

A regularization approach for the analysis of RHESSI X-ray spectra[★]

A. M. Massone¹, M. Piana², A. Conway^{3,5}, and B. Eves⁴

¹ INFN, Unità di Ricerca di Genova, via Dodecaneso 33, 16146 Genova, Italy
e-mail: massone@ge.infn.it

² INFN and Dipartimento di Matematica, Università di Genova, via Dodecaneso 35, 16146 Genova, Italy
e-mail: piana@dima.unige.it

³ Department of Physics, Open University, Milton Keynes MK7 6AA, UK
e-mail: a.j.conway@open.ac.uk

⁴ Department of Physics, Open University, Milton Keynes MK7 6AA, UK
e-mail: b.a.c.eves@open.ac.uk

⁵ Department of Physics and Astronomy, The University, Glasgow G12 8QQ, UK
e-mail: andrew@astro.gla.ac.uk

Received 9 January 2003 / Accepted 1 April 2003

Abstract. The Ramaty High Energy Solar Spectroscopic Imager (RHESSI) mission provides high resolution X-ray spectra emitted by collisional bremsstrahlung of electrons with the ions of the plasma. One of the aims of the mission is to infer information about the acceleration and transport mechanisms of electrons and ions in solar flares. Within the framework of an optically thin model and under the assumption of isotropic conditions in the solar plasma, the emitted photon spectrum is related to the electron flux density spectrum by an integral equation of the first kind. This problem is characterized by a notable numerical instability as a consequence of the experimental errors in the measured photon data. We address the solution of this problem by using the Tikhonov regularization algorithm. The method is applied to the case of three photon spectra registered by RHESSI detectors in the first months of the observational activity of the satellite. Furthermore a comparison of the reconstructions provided by the algorithm is performed in the case of the Kramers, Bethe-Heitler and Bethe-Heitler with the Elwert correction models for the bremsstrahlung cross-section.

Key words. Sun: flares – Sun: X-rays, gamma rays – methods: data analysis

1. Introduction

RHESSI (Ramaty High Energy Solar Spectroscopic Imager) is an imaging spectroscopy providing two-dimensional X-ray images of the solar atmosphere and one-dimensional hard X-ray spectra and γ -rays lines emitted during solar flares. In particular, the RHESSI mission was launched by NASA on February 5 2002 with the aim of providing crucial insights in the modeling of solar flares, by combining 2D imaging with the spectral information given by high-resolution Ge detectors.

Solar flares are transient phenomena characterized by the sudden release of huge amounts of energy. The typical manifestations of a solar flare are:

- acceleration of particles in the solar plasma;
- heating of the solar chromosphere;
- significant electromagnetic emission, particularly in the X-ray range.

One of the most challenging problems in solar physics is the determination of the dynamical origin of flares. In particular, in all the most reliable theoretical models (stochastic processes such as Fermi acceleration, DC electric field acceleration and Maxwellian bulk heating) the shape of the energy spectrum of the accelerated particles is a key to investigate the type of acceleration mechanism involved. In fact the RHESSI scientific objectives are:

- to determine the location and mechanism of particle acceleration in flares with particular emphasis on hard X-rays in the case of electron acceleration and γ -ray lines in the case of proton and ion acceleration;
- to study the relation between particle acceleration and plasma heating;
- to study the particle and energy transport phenomena during flares.

Since the seminal papers by Brown (1971) and Lin (1974), it is well established that the dominant X-ray emission mechanism during solar flares is collisional bremsstrahlung of electrons with the ions of the low density atmospheric plasma.

Send offprint requests to: M. Piana, e-mail: piana@dima.unige.it

[★] Figures 4 and 5 and Appendix A are only available in electronic form at <http://www.edpsciences.org>

Our analysis is focused on the case of optically thin bremsstrahlung radiation (Brown & Emslie 1988), so that absorption can be neglected; furthermore we assume that the electron velocity distribution is characterized by isotropic conditions and, finally, that the plasma is hydrogen dominated, so that the ions are almost all protons. Under these conditions the bremsstrahlung source can be described in terms of the proton density $n(\mathbf{r})$ (cm^{-3}) and the differential flux $F(E, \mathbf{r})$ (electrons $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$) at the non-relativistic energy E , both at position \mathbf{r} within the source volume V . Then the rate of emission of photons per unit ϵ is (Craig & Brown 1986)

$$g(\epsilon) = \bar{n}V \int_{\epsilon}^{\infty} f(E)Q(\epsilon, E)dE, \quad (1)$$

