Mon. Not. R. Astron. Soc. **000**, 1–?? (2002) Printed 4 February 2015 (MN LaTeX style file v2.2)

HELP: Probabilistic De-blender for Herschel SPIRE maps

P.D. Hurley,[1]* S. Oliver,[1]
[1]Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Falmer, Brighton BN1 9QH, UK

Released 2002 Xxxxx XX

ABSTRACT
The Herschel Extragalactic Legacy Project (HELP) will provide ancillary data from other wavelengths alongside the extra Software available at `https://github.com/pdh21/XID_plus/`.

Key words: galaxies: statistics – infrared: galaxies

1 INTRODUCTION

2 DATA

3 XID+ ALGORITHM

3.1 XID and DESPHOT

Although the XID+ algorithm presented in this paper takes different approaches, it builds upon knowledge gained from developing the original XID (a.k.a DESPHOT) algorithm used by HerMES (**???**). Details of the algorithm can be found in the corresponding papers, but for convenience we list some of the main details.

DESPHOT consists of following main steps: map segmentation, source photometry and noise estimation.

3.1.1 Map segmentation

Ideally, source photometry and background estimation would be done on full image, in practice it is often computationally infeasible. DESPHOT segmented the map by locating islands of high SNR pixels enclosed by low SNR pixels. The segmentation algorithm operates thus:

- Locates all pixels with a SNR above some threshold (default value of SNR= 1);
- Takes the first of these high SNR pixel starting in the bottom left corner of the image;
- 'Grows' a region around this pixel by iteratively taking neighbouring high SNR pixels;
- Once there are no more high SNR neighbours jumps to the next high SNR pixel and repeat from step (iii). Each of these independent regions of high SNR pixels is uniquely identified and processed separately by the source photometry component.

3.1.2 Source photometry

DESPHOT assumes the map can be described by the sum of the flux densities from n known sources, a global background level and some unknown noise term

$$\mathbf{d} = \sum_{i=1}^{n} \mathbf{P_i} f_i + B + \delta \qquad (1)$$

where $\mathbf{d}$ is the image, $\mathbf{P_i}$ is the PRF for source i, f_i is the flux density for source i, B is a global background estimate and δ is the noise term. As this is a linear equation, it has a maximum likelihood solution which can be solved directly by matrix inversion or via other linear methods. As discussed in **???**), these approaches ignore prior knowledge that fluxes cannot have negative flux density, which in very degenerative cases can result in any symmetric pairing of positive and negative flux providing a good fit. They are also incapable of discriminating between real and spurious sources, which can result in overfitting. To overcome these issues, **?**) used the non-negative weighted LASSO algorithm (Tibshirani 1996; Zou 2006; ter Braak et al 2010).

LASSO works by treating sources either 'inactive' and 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.

In the first iteration, DESPHOT uses LASSO on each segment, to estimate the source fluxes. It then estimates a value for the background (B) via

$$B = \mathbf{d} - \sum_{i=1}^{n} \mathbf{P_i} f_i \qquad (2)$$

The estimate from B is subtracted, and the LASSO fitting is rerun to get the final flux density estimates.

* Email: p.d.hurley@sussex.ac.uk

3.1.3 Noise estimation

If DESPHOT is assumed to be linear[1], then one can get a lower limit on the noise from $(\mathbf{A^T N_d^{-1} A})^{-1}$. This estimate only includes instrumental noise and degeneracies between sources. An estimate of the remaining residual confusion noise is is calculated by taking the standard deviation of the residual map pixels σres and removing the average instrumental noise in these pixels in quadrature, $\sigma_{conf}^2 = \sigma_{res}^2 - \sigma_{pix}^2$, where σ_{pix} is calculated directly from the exposure time per pixel. The total noise σ_{tot} for a point source is then calculated from both the instrumental noise (and confusion noise from the known sources), $\sigma_i = \sqrt{\mathrm{diag}((\mathbf{A^T N_d^{-1} A})^{-1})}$, and confusion noise from the unknown sources in the residual map σ_{conf} via $\sigma_{tot}^2 = \sigma_i^2 + \sigma_{conf}^2$.

One of the goals of HELP is to extend the use of prior information in order to get below the noise level introduced by source confusion.

3.2 Probabilistic Model

With XID+, we have adopted a Bayesian probabilistic framework to provide us with a natural and transparent way to include prior information. To use this probabilistic framework, we need to create a generative model for the maps. XID+ is the simplest example of a generative model for the Herschel SPIRE maps. Figure **??** shows our probabilistic graphical model for our basic XID+ model, where boxes represent dimensions, open circles as variables, dots as deterministic however one can easily add additional details.

3.3 Stan

3.4 Segmentation

3.5 Uncertainties and Covariances

The uncertainties from the posterior give the uncertainty of the flux given the data. This includes the uncertainty from instrumental noise and confusion (I think). Unlike XID, we are not solving $\mathbf{f} = (\mathbf{A^T N_d^{-1} A})^{-1} \mathbf{A^T N_d^{-1} d}$, we are solving equation **??** and so variations in pixel flux from sources not in the prior list, i.e. from confusion, will directly affect our flux estimates.

4 SIMULATIONS

4.1 XID+ vs. DESPHOT

4.2 Performance with Priors

Adding prior information on fluxes

5 XID+SCIENCE

6 CONCLUSIONS

ACKNOWLEDGEMENTS

[1] introducing LASSO and non-negative priors introduces a non-linearity