDIX A

Stan code for xid+.

```
//Full Bayesian inference fit XID
data {
  int<lower=0> nsrc; //number of sources
  //-----PSW-----
  int<lower=0> npix_psw; //number of pixels
  int<lower=0> nnz_psw; //number of non neg entries in A
  vector[npix_psw] db_psw; //flattened map
  vector[npix_psw] sigma_psw; //flattened uncertainty map (assuming no covariance between pixels)
  real bkg_prior_psw; //prior estimate of background
  real bkg_prior_sig_psw; //sigma of prior estimate of background
  vector[nnz_psw] Val_psw; //non neg values in image matrix
  int Row_psw[nnz_psw]; //Rows of non neg values in image matrix
  int Col_psw[nnz_psw]; //Cols of non neg values in image matrix
  vector[nsrc] f_low_lim_psw; //upper limit of flux
  vector[nsrc] f_up_lim_psw; //upper limit of flux
  //-----PMW-----
  int<lower=0> npix_pmw; //number of pixels
  int<lower=0> nnz_pmw; //number of non neg entries in A
  vector[npix_pmw] db_pmw; //flattened map
  vector[npix_pmw] sigma_pmw; //flattened uncertainty map (assuming no covariance between pixels)
  real bkg_prior_pmw; //prior estimate of background
  real bkg_prior_sig_pmw; //sigma of prior estimate of background
  vector[nnz_pmw] Val_pmw; //non neg values in image matrix
  int Row_pmw[nnz_pmw]; //Rows of non neg values in image matrix
  int Col_pmw[nnz_pmw]; //Cols of non neg values in image matrix
  vector[nsrc] f_low_lim_pmw; //upper limit of flux (in log10)
  vector[nsrc] f_up_lim_pmw; //upper limit of flux (in log10)
  //-----PLW-----
  int<lower=0> npix_plw; //number of pixels
  int<lower=0> nnz_plw; //number of non neg entries in A
  vector[npix_plw] db_plw; //flattened map
  vector[npix_plw] sigma_plw; //flattened uncertainty map (assuming no covariance between pixels)
  real bkg_prior_plw; //prior estimate of background
  real bkg_prior_sig_plw; //sigma of prior estimate of background
  vector[nnz_plw] Val_plw; //non neg values in image matrix
  int Row_plw[nnz_plw]; //Rows of non neg values in image matrix
  int Col_plw[nnz_plw]; //Cols of non neg values in image matrix
  vector[nsrc] f_low_lim_plw; //upper limit of flux
  vector[nsrc] f_up_lim_plw; //upper limit of flux
}

parameters {
  vector<lower=0.0, upper=1.0> [nsrc] src_f_psw; //source vector
  real bkg_psw; //background
  vector<lower=0.0, upper=1.0> [nsrc] src_f_pmw; //source vector
  real bkg_pmw; //background
  vector<lower=0.0, upper=1.0> [nsrc] src_f_plw; //source vector
  real bkg_plw; //background
  real<lower=0.0, upper=8> sigma_conf_psw;
  real<lower=0.0, upper=8> sigma_conf_pmw;
  real<lower=0.0, upper=8> sigma_conf_plw;
}
```

```

model {
  vector[npix_psw] db_hat_psw; //model of map
  vector[npix_pmw] db_hat_pmw; //model of map
  vector[npix_plw] db_hat_plw; //model of map

  vector[npix_psw] sigma_tot_psw;
  vector[npix_pmw] sigma_tot_pmw;
  vector[npix_plw] sigma_tot_plw;

  vector[nsrc] f_vec_psw; //vector of source fluxes
  vector[nsrc] f_vec_pmw; //vector of source fluxes
  vector[nsrc] f_vec_plw; //vector of source fluxes
  // Transform to normal space. As I am sampling variable then transforming I don't need a Jacobian adjustment
  for (n in 1:nsrc) {
    f_vec_psw[n] <- f_low_lim_psw[n] + (f_up_lim_psw[n] - f_low_lim_psw[n]) * src_f_psw[n];
    f_vec_pmw[n] <- f_low_lim_pmw[n] + (f_up_lim_pmw[n] - f_low_lim_pmw[n]) * src_f_pmw[n];
    f_vec_plw[n] <- f_low_lim_plw[n] + (f_up_lim_plw[n] - f_low_lim_plw[n]) * src_f_plw[n];
  }

  //Prior on background
  bkg_psw ~ normal(0,1);
  bkg_pmw ~ normal(0,1);
  bkg_plw ~ normal(0,1);

  // Create model maps (i.e. db_hat = A*f) using sparse multiplication
  for (k in 1:npix_psw) {
    db_hat_psw[k] <- bkg_psw*bkg_prior_sig_psw+bkg_prior_psw;
    sigma_tot_psw[k] <- sqrt(square(sigma_psw[k]) + square(sigma_conf_psw));
  }
  for (k in 1:nnz_psw) {
    db_hat_psw[Row_psw[k]+1] <- db_hat_psw[Row_psw[k]+1] + Val_psw[k]*f_vec_psw[Col_psw[k]+1];
  }

  for (k in 1:npix_pmw) {
    db_hat_pmw[k] <- bkg_pmw*bkg_prior_sig_pmw+bkg_prior_pmw;
    sigma_tot_pmw[k] <- sqrt(square(sigma_pmw[k]) + square(sigma_conf_pmw));
  }
  for (k in 1:nnz_pmw) {
    db_hat_pmw[Row_pmw[k]+1] <- db_hat_pmw[Row_pmw[k]+1] + Val_pmw[k]*f_vec_pmw[Col_pmw[k]+1];
  }

  for (k in 1:npix_plw) {
    db_hat_plw[k] <- bkg_plw*bkg_prior_sig_plw+bkg_prior_plw;
    sigma_tot_plw[k] <- sqrt(square(sigma_plw[k]) + square(sigma_conf_plw));
  }
  for (k in 1:nnz_plw) {
    db_hat_plw[Row_plw[k]+1] <- db_hat_plw[Row_plw[k]+1] + Val_plw[k]*f_vec_plw[Col_plw[k]+1];
  }

  // likelihood of observed map|model map
  db_psw ~ normal(db_hat_psw, sigma_tot_psw);
  db_pmw ~ normal(db_hat_pmw, sigma_tot_pmw);
  db_plw ~ normal(db_hat_plw, sigma_tot_plw);
}

