We address an important inverse problem in solar plasma physics by means of an ‘ad hoc’ implementation of the Tikhonov regularization method. Our approach is validated in the case of different modifications of the mathematical model of the problem based on physical motivations, and for data sets obtained with synthetic experiments. Finally, an application to a real observation recorded by the NASA mission ‘Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI)’ is considered.

Keywords: Inverse problems; solar flares; regularization theory

AMS Subject Classification: 45Q05; 47A52

1 Introduction

Solar flares [26,36,38] are the most dramatic and mysterious events in the solar system. These transient phenomena are characterized by a sudden release of huge amounts of energy of magnetic origin and are revealed by electromagnetic emission mostly in the X-ray energy range. Particularly intense events may release up to $10^{33}$ ergs in a few seconds and accelerate electrons in the solar plasma with a rate up to $10^{37}$ per second. This means that the electrical currents passing through the solar chromosphere during a flare have orders of magnitude of around $10^{18}$ ampère, corresponding to magnetic fields of around $10^4$ tesla. The main theoretical problem with solar flares is that no mathematical modelization so far formulated (based on the well-established
equations of plasma physics and magnetohydrodynamics) is able to fully explain such remarkably large numbers. Most physical interpretations of X-ray emission during solar flares are based on the assumption that a collisional interaction takes place between accelerated electrons and steady ions in the plasma. During such phenomenon, named Bremsstrahlung collision [18], light charged particles (electrons in this case) are deflected and slowed down by the presence of other heavier charged particles (the ions in the solar chromosphere) and loose their energy in form of electromagnetic radiation mainly at X-ray wavelengths. Several solar missions have been designed with the aim of detecting such radiation. For example, the on-orbit solar satellite Yohkoh, launched in August 1991, was able to collect radiation from electronic collisions in the case of both soft (energies in the range 0.25 – 4 keV) and hard X-rays (energies up to nearly 100 keV). The Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI) [27], launched by NASA on February 2002, is currently recording hard X-rays and $\gamma$-rays up to 1 MeV during solar flares.

From a mathematical viewpoint, Bremsstrahlung collision is fully described by a quantity, named Bremsstrahlung cross-section, which represents the probability that a hard X-ray photon of given energy is produced by an electron in the plasma of given (bigger) energy. Therefore, the mathematical description of solar Bremsstrahlung is based on a Volterra integral equation, from now on named the Bremsstrahlung equation, which relates the emitted photon spectrum with the flux spectrum of the electrons in the source and where the integral kernel is represented by the Bremsstrahlung cross-section. The importance of the exciting electron flux spectrum in this theory relies on the fact that its determination from the observed X-rays depends only on the Bremsstrahlung cross-section and does not require any modeling assumption on the acceleration or propagation mechanisms [9]. Furthermore the Bremsstrahlung equation can be easily generalized to account for second order phenomena like albedo [1] or anisotropies in the emission direction [29].

The study of the inverse problem of determining the electron flux spectrum from measurements of the emitted photon flux has been accomplished in several papers from the astronomical community. In particular, in [7] the formulation of the solar Bremsstrahlung problem in terms of a Volterra integral equation of the first kind is introduced for the first time together with an analytical approach to its solution. The ill-posedness of the problem is pointed out in [32] and reconstruction approaches for its solution are discussed and validated in [10, 19, 23, 24, 32, 34]. Finally in [9] some theoretical implications of this formulation of the problem are derived together with some hints on the physical interpretation of reconstructed electron spectra.

The contribution of the present paper to this investigation is two-fold. First we will study the sensitivity of the inverse problem solution to the modifica-
tion of the integral kernel due to different assumptions on the Bremsstrahlung cross-section. Indeed nuclear physics provides several possible different forms for the Bremsstrahlung cross-section, depending, for example, if relativistic or Coulomb corrections are included in the modeling. Some of these forms are easy expressions and in these cases some analytical results can be obtained from the study of the corresponding Bremsstrahlung integral equation. However more general and reliable forms for the cross-section are extremely complicated functions of the photon and electron energies and for them only numerical investigations are reasonably possible. Our study will be based on both a singular value decomposition analysis of the Bremsstrahlung equation and a regularized inversion of the X-ray data based on the Tikhonov method [39].

The second issue we will address is to describe the mathematical content of an inversion code for the analysis of RHESSI data the authors of the present paper have implemented and which has been included in the software of the satellite mission. Although this code accomplishes such analysis by means of the classical Tikhonov approach, here we discuss some specific procedures which have been utilized in the implementation, for determining the error and resolution bars, for choosing the optimal value of the regularization parameter and for handling the notable dynamical range typical of RHESSI data.

The plan of the paper is as follows. Section 2 will perform the mathematical setup of the problem and contain the definition of all the quantities we will use in the equations. Section 3 will provide some analytical results in the case of two possible approximations of the Bremsstrahlung cross-section. In Section 4 a first analysis of the consequences of changes in the integral kernel of the Bremsstrahlung equation will be addressed by means of a singular value decomposition approach. Section 5 contains a description of the inversion software implemented for the RHESSI mission and gives particular emphasis on the issues of choosing the regularization parameter and rescaling the dynamical range of the input data. In Section 6 the regularized inversion of the Bremsstrahlung equation is performed in the case of both significant simulated data and a real spectrum measured by RHESSI. This section will consider effects of variations in the integral kernel by comparing regularized solutions corresponding to different cross-sections. Our conclusions will be offered in Section 7.
2 The Bremsstrahlung equation