where

$$\bar{n} = \frac{1}{V} \int_V n(\mathbf{r})dV, \quad (2)$$

$$f(E) = \frac{1}{\bar{n}V} \int_V n(\mathbf{r})F(E, \mathbf{r})dV \quad (3)$$

and $Q(\epsilon, E)$ is the bremsstrahlung cross-section differential in photon energy. A rather complete survey of possible analytical forms for $Q(\epsilon, E)$ is given in Koch & Motz (1959). Here we will consider three different cross-sections: the Kramers formula

$$Q(\epsilon, E) = \begin{cases} 0 & E \leq \epsilon \\ \frac{Q_0}{\epsilon E} & E > \epsilon \end{cases} \quad (4)$$

and the Bethe-Heitler formula

$$Q(\epsilon, E) = \begin{cases} 0 & E \leq \epsilon \\ \frac{Q_0}{\epsilon E} \log \frac{1 + \sqrt{1 - \epsilon/E}}{1 - \sqrt{1 - \epsilon/E}} & E > \epsilon \end{cases} \quad (5)$$

with

$$Q_0 = \frac{8}{3} \frac{r_0^2}{137} mc^2; \quad (6)$$

and the Bethe-Heitler formula with the Elwert correction (Johns & Lin 1992a,b)

$$Q(\epsilon, E) = \begin{cases} 0 & E \leq \epsilon \\ Q_1 q_1(\epsilon, E) q_2(\epsilon, E) q_3(\epsilon, E) q_4(\epsilon, E) q_5(\epsilon, E) & E > \epsilon \end{cases}, \quad (7)$$

with

$$Q_1 = \frac{Z^2 r_0^2}{137} \frac{16}{3}; \quad (8)$$

$$q_1(\epsilon, E) = \frac{m^2 c^4}{\epsilon E (E + 2mc^2)}; \quad (9)$$

$$q_2(\epsilon, E) = \log \left[\frac{1 + \left(\frac{(E-\epsilon)(E-\epsilon+2mc^2)}{E(E+2mc^2)} \right)^{1/2}}{1 - \left(\frac{(E-\epsilon)(E-\epsilon+2mc^2)}{E(E+2mc^2)} \right)^{1/2}} \right]; \quad (10)$$

$$q_3(\epsilon, E) = \left[E (E + 2mc^2) \right]^{1/2} \left[E - \epsilon + mc^2 \right]; \quad (11)$$

$$q_4(\epsilon, E) = \left\{ 1 - \exp \left[- \frac{2\pi (E + mc^2)}{137 [E (E + 2mc^2)]^{1/2}} \right] \right\}; \quad (12)$$

$$q_5(\epsilon, E) = \left\{ \left[(E - \epsilon) (E - \epsilon + 2mc^2) \right]^{1/2} \times [E + mc^2] \times \left\{ 1 - \exp \left[- \frac{2\pi (E - \epsilon + mc^2)}{137 [(E - \epsilon) (E - \epsilon + 2mc^2)]^{1/2}} \right] \right\} \right\}^{-1}. \quad (13)$$

In Eqs. (8)–(13), Z represents the atomic number of the ion with which the electron collides, r_0 is the classical electron radius, m is the rest mass of the electron and c is the speed of light.

In real applications only a finite set of integrals of the photon spectrum over small photon energy intervals is available. However, if these intervals are small compared with the resolution implied by the smoothing kernel (i.e., the cross-section), the integration effect provided by the detectors can be neglected and the equation describing the emission process is, rather realistically,

$$g(\epsilon_n) = \int_{\epsilon_n}^{\infty} Q(\epsilon_n, E) f(E) dE \quad n = 1, \dots, N. \quad (14)$$

The problem of determining the unknown electron spectrum $f(E)$ from the knowledge of the photon data set $\{g_n\}_{n=1}^{\infty}$ is a typical example of linear ill-posed inverse problem with discrete data (Bertero et al. 1985; Bertero et al. 1988). Without focusing on the technical definition (Hadamard 1923), in this context ill-posedness essentially means two facts: first, the solution of the problem is not unique since there are infinitely many functions fitting the photon data in Eq. (14); second, when the data are affected by experimental noise, a *naive* inverse approach to the solution of the problem induces heavy numerical instabilities which make the restored flux density spectrum completely unreliable from a physical point of view. In order to obtain the maximum benefit from RHESSI data we think that it is essential to deal with this inverse problem by allowing for its ill-posed nature. In particular we believe that the only way to systematically handle both the ambiguity involved by the non-uniqueness of the solution and the numerical instability induced by the presence of noise on the data is to apply numerical techniques based on the regularization theory for linear ill-posed inverse problems. A description of the regularization theory approach in this context can be found in Piana (1994). The basic requirement of such an approach is to solve the minimum problem

