Bayesian flux reconstruction in one and two bands

Switzer, E. R.*[1]; Crawford, T. M.[1]; Holzapfel, W. L.[2]; Reichardt, C. L.[2]; Marrone, D. P.[3]; Vieira, J. D.[4]

1) KICP, 2) Berkeley, 3) Arizona, 4) Caltech

Setting and single-band flux reconstruction

We consider the limit where the angular extent of the astrophysical sources of emission is much smaller than the resolution of the survey, and refer to these sources as “point sources." A multi-band survey instrument produces maps of regions of the sky with many such sources, typically. These maps report the intensity of radiation averaged in bands around some set of wavelengths and can be written as the sum over fluxes S times the point spread function (PSF) **R** plus some noise unrelated to the sources:

$$\mathbf{d} = \mathbf{s} + \mathbf{n} = \sum_{i=1}^{N_s} S_i \mathbf{R}(\mathbf{x}_i) + \mathbf{n} \quad \textbf{Goal: P(catalog of flux and position | noisy map of the sky d)}$$

Prior knowledge of the abundance weighs the interpretation that the observed flux is a bright source versus the interpretation that it is a dimmer source superimposed on a positive noise fluctuation. In typical astrophysical settings, dim sources are more common, favoring the second interpretation, thus “deboosting” the inferred source flux. In a multi-band setting, the prior also represents our knowledge of the correlations of flux between bands.

For an unknown number of sources (potentially in the thousands) the space of catalogs is too large to explore jointly. For simplicity, consider one source at a time (assuming no overlap between sources) and take normally-distributed noise with covariance N. Then the maximum likelihood of the flux at a fixed position[1] is a matched filter **F** applied to the map **d**:

$$\frac{dL}{dS} = \mathbf{R}(\mathbf{x})^T \mathbf{N}^{-1}(\mathbf{d} - S\mathbf{R}(\mathbf{x})) \Rightarrow S_{\rm ML}(\mathbf{x}) = \frac{\mathbf{R}(\mathbf{x})^T \mathbf{N}^{-1}\mathbf{d}}{\mathbf{R}(\mathbf{x})^T \mathbf{N}^{-1}\mathbf{R}(\mathbf{x})} = \mathbf{F}^T(\mathbf{x})\mathbf{d}$$

The denominator of the filter describes the variance of the recovered flux, σ_f^2. At this fixed position, the measured flux S_m is the intrinsic flux S_i plus Gaussian noise n, so that the posterior distribution of the source flux (reflecting prior information $\pi(S)$) is

$$P(S_i|S_m) \propto L(S_m|S_i)\pi(S_i) \propto \exp\left\{-\frac{(S_i - S_m)^2}{2\sigma_f^2}\right\}\pi(S_i)$$

Yet, in many applications in the sub-mm and mm-wavelengths, sources are not isolated, and the measured flux has an added sum over other source fluxes weighed by the PSF at their separations, as

$$S_m = S_i + \sum_{j\neq i} S_j R_f(\Delta\mathbf{x}_j) + n = S_i + S_b + n \qquad \textbf{(*)}$$

In the regime where sources can overlap, the reconstruction problem is not unique. A common approach in literature[2] is to reconstruct the total intrinsic flux that contributes to a given pointing, rather than the flux of an individual source at that pointing. Another popular method[4], the “P(D)” solves this problem by reconstructing the abundance of a *population* as a function of flux which reproduces the measured distribution of S_m across the map pixels. When the goal is to catalog and categorize individual sources[3], one can also break the ambiguity by writing a posterior distribution for the brightest individual source in a given pointing, S_{max}, or

$$P(S_{\rm max}|S_m) = L(S_m|S_{\rm max})\pi(S_{\rm max})$$

The prior probability at S_{max} is the differential count n=dN/dS times the probability that no source in the instrument’s sold angle Ω is brighter than S_{max}

$$\pi(S_{\rm max}) \propto \left.\frac{dN}{dS}\right|_{S=S_{\rm max}} \exp\left(-\Omega\int_{S_{\rm max}}^{\infty}\frac{dN}{dS'}dS'\right)$$

The likelihood can be written as the convolution of the terms in the sum in (*)

$$L(S_m|S_{\rm max}) = \delta(S_{\rm max}) * FT^{-1}\{e^{[r(\omega)-r(0)]}\} * \frac{1}{\sqrt{2\pi\sigma_f^2}}e^{-S_m^2/2\sigma_f^2}$$

exploiting the characteristic function[5] of the differential counts for the PDF of sources with $S < S_{max}$ in the PSF:

$$r(\omega) = FT_q\left\{\int_{<S_{\rm max}} \frac{d\Omega_{\Delta\mathbf{x}}}{|R_f(\Delta\mathbf{x})|} n\left(\frac{q}{R_f(\Delta\mathbf{x})}\right)\right\}$$

The National Science Foundation

Two-band analysis

Extending to two bands, the model of measured fluxes becomes

$$S_{\rm m}^{(1)} = S_{\rm max}^{(1)} + S_b^{(1)} + n^{(1)} \quad \text{and} \quad S_{\rm m}^{(2)} = S_{\rm max}^{(2)} + S_b^{(2)} + n^{(2)}$$

so that the relevant posterior distribution is

$$P(S_{\rm max}^{(1)}, S_{\rm max}^{(2)}|S_{\rm m}^{(1)}, S_{\rm m}^{(2)}) \propto L(S_{\rm m}^{(1)}, S_{\rm m}^{(2)}|S_{\rm max}^{(1)}, S_{\rm max}^{(2)})\pi(S_{\rm max}^{(1)}, S_{\rm max}^{(2)})$$

The posterior distribution must represent two new aspects: 1) the correlation in background source fluxes $S_b^{(1)}$ and $S_b^{(2)}$ by virtue of sharing common physical sources of emission, and 2) that knowledge of the spectral index and the flux in one band informs the flux in the other band. In the absence of correlations, the joint posterior splits into replicas of the single-band case applied independently to each band. To categorize sources, one would also like a posterior in the spectral index α (defined through $S(\lambda_2)=S(\lambda_1)(\lambda_2/\lambda_1)^{-\alpha}$) and can transform $[S^{(1)}_{max}, S^{(2)}_{max}]$ into $[S^{(1)}_{max}, \alpha]$.

