A New Semiblind Deconvolution Approach for Fourier-Based Image Restoration: An Application in Astronomy[*]

S. Bonettini[†], A. Cornelio[‡], and M. Prato[‡]

Abstract. The aim of this paper is to develop a new optimization algorithm for the restoration of an image starting from samples of its Fourier transform, when only partial information about the data frequencies is provided. The corresponding constrained optimization problem is approached with a cyclic block alternating scheme, in which projected gradient methods are used to find a regularized solution. Our algorithm is then applied to the imaging of high-energy radiation emitted during a solar flare through the analysis of the photon counts collected by the NASA Reuven Ramaty High Energy Solar Spectroscopic Imager satellite. Numerical experiments on simulated data show that, in both the presence and absence of statistical noise, the proposed approach provides some improvements in the reconstructions.

1. Introduction. In any image reconstruction problem, the data are the result of an interaction between the desired image and an acquisition system that transforms the image itself into some kind of available information. One of the most popular examples in this framework is the deblurring problem. In this case, the measured information is a degraded version of the real image resulting from both the action of the acquisition system point spread function (PSF) and the presence of statistical noise typically due to measurement errors [5, 22]. On the other hand, in many applications the acquisition system hardware encodes the information needed to restore the unknown image in a different space. This happens, for example, in optical interferometry, where the use of separated telescopes allows one to estimate amplitude and phase of the Fourier components of the image corresponding to some specific spatial frequencies [34]. A similar approach occurs also in several medical applications, such as computerized tomography or magnetic resonance imaging, in which Fourier samples are achieved by means of the Radon transform and the Fourier slice theorem [11]. The common denominator of these problems is that the relationship between the measured data and the desired image is mathematically given by a Fredholm integral equation of the first kind, in

[†]Dipartimento di Matematica e Informatica, Università di Ferrara, 44122 Ferrara, Italy (silvia.bonettini@unife.it).

[‡]Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia, 41125 Modena, Italy (anastasia.cornelio@unimore.it, marco.prato@unimore.it).

which the kernel function is given by the PSF of the imaging system. A vast literature is available on this topic when the PSF is completely known, especially in the case of a space-invariant system, where the integral relation becomes a convolution operator [5, 41].

The problem becomes more difficult when the degraded image has to be found without prior knowledge of the system PSF, which is a frequent situation in several areas, due to restrictions imposed by the system hardware, high costs, or limited time at one's disposal. This problem is typically known as blind deconvolution [27, 28, 29]. In this case, the problem must be reformulated by introducing as far as possible all available a priori information on both the object and the PSF. A halfway situation occurs when the system response is not perfectly known, but a parametrized version of the PSF is provided. For this reason, this kind of problem is also called semiblind (or myopic) deconvolution [3, 13, 14]. In all of these cases, the approaches developed in the literature to reconstruct both the image and the PSF involve either the constrained minimization of regularized functionals [18, 24] or the separate alternate restoration by means of suitable deconvolution algorithms [1, 4, 15].

In this paper, we consider the semiblind deconvolution problem arising in Fourier image reconstruction when the frequencies corresponding to the data samples are not exactly known, but some estimates of them are available. In particular, we reformulate the reconstruction problem as a constrained nonlinear least squares problem and address its solution by means of an alternating minimization scheme, in which at each iteration either the image or the parametrized PSF is kept fixed while the other is updated. In both steps of each iteration, we adopt a projected gradient method to solve inexactly the minimization problem, and regularization is achieved by stopping the iterations before convergence. Moreover, for the PSF recovery steps, we also propose a different approach, achieved by scaling the gradient method with a Gauss–Newton-like diagonal matrix. For the numerical validation of the approaches, we consider a toy problem mimicking a particular application in astronomy, namely the two-dimensional reconstruction of X-rays emitted during a solar flare through the analysis of the radiation collected by the NASA Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI) satellite [26, 31]. We remark that the specific application to RHESSI is used as a motivational example of an application for our more general method, and the considered model is highly simplified with respect to the real one. In this situation, the numerical experiments we performed on synthetic images show that the adopted semiblind scheme provides some improvements with respect to recent reconstruction algorithms developed for the RHESSI mission.

The plan of the paper is the following: in section 2, the nonlinear least squares problem is formulated, together with a description of the alternating algorithm used to solve it. In section 3 the application to the RHESSI imaging problem is described, and our algorithm is tested on some simulated datasets. Finally, our conclusions are offered in section 4.