Following [7] and [9], we first define the mean target proton density in the source as

$$\bar{n} = \frac{1}{V} \int_V n(r)dr$$  \hspace{1cm} (1)

and the mean electron spectrum

$$\overline{F}(E) = \frac{1}{\overline{n}V} \int_V n(r)F(E,r)dr ,$$  \hspace{1cm} (2)

where $V$ is the source volume, $n(r)$ is the local proton density in the plasma and $F(E, r)$ is the local electron distribution function. Then the photon spectrum $I(\epsilon)$ observed at distance $R$ from a source, measured in photons cm$^{-2}$ s$^{-1}$ keV$^{-1}$, is given by

$$I(\epsilon) = \frac{\bar{n}V}{4\pi R^2} \int_{\epsilon}^{\infty} \overline{F}(E)Q(\epsilon, E)dE ,$$  \hspace{1cm} (3)

where $Q(\epsilon, E)$ is the Bremsstrahlung cross-section, which is a function of both the photon energy $\epsilon$ and the electron energy $E$. If we define

$$f(E) = \bar{n}V\overline{F}(E) , \quad g(\epsilon) = 4\pi R^2 I(\epsilon) ,$$  \hspace{1cm} (4)

equation (3) becomes

$$g(\epsilon) = \int_{\epsilon}^{\infty} f(E)Q(\epsilon, E)dE .$$  \hspace{1cm} (5)

Equation (5) is the Bremsstrahlung equation in solar plasma physics: it can be written without any prescription or assumption on the physical processes in the source and just for this reason it is the most general equation describing the X-ray emission mechanism during solar flares. We point out that this model is completely isotropic, i.e. no angular dependency in the mean electron spectrum or in the cross-section is considered, and, furthermore, that secondary effects like albedo are not taken into account. First investigations toward anisotropic modelization of the emission or interpretations of the reconstructed $f(E)$ which involve albedo have been described in [29] and [1], although further generalizations of equation (5) are in progress.

A crucial role in equation (5) is played by the cross-section $Q(\epsilon, E)$. Indeed, this function carries a significant physical meaning, describing the effectiveness
of the Bremsstrahlung process. On the mathematical side, its analytical shape has important consequences on the numerical stability of the solution process and represents a reliable indication of the information content which can be retrieved from the X-ray data. According to the physical conditions where the emission process takes place, many different forms of such a cross-section can be written [22]. Here we consider six possible forms (which are among the most utilized in solar physics) and, for sake of simplicity, group them in two families.

**Family I:** it contains two cross-sections which have a simple analytic expression and, physically, are non- and mildly-relativistic respectively. The two cross-sections are the Kramers (K) cross-section [12] (Figure 1(a))

\[
Q_K(\epsilon, E) = \begin{cases} 
Q_0 & \epsilon \leq E \\
\frac{Q_0}{\epsilon} & \epsilon > E
\end{cases}
\]  

(6)

and the Bethe-Heitler (BH) cross-section [22] (Figure 1(b))

\[
Q_{BH}(\epsilon, E) = \begin{cases} 
\frac{Q_0}{\epsilon} \log \frac{1+\sqrt{1-\frac{\epsilon}{E}}}{1-\sqrt{1-\frac{\epsilon}{E}}} & \epsilon \leq E \\
0 & \epsilon > E
\end{cases}
\]  

(7)

\(Q_0\) is

\[
Q_0 = \frac{8}{3} \frac{Z r_0^2}{137} mc^2 ,
\]  

(8)

where \(Z\) is the atomic number of the target, \(m\) and \(r_0\) are the mass and the classical radius of the electron and \(c\) is the speed of light.

**Family II:** it contains four cross-sections which include physical second order effects. Analytically all forms are extremely complicated and therefore will not be reported here (but appropriate references to papers where these forms can be found are given). In particular, the Bethe-Heitler-Elwert (BHE) cross-section [19] (Figure 1(c)) modifies the Bethe-Heitler formula (7) to account for Coulomb effects at low energies. The 3BN cross-section [22] (Figure 1(d)) includes all relativistic effects. The Haug (H) cross-section [16] (Figure 1(e)) is obtained by interpolating the 3BN formula and its computation is less expensive. Finally, the 3BNee cross-section [17] (Figure 1(f)) is a correction of the 3BN formula which accounts for the electron-electron interaction significant at high energies.

There are essentially two reasons why it is interesting to study equation (5) for different forms of the cross-section. The first one is physical and is the
fact that the differences in the solution $f(E)$ for the different forms of $Q(\epsilon, E)$ previously introduced can provide quantitative information on the incidence of the relativistic, Coulomb, electron-electron interaction corrections on the photon production in the solar plasma. The second reason is computational and is the fact that the numerical solution of equation (5) is notably less onerous in the case of equations (6)-(7) and then it is helpful to verify whether (and to what extent) the second order effects can be neglected in the solution procedure.

3 Analytical results

In this section we present some analytical results concerned with equation (5) obtained when the integral kernels are cross-sections $K$ and BH in equation (6) and (7) respectively. For these kernels, equation (5) can be analytically solved by applying the theory of Mellin transform [6]. The Mellin transform of a function $f : (0, \infty) \rightarrow \mathbb{R}$ such that $\int_0^\infty |f(x)|x^{\sigma-1}dx < \infty$ for some real number $\sigma$ in $(0, 1/2)$, is defined as

$$\tilde{f}(\xi) = \int_0^\infty f(x)x^{-\frac{1}{2}+i\xi}dx,$$

with $\xi \in (-\infty, \infty)$. The corresponding inversion formula is

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{f}(\xi)x^{-\frac{1}{2}-i\xi}d\xi .$$

In the case of linear integral equations of the first kind, when the kernel depends on the variables ratio, i.e.

$$g(x) = \int_0^\infty k\left(\frac{x}{y}\right)f(y)\frac{dy}{y} ,$$

the Mellin transform applied to both sides leads to the diagonalization

$$\tilde{g}(\xi) = \tilde{k}(\xi)\tilde{f}(\xi) .$$

For cross-sections $K$ and BH, the Bremsstrahlung equation (5) assumes the same form as in (11). In fact, if one introduces the new integral kernel

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

(13)
and defines the new data function
\[ J(\epsilon) = \frac{eg(\epsilon)}{Q_0}, \]  
(14)
equation (5) becomes
\[ J(\epsilon) = \int_0^\infty f(E)K(\epsilon, E)\frac{dE}{E}, \]  
(15)
where
\[ K(\epsilon, E) = K_K(\epsilon/E) = \begin{cases} 1 & \epsilon/E \leq 1 \\ 0 & \text{otherwise} \end{cases} \]  
(16)
for formula K and
\[ K(\epsilon, E) = K_{BH}(\epsilon/E) = \begin{cases} \log \frac{1+\sqrt{1-\frac{\epsilon}{E}}}{1-\sqrt{1-\frac{\epsilon}{E}}} & \epsilon/E \leq 1 \\ 0 & \text{otherwise} \end{cases} \]  
(17)
for formula BH. Therefore the following theorem holds:

