MAXIMUM ENTROPY IMAGING OF COMPTEL DATA

A. W. STRONG
Max-Planck Institut für extraterrestrische Physik
Postfach 1603, 85740 Garching, Germany

Abstract. The Maximum Entropy method is a practical technique for generating intensity skymaps from Compton-telescope data. The application of the method to COMPTEL data for point sources and large-scale emission is described. New developments in the method are illustrated.

Key words: gamma-ray astronomy – image processing – maximum entropy

1. Introduction

Compton telescope data provide an interesting challenge in astronomical data analysis. Unlike a spark-chamber telescope it is not possible to generate an image by simply plotting the arrival directions of the photons: 'indirect imaging' methods are necessary if we wish to make intensity maps of the sky. This implies methods in which the 'data-space' and 'image-space' are logically separated from the beginning, the two spaces being related by the instrumental response function.

One method in which this division is explicit is maximum entropy (MaxEnt) which has proved to be very effective in producing images from the COMPTEL instrument on the Compton Gamma Ray Observatory CGRO. In this paper I briefly describe MaxEnt and illustrate applications to COMPTEL data for imaging of continuum emission from point sources and for imaging the full sky. Applications to line emission are covered by R. Diehl in an accompanying paper. I also mention some new developments in Bayesian methods.

2. Why MaxEnt ?

The ultimate aim (not yet achieved) is an image-reconstruction algorithm which is not *ad hoc*, i.e. one which is based on basic principles and is in some sense optimal. One form of the principle of maximum entropy may be stated as follows: given constraints such as $\sum c_i p_i = constant$ on a set of proportions p_i then the only *self-consistent* assignment of p_i in the absence of other information is that which maximises S$=-\sum p_i log p_i$ under the constraints. For image-processing this has been applied to a constraint of form 'χ^2= Number of data points' or its likelihood equivalent for Poisson statistics. (For an extensive introduction see Gull and Skilling 1984). This method was implemented in MEMSYS2 software package (Gull and Skilling) in the early 1980's. More recently a completely *Bayesian approach* has been developed (see Sec 7).

Why use MaxEnt for COMPTEL ? The main reason is that for a COMPTEL-like response with its unusual point-spread-function and 3-D data space, we know of no other way to make true intensity images. Note that other methods such as maximum-likelihood are extremely important in searches for point-sources, but they do not yield intensity distributions, rather they give maps of significance or point-source flux under the assumption of a single or at most a few sources. Especially for *diffuse, extended* emission the MaxEnt method is valuable in providing quantitative intensity distributions.

It should be emphasized that this is a subject in rapid development; the proceedings of the annual workshops under the title of 'Maximum Entropy and Bayesian Methods' (see references) are an invaluable source of information on the current state of the art, and the reader is recommended to consult these.

3. Maximum Entropy Method

The entropy of an image is defined as $S = -\sum f_i log f_i$ where f_i is the intensity in pixel i relative to a default or model value. The statistic used is the Poisson log-likelihood L=log Prob(data|model) = $\sum ln\ e^{-x}x^n/n!$ The MaxEnt algorithm *maximises S over all images which are consistent with the data*; since S is a measure of structure in the image, the method is often characterised as finding the 'flattest' image consistent with the data. The method is basically as follows: define $C = S - \lambda L$; then, starting with a flat map f_i= const corresponding to $\lambda = 0$ progressively increase λ, at each step maximising C. λ is a Lagrangian multiplier so we obtain images with maximum S for a given L. The process stops when L indicates that the model is first 'consistent with the data ' at some 'confidence level'; in practice more subjective criteria, often just the appearance of the image, are applied; when the whole series of iterations is displayed it is usually evident at what point real features give way to artefacts due to over-fitting the data.

4. The COMPTEL Data

4.1 COMPTEL Response representation

The measured quanitites are the scattered photon direction (χ, ψ) and the energy loss in the upper and lower detectors from which is computed the Compton scatter angle $(\bar{\varphi})$. We use the 3-D dataspace $(\chi, \psi, \bar{\varphi})$ to bin the data. The response to a sky intensity distribution $I(\chi', \psi')$ can be represented by the formula:

$$n(\chi,\psi,\bar{\varphi}) = g(\chi,\psi)\int\int I(\chi',\psi')X(\chi',\psi')f(\chi,\psi,\chi',\psi',\bar{\varphi})d\Omega \qquad (1)$$

where $n(\chi,\psi,\bar{\varphi})$ are the expected counts per steradian, $g(\chi,\psi)$ is a geometrical factor related to the upper and lower detector layers, $X(\chi',\psi')$ is an exposure factor, $f(\chi,\psi,\chi',\psi',\bar{\varphi})$ is a point-spread-function and Ω is solid angle. A complete description of the instrument and its reponse can be found in Schönfelder et al. (1993).