2. Problem formulation and reconstruction algorithm. The Fourier imaging problem can be reformulated as the minimization problem

$$(2.1) \qquad \min_{f \in \mathbb{R}^{n^2},\ f \geq 0} J(f) \equiv \frac{1}{2}\|Af - g\|_{\mathbb{C}^N}^2,$$

where the following hold:

(i) The n^2-vector f represents the image to be reconstructed, rearranged in lexicographic order, corresponding to an unknown distribution $f(x, y)$ evaluated at a set of grid points

(x_ℓ, y_ℓ) $(\ell = 1, \dots, n^2)$. Since the pixels' content in general refers to a physical quantity, we require the entries of f to be real and nonnegative.

(ii) The N-vector g contains the measured or estimated complex Fourier samples.

(iii) The linear operator A arises from the discretization of the Fourier transform and, for a given $f \in \mathbb{R}^{n^2}$, is defined as

$$(2.2) \qquad (Af)_k = \sum_{\ell=1}^{n^2} f_\ell e^{2\pi i (x_\ell u_k + y_\ell v_k)}, \quad k = 1, \dots, N,$$

where (u_k, v_k) are the spatial frequencies corresponding to the sample g_k.

If we consider the frequency vectors $u = (u_1, \dots, u_N)^T$ and $v = (v_1, \dots, v_N)^T$ as further unknowns and we make this explicit in the problem formulation, then (2.1) assumes the form

$$(2.3) \qquad \min_{\substack{f \in \mathbb{R}^{n^2}, \ f \geq 0 \\ u \in \mathbb{R}^N, \ u^{\min} \leq u \leq u^{\max} \\ v \in \mathbb{R}^N, \ v^{\min} \leq v \leq v^{\max}}} J(f, u, v) \equiv \frac{1}{2} \|A(u, v)f - g\|_{\mathbb{C}^N}^2,$$

where $u^{\min}$ and $u^{\max}$ (resp., $v^{\min}$ and $v^{\max}$) are N-vectors of problem-dependent lower and upper bounds for u (resp., v).

In several applications, such as the one described in section 3, the data acquisition system makes a polar representation of the spatial frequency more convenient. In this case, (2.3) becomes

$$(2.4) \qquad \min_{\substack{f \in \mathbb{R}^{n^2}, \ f \geq 0 \\ \alpha \in \mathbb{R}^N, \ \alpha^{\min} \leq \alpha \leq \alpha^{\max} \\ \rho \in \mathbb{R}^N, \ \rho^{\min} \leq \rho \leq \rho^{\max}}} J(f, \alpha, \rho) \equiv \frac{1}{2} \|A(\alpha, \rho)f - g\|_{\mathbb{C}^N}^2,$$

where $\alpha = (\alpha_1, \dots, \alpha_N)^T$ and $\rho = (\rho_1, \dots, \rho_N)^T$ are the angular and radial coordinates, respectively, with the related lower and upper bounds $\alpha^{\min}$, $\alpha^{\max}$ and $\rho^{\min}$, $\rho^{\max}$, and

$$(2.5) \qquad (A(\alpha, \rho)f)_k = \sum_{\ell=1}^{n^2} f_\ell e^{2\pi i \rho_k (x_\ell \cos(\alpha_k) + y_\ell \sin(\alpha_k))}, \quad k = 1, \dots, N.$$

The gradient of J is given by

$$\nabla J(f, \alpha, \rho) = (\nabla_f J(f, \alpha, \rho), \nabla_\alpha J(f, \alpha, \rho), \nabla_\rho J(f, \alpha, \rho)),$$

where[1]

$$\nabla_f J(f, \alpha, \rho) = \Re \left[A^H A f - A^H g \right],$$
$$\nabla_\alpha J(f, \alpha, \rho) = \Re \left[(A'_\alpha f) . * \overline{(Af - g)} \right],$$
$$\nabla_\rho J(f, \alpha, \rho) = \Re \left[(A'_\rho f) . * \overline{(Af - g)} \right].$$

---

[1] Here and in the following we omit the dependence of the matrices on (α, ρ) to simplify the notation.