**Theorem 3.1** The analytical form of the solution of equation (5) is given by

\[ f(E) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1 + 2i\xi}{2} \tilde{J}(\xi)E^{-\left(\frac{1}{2} + i\xi\right)}d\xi \]  
(18)
in the case of cross-section K and by

\[ f(E) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1 + 2i\xi}{2\sqrt{\pi}} \frac{\Gamma(1 + i\xi)}{\Gamma\left(\frac{1}{2} + i\xi\right)} \tilde{J}(\xi)E^{-\left(\frac{1}{2} + i\xi\right)}d\xi \]  
(19)
in the case of cross-section BH.

**Proof** Let us introduce the Beta Function \( B(p, q) \) defined as

\[ B(p, q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx \quad \text{Re}(p) > 0 \quad \text{Re}(q) > 0 \]  
(20)
and the relation \( B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \) between this function and the Gamma
Function
\[ \Gamma(z) = \int_0^\infty e^{-t}t^{z-1}dt. \] (21)

Therefore, for case K,
\[ \tilde{K}_K(\xi) = \int_0^1 x^{-\frac{1}{2}+i\xi}dx = B\left(\frac{1}{2} + i\xi, 1\right) = \frac{\Gamma\left(\frac{1}{2} + i\xi\right)\Gamma(1)}{\Gamma\left(\frac{3}{2} + i\xi\right)} = \frac{2}{1 + 2i\xi}. \] (22)

For case BH, integration by parts leads to
\[ \tilde{K}_{BH}(\xi) = \tilde{\nu}(\xi) \frac{1}{\frac{1}{2} + i\xi}, \] (23)

where \( \nu(x) \) is defined as
\[ \nu(x) = \begin{cases} 0 & x \geq 1 \\ \frac{1}{\sqrt{1-x}} & 0 \leq x < 1 \end{cases}. \] (24)

Since
\[ \tilde{\nu}(\xi) = B\left(\frac{1}{2}, \frac{1}{2} + i\xi\right) \] (25)
and \( \Gamma(1/2) = \sqrt{\pi} \) we have
\[ \tilde{K}_{BH}(\xi) = \sqrt{\pi} \frac{\Gamma\left(\frac{1}{2} + i\xi\right)}{\frac{1}{2} + i\xi} \frac{1}{\Gamma(1 + i\xi)}. \] (26)

\[ \square \]

Equations (22) and (26) show that both \( \tilde{K}_K(\xi) \) and \( \tilde{K}_{BH}(\xi) \) are entire functions which tend to zero continuously for \( |\xi| \to \infty \). This implies that the two analytical solutions (18) and (19) do not depend continuously on the data \( J(\epsilon) \).

In real observations the photon energies involved in the flare process belong to a bounded interval \( (\epsilon_{\text{min}}, \epsilon_{\text{max}}) \). Therefore a more realistic model for the Bremsstrahlung emission is given by the linear inverse problem
\[ J = A_1f, \] (27)
with \( A_1 : L^2(\epsilon_{\min}, \infty) \rightarrow L^2(\epsilon_{\min}, \epsilon_{\max}) \) the linear operator defined as

\[
(A_1 f)(\epsilon) = \int_{\epsilon_{\min}}^{\epsilon_{\max}} f(E) K \left( \frac{\epsilon}{E} \right) \frac{dE}{E}.
\]

(28)

**Theorem 3.2** When the integral kernel \( K(\frac{\epsilon}{E}) \) is given by formula K or formula BH the operator (28) is compact.

**Proof** Let us consider first the case of cross-section K; we have to show that the function \((\epsilon, E) \mapsto \frac{1}{E} K(\frac{\epsilon}{E})\) belongs to \( L^2([\epsilon_{\min}, \epsilon_{\max}] \times [\epsilon_{\min}, \infty)) \). We have

\[
\int_{\epsilon_{\min}}^{\epsilon_{\max}} \int_{\epsilon_{\min}}^{\epsilon} \frac{1}{E^2} \left| K \left( \frac{\epsilon}{E} \right) \right|^2 dE d\epsilon = \int_{\epsilon_{\min}}^{\epsilon_{\max}} \int_{\epsilon_{\min}}^{\epsilon} \frac{1}{E^2} dE d\epsilon \\
\leq \int_{\epsilon_{\min}}^{\epsilon_{\max}} \int_{\epsilon_{\min}}^{\epsilon} \frac{1}{E^2} dE d\epsilon \\
= \frac{\epsilon_{\max} - \epsilon_{\min}}{\epsilon_{\min}} \\
< \infty
\]

and therefore the operator \( A_1 \) is Hilbert-Schmidt.

In the case of the Bethe–Heitler cross section, we make the usual change of variable \( x = \epsilon/E \) and we have

\[
\int_{\epsilon_{\min}}^{\epsilon_{\max}} \int_{\epsilon_{\min}}^{\epsilon} \frac{1}{E^2} \left| K \left( \frac{\epsilon}{E} \right) \right|^2 dE d\epsilon = \int_{\epsilon_{\min}}^{\epsilon_{\max}} \int_{\epsilon}^{\epsilon_{\max}} \left( \log \frac{1 + \sqrt{1-x}}{1 - \sqrt{1-x}} \right)^2 dE d\epsilon \\
= \int_{\epsilon_{\min}}^{\epsilon_{\max}} \frac{1}{\epsilon} d\epsilon \int_{0}^{1} \left( \log \frac{1 + \sqrt{1-x}}{1 - \sqrt{1-x}} \right)^2 dx \\
= \log \frac{\epsilon_{\max}}{\epsilon_{\min}} \int_{0}^{1} \left( \log \frac{1 + \sqrt{1-x}}{1 - \sqrt{1-x}} \right)^2 dx \\
< \infty
\]

and therefore even in this case the operator \( A_1 \) is Hilbert-Schmidt. \( \square \)

