RECONSTRUCTION OF RHESSI SOLAR FLARE IMAGES WITH A FORWARD FITTING METHOD

MARKUS J. ASCHWANDEN[1], ED SCHMAHL[2] and the RHESSI TEAM[1,2,3]
[1]*Lockheed Martin ATC, Solar & Astrophysics Laboratory, Dept. L9-41, Bldg.252, 3251 Hanover St., Palo Alto, CA 94304, U.S.A. (e-mail: aschwanden@lmsal.com)*
[2]*NASA Goddard Space Flight Center, Solar & Astrophysics Laboratory, Greenbelt, MD 20771, U.S.A.*
[3]*Space Sciences Laboratory, University of California Berkeley, Berkeley, CA 94720; U.S.A.*

(Received 7 August 2002; accepted 19 August 2002)

Abstract. We describe a forward-fitting method that has been developed to reconstruct hard X-ray images of solar flares from the Ramaty High-Energy Solar Spectroscopic Imager (RHESSI), a Fourier imager with rotation-modulated collimators that was launched on 5 February 2002. The forward-fitting method is based on geometric models that represent a spatial map by a superposition of multiple source structures, which are quantified by circular gaussians (4 parameters per source), elliptical gaussians (6 parameters), or curved ellipticals (7 parameters), designed to characterize real solar flare hard X-ray maps with a minimum number of geometric elements. We describe and demonstrate the use of the forward-fitting algorithm. We perform some 500 simulations of rotation-modulated time profiles of the 9 RHESSI detectors, based on single and multiple source structures, and perform their image reconstruction. We quantify the fidelity of the image reconstruction, as function of photon statistics, and the accuracy of retrieved source positions, widths, and fluxes. We outline applications for which the forward-fitting code is most suitable, such as measurements of the energy-dependent altitude of energy loss near the limb, or footpoint separation during flares.

1. Introduction

Image reconstruction from RHESSI data can be performed with a number of methods, including *backprojection*, CLEAN, the *maximum entropy method*, *pixon mapping*, *polar mapping*, and *forward-fitting*. In this paper we describe the principles of a *forward-fitting algorithm* that is currently implemented in the software. We simulate RHESSI observations, perform image reconstructions, and compare the test results with the input models, in order to quantify the fidelity of the reconstructed images as a function of the count rate. These tests provide error bars on the accuracy of retrieved source positions, sizes, or relative displacements in energy and time.

This paper includes a brief definition of relevant instrumental parameters (Section 2), the parameterization of images (Section 3), a description of the forward-fitting code, numerical test results of image reconstruction from simulated data (Section 4), and a summary (Section 5). More detailed information to some sec-

Solar Physics **210:** 193–211, 2002.

tions is provided on the attached CD-ROM, also available at *http://www.lmsal.com/ ~aschwand/ eprints/2002_fwdfit*.

2. Instrument

2.1. Instrumental description of RHESSI

Technical descriptions of the RHESSI instrument can be found in Lin *et al.* (1993, 1994, 1998) and Dennis *et al.* (1996). The RHESSI webpages (*http://hessi.ssl. berkeley.edu, http://hesperia.gsfc.nasa.gov/ rhessidatacenter/*), contain the most up-to-date information. Here we summarize the instrumental specifications that are most relevant for image reconstruction. For a detailed description of the imaging capabilities of RHESSI, the grid characteristics, the modulation patterns, and the modulation profiles we refer to Hurford *et al.* (2002).

RHESSI employs the technique of Fourier-transform imaging with rotating modulation collimators, using a set of 9 bi-grid collimators with a grid period (pitch) ranging from 34 μm to 2.6 mm, separated by a distance of 1550 mm between front grids and rear grids. The grid pitch increases by a factor of $\sqrt{3}$ from one grid to the next coarser one, yielding a FWHM resolution ranging from $2.26''$ to $2.26'' \times \sqrt{3}^{\,8} = 183''$. This defines the range of spatial scales over which RHESSI is able to map. The full angular resolution of $\approx 2''$ is achieved at the lowest energies, up to $\lesssim 100$ keV. At higher energies, where the thickness of the tungsten grids (Nos. 2–9) or molybdenum (Mb) grid (No. 1) is not sufficient to completely absorb high-energy photons, the modulational efficiency is thus reduced, restricting the spatial resolution to the spacing of the thicker, fully-absorbing grids, which is $\approx 35''$ above 1 MeV. The angular coverage of an image is thus $2''$–$183''$, while the field of view includes the full Sun, as long as the spacecraft spin axis points within $\approx 12'$ of the Sun center. The temporal resolution for highest quality images is half a spacecraft rotation period (providing maximum uv-coverage in Fourier space), i.e., ≈ 2.0 s, but 'snapshot images' with poorer uv-coverage can be obtained down to tens of ms. The pointing of the spacecraft is determined by the Solar Aspect System (SAS) which measures the direction to the Sun center to sub-second accuracy, and by two redundant systems that measure the roll angles to $< 3'$ – a CCD system and a PMT-based Roll Angle Aspect System (PMTRAS). Great care has been taken to warrant that the relative twist of the finest grid pair is maintained to within $1'$ by the metering tube structure. Spectroscopy of hard X-rays and gamma rays is achieved with internally segmented germanium detectors (GeDs), placed behind the bi-grid collimators, having diameters of 71 mm and a length of 85 mm. The germanium detectors are cooled down to 75 K, providing < 1 keV FWHM energy resolution at 10 keV, increasing to ≈ 5 keV at 7 MeV (see Smith *et al.*, 2002).

2.2. Data structure of observations

RHESSI records every photon above a threshold energy of $\gtrsim 3$ keV as a time-tagged event up to a maximum rate of $\approx 50\,000\ \mathrm{s}^{-1}$, while attenuation and decimation schemes take place at higher rates to avoid pulse pileup effects. Thus, the telemetry data contain photon event information for each of the 9 detectors, tagged with a time accuracy of 1 μs (called Photon Event List).

For standard image processing, the time-tagged photon events are binned into time intervals that resolve the grid-induced modulation due to the spinning spacecraft. The number of modulations per rotation depends on the angular distance r of the observed source from the spacecraft spin axis, divided by the angular pitch $p_l = 4.52'' \sqrt{3}^{\,(l-1)}$ of the collimator ($l = 1, \ldots, 9$),

$$N_l^{\mathrm{mod}} = \frac{4r}{p_l} = \frac{0.88\ r''}{\sqrt{3}^{\,(l-1)}}\ , \quad l = 1, \ldots, 9\ . \tag{1}$$