A good computation of the response is now available based on extensive Monte-Carlo simulations of the COMPTEL instrument for various assumed input spectra. The convolution in equation (1) is time-consuming; it can be done by FFT methods but only when the field studied is not too large (see discussion on large-scale imaging below).

4.2 Background estimation

COMPTEL data are background-dominated, this component contributing at least 95% of the counts. For the case of *continuum emission*, which is the subject of this paper, the problem of estimating the background is difficult and the solutions not entirely satisfactory; in contrast for *line* emission more satisfactory methods are available, as described by R. Diehl in these proceedings. Two different methods for estimating the continuum background have been applied. Method 1 uses a smoothed version of the data themselves, so that source-like signals are removed from the background dataset. This has the advantage that it is not affected by time-dependence of background and is easy to apply, but the disadvantage that part of the signal is discarded and extended emission is suppressed. Method 2 uses independent data in the form of an average over many 'empty field' observations (e.g. high Galactic latitudes); this has the advantage that no signal is lost and is therefore preferred for extended emission and for model-fitting. Its disadvantages are that it is affected by time-dependence of background, and that it is quite complicated to apply.

5. Point sources

Fig 1 shows an image of the Crab in the 1-3 MeV range, illustrating the very good image attainable for a strong source. Fig 2 shows the same region in 10-30 MeV showing in addition to the Crab the QSO PKS05128+134 which was active in early 1991 (see Collmar et al. 1994). At this time the QSO was on average only about 20% of the Crab flux, showing that MaxEnt is capable of resolving very unequal sources at a few degrees separation. Both these images were obtained with background method 1.

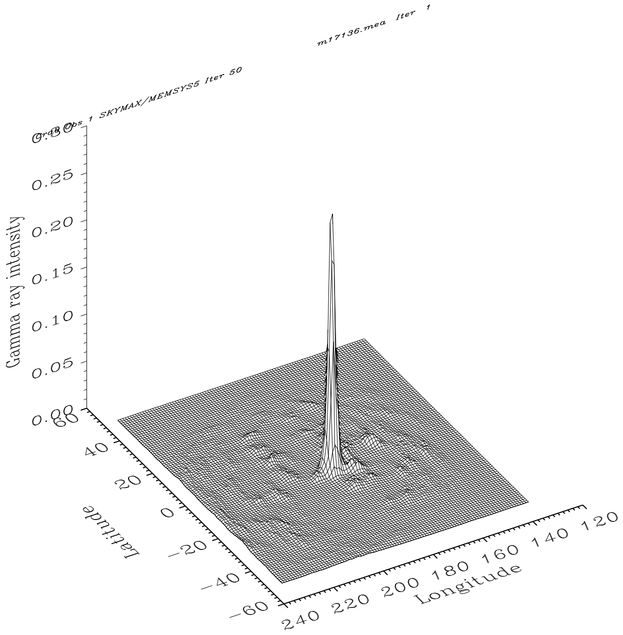

Fig. 1. Crab image in energy range 1-3 MeV

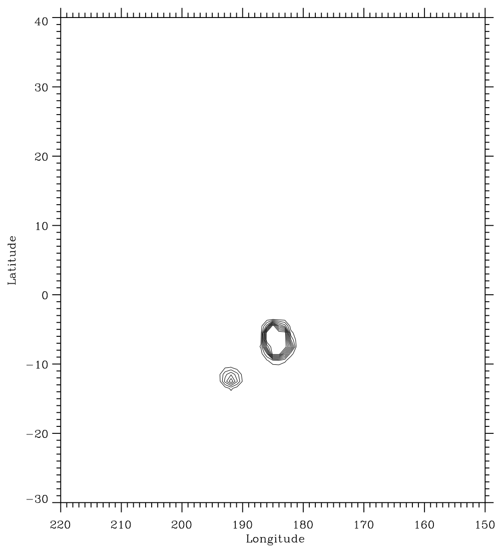

Fig. 2. Crab and PKS05128+134 in energy range 10-30 MeV

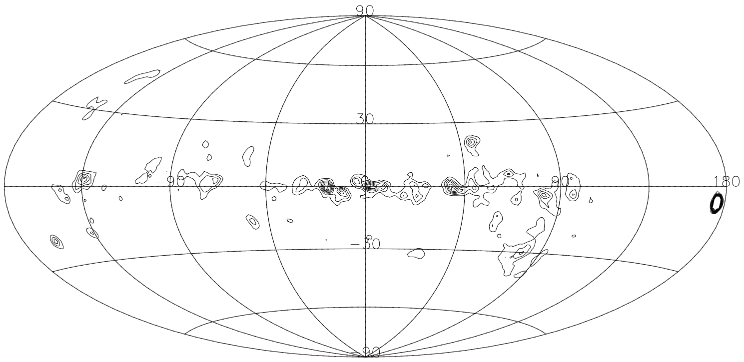

Fig. 3. Full-sky image in 3-10 MeV energy range, in Galactic coordinates centred on l=0.