Theorem 3.2 implies that the linear inverse problem (27) is ill-posed in the sense of Hadamard and therefore that even in this case the solution does not depend continuously on the data. Real measurements provide sampled values of the photon spectrum at a finite number of photon energies. In [32] the
problem

\[ g(\epsilon_n) = \int_{\epsilon_n}^{\infty} Q(\epsilon_n, E) f(E) dE \quad n = 1, \ldots, N \]  

(29)

is addressed when \( Q(\epsilon, E) \) is the BH cross-section, by maintaining \( f \) in \( L^2(\epsilon_{\text{min}}, \infty) \). This approach is realistically possible only when the cross-section has a relatively simple form. In the present paper we want to consider very general forms of \( Q(\epsilon, E) \) and therefore the discretization of the Bremsstrahlung equation becomes a necessary preliminary step for inversion. However the impact of ill-posedness on the discretized equation is notable, since any discretization of equation (27) must account for the numerical instabilities due to the presence of noise on the observed spectrum.

4 Singular systems

The kind of data that real observations put at disposal of the mathematical analysis are vectors \( g \) whose components are the sampled values of the observed spectrum at \( N \) discrete photon energies \( \epsilon_1, \ldots, \epsilon_N \), i.e.

\[ (g)_n = g(\epsilon_n) \quad n = 1, \ldots, N \]  

(30)

We choose as data space the finite dimensional vector space \( Y \) equipped with the inner product

\[ (g, h)_Y = \sum_{l, n=1}^{N} (g)_l w_{ln} (h)_n \]  

(31)

where the \( w_{ln} \) \( (l, n = 1, \ldots, N) \) are appropriate weights depending on the energy sampling. The discretized inverse problem we have to solve in practice is therefore the linear system

\[ g = A f \]  

(32)

where \( f \) is the solution vector with components

\[ (f)_m = f(E_m) \quad m = 1, \ldots, M \]  

(33)
$E_1, \ldots, E_M$ are the discretized electron energies. We assume that the solution space $X$ is the canonical $\mathbb{R}^M$. $A$ is the rectangular matrix with entries

$$A_{nm} = Q(\epsilon_n, E_m)\eta_{nm},$$

(34)

$\eta_{nm}$ being the quadrature coefficients. For reasons related to the physics of the problem and, in particular, to the characteristics of the acquisition procedure applied by RHESSI, in the following we will always assume $M > N$. Of course the explicit form of the entries (34) will be different for the six different possible choices of the cross-section.

The generalized solution of the linear system (32), i.e. the vector solving the constrained least-squares problem

$$\|g - Af\|_Y = \min \quad \|f\|_X = \min,$$

(35)

is

$$f^\dagger = \sum_{k=1}^p \frac{1}{\sigma_k} (g, v_k)_Y u_k,$$

(36)

where $p$ is the rank of the matrix $A$ and $\{\sigma_k, v_k, u_k\}_{k=1}^p$ is its singular system defined as

$$Au_k = \sigma_k v_k \quad A^* v_k = \sigma_k u_k.$$  

(37)

Problem (35) is well-posed in the sense of Hadamard but, owing to the ill-posedness of the original continuous problem (27), it is ill-conditioned. The dependence of the degree of ill-conditioning with respect to the cross-section can be quantitatively assessed by means of the singular system of the matrix $A$ when the six different formulas for the integral kernel $Q(\epsilon, E)$ are adopted. The setup of this numerical experiment is as follows. We consider $N = 240$ photon energies from $\epsilon_1 = 10$ keV to $\epsilon_N = 249$ keV, with a uniform sampling distance of 1 keV. The electron energies are $M = 480$, from $E_1 = 10$ keV to $E_M = 489$ keV and are again uniformly sampled with 1 keV binning (these choices are realistic in the framework of typical RHESSI observations). Therefore both the weights in the vector data space and the quadrature weights in (34) are constant and equal to 1. In Figure 2 we plot the normalized singular spectra corresponding to all six cross-sections, where the normalization is performed by dividing all singular values of each matrix by the first one. In particular Figure 2(a) contains the first 100 singular values corresponding to formulas K and BH, while Figure 2(b) contains the first 100 singular values corresponding to formulas BHE, 3BN, H and 3BNee. The general trend of these singular
spectra shows that the introduction of relativistic corrections in the integral kernels increases the ill-conditioning of the problem: indeed the singular spectra corresponding to the non-relativistic and mildly-relativistic cross-sections $K$ and BH decrease more slowly with the singular value order than in the case of the other four cross-sections. In particular, the matrix $A$ corresponding to cross-section $K$ is the most well-conditioned one, in accordance with the fact that $Q_K(\epsilon, E)$ in (6) is less smooth than the other cross-section forms (see Figure 1). From Figure 2(b) we point out that the singular spectrum corresponding to formula H cannot be distinguished from the singular spectrum corresponding to formula 3BN, in accordance with the fact that formula H is an accurate interpolation of formula 3BN. Moreover, the electron-electron correction (accounted for in formula 3BNee) reflects into a slight increase of the ill-conditioning of the problem. Figure 3 shows the first three singular functions for each integral kernel used in the computation. A general feature for all cases is that the number of zeros increases with the order of the singular function. More in particular, we have that: the singular functions corresponding to the integral kernel $K$ present the fastest oscillations; formulas H and 3BN provide essentially the same singular functions; the differences due to the use of formula BHE are more significant at small $E$.

5 Inversion method

‘SolarSoftWare’ (SSW) is the IDL package officially used for the analysis of RHESSI data. Within this software, 19 routines have been implemented by the authors of the present paper for the regularized inversion of photon flux spectra emitted in correspondence with flare events observed by RHESSI. In order to describe the mathematical basis of this set of routines we first discuss the regularized inversion of the solar Bremsstrahlung problem in a continuous setting, i.e. in the case of

$$g = Af,$$  \hspace{1cm} (38)

where $A$ is the integral operator defined at the right hand side of (5).