For a maximum field of view of $r \leq 1^\circ$ we record $N_1^{\mathrm{mod}} = 3185$ modulations for the finest grid ($l = 1$) and $N_9^{\mathrm{mod}} = 39$ modulations for the coarsest grid ($l = 9$). The required time binning to resolve at least 2 data points per modulation period is thus

$$\Delta t_l^{\mathrm{mod}} = \frac{T^{\mathrm{rot}}}{2N_l^{\mathrm{mod}}} = 0.6\Big(\frac{r''}{3600''}\Big)^{-1} \sqrt{3}^{\,(l-1)}\ \mathrm{ms}\ , \tag{2}$$

ranging from $\Delta t_1^{\mathrm{mod}} \gtrsim 0.6$ ms for the finest grid to $\Delta t_9^{\mathrm{mod}} \gtrsim 50$ ms for the coarsest grid, for a spacecraft spin period of $T^{\mathrm{rot}} = 4.0$ s. A typical binning of the data is done by the nearest numbers of powers of 2, i.e., the default number of binning intervals varies from 128 (Detector 9) to 8192 (Detector 1), and the corresponding time intervals range from 0.5 ms (Detector 1) to 32 ms (Detector 9). The values of the detector pitch angles p_l and the default time binning Δt_i are listed in Table I. Thus the typical data structure used for image processing consists of 9 arrays of *modulation profiles* $N_l^{\mathrm{obs}}(t_k)$ for the detectors $l = 1, \ldots, 9$ with variable number of time bins, $t_k =, 1, \ldots, N_l$, ranging from $N_1 = 8192$ to $N_9 = 128$, amounting to a total of $\sum_{l=1}^{9} N_l = 17,792$ data points per spacecraft rotation. This data set of binned photon counts, along with the associated aspect variables, instrumental transmission parameters, amplitudes and phases, livetimes and other quantities required for mapping, is called the Calibrated Event List in the RHESSI jargon.

3. Parameterization of Maps

The simulation of maps as well as the forward-fitting method require a parameterization of the geometry of source morphologies. Hard X-ray sources in solar flares are known to appear as (often unresolved) point sources, double sources, multiple sources, or parts of flare loops (Sakao, 1994; see also *Yohkoh/HXT Image*

TABLE I
RHESSI data structure.

Detector No. l	Angular pitch p_l	FWHM resolution α_l	Modulation per rotation[a] N_l^{mod}	Data time bins[b] N_l^{bin}	Time bin Δt_l(ms)	Minimum count rate[c] R_l^{min} (s^{-1} SC^{-1})
1	4.52″	2.26″	3168	$2^{13} = 8192$	0.512	900
2	7.84″	3.92″	1829	$2^{12} = 4096$	1.024	450
3	13.57″	6.79″	1056	$2^{11} = 2048$	2.048	220
4	23.5″	11.8″	609	$2^{10} = 1024$	4.096	110
5	40.7″	20.4″	352	$2^{10} = 1024$	4.096	110
6	70.5″	35.3″	203	$2^9 = 512$	8.192	54
7	122.2″	61.1″	117	$2^9 = 512$	8.192	72
8	211.6″	105.8″	68	$2^8 = 256$	16.384	38
9	366.5″	183.2″	39	$2^7 = 128$	32.768	17

[a]For a source with a distance of 1° from spin axis.
[b]Default values for time binning.
[c]Minimum count rate to produce detectable modulation, producing r.m.s. fluctuations of the modulation profiles that exceed the Poisson fluctuations.

Catalogue by Kosugi *et al.*, 1995, and Sato *et al.*, 1998). We adopt the philosophy that such maps can mathematically be represented as a superposition of a few elementary geometric shapes.

3.1. Geometry of elementary source structures

The simplest shape is a point source, placed at location (x_i, y_i), with a flux amplitude f_i. Every point source has also a finite width, regardless whether this width represents the physical extent of the source or the apparent size of an unresolved source, as rendered by the instrumental point-spread function. Thus, it is most natural to characterize a point source with a gaussian width w_i, so that its 2-dimensional distribution in a map is defined by 4 parameters (x_i, y_i, f_i, w_i),

$$F(x, y) = f_i \ \exp\left[-\frac{(x - x_i)^2 + (y - y_i)^2}{2w_i^2}\right]. \tag{3}$$

In the next step, we generalize a circular gaussian geometry to an elliptical shape, which can be defined by two additional parameters, the eccentricity e_i and the tilt angle α_i^{tilt}. We define the departure from circularity by the ratio of the ellipse half-axis w_i^x along the x-axis to the half-axis w_i^y along the y-axis (Figure 1, top), and subtract the value of one to render the parameter zero for circular symmetry ($w_i^x = w_i^y$), i.e.,

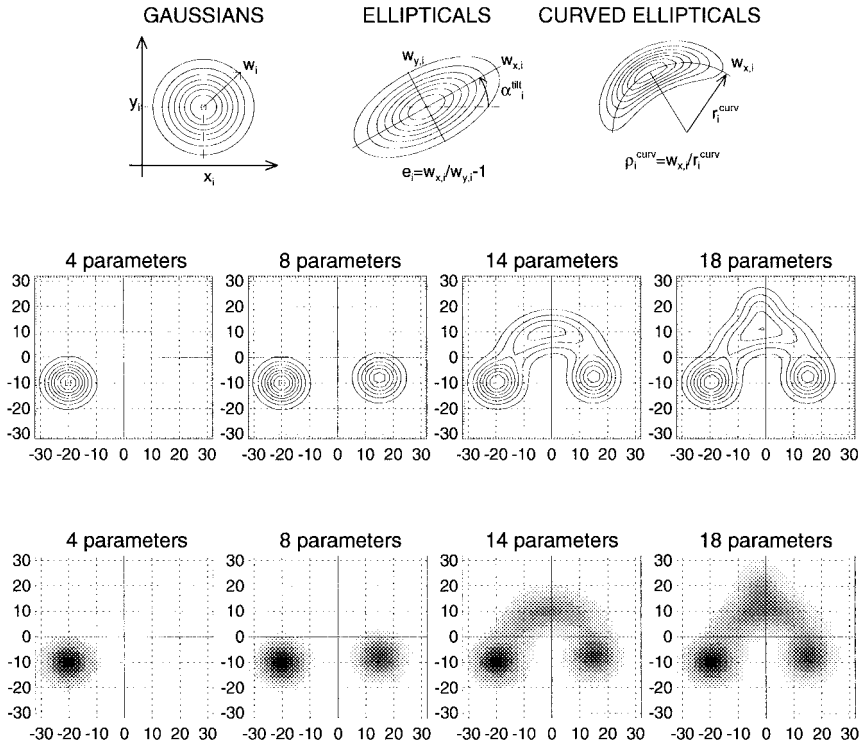

Figure 1. The parameterized model maps, using circular gaussians (4 parameters, x_i, y_i, w_i, f_i), elliptical gaussians (6 parameters, ..., e_i, α_i^{tilt}), and curved ellipticals (7 parameters, ..., ρ_i^{curv}), are defined in the *upper panels*. The *middle* and *lower rows* show 4 examples with superpositions of the basic geometric elements defined above (with 1, 2, 3, 4 elements, requiring 4, 8, 14, 18 free parameters, respectively).

$$e_i = \left(\frac{w_i^x}{w_i^y} - 1 \right) . \tag{4}$$

The 2-dimensional distribution of an elliptical gaussian is then

$$F(x'', y'') = f_i \ \exp\left[-\frac{(x'' - x_i)^2}{2w_i^2} - \frac{(y'' - y_i)^2}{2w_i^2}(e_i + 1)^2 \right] . \tag{5}$$