Here $\Re[z]$ is the real part of a complex number z, A^H is the adjoint matrix of A, $w_1. * w_2$ denotes the componentwise multiplication between two vectors w_1, w_2, and A'_α, A'_ρ are the two $N \times n^2$ matrices defined by $(k = 1, \ldots, N, \ell = 1, \ldots, n^2)$

$$
\begin{aligned}
(A'_\alpha)_{k\ell} &= 2\pi i \rho_k(-\sin(\alpha_k)x_\ell + \cos(\alpha_k)y_\ell)e^{2\pi i \rho_k(\cos(\alpha_k)x_\ell + \sin(\alpha_k)y_\ell)}, \\
(A'_\rho)_{k\ell} &= 2\pi i(\cos(\alpha_k)x_\ell + \sin(\alpha_k)y_\ell)e^{2\pi i \rho_k(\cos(\alpha_k)x_\ell + \sin(\alpha_k)y_\ell)}.
\end{aligned}
$$
(2.6)

The calculation of the Hessian matrix of J is more complicated and leads to a block matrix that is difficult to analyze and compute. Here we report only the diagonal blocks $\nabla^2_{\alpha\alpha}J$ and $\nabla^2_{\rho\rho}J$ (since $\nabla^2_{ff}J$ is trivially $\Re[A^H A]$), which are $N \times N$ diagonal matrices whose kth diagonal entries $(k = 1, \ldots, N)$ are given by

$$
\begin{aligned}
[\nabla^2_{\alpha\alpha}J(f, \alpha, \rho)]_{kk} &= \Re\left[(A''_\alpha f). * (\overline{Af - g})\right]_k + |(A'_\alpha f)_k|^2, \\
[\nabla^2_{\rho\rho}J(f, \alpha, \rho)]_{kk} &= \Re\left[(A''_\rho f). * (\overline{Af - g})\right]_k + |(A'_\rho f)_k|^2,
\end{aligned}
$$
(2.7)

where A''_α, A''_ρ are the two $N \times n^2$ matrices defined by $(k = 1, \ldots, N, \ell = 1, \ldots, n^2)$

$$
\begin{aligned}
(A''_\alpha)_{k\ell} &= 2\pi i \rho_k(-\cos(\alpha_k)x_\ell - \sin(\alpha_k)y_\ell)e^{2\pi i \rho_k(\cos(\alpha_k)x_\ell + \sin(\alpha_k)y_\ell)} \\
&\quad - 4\pi^2\rho_k^2(-\sin(\alpha_k)x_\ell + \cos(\alpha_k)y_\ell)^2 e^{2\pi i \rho_k(\cos(\alpha_k)x_\ell + \sin(\alpha_k)y_\ell)}, \\
(A''_\rho)_{k\ell} &= -4\pi^2(\cos(\alpha_k)x_\ell + \sin(\alpha_k)y_\ell)^2 e^{2\pi i \rho_k(\cos(\alpha_k)x_\ell + \sin(\alpha_k)y_\ell)}.
\end{aligned}
$$
(2.8)

From the diagonal elements of the Hessian $\nabla^2 J(f, \alpha, \rho)$ we can see that problem (2.4) is convex if restricted to f only, but is nonconvex with respect to α and ρ and, even more so, with respect to (f, α, ρ), thus leading to the possible presence of several local minima. In the following section we describe the block-iterative optimization method proposed for the solution of (2.4). The analogous scheme for the minimization problem (2.3) can be easily derived in the same way.

2.1. Block-iterative optimization method. The problem in (2.4) is a nonlinear, in general nonconvex, optimization problem. A remarkable feature of (2.4) is that the optimization variables are naturally grouped in three separate convex sets and the function $J(\cdot, \alpha, \rho)$ is convex for any (α, ρ). The separable structure of the domain is typically exploited by means of projection methods [18] or alternating optimization strategies [19, 20]. In this paper we adopt the latter approach, consisting in one of the following iterative minimization schemes:

(a)

$$
f^{(h+1)} = \underset{f \geq 0}{\arg\min}\, J(f, \alpha^{(h)}, \rho^{(h)}),
$$
(2.9)

$$
\alpha^{(h+1)} = \underset{\alpha^{\min} \leq \alpha \leq \alpha^{\max}}{\arg\min}\, J(f^{(h+1)}, \alpha, \rho^{(h)}),
$$
(2.10)

$$
\rho^{(h+1)} = \underset{\rho^{\min} \leq \rho \leq \rho^{\max}}{\arg\min}\, J(f^{(h+1)}, \alpha^{(h+1)}, \rho).
$$
(2.11)