Regularization theory [5, 11, 14, 21, 40] is certainly the most effective tool for the solution of such a linear ill-posed inverse problem. In this framework a stable and reliable approximate solution is obtained by exploiting a priori information on both the source function $f$ and the noise affecting the data $g$. However, there exist several different regularization methods and in order to choose the best approach for the present problem it is necessary to account for its mathematical and physical properties. In particular:

- Tikhonov-like methods [4, 39] look for the one-parameter family of solutions
of the minimum problem

$$\|Af - g\|_Y^2 + \lambda \| Cf\|_X^2 = \text{minimum},$$

(39)

where $X$ and $Y$ are Hilbert spaces and $C$ is a closed operator in $X$. The only a priori information on $f$ available from physical considerations is that $f$ is always positive and characterized by a huge dynamical range (five or six orders of magnitude of the electron flux can be covered within one order of magnitude in the energy $E$ axis). At low energies a possibly strong thermal component may induce the presence of a non-monotonic behaviour while short-wavelengths features (related, by instance, to pile-up corrections) are possible at intermediate energies. Therefore in general utilizing $f'$ or even $f''$ in the penalty term (i.e., utilizing the first or second derivative for $C$) may result in loosing these details. On the other hand, the choice $C = I$ realizes the search for the solution of minimal energy compatible with the data, which is a physically sound condition. Furthermore the supports of physically realistic electron flux spectra are certainly too wide for the application of a support constraint to be advantageous. This implies that the choices $X = L^2(\epsilon_{\text{min}}, \infty)$, $Y = L^2(\epsilon_{\text{min}}, \epsilon_{\text{max}})$ and $C$ equal to the identity operator are probably the most reliable.

- Deterministic iterative approaches like the Landweber method [5] present the advantage that the positivity constraint can be rather easily implemented into the algorithm [25]. However such approaches are typically computationally expensive and preconditioned versions of the schemes sometimes do not preserve a satisfactory reconstruction accuracy [33].

- A priori knowledge on the statistical properties of the noise can be exploited by iterative schemes of the Expectation-Maximization family [20]. These methods have been formulated in the case of Poisson noise [37] and generalized to Gaussian noise [3, 13]. However the aim of our code is the analysis of X-ray spectra recorded by RHESSI and in this mission the noise affecting the measurements is not purely Poissonian, but has a systematic component of difficult modeling.

Given the previous items, our approach to the regularized solution of the solar Bremsstrahlung inverse problem has been based on Tikhonov regularization with $X = L^2(\epsilon_{\text{min}}, \infty)$, $Y = L^2(\epsilon_{\text{min}}, \epsilon_{\text{max}})$ and $C$ the identity operator (anyway, the behaviour of first-order Tikhonov regularization, i.e. when $C$ is the first derivative, has been tested in [23] while the effect of Landweber iterations with positivity constraint has been studied in the case of the Bremsstrahlung problem under thermal conditions in [35]). In particular, in a discretized set-
ting, our routines address the minimum problem

$$\|A f - g\|^2_Y + \lambda \|f\|^2_X = \min,$$  \hspace{1cm} (40)

where the regularization parameter $\lambda$ is real and positive. By using a classical variational approach and the singular system of $A$, it is straightforward to show that the solution of (40) is

$$f_\lambda = \sum_{k=1}^{p} \frac{\sigma_k}{\sigma_k^2 + \lambda} (g, v_k) Y u_k.$$  \hspace{1cm} (41)

We point out that equation (41) does not necessarily provide a positive solution. In order to realize the positivity condition in Tikhonov regularization a box constraint should be applied within the optimization procedure (40), which, nevertheless, would notably slow down the computation. However, in our problem, many experiments with both real and synthetic data have shown that the Tikhonov functions (41) are always positive, in accordance with the physics of the problem.

Three main issues are involved with the actual computation of formula (41): 1) the determination of an optimal value of the regularization parameter; 2) the assessment of the propagation error from the data to the regularized solution; 3) the assessment of the spectral resolution associated with the regularized solution, i.e. the size of the electron energy range whereby $f$ presents detectable features (once the optimal value of the regularization parameter has been fixed). In the following we will describe how these issues have been addressed in our inversion software. Then, in the final part of this section, we will describe a rescaling procedure we have adopted in order to reduce the effects of the huge dynamical range of RHESSI data onto the regularization process.

5.1 Choice of the regularization parameter

The optimal choice of the regularization parameter realizes the perfect trade-off between the complexity of the solution and its ability to reliably reproduce the experimental data. In current literature many criteria exist for addressing this optimality problem [15] and a useful hint for validating them would be to formulate a statistical test quantitatively assessing the over- or under-smoothing properties of the selected $\lambda$. For the particular case of the Bremsstrahlung inverse problem several applications [28, 32] have shown that Morozov discrepancy principle [31] has often oversmoothing properties, while Generalized Cross Validation (GCV) [41] has often undersmoothing properties.
(these trends are confirmed for other applications [5]). In order to control the smoothing properties of the selection algorithm we have applied a new approach where the optimal choice of $\lambda$ is accomplished by a statistical analysis of the regularized cumulative residuals for problem (32) [34]. In a stochastic framework, the data and the solution of (32) are random variables and the model

$$ g = A f + \delta g $$

(42)

for the observed data can be assumed, where $\delta g$ is the random variable representing the noise on the data. For each sampled photon energy $\epsilon_k$ we define the normalized residuals

$$ r_k = \frac{(A f)_k - (g)_k}{(\delta g)_k}, \quad k = 1, \ldots, N \quad . $$

(43)

The set $\{c_j\}_{j=1}^N$ of the cumulative normalized residuals is defined as

$$ c_j = \frac{1}{j} \sum_{k=1}^{j} r_k, \quad j = 1, \ldots, N \quad . $$

(44)

From a statistical viewpoint the normalized residuals must be consistent with the distribution of the (normalized) noise and therefore are approximately distributed according to a Gaussian function of zero mean and standard deviation equal to 1. It follows that the sum $\sum_{k=1}^{j} r_k$ represents a random walk whose variance is $j$ and therefore $c_j$ has standard deviation $1/\sqrt{j}$. When the regularization switch is on, i.e. if $\lambda \neq 0$, we can introduce the regularized normalized residuals

$$ r_k^\lambda = \frac{(A f_\lambda)_k - (g)_k}{(\delta g)_k}, \quad k = 1, \ldots, N \quad , $$