$$\|Af - \mathbf{g}\|^2 + \lambda \|f\|^2 = \text{minimum}, \quad (15)$$

where (Af) is a vector with components

$$(Af)_n = \int_{\epsilon_n}^{\infty} Q(\epsilon_n, E) f(E) dE, \quad (16)$$

\mathbf{g} is a vector with components $g_n = g(\epsilon_n)$ and the norms are taken in the appropriate spaces. On the left hand side of Eq. (15), the first term measures the fitting accuracy while the second term introduces a constraint on the size of the solution, suppressing large departures from a smooth electron spectrum

and therefore is concerned with stability. The regularization parameter $\lambda > 0$ tunes the trade-off between approximation and stability. In Piana (1994) the Tikhonov (1963) regularization method has been applied for the solution of Eq. (14) only with the Bethe-Heitler cross-section in the case of the X-ray spectrum registered during the June 27 1980 flare by Ge detectors placed on a balloon. In the present paper we will follow an analogous approach in the case of the high-resolution RHESSI spectra, this time by considering the Kramers cross-section and the Elwert correction, too. More precisely, in Sect. 2 three different RHESSI spectra, registered in the first months of the satellite's flight, will be inverted by using the Tikhonov algorithm. Then, in Sect. 3, we will discuss the main advantages of using a regularized inversion technique. Finally Sect. 4 will contain a discussion of the results. In Appendix A an overview of the Tikhonov method within the framework of the linear inverse problem theory will be given.

2. RHESSI spectra: Application of Tikhonov method

Tikhonov regularization theory is a two-step inversion technique providing an optimal trade-off between stability and best-fitting. In the first step, many approximate reconstructions are obtained by solving problem (15) for different values of the regularization parameter λ . The second step consists in selecting one particular approximate reconstruction by means of an optimal choice of the parameter in the quadratic functional. The optimality criterion at the basis of this choice will be to tune this real positive number λ in order that the numerical oscillations are reduced but the compatibility with the observed data is preserved. A concise overview of the formalism of Tikhonov regularization theory is in Appendix A. The most notable advantages of this technique are that, first of all, very few a priori assumptions on the solution are necessary in the inversion procedure (namely, only that the electron flux density spectrum is as smooth as possible) and, moreover, that if conveniently implemented, the algorithm is efficient: our IDL code produces reconstructions of the electron spectrum from 90 point photon spectra in a few seconds.

We consider now the three photon spectra registered by RHESSI detectors on February 20 at 11:06, March 17 at 19:30 and August 6 at 12:56. The plot of these three data vectors is in Fig. 1. These data are characterized by RHESSI's highest spectral resolution of 1 keV bins in the 10–100 keV range (to obtain such a resolution, data from detectors 2 and 7 have been omitted since their resolution is poorer than that of RHESSI's other seven detectors). Furthermore we chose time bins equal to RHESSI's rotation period in order that there would be no modulation from the imaging grids. The standard software has been used to perform some pre-processing steps: detector energy response, detector lifetime, attenuator transmission and imaging grid transmission. The background has been subtracted by using the SPEX package to interpolate between two background intervals, one before and one after the flare. No thermal component has been subtracted, since in Eqs. (1)–(3) $f(E)$ includes both the thermal and non-thermal contributions (Piana et al. 2003). Finally a pulse pile-up correction has been applied to