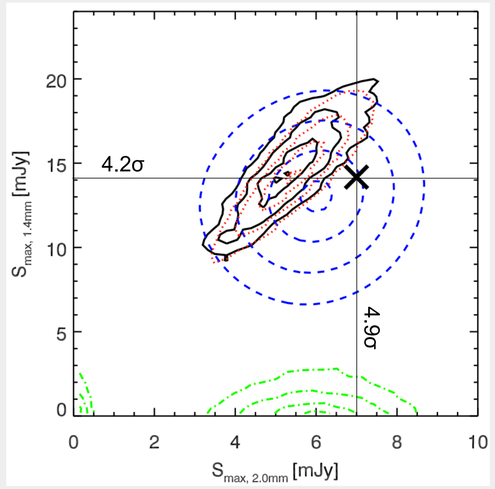
The formalism for finding S_{max} posterior in the single-band calculation becomes challenging in the multiband case. Fortunately, many sub-mm and mm-wave instruments[6] operate in a regime where the sources subdominant to the brightest source in the resolution element contribute an effectively normal distribution of fluxes to the pixel. In this case, the log-likelihood greatly simplifies the calculation to Gaussian functions and is

$$\ln L(S_{\rm m}^{(1)}, S_{\rm m}^{(2)}|S_{\rm max}^{(1)}, \alpha) = C' - \frac{1}{2}\,\mathbf{r}^T\mathbf{C}^{-1}\mathbf{r}$$

with:
$$\mathbf{r} = \left\{S_{\rm m}^{(1)} - S_{\rm max}^{(1)} - \overline{S_b^{(1)}}, S_{\rm m}^{(2)} - S_{\rm max}^{(2)}(\alpha) - \overline{S_b^{(2)}}\right\}$$

Under the Gaussian likelihood assumption, this method can be trivially extended to multiple bands.

One can explicitly test the flux reconstruction by simulating realizations of maps in the various bands and comparing the reconstructed fluxes with the known flux which entered the simulations. By applying the single-band method to each band individually (equivalently, assuming a flat prior on the spectral index) the figure below shows that the posterior could be dramatically incorrect. This is because the signal in the weaker band can be effectively deboosted to the background flux level. This can lead to an incorrect categorization based on the spectral index. In contrast, in the multiband case, information from a stronger detection in another band is carried over to the weak band.



Comparison of the true, underlying posterior two-band flux PDF extracted from simulated observations and calculated values for that PDF. Black X: value of measured flux in the two bands. Black contours: true, underlying posterior two-band flux PDF, extracted from simulated observations. Blue contours: posterior PDF assuming a flat prior on spectral index with cutoff values of −3 and 5. Red contours: posterior PDF assuming a prior on the spectral index equal to the same distribution used in the simulated observations. Green contours: posterior PDF obtained using the single-band procedure ignoring the correlations between the priors in the two bands. All posterior PDF calculations use the Negrello et al. (2007) source counts model to construct the flux prior. All contours are drawn at 0.5σ intervals.

Bigger picture: The fluxes inferred from the procedures described here usually enter into three sorts of subsequent products: population abundance, spectral energy distributions, and categorized source catalogs. Source categorization is central to many of these goals, and here one can apply a cut that the spectral index exceed a discrimination threshold α_d for one population versus another, $P(\alpha > \alpha_d) > P_d$. In developing the population abundance, one now has a PDF of fluxes for each source, and would like to determine the underlying abundance as a function of flux which is consistent with that sample. At high flux, most of the abundance information comes from the data, while a lower flux, progressively more information comes from the prior. More work needs to be done to quantify this tradeoff. Further rigorous methods should also be developed to combine counts data from a variety of experiments and to properly include prior information in this setting.

References:

1. P. Carvalho, G.Rocha, and M. P. Hobson, “A fast Bayesian approach to discrete object detection in astronomical data sets - PowellSnakes I” MNRAS, 393:681–702, March 2009.
2. K. Coppin, M. Halpern, D. Scott, C. Borys, and S. Chapman, “An 850-μm SCUBA map of the Groth Strip and reliable source extraction” MNRAS, 357:1022–1028, March 2005.3.
3. T. M. Crawford, E. R. Switzer, W. L. Holzapfel, C. L. Reichardt, D. P. Marrone, and J. D. Vieira, “A Method for Individual Source Brightness Estimation in Single- and Multi-band Data” ApJ, 718:513–521, July 2010.
4. Patanchon et al., “Submillimeter Number Counts from Statistical Analysis of BLAST Maps” ApJ, 707:1750–1765, December 2009.
5. P. A. G. Scheuer, “A statistical method for analysing observations of faint radio stars” In Proceedings of the Cambridge Philosophical Society, volume 53 pages 764–773, 1957.
6. J. D. Vieira, T. M. Crawford, E. R. Switzer, et al. Extragalactic Millimeter-wave Sources in South Pole Telescope Survey Data: Source Counts, Catalog, and Statistics for an 87 Square- degree Field. ApJ, 719:763–783, August 2010.

Two-band Bayesian reconstruction code is publicly maintained at:
https://github.com/eric-switzer/bayes_flux

THE KAVLI FOUNDATION