To obtain the 2-dimensional distribution of an elliptical gaussian in an arbitrary orientation, the coordinates (x'', y'') aligned with the elliptical major and minor axes have to be rotated by the tilt angle α_i^{tilt} (defined anti-clockwise with respect to the x-axis),

$$x = x'' \ \cos(\alpha_i^{\text{tilt}}) - y'' \ \sin(\alpha_i^{\text{tilt}}), \tag{6}$$

$$y = x'' \ \sin(\alpha_i^{\text{tilt}}) + y'' \ \cos(\alpha_i^{\text{tilt}}). \tag{7}$$

In order to provide geometric elements that can represent loop-like shapes we go a step further and introduce curved gaussians, which require only one additional parameter, the curvature radius $r^{\rm curv}$. It is mathematically more convenient to choose the reciprocal curvature radius, because this way we have no singularity $r^{\rm curv} = \infty$ for uncurved ellipticals or gaussians, and this additional parameter is then zero if no curvature is wanted. We define a dimensionless parameter $\rho_i^{\rm curv}$ by the ratio of the gaussian width w_i along the x-axis to the curvature radius $r_i^{\rm curv}$,

$$\rho_i^{\rm curv} = \frac{w_i}{r_i^{\rm curv}} \; . \tag{8}$$

For coordinate transformations from curved ellipse coordinates (x'', y'') into cartesian coordinates (x, y), it is most convenient to use the rotation angle $\alpha^{\rm curv}$ that measures the x-position on the curved x-axis with respect to the curvature center (Figure 1, top right),

$$\alpha^{\rm curv} = \frac{x''}{r_i^{\rm curv}} = \frac{x''}{w_i} \rho_i^{\rm curv} . \tag{9}$$

The transformation from curved ellipse coordinates (x'', y'') into cartesian coordinates (x', y') is then simply

$$x' = (r_i^{\rm curv} + y'') \sin(\alpha^{\rm curv}), \tag{10}$$

$$y' = (r_i^{\rm curv} + y'') \cos(\alpha^{\rm curv}). \tag{11}$$

With a subsequent rotation by the orientation angle $\alpha_i^{\rm tilt}$, the cartesian ellipse coordinates (x', y') can then be rotated into the Cartesian reference system (x, y) (by substituting (x', y') from Equations (10) and (11) into (x'', y'') of Equations (6) and (7)).

3.2. MAP PARAMETERIZATION

We have now defined the general geometry of a single source with 7 free parameters, which has the morphology of a curved gaussian in the most general form, but can be reduced to 6 free parameters (in the form of an elliptical gaussian without curvature), or to 4 parameters (in the form of a circular gaussian). The geometry of these three shapes can be described with the same general 7-parameter formalism $(x_i, y_i, f_i, w_i, e_i, \alpha_i^{\rm tilt}, \rho_i^{\rm curv})$, while setting a subset of the last three parameters to zero will automatically reproduce the simpler shapes.

We define a map simply by a superposition of the 2-dimensional distributions $F_i(x, y; P_i)$ of the elementary shapes, defined by the parameter sets $P_i = [x_i, y_i, f_i, w_i, e_i, \alpha_i^{\rm tilt}, \rho_i^{\rm curv}]$,

$$F(x, y) = \sum_{i=1}^{n^g} F_i(x, y; P_i), \tag{12}$$

with n^g the number of gaussian sources (or elementary shapes). The total number of free parameters of such a parameterized map is

$$n^{\text{free}} = n^g \times n^{\text{par}} - 1, \tag{13}$$

where n^{par} represents the number of parameters per source (i.e., $n^{\text{par}} = 4$ for circular gaussians, $n^{\text{par}} = 6$ for elliptical gaussians, and $n^{\text{par}} = 7$ for curved gaussians). The total number of free parameters is reduced by 1, because the total flux of the map can be normalized to unity [$I^{\text{norm}}(x, y) = I(x, y)/\int I(x, y)\, \mathrm{d}x\, \mathrm{d}y$], having the physical unit of 1 photon count per map. The total number of photons can be obtained by summing the time profiles, $N_l^{\text{phot}} = \sum_{k=1}^{N_l} N_l^{\text{obs}}(t_k)$ for each detector $l = 1, \ldots, 9$. A map in physical units of photons per pixel and per detector can then simply be retrieved by scaling the normalized map with the averaged photon number per detector, i.e., $I(x, y) = I^{\text{norm}}(x, y) \times \langle N_l^{\text{phot}} \rangle$. The standard unit used in RHESSI software is also normalized by time, the detector area, and arc sec^2 of a pixel, i.e., photons s^{-1} cm^{-2} arc sec^{-2}.

Some examples of such parameterized maps are shown in Figure 1, including a single gaussian source (4 parameters), double gaussian sources (8 parameters), a double source with a loop (14 parameters), and a Masuda-type flare (Masuda *et al.*, 1994) with an additional above-the-loop-top source (18 parameters).

4. Forward-Fitting Algorithm

4.1. Pixelized and unpixelized approach

We describe now the mathematical and numerical formalism of the basic *forward-fitting method*, as currently implemented in the RHESSI software (since release on 4 January 2000), a package in the Interactive Data Language (IDL)-based Solar SoftWare (SSW) (Freeland and Handy, 1998). This approach is also called the *pixelized forward-fitting method* because the spatial distribution of the photon flux $F(x, y)$ is modeled as a function of a coordinate grid (x, y) with regular spacing, either in cartesian (x, y) or polar coordinates (r, ϑ). An *unpixelized forward-fitting method* has been proposed, where an image of the photon flux can be reconstructed with a set $(i = 1, \ldots, n)$ of n point or gaussian sources at positions (x_i, y_i), without using the parameterization of a regular grid (x, y). This second method is theoretically more efficient for image reconstruction, but is currently not implemented in the RHESSI software.

4.2. The basic steps of the algorithm

The algorithm is basically defined by the input, which consists of a list of time-tagged photon events from the 9 detectors (packed in telemetry files) and the desired output in the form of a reconstructed map $I(x, y)$. This map should be