(45)

and the regularized cumulative normalized residuals

$$ c_j^\lambda = \frac{1}{j} \sum_{k=1}^{j} r_k^\lambda, \quad j = 1, \ldots, N \quad . $$

(46)

Now the sum $\sum_{k=1}^{j} r_k^\lambda$ is not a random walk (here also the regularized solution $f_\lambda$ is considered as a random variable). However, if $\lambda$ is fixed in such a way that a specific realization of $c_j^\lambda$ is in the range $[-3/\sqrt{j}, 3/\sqrt{j}]$ for all $j$, then such realization can be considered as one of the realizations of $c_j$ oscillating in this
range with a probability larger than 99%. This suggests that the comparison between $c_\lambda^j$ as a function of $j$ and the bounds $\pm 3/\sqrt{j}$ can be considered a reliable statistical test to evaluate the under- or over-smoothing effects of the optimally selected $\lambda$. And, even more, a heuristical regularization algorithm can be formulated on the basis of this statistical test. In fact, if, by instance, the selection criterion adopted provides an oversmoothing value for $\lambda$, i.e. if $|c_\lambda^j| > 3/\sqrt{j}$ for some $j$, then $\lambda$ can be decreased up to the maximum value where

$$|c_\lambda^j| \leq \frac{3}{\sqrt{j}} \quad \forall j$$

(when the choice criterion is undersmoothing the $\lambda$ value should be increased up to the maximum value where (47) is verified). In our code this heuristic procedure is used to tune the value of the regularization parameter provided by the discrepancy principle [31]. We finally point out that several stochastically-based criteria for the choice of the regularization parameter have been recently introduced in the literature for both linear and non-linear ill-posed problems (see, for example, [2, 30]). The main characteristics of our semi-heuristic approach are that, first of all, its implementation is straightforward since it simply requires a comparison between the experimental and the regularized residuals. Furthermore, the method is quite general, since for its application it is not necessary to introduce strict hypothesis on the statistical properties of the noise distribution or to a priori know overall bounds on the noise amount (we use the discrepancy principle only as an initialization but any other initial guess can be applied in order to initialize the iteration on $\lambda$).

### 5.2 Errors and resolution bars

The determination of the propagation error on the regularized solution is performed by means of the computation of the confidence strip [4]. Many different random realizations of the input photon spectrum are realized and the corresponding regularized solutions are superimposed to determine a strip of reconstructed electron flux spectra all with the same regularization parameter determined by means of the heuristic approach described in the previous section. The width of the confidence strip provides the error bar on the regularized solution at each sampled electron energy.

In order to assess the spectral resolution, we have already observed in Section 4 that the singular vectors $u_k$ are characterized by a number of oscillations around zero which increases with their order. On the other hand, the regularization parameter damps the contribution of high order singular components so that only a few of them are really effective in the reconstruction of the elec-
tron spectrum. In order to determine the last significant singular component in the expansion (41) we introduce the power spectra \( P_1 \) and \( P_2 \) of \( f \) and \( \delta g \) respectively defined as

\[
< (f)_k, (f)_j > = P_1^2 \delta_{kj}
\]  

(48)

and

\[
< (\delta g)_k, (\delta g)_j > = P_2^2 \delta_{kj}
\]  

(49)

where \(< \cdot, \cdot >\) denotes the average over all possible data realizations. From equations (42) and (37) we easily obtain the relation

\[
< (g, v_k)_Y > = \sigma_k^2 P_1^2 + P_2^2.
\]  

(50)

The crucial point in equation (50) is that the first term at the r.h.s., i.e. the signal component, decreases with \( k \) while the second one, i.e. the noise component, is constant. In particular, the signal part stays larger than the noise part until

\[
\sigma_k^2 \geq \left( \frac{P_2}{P_1} \right)^2.
\]  

(51)

Now \( P_2 / P_1 \) is not explicitly known in general but regularization theory assures that \( \sqrt{\lambda} \) is a reliable estimate of this ratio, so that the truncation condition [4]

\[
\sigma_k \geq \sqrt{\lambda}
\]  

(52)

for the expansion (41) is reasonable. Therefore, if \( \sigma_k \) is the first singular value satisfying (52), a measure of the resolution limit achievable by Tikhonov method is given by the distance between two successive zeros of the corresponding last significant singular vector \( u_{k^*} \). Since in our implementation the electron energies are discretized before the inversion, the software provides, as resolution limit, the distance between the two sampled electron energies closest to the successive zeros.

5.3 Rescaling

A peculiar feature of solar photon spectra is its huge dynamical range: for a photon energy range between 10 keV and 100 keV the most energetic flares present photon fluxes which decrease from around \( 10^5 \) to 1 photons cm\(^{-2}\) keV\(^{-1}\) s\(^{-1}\). From the viewpoint of regularization this implies that the estimate
of the regularization parameter, which is a global parameter tuning stability and fitting over the whole energy range, may be unreliable. To overcome this difficulty, in our approach a rescaling procedure is applied, which flattens the input photon spectrum together with a coherent modification of the electron spectrum and of the integral kernel. The idea is to rescale the fundamental equation (5) to the form

\[ g'(\epsilon) = \int_{\epsilon}^{\infty} Q'(\epsilon, E) E^p f(E) dE \]  

(53)

with

\[ g'(\epsilon) = \epsilon^p g(\epsilon) \]  

(54)

and

\[ Q'(\epsilon, E) = \frac{\epsilon^p}{E^p} Q(\epsilon, E) \]  

(55)

Now the problem is how to optimally determine the value of the rescaling parameter. To this aim we first observe that a typical solar photon spectrum is a slight modification of an overall double power-law behaviour

\[ g(\epsilon) \sim \begin{cases} A\epsilon^{-\gamma_1} & \epsilon \leq \tau \\ B\epsilon^{-\gamma_2} & \epsilon > \tau \end{cases} \]  

