

University of New Hampshire University of New Hampshire Scholars' Repository

Space Science Center

Institute for the Study of Earth, Oceans, and Space
(EOS)

2001

Bayesian multiscale deconvolution applied to gamma-ray spectroscopy

C A. Young

Goddard Space Flight Center

A Connors

Eureka Scientific

E Kolaczyk

Boston University

Mark L. McConnell

University of New Hampshire - Main Campus, mark.mcconnell@unh.edu

G Rank

Max-Planck-Institut für extraterrestrische Physik

See next page for additional authors

Follow this and additional works at: <https://scholars.unh.edu/ssc>

 Part of the [Astrophysics and Astronomy Commons](#)

Recommended Citation

Bayesian multiscale deconvolution applied to gamma-ray spectroscopy Young, C. A. and Connors, A. and Kolaczyk, E. and McConnell, M. and Rank, G. and Ryan, J. M. and Schönfelder, V., AIP Conference Proceedings, 587, 701-705 (2001), DOI:<http://dx.doi.org/10.1063/1.1419486>

This Conference Proceeding is brought to you for free and open access by the Institute for the Study of Earth, Oceans, and Space (EOS) at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Space Science Center by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.

Authors

C A. Young, A Connors, E Kolaczyk, Mark L. McConnell, G Rank, James M. Ryan, and V Schonfelder



Bayesian multiscale deconvolution applied to gamma-ray spectroscopy

C. A. Young, A. Connors, E. Kolaczyk, M. McConnell, G. Rank, J. M. Ryan, and V. Schönfelder

Citation: [AIP Conference Proceedings](#) **587**, 701 (2001); doi: 10.1063/1.1419486

View online: <http://dx.doi.org/10.1063/1.1419486>

View Table of Contents: <http://scitation.aip.org/content/aip/proceeding/aipcp/587?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

[AGILE: A gamma-ray mission](#)

AIP Conf. Proc. **510**, 746 (2000); 10.1063/1.1303299

[Study of detection limit in spectroscopy of coincident gamma-rays](#)

AIP Conf. Proc. **447**, 151 (1998); 10.1063/1.56699

[Quantitative analysis for gamma-ray spectroscopy](#)

Comput. Phys. **12**, 356 (1998); 10.1063/1.168698

[Spectroscopy of bright bursts with the transient gamma-ray spectrometer \(TGRS\)](#)

AIP Conf. Proc. **428**, 339 (1998); 10.1063/1.55345

[Search for millisecond periodic pulsations in BATSE gamma-ray bursts](#)

AIP Conf. Proc. **428**, 216 (1998); 10.1063/1.55325

Bayesian Multiscale Deconvolution Applied to Gamma-ray Spectroscopy

C. A. Young^a, A. Connors^b, E. Kolaczyk^c, M. McConnell^d, G. Rank^e,
J. M. Ryan^d, V. Schönfelder^e

^a*Emergent-IT, Inc., Goddard Space Flight Center*

^b*Eureka Scientific*

^c*Boston University*

^d*University of New Hampshire*

^e*Max Planck Institute for Extraterrestrial Physics*

Abstract. A common task in gamma-ray astronomy is to extract spectral information, such as model constraints and incident photon spectrum estimates, given the measured energy deposited in a detector and the detector response. This is the classic problem of spectral “deconvolution” or spectral inversion. The methods of forward folding (i.e. parameter fitting) and maximum entropy “deconvolution” (i.e. estimating independent input photon rates for each individual energy bin) have been used successfully for gamma-ray solar flares (e.g. Rank, 1997; Share and Murphy, 1995). These methods have worked well under certain conditions but there are situations where they don't apply. These are: 1) when no reasonable model (e.g. fewer parameters than data bins) is yet known, for forward folding; 2) when one expects a mixture of broad and narrow features (e.g. solar flares), for the maximum entropy method; and 3) low count rates and low signal-to-noise, for both. Low count rates are a problem because these methods (as they have been implemented) assume Gaussian statistics but Poisson are applicable. Background subtraction techniques often lead to negative count rates. For Poisson data the Maximum Likelihood Estimator (MLE) with a Poisson likelihood is appropriate. Without a regularization term, trying to estimate the “true” individual input photon rates per bin can be an ill-posed problem, even without including both broad and narrow features in the spectrum (i.e. a *multiscale* approach). One way to implement this regularization is through the use of a suitable Bayesian prior. Nowak and Kolaczyk (1999) have developed a fast, robust, technique using a Bayesian multiscale framework that addresses these problems with added algorithmic advantages. We outline this new approach and demonstrate its use with time resolved solar flare gamma-ray spectroscopy.