```

APPENDIX B

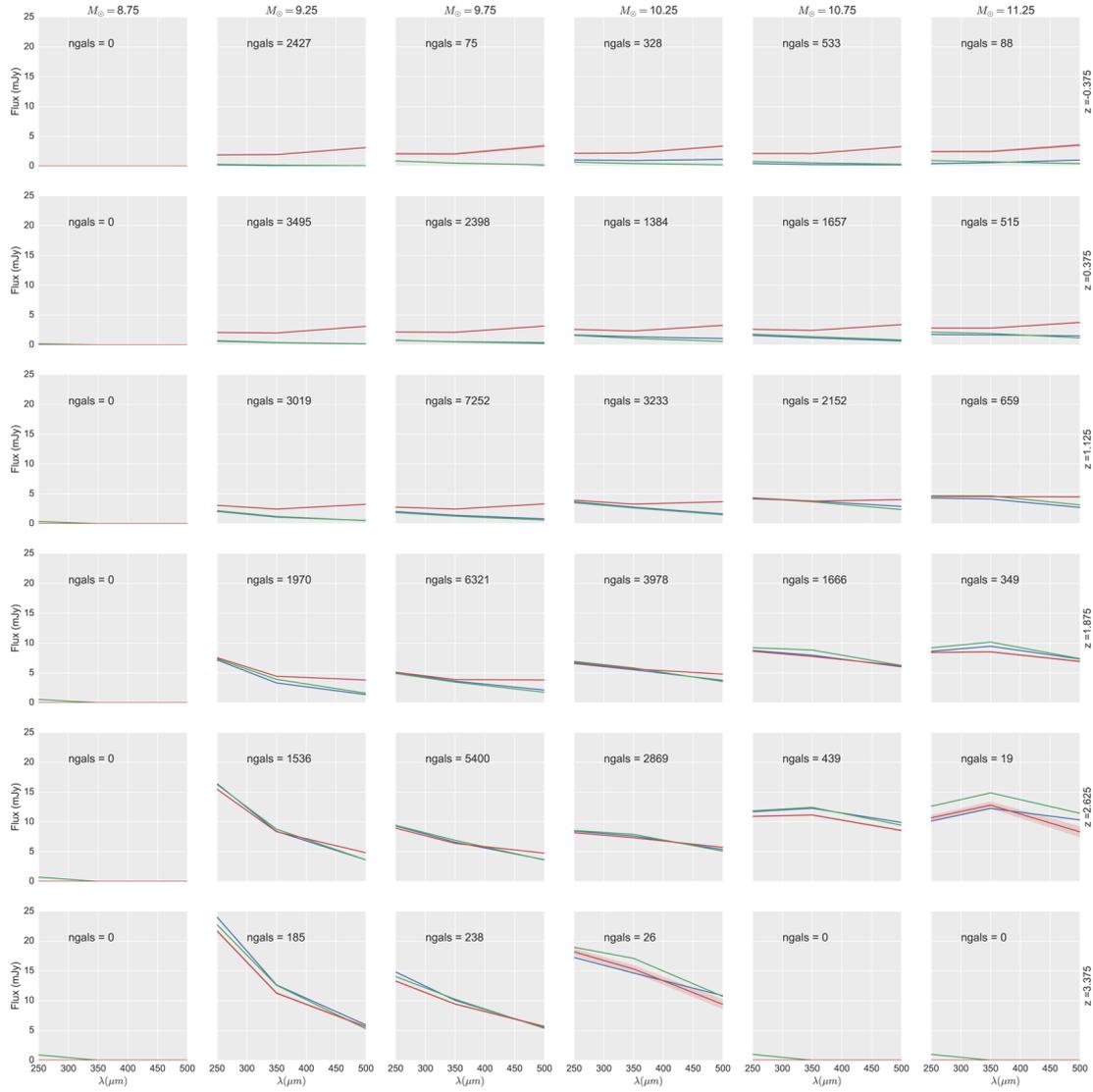


Figure B1. The average flux from simulated objects, grouped by stellar mass and redshift. The flux from stacking is shown in blue, the average true flux for each group in green, and the average $\chi\text{ID}+$ flux for the group in red. For average fluxes above the detection limit, $\chi\text{ID}+$ is just as capable of returning true average flux as stacking, but with the added advantage of having the flux for each object.

This paper has been typeset from a $\text{T}_{\text{E}}\text{X}/\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$ file prepared by the author.