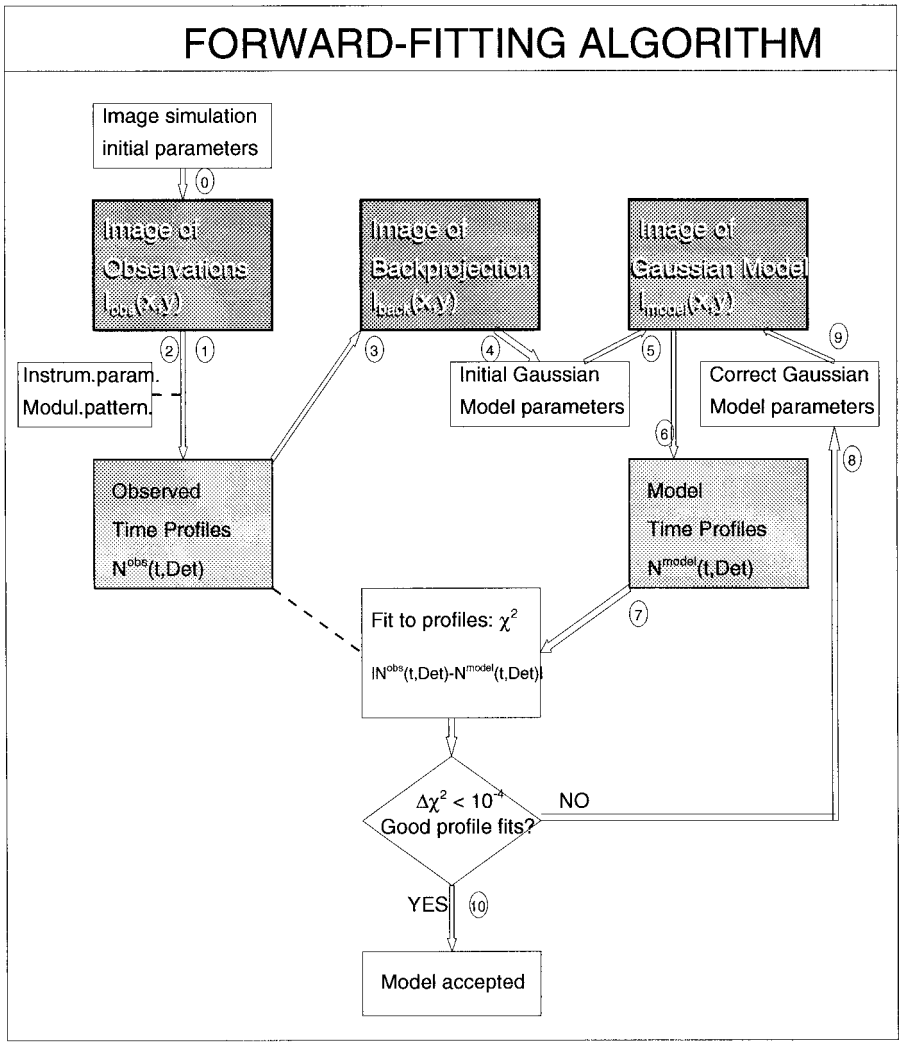

Figure 2. Flow chart of the numeric algorithm of forward fitting (see description of 10 basic steps in Section 4.2).

consistent with the observed count-rate time profiles, within the expected Poisson fluctuations. The algorithm consists of 10 basic steps (Figure 2). For the simulations here, an additional step is required beforehand, which includes the simulation of telemetry files from a model map (numbered as 'step 0' in the following scheme).

Step 0: selection of a set of model parameters $P_i, i = 0, \ldots, n^{\text{free}}$ that define a model map $I^{\text{model,sim}}(x, y)$. Simulation of a list of photon events that are randomly distributed in time and reproduce the spatial 2D distribution prescribed by the

model map. The time-tagged photon events are written out into telemetry files that have the same format as observations telemetered down from the spacecraft.

Step 1: time intervals $[t_1, t_2]$, energy intervals $[E_1, E_2]$, image size $[N_x, N_y]$, image offset from Sun center $[X_{\rm offset}, Y_{\rm offset}]$, and pixel resolution Δx are selected for desired image reconstruction.

Step 2: from the telemetry files, a binned event list is created, and then from the aspect solution, livetime calculations and instrumental calibration tables, a calibrated event list is created, $N^{\rm obs}(t_k, D_l)$ (for time bins t_k and detectors $D_l, l = 1, \ldots, 9$), and the corresponding modulation patterns are calculated, M_{ijkl} (for each pixel x_i, y_j).

Step 3: a *back-projection* (Hurford *et al.*, 2002) is performed from the calibrated event list to obtain an initial first-guess map $I^{BP}(x, y)$ for the forward-fitting procedure.

Step 4: the back-projection map $I^{BP}(x, y)$ is decomposed into circular source components (see Figure 4 and description on CD-ROM) in order to determine a first-guess value of the number (n_g) of gaussian components and their spatial positions. These are used to specify an initial set of gaussian parameters $P_i, i = 0, \ldots, n^{\rm free}$ (for a small number of gaussian source components).

Step 5: based on the first-guess gaussian parameters $P_i, i = 0, \ldots, n^{\rm free}$, a first model map $I^{\rm model1}(x, y)$ is composed. It provides the input for the first iteration of the forward-fitting procedure. We normalize the model map to 1 photon count per image, i.e., by $I^{\rm model1,norm}(x, y) = I^{\rm model,1}(x, y)/\int I^{\rm model1}(x, y)\,{\rm d}x\,{\rm d}y$.

Step 6: the first model map $I^{\rm model1}(x_i, y_j)$ is multiplied with the modulation pattern M_{ijkl} to obtain the modulation time profile P_{kl},

$$P^{\rm model1}(t_k, D_l) = M_{ijkl}\, I^{\rm model1,norm}(x_i, y_j)\ . \tag{14}$$

This is then scaled into the same physical units as the observed time profiles, i.e., counts per time bin,

$$N^{\rm model1}(t_k, D_l) = R^{\rm phot}\Delta t(D_l) q^{\rm trans}(D_l) P^{\rm model,1}(t_k, D_l)\ , \tag{15}$$

where $R^{\rm phot}$ represents the photon counts per s and per subcollimator (s^{-1} SC^{-1}), $\Delta t_k(D_l)$ is the time bin used for detector D_l, and $q^{\rm trans}(D_l)$ is the bi-grid transmission factor of detector D_l.

Step 7: the values of the reduced χ^2 (or C-statistic, which is currently implemented in the code, see Appendix A for a description) is calculated between the observed time profiles $N^{\rm obs}(t_k, D_l)$ and the model time profiles $N^{\rm model1}(t_k, D_l)$,

$$\chi^2 = \frac{1}{n_{t'}} \sum_{k'=1}^{n_{t'}} \frac{[N^{\rm obs}(t_{k'}) - N^{\rm model}(t_{k'})]^2}{[\sigma^{\rm model}(t_{k'})]^2}\ , \tag{16}$$

where the time series t_k from each detector D_l are concatenated $[t_{k'} = t_k\,(D_1), t_k\,(D_2), \ldots, t_k\,(D_9)], k' = 1, \ldots, n_{t'}$ to obtain a single χ^2 value for a given image.

The currently implemented method gives each spatial scale the same weight, by averaging the χ^2's from each detector.

Step 8: the parameters $P_i, i = 0, \ldots, n^{\text{free}}$ are varied in all (n^{free}) dimensions and the change of χ^2 (or C-statistic), $\Delta C = C(P_i + \Delta P_i) - C(P_i)$, is calculated. A new parameter set $P_i^{\text{new}} = P_i - \Delta P_i(\partial C/\partial P_i)$ is calculated from the steepest gradient of χ^2 in the n^{free}-dimensional parameter space, so that the new parameter set P_i^{new} points in the direction where C decreases most rapidly. The search for a local minimum in the direction of the steepest gradient is performed by increasing the steps ΔP_i by powers of 2 and using a cubic spline fit of $C(P_i^{\text{new}})$. The set of parameters $P_i^{\text{new}}(\min[C])$ associated with the local minimum of C is used as the starting point for the next iteration cycle of the forward-fitting procedure.