6. Maps of the whole sky

Maps of the whole Galactic plane are useful especially in connection with diffuse Galactic emission; such a map, generated using background method 2, can be found in Strong et al. (1994). More recently it has been possible for the first time to produce maps of the entire sky. The advantages of this method are that it uses *all* available data, there are no 'edge-effects' and the result can be transformed to other projections from the original Mercator. A disadvantage is that it is very computer-time consuming; also any time-variations are smeared out so that transient sources may not be detectable. In a 'Cartesian' system one can do the convolution by FFT since the point-spread-function is shift-invariant. This works only for maps which are restricted in size at least in latitude. For *full-sky imaging* we have to use spherical coordinates; the convolution cannot then be done by FFT but has to be done explicitly. This is however much more time-consuming. For a full-sky image with 1^o binning and 20 $\bar{\varphi}$ values we require $360 \times 180 \times 60^2 \times 20 \approx 4\ 10^9$ operations per convolution. The method requires typically 36 such convolutions to reach a reasonable image. For the 3-10 MeV full-sky image shown in Fig 3, a large fraction of the sky survey data were combined into one dataset, and the background was estimated by method 1. The most striking feature is the band of emission following the Galactic plane; some point sources are visible including the Crab, Cen A and 2CG135+1.

7. Quantified Maximum Entropy

Among the defects in MaxEnt as described above is the fact that the stopping criterion $\chi^2 = N$ (see Section 2) lacks justification and there is no way to judge significance of features (there are no 'error bars'). A new development in the framework of Bayesian statistics (Skilling 1989, 1991;

Gull 1989) addresses these problems: 'classical' or 'quantified' MaxEnt (in contrast to the previously described method, now dubbed 'historical'). It has been implemented in the MEMSYS5 package. Imaging is treated as a Bayesian parameter estimation problem for the pixel intensities. The prior probability has the form $P(\mathbf{f}\,) \propto e^{\alpha S(\mathbf{f}\,)}$, the posterior probability for data D is given by $P(\mathbf{f}\,) \propto e^{\alpha S(\mathbf{f}\,)+L(D|\mathbf{f}\,)}$, with the notation of Section 3. α is a parameter controlling the relative effect of entropy and data; for $\alpha \to \infty$ entropy dominates giving a flat image, while for $\alpha \to 0$ the data dominates giving a maximum likelihood image. MEMSYS5 determines the full posterior probability P($\mathbf{f}$|D) from which error bars on intensities can be derived. The only unknown is α; in general this can be treated as a parameter with prior P(α). In practice we can choose one best α determined by maximising the 'evidence' P($D|\alpha$).

In view of these advances, some efforts have been made to evaluate this method for COMPTEL data using MEMSYS5. The same data as used for the 1-3 MeV Crab image in Fig 1 were used to make an intensity scan with error estimates; the intensities are for 10^o bins centred on the Crab:

b / l	195^o	185^o	175^o	
+ 5^o	104±83	99±80	229±106	Average intensity
- 5^o	144±86	3148±117	261±109	$10^{-5} cm^{-2} sr^{-1} s^{-1}$
-15^o	11±40	210±93	221±96	

The Crab flux and error are in good agreement with those obtained by direct fitting using maximum likelihood (Strong et al. 1993). This only illustrates the technique: it will be most useful for deriving errors on extended emission, for example in longitude or latitude profiles.

References

Collmar W. et al.: 1994, 'COMPTEL Observations of the Quasar PKS 0528+134' in C.E. Fichtel, N. Gehrels, J.P. Norris, ed(s)., *Proc. 2nd Compton Symposium, AIP Conf. Proc. 304*, AIP Press: New York, 659

Gull, S.F.: 1989, 'Developments in maximum entropy data analysis' in J. Skilling, ed(s)., *Maximum Entropy and Bayesian Methods*, Kluwer:Dordrecht ISBN 0-7923-0224-9, 53

Gull, S.F., Skilling J: 1984, ' Maximum entropy method in image processing', *IEE Proc.* **131(F)**, 646

Schönfelder V. et al.: 1993, *ApJ Supp* **86**, 657

Skilling, J.: 1989, 'Classic maximum entropy ' in J. Skilling, ed(s)., *Maximum Entropy and Bayesian Methods*, Kluwer:Dordrecht ISBN 0-7923-0224-9, 45

Skilling, J.: 1991, 'On parameter estimation and quantified MaxEnt ' in W.T. Grandy, Jr. and L.H. Schick , ed(s)., *Maximum Entropy and Bayesian Methods*, Kluwer:Dordrecht ISBN 0-7923-1140-X, 267

Strong, A. W.: 1993, 'The Crab and Galactic anticentre region observed by COMPTEL', *A& A Supp.* **97**, 133

Strong, A. W.: 1994, 'Diffuse Galactic continuum emission measured by COMPTEL and the cosmic-ray electron spectrum', *A& A* **292**, 82