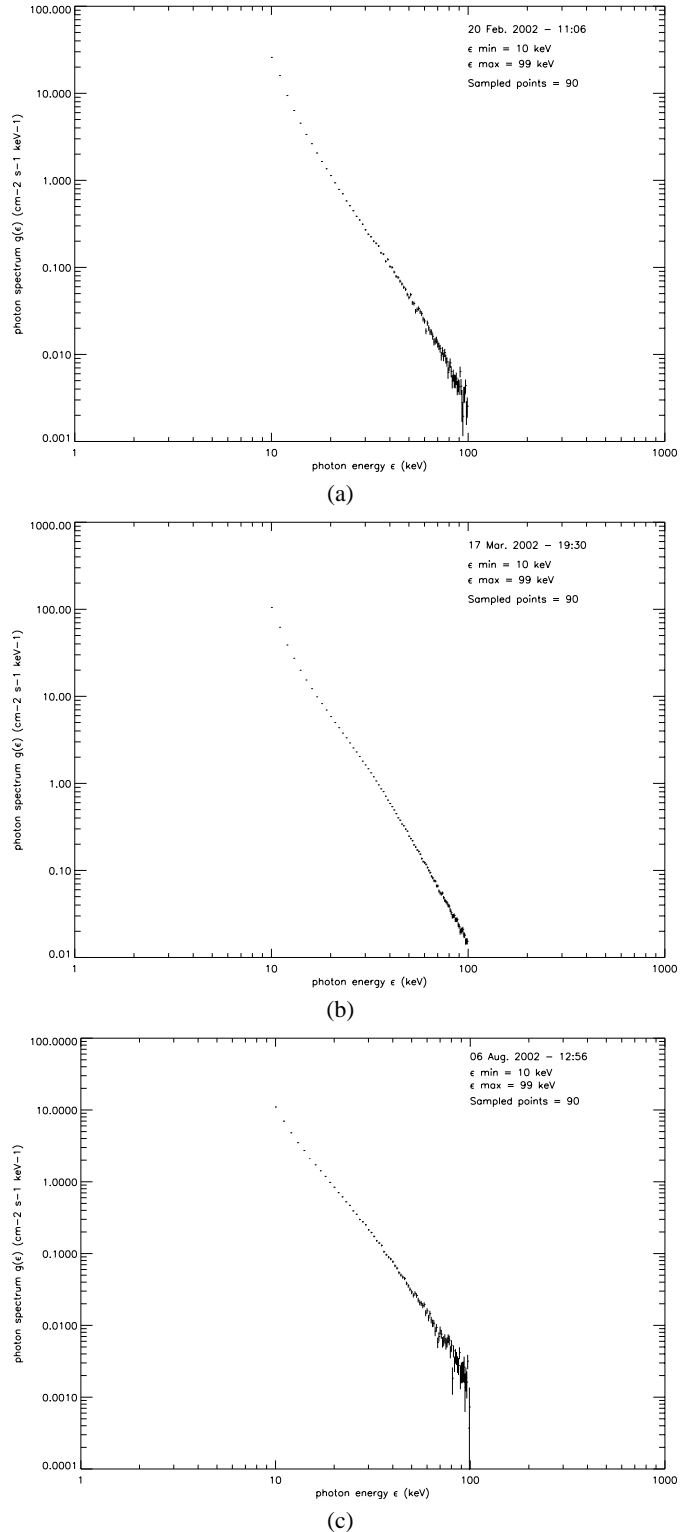


Fig. 1. The RHESSI spectra: **a)** data set registered on February 20 2002 at 11:06; **b)** data set registered on March 17 2002 at 19:30; **c)** data set registered on August 6 2002 at 12:56.

the 17 March 2002 event, which is the largest flare in terms of GOES class (the software and data used were the most up-to-date available on 23rd January 2003). We have applied the Tikhonov regularization method to these three data sets in

Table 1. Optimal values of the regularization parameter for the three events under examination in the case of three different bremsstrahlung cross-sections. The choice criterion is the Morozov's discrepancy principle.

	Kramers	Bethe-Heitler	Bethe-Heitler-Elwert
Feb. 20	$\lambda = 2.8 \times 10^{-6}$	$\lambda = 1.2 \times 10^{-6}$	$\lambda = 3.4 \times 10^{-12}$
Mar. 17	$\lambda = 1.3 \times 10^{-6}$	$\lambda = 4.3 \times 10^{-7}$	$\lambda = 1.7 \times 10^{-12}$
Aug. 6	$\lambda = 5.5 \times 10^{-6}$	$\lambda = 2.1 \times 10^{-6}$	$\lambda = 7.8 \times 10^{-12}$

the case of Kramers, Bethe-Heitler and Bethe-Heitler with the Elwert correction cross-sections. Figure 2a contains the reconstruction of $\bar{n}Vf(E)$ for the February 20 flare and the three models. In Figs. 2b, c the same quantity is reconstructed for the March 17 and August 6 events respectively. In the restoration procedure the regularization parameter has been fixed by means of the discrepancy principle (Tikhonov et al. 1995; see Appendix A). The optimal values for the three events and the three cross-sections are given in Table 1. Clearly, the selected values for the Bethe-Heitler-Elwert model are systematically smaller than the ones for the other two cross-sections. This result indicates that when this cross-section is adopted, the problem is numerically more stable than in the other two cases, since, from Eq. (15), smaller values of the regularization parameter correspond to a smaller weight given to the stability term at the left hand side of the equation. As a confirmation of this fact, we have that the condition number associated to the linear Eq. (14) (see Appendix A) is smaller when $Q(\epsilon, E)$ is the Bethe-Heitler-Elwert cross-section ($\sim 10^2$) than when $Q(\epsilon, E)$ are the Kramers ($\sim 10^3$) or the Bethe-Heitler ($\sim 10^5$) cross-sections.