Step 9: the next model map $I^{\text{model2}}(x_i, y_j)$ is calculated based on the new parameter set $P_i^{\text{new}}, i = 0, \ldots, n^{\text{free}}$ and steps 6–8 are repeated iteratively until the change in C converges, e.g., $\Delta C < 10^{-4}$ per iteration cycle.

Step 10: after n iterations, the final map $I^{\text{model},n}(x, y)$ becomes the desired reconstructed map of the sources. In the case of the simulations, the final map $I^{\text{model},n}(x, y)$ can be compared with the initial model $I^{\text{model,sim}}(x, y)$ used for creating the telemetry files (Step 0) to verify the fidelity of the reconstruction. In the case of real observations, the final solution may be compared with maps reconstructed using other methods such as CLEAN or the maximum entropy method (MEM). A self-consistency criterion between the reconstructed map and the observations is the convergence value of χ^2 or C-statistic, which should have values near unity, say within $\lesssim 10\%$. If the χ^2 or C-values are significantly above unity (provided there is no convergence problem to the global minimum), this would indicate that the model parameterization is not adequate to represent the data, and thus requires a more complex model.

This outline is only a formal description of the basic steps of the forward-fitting code (Figure 2). To understand the most efficient implementation, the accuracy and limitations of the solutions, and the convergence behavior of the algorithm, we provide more detailed information in the following subsections.

4.3. Limits set by Poisson fluctuations

It is a property of RHESSI image reconstruction algorithms that parameter changes are most significant in the first iteration steps, while later iterations near the convergence limit are less efficient and essentially are dominated by Poisson fluctuations. (For the purposes of this paper, we ignore background, cosmic rays and other types of noise, which in many cases have been found to be smaller than the Poisson fluctuations.) In the case of forward-fitting in particular, it is desirable to suppress those computation-expensive parts of the calculations that bluntly fit the Poisson fluctuations. A first important efficiency consideration is therefore to select only those detectors with time bins large enough to have a reasonable count-to-fluctuation ratio to be sensitive to the fitting algorithm.

The time profiles of the finest collimators have the greatest Poisson fluctuations, because with the standard RHESSI time-binning, the same number of photons (per detector) is spread out over a larger number of time bins (by the requirement of resolving the rotational modulation). Ignoring background, the amplitude of the fluctuations is entirely Poisson, i.e., count/fluctuations $= \sqrt{N^{\rm phot}}$, and the ratio (for standard binning) is smaller for the finer grids, because it decreases as $C/F \sim \sqrt{\Delta t_i^{\rm bin}}$, where $\Delta t_i^{\rm bin}$ is the time bin length for detector i. For instance, if the coarsest grid gives a count/fluctuations ratio of 100:1, it will be 12:1 for the finest grid, because the default time bins vary by a factor of $\Delta t_9^{\rm bin}/\Delta t_1^{\rm bin} = 64$.

In order to decide which time profiles can be ignored in the forward-fitting procedure without significant loss of information, we define a significance criterion that is expressed by the ratio of its standard deviation $\sigma^{\rm r.m.s.}$ to the expected Poisson fluctuations $\sigma^{\rm Poisson}$. The r.m.s. fluctuation of a time profile $N^{\rm obs}(t_k, D_l)$ is composed of the variability due to the rotational modulation $\sigma^{\rm mod}$ and the photon count statistic $\sigma^{\rm Poisson}$,

$$(\sigma^{\rm r.m.s.})^2 = (\sigma^{\rm mod})^2 + (\sigma^{\rm Poisson})^2. \tag{17}$$

For instance, when the variability due to rotational modulation, $\sigma^{\rm mod}$, falls below 20% of the counting fluctuations, $\sigma^{\rm Poisson}$, the combined r.m.s. variability, $\sigma^{\rm r.m.s.} = \sigma^{\rm Poisson} \times \sqrt{1 + 0.2^2} = 1.02\ \sigma^{\rm Poisson}$, exceeds the Poisson fluctuations by 2% only. Any fitting method based on a χ^2-criterion becomes quite insensitive at this level, because $\chi^2 \sim (\sigma^{\rm Poisson})^{-2}$ (Equation (16)) varies then by 4% only, which has little significance in judging the goodness of a fit. Therefore, we adopt a threshold criterion of

$$\left(\frac{\sigma^{\rm r.m.s.}}{\sigma^{\rm Poisson}}\right) \geq 1.02, \tag{18}$$

to select detectors that have time profiles with modulation large enough to be usable by the forward-fitting procedure. This criterion is equivalent to

$$\left(\frac{\sigma^{\rm mod}}{\sigma^{\rm Poisson}}\right) \geq \sqrt{(1.02^2 - 1)} = 0.2\ . \tag{19}$$

The average counts per time bin is $N_l^{\rm obs} = R^{\rm phot} \Delta t_l q_l^{\rm trans}$ (see Equation (15)), while the total number of events in the time interval $(t_2 - t_1)$ is $N_l^{\rm tot} = R^{\rm phot}(t_2 - t_1) q_l^{\rm trans}$. For Poisson statistics, with $n = N_l^{\rm tot}$ events and a mean of $\langle x \rangle = N_l^{\rm obs}$ the standard deviation from the mean is $\sigma = \sqrt{\langle x \rangle}$ (Bevington, 1969, p. 33), and thus

$$\sigma_l^{\rm Poisson} = \sqrt{N_l^{\rm obs}} = [R^{\rm phot} q_l^{\rm trans} \Delta t_l]^{1/2}. \tag{20}$$

In Figure 3 we compare the r.m.s. $\sigma_l^{\rm Poisson}$ with the fluctuations due to rotational modulation $\sigma_l^{\rm mod}$, which is

$$\sigma_l^{\rm mod} \approx 0.5\ \varepsilon_l^{\rm mod} N_l^{\rm obs} \tag{21}$$