Given a measured energy deposit spectrum we can extract spectral information such as constraints on model parameters and an estimate of the incident photon spectrum. This is the classic problem of spectral deconvolution or spectral inversion. In the ideal case this is equivalent to solving an integral equation for the photon spectrum $f(E')$ given the measured energy loss spectrum $C(E)$ and the energy response of the detector $R(E', E)$,

$$C(E) = \int_0^{E'} R(E, E') f(E') dE' \quad (1)$$

The solution to this deceptively simple equation is part of a larger class of problems commonly referred to as inverse problems. If we had the exact continuous forms of C and R we could, in principle, solve the integral equation that is just a Fredholm equation of the first kind (Craig and Brown 1986; Hansen 1998). The count data $C(E)$ are only known for a discrete set of energies, E_i , where $i=0, \dots, n-1$. We are then solving a set of n integral equations but many functions can satisfy a given discrete set of equations without satisfying our original

equation. The instrument response is seldom described analytically so that we then must replace the set of integral equations with the matrix equation,

$$C_i = R_{ij}f_j \text{ for } i = 0, \dots, n-1 \text{ and } j = 0, \dots, m-1. \quad (2)$$

The formal solution of this equation is $f = R^{-1}C$ but usually R^{-1} is unbounded and its computation is sensitive and unstable to small perturbations in the data. The inverse problem is then termed “ill-posed”.

There exist a number of “classical” methods for numerical inversion of matrix equations, which were discussed in detail by Craig and Brown (1986) and Hansen (1998). These methods are classical in the sense that they explicitly use prior information or make assumptions about the source function (Craig and Brown 1986). Due to many shortcomings we do not use any of these methods but instead use “non-classical” or regularization methods. The general idea of regularization is to introduce an extra term (or regularization function ϕ) to minimize irregular solutions, i.e.,

$$\mathcal{L}(f) = \|C - Rf\|^2 + \phi(f, \alpha) \quad (3)$$

The free parameter α is chosen to balance the minimization of the norm versus the suppression of noise. Below, we will discuss two commonly used techniques and we present the use of a novel new method applied to γ -ray spectroscopy.

The method of forward folding is commonly used in γ -ray and X-ray astronomy. A model μ is convolved (folded) with the detector response R yielding a model set of data d . A maximum-likelihood fit of the real data C with the model data is then performed maximizing the log of the probability (or in the case of Gaussian statistics minimizing χ^2). This procedure has the disadvantage that it is inherently restricted to the assumed model. One does not obtain an estimate of the real spectrum but rather a set of parameters associated with that model. However, when the model choice is appropriate and realistic and the response is dominated by diagonal elements the inferred spectrum is robust.

Maximum Entropy methods are also popular for the deconvolution of image data and have proven to be successful in spectral deconvolution (Gull and Skilling 1990). In addition to maximizing a likelihood, prior information in the form of information entropy S is maximized. Many versions of S have been defined in the literature.

The methods of forward folding and maximum entropy “deconvolution” have been used successfully for gamma-ray solar flares (e.g. Rank, 1996; Share and Murphy, 1995). These methods worked well under certain conditions but there are situations where they do not apply. These are: 1) when no reasonable model (e.g. fewer parameters than data bins) is yet known, for forward folding; 2) when one expects a mixture of broad and narrow features (e.g. solar flares), for the maximum entropy method; and 3) low count rates and low signal-to-noise, for both. For Poisson data the Maximum Likelihood Estimator (MLE) with a Poisson likelihood is appropriate. Without a regularization term, trying to estimate the “true” individual input photon rates per bin can be an ill-posed problem. One way to implement this regularization, though, is through the use of a suitable Bayesian prior. Nowak and Kolaczyk (1999) developed a fast, robust, technique using a Bayesian multiscale framework that addresses these problems with added algorithmic advantages. We outline this new approach so that we can apply it to solar flare gamma-ray spectroscopy.