(b)

$$(2.12) \qquad f^{(h+1)} = \underset{f \geq 0}{\operatorname{argmin}} \ J(f, \alpha^{(h)}, \rho^{(h)}),$$

$$(2.13) \qquad (\alpha^{(h+1)}, \rho^{(h+1)}) = \underset{\substack{\alpha^{\min} \leq \alpha \leq \alpha^{\max} \\ \rho^{\min} \leq \rho \leq \rho^{\max}}}{\operatorname{argmin}} \ J(f^{(h+1)}, \alpha, \rho).$$

The more general convergence result about scheme (b) can be found in [20, Corollary 2], where the authors prove, without any convexity assumption, that any limit point of the sequence $\{(f^{(h)}, \alpha^{(h)}, \rho^{(h)})\}$ generated by the alternating algorithm is a stationary point for problem (2.4). The convergence of scheme (a) to a solution of (2.4), instead, is not guaranteed (see, e.g., the counterexample shown by Powell in [40]). We further point out that, in both cases, computing the exact minimum points in each subproblem is impractical. A typical way to overcome this issue, which is popular in the blind deconvolution framework [15, 16] but also in other contexts [12, 30], consists in computing an approximation of these minimum points by applying an iterative method to each subproblem, stopping the iterations when some criterion is satisfied. In this case, the convergence properties of the alternating scheme also depend on the features of the inner solver, and the convergence result shown in [20] does not apply to these inexact versions of the alternating scheme.

However, convergence results can be proved when a suitable descent method is applied to each partial minimization problem [7, 19]; in particular, this occurs when the approximation of each solution of (2.9)–(2.11) or (2.12)–(2.13) is achieved by performing a finite number of iterations of the scaled gradient projection (SGP) method [9].

2.2. The scaled gradient projection method. The SGP method applies to any minimization problem over a convex set Ω of the form

$$\min_{x \in \Omega} \psi(x)$$

and combines a gradient projection step with variable stepsize and scaling with the well-known Armijo rule to achieve a sufficient decrease in the objective function [6]. Recent applications in deblurring and denoising problems show that this combination can lead to significant accelerations in the convergence speed with respect to traditional first-order methods [21, 32, 38, 44].

The main SGP steps are given in Algorithm 1, where $\mathcal{D}$ denotes the set of positive definite matrices whose eigenvalues and their inverses are bounded below and above by positive constants, and $P_{\Omega, D^{-1}}(z)$ denotes the projection of the point z onto Ω with respect to the norm induced by the matrix D^{-1}, i.e.,

$$P_{\Omega, D^{-1}}(z) = \arg\min_{x \in \Omega}(x - z)^T D^{-1}(x - z).$$

Then, problem (2.4) can be approximately solved by applying a finite number of SGP steps to the minimum problems (2.9)–(2.11): the resulting scheme is detailed in Algorithm 2, and any limit point of the sequence $\{(f^{(h)}, \alpha^{(h)}, \rho^{(h)})\}$ is stationary (see [7, Theorem 4.2]). Analogous results hold true if a finite number of SGP iterations are applied to subproblems (2.12)–(2.13).

Algorithm 1. Scaled gradient projection (SGP) method.

Choose the starting point $x^{(0)} \in \Omega$, and set the parameters $\nu, \mu \in (0, 1)$, $0 < \sigma_{\min} < \sigma_{\max}$.

FOR $k = 0, 1, 2, \ldots$ DO THE FOLLOWING STEPS:

STEP 1. Choose the steplength parameter $\sigma_k \in [\sigma_{\min}, \sigma_{\max}]$ and the scaling matrix $D_k \in \mathcal{D}$;

STEP 2. Projection: $y^{(k)} = P_{\Omega, D_k^{-1}}(x^{(k)} - \sigma_k D_k \nabla \psi(x^{(k)}))$;

STEP 3. Descent direction: $\Delta x^{(k)} = y^{(k)} - x^{(k)}$;

STEP 4. Set $\lambda_k = 1$;

STEP 5. Backtracking loop:

Let $\psi_{new} = \psi(x^{(k)} + \lambda_k \Delta x^{(k)})$;

IF $\psi_{new} \leq \psi(x^{(k)}) + \nu \lambda_k \nabla \psi(x^{(k)})^T \Delta x^{(k)}$ THEN

go to Step 6;

ELSE

set $\lambda_k = \mu \lambda_k$ and go to Step 5.

ENDIF

STEP 6. Set $x^{(k+1)} = x^{(k)} + \lambda_k \Delta x^{(k)}$.