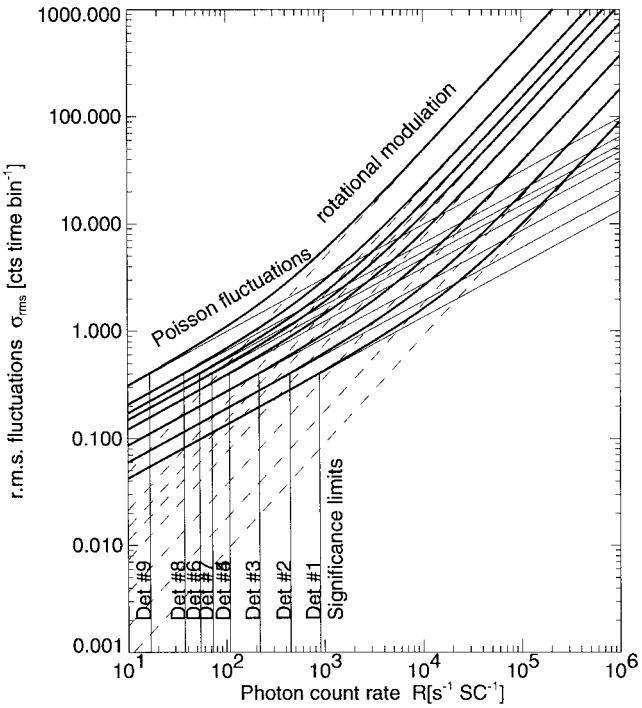

Figure 3. The r.m.s. fluctuations of observed time profiles $N^{\rm obs}(t_k, D_l)$ as a function of the photon count rate R (*thick lines*), modeled in terms of two components: the photon counting r.m.s. fluctuations $\sigma^{\rm Poisson} \sim \sqrt{R}$ (*thin lines*), and the r.m.s. fluctuations due to the rotational modulation $\sigma^{\rm mod} \sim R$ (*dashed lines*). Each component is shown for each detector l, $D_1, \ldots, D_9$, where the coarsest detector D_9 represents the topmost lines. Note that the photon Poisson fluctuations dominate for small photon count rates R_l. The limit $R_l^{\rm min}$ where the rotational modulation exceeds 20% of the Poisson fluctuations is indicated with a *vertical line* for each detector. These values of $R_l^{\rm min}$, which represent lower limits for the photon rate where rotational modulation is detectable in the time profiles, are also listed in Table I.

for the 9 detectors ($l = 1, \ldots, 9$) and modulation efficiency $\varepsilon_l^{\rm mod} (\leq 1.0)$. From the requirement that the variability due to rotational modulations exceeds the Poisson fluctuations by 20% (Equation (19)), we can infer a requirement for a minimum photon rate to produce significant modulations in the detectors. Combining Equations (19)–(22) we find

$$N_l^{\rm obs,min} \geq \left(\frac{0.2}{0.5 \varepsilon_l^{\rm mod}} \right)^2 = \frac{0.16}{(\varepsilon_l^{\rm mod})^2} \,, \tag{23}$$

which yields a minimum photon rate $R_l^{\rm min}$ of

$$R_l^{\rm min} \geq \frac{N^{\rm obs,min}}{q_l^{\rm trans} \Delta t_l} \approx \frac{10^3}{\sqrt{3}^{\,l-1}}. \tag{23}$$

These limits are indicated in Figure 3 and tabulated in Table I, ranging from $R_1^{\min} \geq 900$ counts s^{-1} SC^{-1} for the finest detector No. 1 to $R_9^{\min} \geq 17$ counts s^{-1} SC^{-1} for the coarsest detector No. 9. These count rate limits are derived for 100% modulation efficiency, but would be higher for source sizes that are comparable or larger than the collimator resolution, where the modulation efficiency drastically drops. They would also be higher at higher energies because the grids become transparent, and at lower energies for the finest grids because of diffraction effects.

The count-rate limits $R_l^{\min}$ for different detectors l (indicated in Figure 3 and listed in Table I) save a lot of computational effort. For instance, if a weak flare with a relatively low count rate of $R^{\text{phot}} = 100$ counts s^{-1} SC^{-1} is observed, we see immediately (Table I) that this count rate is below the requirement $R_l^{\min}$ of significant modulation for detectors Nos. 1–5. We need therefore only to consider detector Nos. 6–9 for image reconstruction. The number of time bins from detectors Nos. 6–9 is $\sum_{i=6}^{9} N_i^{\text{bin}} = 1408$, and thus a factor of $(17\,792/1408) = 12.6$ smaller compared with the number of time bins from all detectors Nos. 1–9, making the computation a factor of ≈ 12 faster (since the computation of modulation profiles scales linearly with the number of time bins).

4.4. Decomposition of backprojection maps

For the first guess of source positions the forward-fitting algorithm makes use of a backprojection map (Hurford *et al.*, 2002), by iterative decomposition of the strongest peaks in the backprojection map. The exact peak location (x_i, y_i) is determined with sub-pixel accuracy by parabolic interpolation of the maximum in both x- and y-direction. An example with a detailed description of the procedure is given in Figure 4 on the CD-ROM.

4.5. Parameter optimization in forward-fitting

The forward-fitting method requires the optimization of n^{free} parameters until the modulation time profiles resulting from the model map are consistent with the observed time profiles $N^{\text{obs}}(t_k, D_l)$. For simple maps, e.g., the examples shown in Figure 1, this task consists in the optimization of 3–18 free parameters. The minimization of the χ^2 or C-statistic of a parameter set (P_i), $i = 1, \ldots, n^{\text{free}}$ requires the calculation of a modulation profile $N^{\text{model}}(t_k, D_l)$ for each new trial of a model map $I(x, y; P_i)$, in order to calculate the χ^2 or C-statistic of a new parameter set (P_i). We developed a variant of the multi-dimensional *downhill simplex method* (Press *et al.*, 1986, p. 289), which is described in some detail in Figure 6 on the CD-ROM.

4.6. Numerical simulations and tests

In this study we performed tests using some 500 simulations with image reconstructions by the forward-fitting method. To obtain statistics on the accuracy of

reconstruction we simulated basic maps that contain one or two circular gaussian sources, which can be quantified by 4 or 8 free parameters. A few simulations with more complex source topologies were also performed as a feasibility test, i.e. with triple, quadruple, elliptical, and curved elliptical sources. The results of the tests are reported in more detail in the accompanying CD-ROM (Figures 5–13). We tested mainly the accuracy Δr_i of the reconstructed source positions (x_i, y_i), widths (w_i), and flux amplitudes (a_i) for model maps with 1 or 2 gaussian sources, and obtained uncertainties that scale with the square root of the count rate R,

$$\frac{\Delta r_i}{w_i} = \frac{\sqrt{(x_i^{\rm fit} - x_i^{\rm sim})^2 + (y_i^{\rm fit} - y_i^{\rm sim})^2}}{w_i^{\rm sim}} = \sigma_r \left(\frac{R}{10^4}\right)^{-1/2}, \tag{24}$$

$$\frac{\Delta w_i}{w_i} = \frac{(w_i^{\rm fit} - w_i^{\rm sim})}{w_i^{\rm sim}} = \sigma_w \left(\frac{R}{10^4}\right)^{-1/2}. \tag{25}$$

The results can be summarized as follows. For double-source simulations, a forward-fitting solution with two gaussians was found in 90 cases, using a fully automated computation mode without human interaction. For uncontaminated statistics of errors, we excluded cases with insufficient separation and cases with weak secondary sources. For a count rate of $R = 10^4$ counts s^{-1} SC^{-1}, two gaussian centroid positions could be reconstructed with automated forward-fitting with a typical positional accuracy of $\sigma_r \approx 10$–20% of their gaussian width, and a width accuracy of $\sigma_w \approx 5\%$. The total photon flux was typically retrieved with an accuracy of $\approx 1\%$. Overlapping double sources or weak sources with a contrast of less than 1:20 are less reliably reconstructed. The details of the accuracy tests are provided in Figures 8–13 on the accompanying CD-ROM.