Recent treatments of Poisson inverse problems have augmented the likelihood equations with a regularization or penalization term as discussed above. This regularization term stabilizes the otherwise ill posed ML problem. The regularization term can take the form of a Bayesian prior so that the MLE is replaced with the Maximum *a posteriori* (MAP) estimator. If we wish to use a MAP estimator, we first apply Bayes’ theorem,

$$p(\lambda|y) = \frac{p(y|\lambda)p(\lambda)}{p(y)} \quad (4)$$

This equation relates the likelihood to the posterior with the prior $p(\lambda)$ and $p(y)$ being a normalization based on the data. The prior can also be interpreted as a penalizing function

giving the terminology “Penalized MLE”. The MAP estimate is then the value of λ that maximizes the log of the posterior.

Multiscale analysis is the study of behavior or structure in data at various spatial and/or temporal scales. One way to address our ill-posed problem is through a multi-scale framework (Starck, Murtagh, and Bijaoui, 1998). The usual multi-scale model is formulated with a wavelet decomposition but wavelets and Poisson data are somewhat incompatible (Kolaczyk 1999). Nowak and Kolaczyk (1999) developed a deconvolution technique that uses a Bayesian multiscale framework that addresses these problems with other advantages. The technique uses an Estimator Maximization (EM) algorithm that has a closed-form step. Under reasonable choice of the multiscale priors, the EM algorithm converges to a unique, global MAP estimate.

The problem at hand is to estimate the photon flux, λ , from the observed count data y . The counts are related to the flux by the relation $y_n = P(\mu_n)$, $n=0, \dots, N-1$, where $P(\mu_n)$ is the Poisson distribution with mean counts μ_n . The mean counts μ_n are related to the flux by the relation $\mu = \mathbf{R} \cdot \lambda$, where \mathbf{R} is an $N \times M$ matrix (the response) of transition probabilities.

It will be useful later in this discussion to introduce the idea of the “complete data” $z(n,m)$. This is the total number of m to n (emission to detection) events, $z(n,m) = P(\lambda_m R_{n,m})$. The indirectly observed count data is then given by summing the complete data over m , $y_n = \sum_m z(n,m)$. Also, were we able to detect the photons directly without the detector we would have the direct data, $x_m \equiv \sum_n z(n,m)$, from which it follows that $x_m = P(\lambda_m)$.

To seek a solution of the general inverse problem we must first solve the direct-data Poisson estimation problem, $x_m = P(\lambda_m)$. The simplest multiscale data analysis is the unnormalized Haar analysis. The data $\{x_{j,k}\}$ are the unnormalized Haar scaling coefficients of \mathbf{x} . The index j refers to the resolution of the analysis, 2^j , where $j = J$ is the index for the highest or finest scale and $j = 0$ is the lowest or coarsest scale.

Using conditional probability relationships, the joint probability of the data in a multiscale representation can be expressed with the factorized form,

$$\rho(\mathbf{x}) = P(x_{0,0}) \prod_{j=0}^{J-1} \prod_{m=0}^{2^j-1} \Pr(x_{j+1,2m} | x_{j,m}). \quad (5)$$

The parent ($x_{j,k}$), child ($x_{j+1,2k}$) relationship is expressed by the conditional likelihood, $\Pr(x_{j+1,2m} | x_{j,m})$. The MAP estimation of λ requires the likelihood function of \mathbf{x} , $\rho(\mathbf{x} | \lambda)$. The multiscale expansion of $\rho(\mathbf{x} | \lambda)$ requires that we define the multiscale analysis of the intensity λ , analogous to the analysis of \mathbf{x} .

The parameters $\{\lambda_{j,m}\}$ are the unnormalized Haar scaling coefficients of the intensity λ .

Using the definitions for the multiscale analysis of \mathbf{x} and λ and the multiscale factorization of $\Pr(\mathbf{x})$ we can express the parent-child conditional likelihood as,

$$\rho(x_{j+1,2m} | x_{j,m}, \lambda_{j,m}) = \mathcal{B}(x_{j+1,2m} | x_{j,m}, \rho_{j,m}), \quad (6)$$

where $\mathcal{B}(x | n, \rho) = \binom{n}{x} \rho^x (1-\rho)^{n-x}$, is the binomial distribution with parameters n and ρ . The parameters $\rho_{j,m} = \frac{\lambda_{j+1,2m}}{\lambda_{j,m}}$ are the canonical multiscale parameters for the Poisson model and can

be viewed as “splitting” factors, governing the multiscale refinement of the intensity. We can represent the multiscale analysis of \mathbf{x} and λ as a binary tree where the splitting factors are multiplicative weights in the tree’s links. The complete factorization of the likelihood is then

$$\rho(\mathbf{x} | \lambda) = P(x_{0,0} | \lambda_{0,0}) \times \prod_{j=0}^{J-1} \prod_{m=0}^{2^j-1} \mathcal{B}(x_{j+1,2m} | x_{j,m}, \rho_{j,m}) \quad (7)$$

where $P(x_{0,0} | \lambda_{0,0})$ is just the Poisson probability function of $x_{0,0}$ with mean $\lambda_{0,0}$.