Figure 3 contains the electron flux density spectrum corresponding to the February 20 event for the three cross-sections. The vertical bars at each point have been obtained by computing the confidence strip corresponding to the photon spectrum, i.e. from repeated inversions of different random realizations of the photon spectrum for a fixed value of the regularization parameter. The width of the horizontal bars in the case of Kramers and Bethe-Heitler cross-sections (Figs. 3a, b) have been obtained as described in Appendix A. If the Elwert correction is adopted (Fig. 3c), the resolution bar is given by the discretization interval in the electron energy range. Figures 4 and 5 contain the electron flux density spectra corresponding to the March 17 and August 6 events.

3. Direct and inverse approaches

In solar physics literature the solution of Eq. (14) can be addressed without allowing for the ill-posedness of the inverse problem, for example according to the following two strategies:

1. a direct approach, whereby the photon spectrum is best-fitted by using a priori chosen parametrized forms for the electron spectrum
2. an inverse approach based on the discretization of analytical inversion formulas.

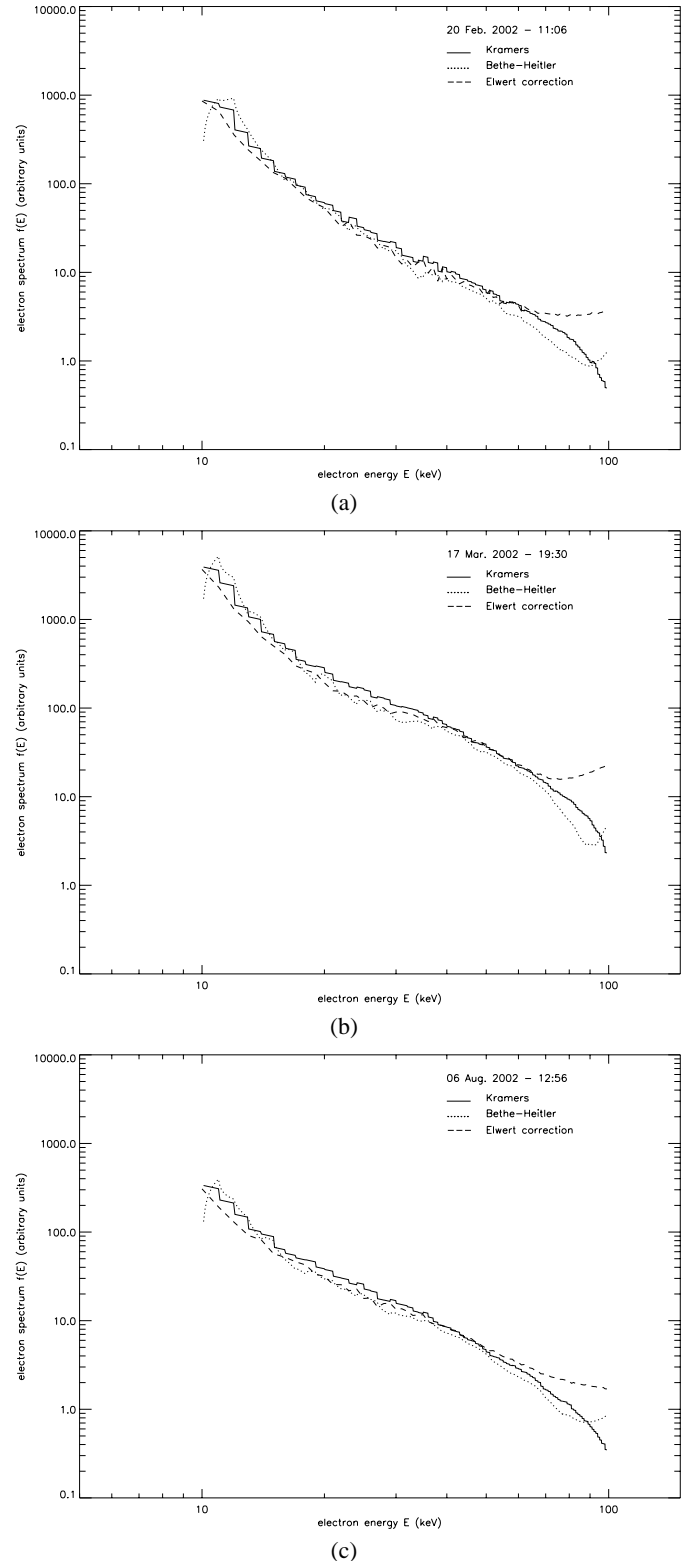
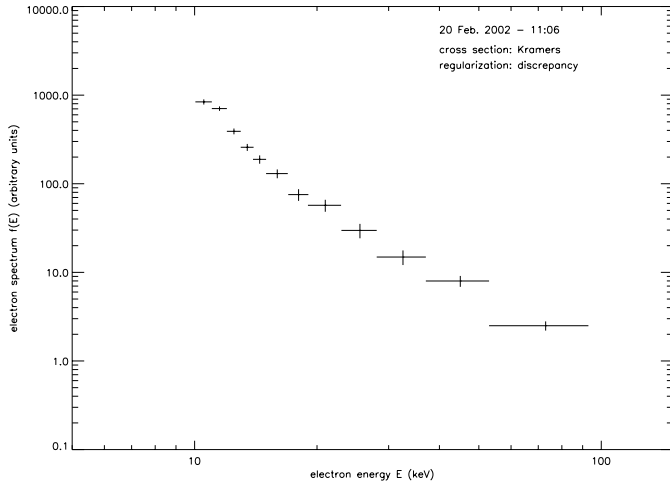
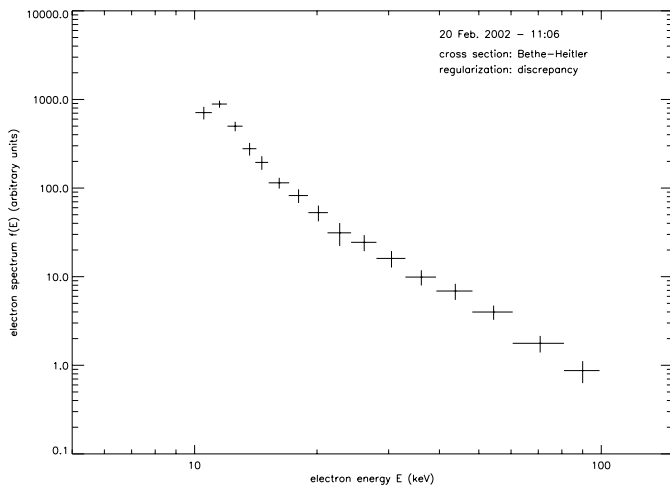