4.7. APPLICATIONS

The six different image reconstruction algorithms currently supported by the RHESSI software each have different advantages and disadvantages. Image reconstruction with forward-fitting has the specific property that it places all the flux into the prescribed number of gaussian source components, and thus has a high photometric accuracy because no flux is spread into noisy residuals (ignoring the detector background). Another specific property is that the determined source positions are more stable than in other imaging algorithms. This is because the fitted gaussians are sensitive to the centroid position. For these reasons, forward-fitting is especially well suited to measure small differential changes of source positions in time or energy. A possible application for small differential changes in source positions as a function of time is the separation of flare loop footpoints. This is expected to increase systematically in the Kopp–Pneuman model (Kopp and Pneuman, 1976) due to the rise of the X-type reconnection point. Previous measurements, based on *Yohkoh* HXT images, showed a mixed trend, where only half of the flares showed

evidence for the predicted increase in footpoint separation with time (Sakao, 1998). Another possible application for small differential changes in source positions is the energy-dependent altitude of footpoint sources, which can be measured most favorably near or above the limb. Such accurate height measurements can constrain the energy loss of precipitating electrons in the chromosphere and can be used to infer a chromospheric density model (Brown, Aschwanden, and Kontar, 2002; Aschwanden, Brown, and Kontar, 2002).

5. Summary and Conclusions

We described a newly-developed forward-fitting method to reconstruct images from the RHESSI telescope for simple source morphologies that can be represented by a superposition of multiple gaussians. We performed some 500 numerical simulations of model source geometries and compared the forward-fitted images with the model images in order to quantify the fidelity of the image reconstruction, the accuracy of the retrieved physical parameters, the maximum spatial resolution, and the convergence behavior. These tests, covered in detail in the accompanying RHESSI CD-ROM, represent a basic reference for all other RHESSI image reconstruction algorithms, providing error bars and quantitative uncertainties that are important for the scientific interpretation of hard X-ray images from RHESSI. The results of the forward-fitting method can be summarized as follows:

(1) Using circular, elliptical, and curved gaussians, most of the so far observed hard X-ray maps of solar flares can be represented by a relatively small number of free parameters, e.g., $n^{\rm free} = 3$ for a single source, $n^{\rm free} = 7$ for double sources, $n^{\rm free} = 14$ for double footpoint flares with loop-top hard X-ray emission, or $n^{\rm free} = 18$ for Masuda-type flares. The current implementation of the forward-fitting code is designed to reconstruct hard X-ray images for this variety of source morphologies.

(2) The minimum count rate required for a detector to produce modulation large enough for the current forward-fitting algorithm to work is approximately $R_l^{\rm min} = 1000/\sqrt{3}^{\,(i-1)}$ counts s^{-1} SC^{-1} for detector $l = 1, \ldots, 9$, e.g., $R_1^{\rm min} \approx 1000$ counts s^{-1} SC^{-1} for the finest grid (detector No. 1), and $R_9^{\rm min} \approx 10$ counts s^{-1} SC^{-1} for the coarsest grid (detector No. 9). For a given flare event with count rate R, detectors that have a count rate below this limit, $R_l^{\rm min} < R$, should be discarded for image reconstruction by this method. This saves considerable computational effort.

(3) Test runs of forward-fitting with model maps containing a single gaussian source show that the relative accuracy of most gaussian parameters scales with the count/Poisson-fluctuation ratio, i.e., $\sim \sqrt{R}$, with R being the photon count rate. The average accuracy in the reconstruction of single gaussian sources at a count rate of $R = 10^4$ counts s^{-1} SC^{-1} is $\approx 5\%$ (of the source width) for the position, and $\approx 3\%$ for the gaussian source width. The accuracy depends primarily on the photon count rate and not at all on the chosen pixel size (as long as the source

is larger than the half pixel size). Subpixel accuracies in source position down to 1% can be achieved for count rates of $R \approx 10^5$ counts s^{-1} SC^{-1}, slightly above the value at which the RHESSI attenuators automatically switch in. This allows efficient mapping to be performed with relative large pixel sizes without losing accuracy.

(4) Test runs of forward-fitting with model maps containing double gaussian sources demonstrate that the positions and widths can be reconstructed with almost the same accuracy as single source maps, unless there are source confusion problems or strong contrast ($\lesssim$ 1:20 for the weaker source). We find that the forward-fitting code converges to a χ^2 or C-statistic value that is equally consistent with the data as the model.

(5) The maximum spatial resolution that can be obtained with the forward-fitting technique for high count rates is $\mathrm{FWHM}^{\mathrm{min}} \approx 0.8''$, which is about a factor of 3 below the nominal resolution of the finest grid, i.e., $\mathrm{FWHM}_1^{\mathrm{res}} = 2.3''$.

(6) We find that the forward-fitting algorithm virtually always converges to the same solution as the model (used for the simulation of the telemetry data). The final χ^2- and C-statistic values reached at the convergence limit, which are found in the range of $0.95 \lesssim C \lesssim 1.1$, are constrained by the Poisson fluctuation statistics of a particular event, and not by the convergence behavior of the forward-fitting code.

(7) Forward-fitting is especially well suited to measure small differential changes of source positions in time or energy. Possible applications are energy-dependent altitude measurements of energy loss near the limb, or footpoint separations during flares.

Future work with the forward-fitting method will include (1) the simulation of RHESSI telemetry data with more complex source morphologies (i.e., curved loops, Masuda-type flares), (2) combined spatial-spectral imaging, (3) combined spatial-temporal imaging, (4) fast imaging on time intervals shorter than a half spacecraft spin period, and quantitative comparisons with other RHESSI imaging methods.

Acknowledgement

Support for this work was provided by the NASA SMEX grant NAS5-98033 through University of California, Berkeley (subcontract SA2241-26308PG).

Appendix
χ^2 and C-Statistic for Sparse Sampling

An observed time series of photon counts, $f^{\mathrm{obs}}(t)$, which is binned into time intervals of length Δt, so that the number $n_i = f^{\mathrm{obs}}[t_i < t < t_i + \Delta t]$ represents the photon counts in bin $[t_i, t_i + \Delta t]$, is modeled by an analytical function, $f^{\mathrm{model}}(t)$,

which predicts an expectation value $e_i = f^{\rm model}[t_i < t < t_i + \Delta t]$ of photon counts in time bin i. In the limit of large-number counts, $n_i \gg 0$, Poisson statistics can be approximated by a gaussian distribution, which predicts an r.m.s. fluctuation of $\sigma_i = \sqrt{e_i} \approx \sqrt{n_i}$. The goodness of fit of a model can then be evaluated by the well-known χ^2-statistic,

$$\chi^2 = \frac{1}{N} \sum_{i=1}^{N} \frac{(n_i - e_i)^2}{\sigma_i^2} = \frac{1}{N} \sum_{i=1}^{N} \frac{(n_i - e_i)^2}{e_i} \ , \tag{A1}$$

yielding a normalized value of $\chi^2 \approx 1$ in the case when the model is consistent with the data, within the uncertainties of the expected fluctuations.