A maximum likelihood analysis of the binomial conditional likelihood leads to a MLE estimation of the splitting parameters (Nowak and Kolaczyk 1999) of

$\hat{\rho}_{j,m} = \frac{\lambda_{j+1,2m}}{\lambda_{j,m}}$. There is a one-to-one mapping from (ρ, λ_{00}) to $\hat{\lambda}$ so using the multiscale synthesis equation (Nowak and Kolaczyk 1999) and the estimate $\hat{\rho}_{j,m}$ we find the MLE of each intensity element of the finest scale to be $\hat{\lambda}_{j,m} = x_{j,m} \equiv x_m$. The MLE returns the raw data as our MLE intensity estimate, an expected result (Nowak and Kolaczyk 1999). The next step is the MAP estimation.

The crucial ingredient in moving from a MLE estimation to a Bayesian estimation is the choice of a prior distribution $p(\lambda)$. A good choice of prior reflects known or assumed attributes of the intensity and matches the functional form of the likelihood (in our case Poisson and Poisson-binomial). Conjugate priors have the computational advantage that they are obtained by updating the parameters of the prior based on the measurements (Nowak and Kolaczyk 1999). The natural choice of the conjugate prior for the total intensity λ_{00} is the gamma probability density. The choice for modeling the splitting parameter is as an independent beta distributed random variable $Be(\rho|\alpha, \beta)$. We have no a priori knowledge of asymmetry therefore we use only a symmetric beta prior of mean 1/2 with $\alpha = \beta$. In our case the gamma prior has negligible effect so the important parameter is the beta prior in the splitting parameters (Nowak and Kolaczyk 1999). The beta priors $\{\alpha_j\}$ reflect our belief or prior knowledge of the intensities regularity. Combining the prior and the likelihood and using the conjugacy of the prior with the likelihood produces a posterior density (Nowak and Kolaczyk 1999). As with the MLE, the synthesis equations can be used to obtain a MAP estimate of λ .

Moving back to the more difficult Poisson inverse problem: As we showed for the analysis of the directly observed Poisson data, a multiscale factorization of the data likelihood played a key role. For the analysis of the indirectly observed data, the complete data likelihood plays a key role through the EM algorithm. Nowak and Kolaczyk (1999) show that the complete data likelihood is proportional to the direct data likelihood. The log complete data posterior is then just a combination of the log complete data likelihood and the log prior,

$$\begin{aligned} \mathcal{L}(\lambda) &\equiv \log p(\lambda_{0,0}, \rho | z) \\ &= \log \mathcal{P}(x_{0,0} | \lambda_{0,0}) + \log G(\lambda_{0,0} | \gamma, \delta) + \\ &\quad \sum_{j=0}^{J-1} \sum_{m=0}^{2^j-1} \log B(x_{j+1,2m} | x_{j,m}, \rho_{j,m}) + \log Be(\rho_{j,m} | \alpha_j, \alpha_j) + C. \end{aligned} \tag{8}$$

The key to the EM algorithm for MLE (Nowak and Kolaczyk 1999) is the introduction of the complete data, $z(n,m)$. If we could observe these complete data, we have shown that a closed-form maximizer of the complete data posterior exists. The EM algorithm iteratively alternates between computing the expected complete data log-posterior and a maximizer of this function leading to a MAP estimate of the log-posterior. The problem with the MAP-EM algorithm is that now the M-step does not have a closed-form solution. Fortunately, Nowak and Kolaczyk (1999) solved this problem by taking a multiscale approach that does have a closed-form M-step. As an EM algorithm it has the standard property that the posterior probability does not decrease with subsequent iterations and the estimate is non-negative. The algorithm also has the feature that it is computationally simple and for a certain choice of prior parameters the MAP converges to a global solution.