Fig. 2. Tikhonov regularized solutions in the case of Kramers (solid), Bethe-Heitler (dashes) and Bethe-Heitler with Elwert correction (dots) cross-sections: **a)** the February 20 event; **b)** the March 17 event; **c)** the August 6 event.

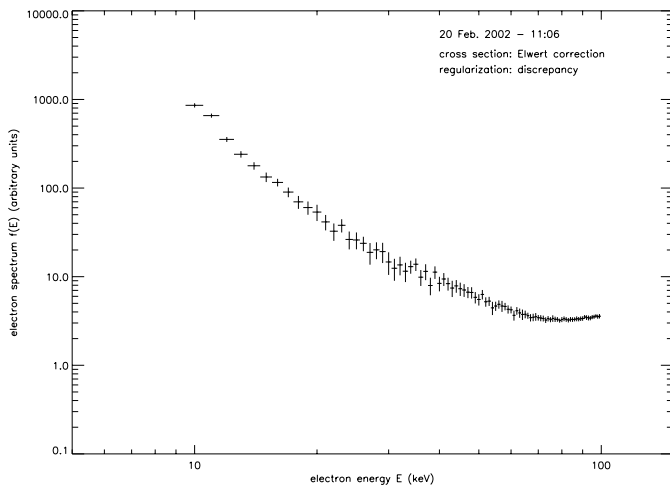
In the forward-fitting approach the photon spectrum is fitted by a simple function such as a power law to obtain a photon spectrum index γ and then either a thin- or a thick-target



(a)