For sparse sampling, in particular when the observed time series n_i contains zero's or few counts, say $n_i \lesssim 10$, e.g., as it is the case for the default timebins of the finest RHESSI collimators, where the total photon counts are binned into some $2^{12} = 4096$ time bins (for the finest grids) to resolve the rotational modulation, the gaussian χ^2-statistic is not appropriate anymore. Photon counting statistics for sparse sampling can be derived from the general probability function P as outlined by Cash (1979), which is

$$P = \prod_{i=1}^{N} \frac{e_i^{n_i} e^{-e_i}}{n_i!} \tag{A2}$$

for a particular result n_i given the correct set of e_i. The likelihood ratio, called *C-statistic* by Cash (1979), can be expressed in logarithmic form from Equation (A2),

$$C_{\rm Cash} = -2 \ln P = -2 \sum_{i=1}^{N} (n_i \ \ln \ e_i - e_i - \ln n_i!) \ . \tag{A3}$$

Let the theoretical model $e_i(\Theta_1, \ldots, \Theta_p)$ be defined by p free parameters Θ_i, $i = 1, \ldots, p$. The best-fitting model is found by varying all p parameters $\Theta_1, \ldots, \Theta_p$ until the C-statistic reaches a minimum, which we denote by $(C_{\rm min})_p$. Now, during the iterative fitting procedure, only a partial subset of model parameters, say q parameters (with $q < p$), may have already converged to the true solution, $\Theta_1^T, \ldots, \Theta_p^T$, so that $\Theta_1, \ldots, \Theta_q$ are set to $\Theta_1^T, \ldots, \Theta_q^T$, while the remaining $p - q$ parameters, $\Theta_{q+1}, \ldots, \Theta_p$, still need to be varied until a global minimum of C is reached. We denote this partial solution, where $p - q$ parameters have already converged to the true solution Θ_i^T, with the value $(C_{\rm min})_{p-q}^T$. According to Cash, quoting the theorem of Wilks (1938, 1963), the difference

$$\Delta C = (C_{\rm min})_{p-q}^T - (C_{\rm min})_p \tag{A4}$$

will be distributed as χ^2 with q degrees of freedom. Therefore, the quantity ΔC can be used to establish a confidence criterion of the model e_i to the data n_i. Following Cash (1979), the term $\ln(n_i!)$ of Equation (A3) drops out in the difference ΔC, because only the parameters $e_i(\Theta_1, \ldots, \Theta_p)$ are varied during the fitting procedure. Thus, it is more convenient to use the simplified statistic

$$C_{\text{simplified}} = -2 \ln P = -2 \sum_{i=1}^{N} (n_i \ \ln \ e_i - e_i) \ . \tag{A5}$$

For the evaluation of the difference ΔC in Equation (A4) we have the partially optimized term $(C_{\min})_{p-q}^{T}$,

$$(C_{\min})_{p-q}^{T} = -2 \sum_{i=1}^{N} (n_i \ \ln \ (e_i)_{p-q}^{T} - (e_i)_{p-q}^{T}) \ , \tag{A6}$$

and the absolute minimum $(C_{\min})_p$, which represents the asymptotic limit when the data n_i perfectly match the model, i.e., $e_i = n_i$,

$$(C_{\min})_p = -2 \sum_{i=1}^{N} [n_i \ \ln (n_i) - n_i] \ . \tag{A7}$$

The combination of Equation (A4)–(A6) yields then

$$\Delta C = -2 \sum_{i=1}^{N} [n_i \ \ln (e_i) - e_i - n_i \ \ln (n_i) + n_i] \ . \tag{A8}$$

Instead of the reduced χ^2-statistic (A1), we can therefore use the equally simple C-statistic (where we drop the symbol Δ for brevity),

$$C := \frac{1}{N} \Delta C = \frac{2}{N} \sum_{i=1}^{N} \left[n_i \ \ln \left(\frac{n_i}{e_i} \right) - (n_i - e_i) \right] \ . \tag{A9}$$

This form has the advantage that, in addition to being asymptotic to χ^2, C vanishes identically when the model fits the data exactly. Numerical care has to be taken for the time bins that contain zeros in the data, $n_i = 0$, in which case the mathematical relation $n_i \ln (n_i) = 0$ has to be used to avoid the singularity $\ln (n_i) \mapsto \infty$. On the other hand, a singularity could arise when the model predicts zero counts, $e_i = 0$, but the observed counts are not zero, $n_i \neq 0$, because the term $- n_i \ \ln (e_i)$ yields then infinity. It seems therefore to be recommendable to restrict the model fit to time intervals with a finite probability for photon counts, i.e., $e_i > 0$.

References

Aschwanden, M. J., Brown, J. C. and Kontar, E. P.: 2002, *Solar Phys.*, this volume.

Bevington, P. R.: 1969, *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill Book Co., New York.

Brown, J. C., Aschwanden, M. J., and Kontar, E. P.: 2002, *Solar Phys.*, this volume.

Cash, W.: 1979, *Astrophys. J.* **228**, 939.

Dennis, B. R. *et al.*: 1996, *Proc. SPIE* **2804**, 228.

Freeland, S. L. and Handy, B. N.: 1998, *Solar. Phys.* **182**, 497.

Hurford, G. J. *et al.*: 2002, *Solar Phys.*, this volume.
Kopp, R. A. and Pneuman, G. W.: 1976, *Solar Phys.* **50**, 85.
Kosugi, T., Sawa, M., Sakao, T., Masuda, S., Inda-Koide, M., Yaji, K., and Sato, J.: 1995, *The Yohkoh HXT Databook October 1991 – December 1994*, Natl. Astron. Obs. Tokyo, Japan.
Lin, P. R., Dennis, B. R., Emslie, A. G., Ramaty, R., Canfield, R., and Doschek, G.: 1993, *Adv. Space Res.* **13**, 401.
Lin, P. R., Dennis, B. R., Ramaty, R., Emslie, A. G., Canfield, R., and Doschek, G.: 1994, *Geophys. Monogr.* **84**, 283.
Lin, P. R. *et al.*: 1998, *Proc. SPIE* **3442**, 2.
Masuda, S., Kosugi, T., Hara, H., Tsuneta, S., and Ogawara, Y.: 1994, *Nature* **371**, 495.
Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T:. 1986, *Numerical Recipes, the Art of Scientific Computing*, Cambridge University Press, Cambridge.
Sakao, T.: 1994, Ph.D. Thesis, University of Tokyo.
Sakao, T., Kosugi, T., and Masuda, S.: 1998, in T. Watanabe, T. Kosugi, and A. C. Sterling (eds.), *Observational Plasma Astrophysics: Five Years of Yohkoh and Beyond*, Kluwer Academic Publishers, Dordrecht, Holland; p. 273.
Sato, J., Sawa, M., Masuda, S., Sakao, T., Kosugi, T., and Sekiguchi, H.: 1998, *The Yohkoh HXT Image Catalogue, October 1991 – August 1998*, Nobeyama Radio Observatory/NAO, Japan.
Smith D. M. *et al.*: 2002, *Solar Phys.*, this volume.
Wilks, S. S.: 1938, *Ann. Math. Stat.* **9**, 60.
Wilks, S. S.: 1963, *Mathematical Statistics*, Ch. 13, Princeton University Press, Princeton.