(56)

Hence we implemented the following computational procedure:

- in simulation and for a single power-law \( g(\epsilon) \sim \epsilon^{-\gamma} \) we fixed a value of \( \gamma \), applied the regularized inversion for different values of \( p \) and selected the \( p \) value for which the regularized solution best fit (according to \( \chi^2 \) statistics) the original data;
- we repeated the process at the previous step for many values of \( \gamma \) in a realistic range and built up a look-up table which provides the optimal value of \( p \) for each \( \gamma \) value;
- the number of values of the look-up table is increased by means of a spline interpolation;
- given a real spectrum, we fit it with a double power-law, determine the fitting parameter \( \gamma_1 \) and \( \gamma_2 \), compute \( \gamma = 0.5(\gamma_1 + \gamma_2) \), find the closest \( \gamma \) in the interpolated look-up table and assign the corresponding \( p \) value.
6 Regularized reconstructions

The effectiveness of this implementation of Tikhonov regularization for solving the solar spectroscopy inverse problem is tested here in the case of both synthetic and real data. In particular, simulations are concerned with the determination of the so-called spectral index in the easiest case of a monotonically decreasing photon spectrum and with the reconstruction of a realistic form of the averaged electron flux spectrum. The real case refers to the inversion of a photon spectrum recorded by RHESSI in correspondence of the solar flare occurred on November 3, 2003.

6.1 Spectral index

The local spectral indices associated with a photon and an electron spectrum are defined as

\[ \gamma(\epsilon) = -\epsilon \frac{dg(\epsilon)}{d\epsilon} = -\frac{d\log g(\epsilon)}{d\log \epsilon} \]  

and

\[ \delta(E) = -E \frac{df(E)}{dE} = -\frac{d\log f(E)}{d\log E} \]  

(57)

(58)

One of the main reasons why spectral indeces are introduced in solar physics [8] is because they provide information on the X-ray emission without depending on the absolute magnitude of the event and therefore allow a quantitative comparison of flares of different size. As discussed in Section 5.3, many flares produce photon spectra with overall behaviours of a power-law or double power-law and therefore are characterized by constant or piecewise constant spectral indeces. Due to equation (5) the local electron spectral index must be quantitatively related to the local photon spectral index. In the case of cross-sections K and BH integration by parts immediately shows that the electron spectral index corresponding to a constant photon spectral index \( \gamma \) is constant and is given by

\[ \delta(E) = \delta = \gamma - 1 \]  

(59)

However this relation is not true in general for the other integral kernels. In order to study how relation (59) is modified by the introduction of more complicated cross-sections we consider the photon spectrum in Figure 4(a)
which is obtained by affecting the function

\[ g(\epsilon) = A\epsilon^{-5} \]  

(60)

with a realistic Poisson noise. The spectrum has been simulated by using 996 samples uniformly distributed in the energy range [1, 200] keV. The inversion has been done by considering electron energies with the same uniform binning but over the range [1, 400] keV. Tikhonov regularization method is applied for all six cross sections and the corresponding six electron spectral indeces are computed in the two panels of Figure 4(b). As expected, \( \delta(E) \) is very close to 4 when formulas K and BH are used (the deviations are due to numerical reasons and decrease if the computation precision is increased) while significant modifications from the constant values are introduced by the Elwert correction (formula BHE) and by accounting for relativistic effects (formulas 3BN and H) or effects due to electron-electron interaction (formula 3BNee).

6.2 Reconstruction of the averaged electron flux spectrum

A realistic synthetic data set is constructed by considering first the theoretical averaged electron flux spectrum

\[
f(E) = \begin{cases} 
C_1 \cdot \left( \frac{E}{50} \right)^{-\delta_1} + C_2 \cdot \exp\left(-\frac{E}{20}\right) & E \leq 100 \\
C_1 \cdot \left( \frac{100}{50} \right)^{-\delta_1} + C_2 \cdot \exp\left(-\frac{100}{20}\right) \cdot \left( \frac{E}{100} \right)^{-\delta_2} & E > 100
\end{cases}
\]  

(61)

where \( C_1 = C_2 = 5 \cdot 10^5, \delta_1 = 4 \) and \( \delta_2 = 5 \), represented in Figure 5(a). This function presents two features which are predicted by most theoretical models for solar flare physics: a spectral ‘knee’ at intermediate energies, typically representing a signature of the slowing down of the emission at higher energies, and a Maxwellian component which becomes significant at low energies and which accounts for possible thermal emission. Then equation (61) has been inserted into equation (5) when the integral kernel is the most general cross section, i.e. formula 3BNee, and realistic Poisson noise is added for each data component. In the case of Poisson noise the relative error depends on the energy: in our simulation, it is around 0.1% at low \( \epsilon \) and grows up to more than 30% at high energies. The result is the simulated photon flux spectrum in Figure 5(b) uniformly sampled with 240 points from 10 keV to 249 keV. This vector is inverted by means of our regularization code. In order to avoid an ‘inverse crime’ procedure the inversion has been performed by using the five cross-sections different than the one utilized to produce the data in the forward procedure (namely formula 3BNee). The reconstructions are presented in Figure 5(c) in the case of formulas K and BH and in Figure 5(d) in the
case of formulas BHE, 3BN and H. Here the electron flux spectra are uniformly sampled with 489 points from 10 keV to 498 keV. From a quantitative viewpoint the inversion outcome can be assessed by some reconstruction error whose computation is based on the comparison with the theoretical form \( (61) \). In this application, for each one of the five inversions we have computed the relative root mean square error

\[
\text{err} = \sqrt{\frac{\sum_{k=1}^{M} |(f_{\lambda})_k - f(E_k)|^2}{\sum_{k=1}^{M} |f(E_k)|^2}}.
\] (62)

The results given in Table 1 and Figure 5 show that:

- the use of the classical formulas K and BH produces the most inaccurate results;
- formulas 3BN and H provide very similar (and accurate) reconstructions, which is reasonable, since H is an interpolation of 3BN;
- BHE works fine at small energies but rather badly at high energies. The fact that the spectrum is very steeply decreasing explains the small reconstruction error in the table for this cross-section.