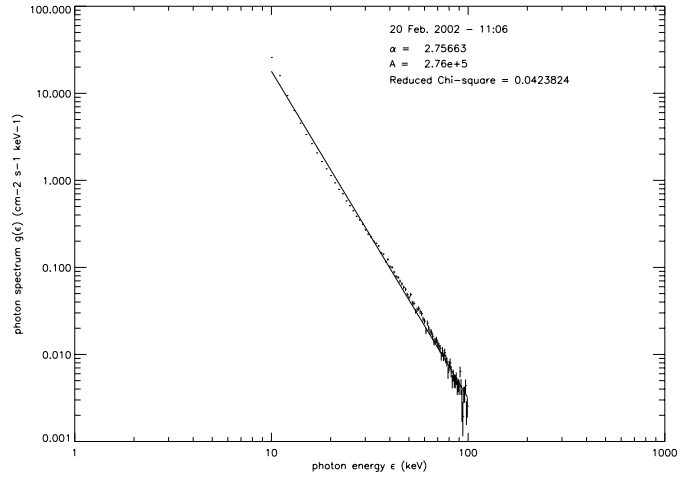
(b)



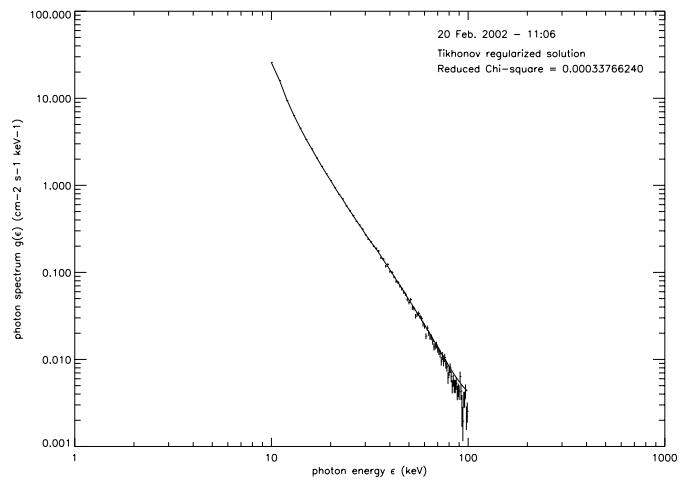
(c)

Fig. 3. Tikhonov reconstructions in the case of the February 20 event. Also the error and resolution bars are shown: **a)** the Kramers cross-section is adopted; **b)** the Bethe-Heitler cross-section is adopted; **c)** the Bethe-Heitler with Elwert correction is adopted.

model is assumed to obtain an electron spectrum index δ . Such an approach gives the best power-law fit to the behaviour at high energies, but obviously cannot detect possible detailed



(a)



(b)

Fig. 6. Data fit of the February 20 event: **a)** a power law is used in the forward-fitting approach; **b)** the fit is obtained by inserting the Tikhonov regularized solution in Eq. (14) when the Kramers cross-section is adopted.

features which may have a physical meaning or may be produced by systematic defects in the measurement procedure. Regularization techniques preserve such features and, in the meantime, reproduce the data vector with great accuracy. As an example, in Fig. 6a the data set corresponding to the February 20 flare (Fig. 1a) has been best-fitted by a one-parameter power law with index $\delta \approx 2.8$. In Fig. 6b the same spectrum is fitted with the regularized solution represented in Fig. 2a (solid line) obtained by using the Kramers cross-section. Clearly, the data fit provided by the Tikhonov regularized solution is more accurate than the one provided by the power law.

Simple inversion schemes for the solution of the solar bremsstrahlung problem can be obtained by analytically inverting Eq. (1) and then by discretizing the inversion formula. However, it is extremely difficult to obtain an analytical inversion formula in the case of a general cross-section. For example in Brown (1971), Eq. (1) is analytically solved for the Bethe-Heitler scattering, by showing that an appropriate change of variables leads to an Abel-type equation

(Riesz & Nagy 1953); nevertheless, to our knowledge, there is no inversion formula for the same equation when also the Elwert correction is allowed for, as in Eqs. (7)–(13). Furthermore, analytical inversion formulas typically involve expressions containing derivatives of the photon spectrum and therefore cannot be applied when the data are not known very accurately. In Johns & Lin (1992), an inversion method is proposed where an iterative procedure is obtained by exploiting the facts that bremsstrahlung spectra from optically thin sources always decrease with increasing photon energy, and electrons of a given energy can only produce photons of equal or lower energy. This method exploits the Volterra character of the original integral equation and is based on a standard row elimination technique, where a “local smoothing” procedure is applied via adjustable bin widths. This discretization-rebinning operation introduces a certain regularization effect whose intensity, however, is not chosen by means of some objective criterion, as in the Tikhonov method. Therefore, if applied to real data affected by observational noise, this scheme may produce oscillating solutions, as shown in (Emslie et al. 2001).