Unfortunately, errors or confidence intervals in the traditional sense do not follow (Kolaczyk 1999). In order to produce spectra with which we can then calculate line fluxes and physical parameters; we must be able to produce errors or uncertainties in our estimates. The most straightforward method for this is to use a parametric bootstrap (Connors 2000; Kolaczyk 2000).

Figure 1 shows the light curve for a gamma-ray solar flare divided into 3 time intervals and the deconvolved spectra for each of the time intervals.

We have shown this technique to be very useful for the analysis of solar flare gamma-ray spectra. This technique is fast, robust, and requires no knowledge of the underlying physics. This technique holds great promise for the deconvolution of gamma-ray spectra as well as spectra in other energy ranges.

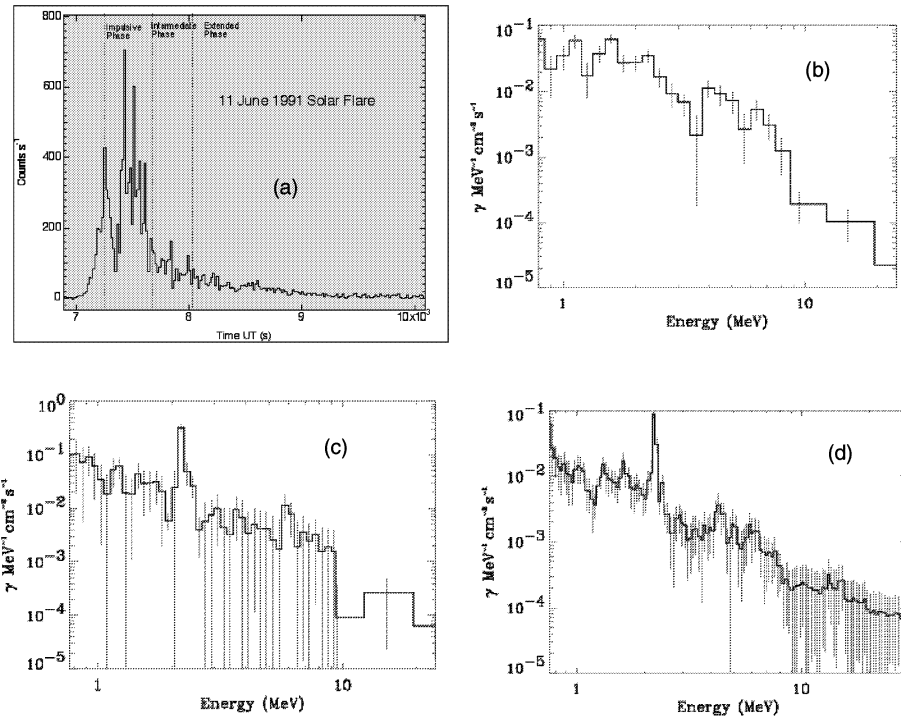


FIGURE 1. The light curve for a gamma-ray solar flare (a). The light curve is divided into 3 time intervals. Also shows the deconvolved spectra for the time intervals defined in the light curve (b-d). (Error bars extending to the bottom of the plots are 1σ upper limits.)

ACKNOWLEDGMENTS

This work was supported through NASA contract NAS5-26645 and NASA's Supporting Research and Technology program.

REFERENCES

1. Connors, A., *Confidence Intervals for MAP estimation method*, Private Comm. (2000).
2. Craig, I. J. D., and Brown, J. C., *Inverse Problems in Astronomy*, Boston: Adam Hilger, Ltd., 1986.
3. Gull, S. F., and Skilling, J., *the MEMSYS User's Manual*, Royston: Max. Entropy Data Consultants Ltd., 1990.
4. Hansen, P. C., *Rank-Deficient and Discrete Ill-Posed Problems*, Philadelphia: Siam, 1998.
5. Kolaczyk, E. D., "Some Observations on the Tractability of Certain Multiscale models," in *Bayesian Inference in Wavelet-based Models*, edited by M. A. Vidkovic, New York City: Springer-Verlag, 1999.
6. Kolaczyk, E. D., *Bootstrapping Poisson Estimators*, Private Comm. (2000b).
7. Nowak, R. D., and Kolaczyk, E. D., "A Bayesian Multiscale Framework for Poisson Inverse Problems" in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1999.
8. Share, G.H. and Murphy, R. J., *Ap. J.* **452**, 933 (1995).
9. Starck, J.-L. et al., *Image Processing and Data Analysis: The Multiscale Approach*, New York City: Cambridge University Press, 1998.