We have performed numerical experiments like this one for many theoretical forms of the input electron spectrum and the results are coherent with the ones in the table and in the figure.

6.3 Real data

As an example of application to real data we consider the flare event recorded by RHESSI on November 3, 2003. The number of counts detected by the satellite for different energy ranges is given in Figure 6(a) for the time interval 09:41:44 - 10:03:40 UT. The photon flux spectrum considered in the inversion has been built up by processing these counts in the time interval 09:57:00 - 09:57:20 UT and the resulting spectrum, uniformly sampled with 397 points from 9 keV to 223 keV, is given in Figure 6(b). This data vector has been inverted by means of our code for all cross-sections to give the reconstructed electron flux spectra, uniformly sampled with 437 points from 9 keV to 446 keV, in Figure 6(c) and Figure 6(d). For each reconstruction we give the confidence strip and the resolution bars computed as explained in Section 5.2. We point out that the overall behaviour is approximately the same for all regularized solutions, which denotes a notable robustness of the inversion procedure with respect to the modification of the model. As an example of a more quantitative analysis, we computed the photon and electron spectral indeces for all cases. In the photon energy range \([70, 200]\) keV the best fit for
\( \gamma(\epsilon) \) is \( \gamma \simeq 3.83 \); in the same range of electron energies we found that equation (59) is approximately satisfied for cross-sections K, BH, 3BN, H and 3BNee (in particular, \( \delta \simeq 2.63 \) for formula K) while a more significant displacement occurs in the case of cross section BHE (\( \delta \simeq 2.23 \)).

7 Conclusions and open problems

The main content of this paper was computational and aimed at discussing an implementation of Tikhonov regularization method which is ‘ad hoc’ for an important inverse problem in solar X-ray physics. The scientific motivation of this effort was the launch performed by NASA in 2002 of the solar satellite RHESSI with the precise intent of recording X-ray emissions from the Sun during flares. In particular, we discussed the incidence of changes in the cross-section, i.e., mathematically, in the integral kernel, on the regularized solution, showing that the overall behaviour of the reconstructions is preserved during the inversion but that significant differences occur if an analysis of more quantitative features like spectral indices or slope changes is performed. From a methodological viewpoint, we believe that some issues of our implementation (like the choice criterion for the regularization parameter based on the analysis of cumulative residuals or the data rescaling procedure) have a general potentiality and can be applied to other linear inverse problems.

Possible further applications of this approach to the Bremsstrahlung inverse problem may concern:

- a statistical study of a large number of flares, possibly classified on the basis of their size and intensity, in order to point out general trends on the emission mechanism;
- an application to spatially resolved spectra which can be able to determine hierarchies in the emission mechanism between different regions of the solar chromosphere involved in the flare.

As a further step in this work, we want to compare the results provided by Tikhonov regularization with reconstructions obtained by means of statistical iterative algorithms which explicitly account for the statistical properties of the noise affecting the data. Finally a difficult mathematical problem would be the one to simultaneously identify both the source function \( f \) and the kernel \( Q \) from data in a blind deconvolution approach to the Bremsstrahlung problem.

Acknowledgments

The Swiss International Space Science Institute (ISSI) is kindly acknowledged.
References


(Verlag).
Table 1. Reconstruction errors (%) in the restoration of equation (61) in the case of five cross-sections considered in the paper. Formula (62) has been used to assess the reconstruction errors.

<table>
<thead>
<tr>
<th>Cross-section</th>
<th>err</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>20.7</td>
</tr>
<tr>
<td>BH</td>
<td>30.8</td>
</tr>
<tr>
<td>BHE</td>
<td>2.6</td>
</tr>
<tr>
<td>3BN</td>
<td>9.0</td>
</tr>
<tr>
<td>H</td>
<td>8.5</td>
</tr>
</tbody>
</table>
Figure 1. The six cross-sections considered in the paper: (a) Kramers cross-section (formula K in the text); (b) Bethe-Heitler cross-section (formula BH in the text); (c) Bethe-Heitler cross-section with Elwert correction for Coulomb effects (formula BHE in the text); (d) 3BN cross-section with highly relativistic effects (formula 3BN in the text); (e) Haug interpolated cross-section (formula H in the text); (f) 3BN cross-section with electron-electron interaction effects (formula 3BNee in the text).
Figure 2. Singular spectra for the six (discretized) integral kernels: (a) formulas K and BH; (b) formulas BHE, 3BN, H and 3BNee.
Figure 3. The first three singular functions for the six (discretized) integral kernels: (a) \( u_1(E) \) for formulas K and BH; (b) \( u_1(E) \) for formulas BHE, 3BN, H and 3BNee; (c) \( u_2(E) \) for formulas K and BH; (d) \( u_2(E) \) for formulas BHE, 3BN, H and 3BNee; (e) \( u_3(E) \) for formulas K and BH; (f) \( u_3(E) \) for formulas BHE, 3BN, H and 3BNee.
Figure 4. Estimate of the spectral index by means of regularized inversion: (a) input photon spectrum, a power-law with constant spectral index 5; (b) reconstructed electron spectral index obtained in the case of the six cross-sections considered in the paper.

Figure 5. Reconstruction of a theoretical input averaged electron flux spectrum by means of regularized inversion: (a) theoretical input electron spectrum, a broken power-law with thermal component; (b) simulated photon flux spectrum; (c) regularized solutions in the case of formulas K and BH; (d) regularized solutions in the case of formulas BHE, 3BN and H.
Figure 6. Application to RHESSI data: (a) count flux in the case of the solar flare of November 3, 2003. The counts detected in five increasing energy channels are plotted without background removal (solid: 6 – 12 keV, dotted: 12 – 25 keV, dashed: 25 – 50 keV, dash dot: 50 – 100 keV, dash dot dot: 100 – 300 keV); (b) the photon flux spectrum constructed from the counts in (a); (c) reconstructed averaged electron flux spectrum for formulas K and BH; (d) reconstructed averaged electron flux spectrum for formulas BHE, 3BN, H and 3BNee.