4. Comments and open problems

In the present paper we have shown that the Tikhonov regularization algorithm is a potentially effective method for the reconstruction of electron flux density spectra from the photon spectra provided by the RHESSI Mission. In particular, this approach allows the suppression of the noise amplification due to the ill-posedness of the original inverse problem. Furthermore, many interesting features in the recovered $\bar{n}Vf$, which cannot be pointed out by means of forward-fitting approaches, are clearly preserved. For example, the March 17 spectrum presents a pronounced concavity for $E \geq 30$ keV while the February 20 spectrum presents a detailed feature just around 30 keV (more evident in the Bethe-Heitler and Bethe-Heitler-Elwert reconstructions). The origin and physical meaning of such features is a crucial problem currently under investigation (first steps in this direction are in Piana et al. 2003).

The problem of the dependence of the reconstructed electron spectrum from the cross-section is still open and certainly deserves a systematic study. For the three data sets considered in the present paper, there is a substantial independence from the model in the middle energy range. The upturn in solution f at highest E for some cross-sections is a numerical artefact of the truncation of the data and hence of our representation above 100 keV. This should be ignored until a dependable fix

is added. At low energies, the Bethe-Heitler cross-section produces a downturn in $\bar{n}Vf$ below 15 keV, which occurs also when the Haug (1997) model is adopted (Piana et al. 2003). This effect vanishes if the Elwert correction is introduced. The physical reliability of this feature is currently under investigation. From a more technical point of view, we would also like to investigate the sensitivity of the reconstructions to the values of the regularization parameters, by applying optimality criteria different from the discrepancy principle used in the present paper.

Acknowledgements. Part of this work was supported by a Royal Society grant. We want to thank Dr. E. Kontar for his help with the pile-up correction to the data.

References

- Bertero, M., De Mol, C., & Pike, E. R. 1985, *Inverse Problems*, 1, 301
 Bertero, M., De Mol, C., & Pike, E. R. 1988, *Inverse Problems*, 4, 573
 Brown, J. C. 1971, *Sol. Phys.*, 18, 489
 Brown, J. C., & Emslie, A. G. 1988, *ApJ*, 331, 554
 Craig, I. J. D., & Brown, J. C. 1986, *Inverse Problems in Astronomy* (Bristol: Adam Hilger)
 Crannell, C. J., Frost, K. J., Mätzler, C., Ohki, I., & Saba, J. L. 1978, *ApJ*, 223, 620
 Elcan, M. J. 1978, *ApJ*, 226, L99
 Emslie, A. G., Barrett, R. K., & Brown, J. C. 2001, *ApJ*, 557, 1
 Hadamard, J. 1923, *Lectures on Cauchy’s Problem in Linear Partial Differential Equation* (New Haven: Yale University Press)
 Haug, E. 1997, *A&A*, 326, 417
 Johns, C. M., & Lin, R. P. 1992a, *Sol. Phys.*, 137, 121
 Johns, C. M., & Lin, R. P. 1992b, *Sol. Phys.*, 142, 219
 Kane, S. R., & Anderson, K. A. 1970, *ApJ*, 162, 1002
 Kane, S. R., Crannell, D., Datlowe, D., et al. 1980, in *Solar Flares*, ed. P. A. Sturrock (Boulder: Associated University Press)
 Koch, H. W., & Motz, J. W. 1959, *Rev. Mod. Phys.*, 31, 920
 Lin, R. P. 1974, *Space Sci. Rev.*, 16, 189
 Lin, R. P., & Schwartz, R. A. 1987, *ApJ*, 312, 462
 Piana, M. 1994, *A&A*, 288, 949
 Piana, M., Brown, J. C., & Thompson, A. M. 1995, *Sol. Phys.*, 156, 315
 Piana, M., Massone, A. M., Kontar, E. C., et al. 2003, *ApJ*, in press
 Riesz, F., & Nagy, B. 1953, *Leçons d’analyse fonctionnelle* (Budapest: Akademiai Kiado)
 Tikhonov, A. N. 1963, *Sov. Math. Dokl.*, 4, 1035
 Tikhonov, A. N., Goncharsky, A. V., Stepanov, V. V., & Yagola, A. G. 1995, *Numerical Methods for the Solution of Ill-posed Problems* (Dordrecht: Kluwer)