END

We underline that any steplength σ_k in the range $[\sigma_{\min}, \sigma_{\max}]$ and scaling matrix D_k in $\mathcal{D}$ are allowed; this freedom of choice can then be fruitfully exploited in order to achieve significant improvements in terms of convergence speed. With regard to the choice of σ_k, notable results have been obtained in denoising and deblurring problems [32, 38, 44] by alternating the two Barzilai–Borwein rules [2]

$$(2.14) \qquad \sigma_k^{(BB1)} = \frac{s^{(k-1)^T} D_k^{-1} D_k^{-1} s^{(k-1)}}{s^{(k-1)^T} D_k^{-1} z^{(k-1)}}, \qquad \sigma_k^{(BB2)} = \frac{s^{(k-1)^T} D_k z^{(k-1)}}{z^{(k-1)^T} D_k D_k z^{(k-1)}},$$

with $s^{(k-1)} = x^{(k)} - x^{(k-1)}$ and $z^{(k-1)} = \nabla \psi(x^{(k)}) - \nabla \psi(x^{(k-1)})$. The values produced by these rules are constrained into the interval $[\sigma_{\min}, \sigma_{\max}]$ in the following way:

IF $s^{(k-1)^T} z^{(k-1)} \leq 0$ THEN

$\sigma_k^{(1)} = \min\{\sigma_{k-1}, \ \sigma_{\max}\}$;

ELSE

$\sigma_k^{(1)} = \min\left\{\sigma_{\max}, \ \max\left\{\sigma_{\min}, \ \sigma_k^{(BB1)}\right\}\right\}$;

ENDIF

IF $s^{(k-1)^T} z^{(k-1)} \leq 0$ THEN

$\sigma_k^{(2)} = \min\{\sigma_{k-1}, \ \sigma_{\max}\}$;

ELSE

$\sigma_k^{(2)} = \min\left\{\sigma_{\max}, \ \max\left\{\sigma_{\min}, \ \sigma_k^{(BB2)}\right\}\right\}$;

ENDIF

In particular, the final parameter σ_k is chosen by exploiting the alternation strategy [17]

IF $\sigma_k^{(2)}/\sigma_k^{(1)} \leq \tau_k$ THEN

Algorithm 2. Cyclic block scaled gradient projection (CBSGP) method.

Choose the starting points $f^{(0)} \geq 0$, $\alpha^{\min} \leq \alpha^{(0)} \leq \alpha^{\max}$, $\rho^{\min} \leq \rho^{(0)} \leq \rho^{\max}$, and three integers $N_f, N_\alpha, N_\rho \geq 1$.

FOR $h = 0, 1, 2, \ldots$ DO THE FOLLOWING STEPS:

      STEP f. Choose an integer $1 \leq N_f^{(h)} \leq N_f$ and compute $f^{(h+1)}$ by applying $N_f^{(h)}$ iterations of Algorithm 1 to problem (2.9) starting from the point $f^{(h)}$.

      STEP α. Choose an integer $1 \leq N_\alpha^{(h)} \leq N_\alpha$ and compute $\alpha^{(h+1)}$ by applying $N_\alpha^{(h)}$ iterations of Algorithm 1 to problem (2.10) starting from the point $\alpha^{(h)}$.

      STEP ρ. Choose an integer $1 \leq N_\rho^{(h)} \leq N_\rho$ and compute $\rho^{(h+1)}$ by applying $N_\rho^{(h)}$ iterations of Algorithm 1 to problem (2.11) starting from the point $\rho^{(h)}$.

END

$$\sigma_k = \min_{j = \max\{1, k+1-M_\sigma\}, \ldots, k} \sigma_j^{(2)};$$
$$\tau_{k+1} = 0.9 \cdot \tau_k;$$

ELSE

$$\sigma_k = \sigma_k^{(1)};$$
$$\tau_{k+1} = 1.1 \cdot \tau_k;$$

ENDIF

where M_σ is a prefixed positive integer and $\tau_1 \in (0, 1)$.

On the contrary, the choice of a suitable scaling matrix strongly depends on both the objective function and the constraints required by the specific application [9]. Therefore, we will discuss our choices in the following section.

3. Numerical experiments: Imaging with RHESSI. The solar satellite RHESSI [31] was launched by NASA on February 5, 2002 to study solar flares and other energetic solar phenomena [26]. Thanks to the spacecraft's imaging system, RHESSI may provide both X-ray two-dimensional images with an angular resolution from 2 to 7 arcseconds and X-ray one-dimensional spectra with a spectral resolution from 0.5 to 2 keV. Instead of the typical approaches for focusing optics wavelengths, which are impractical when dealing with X-rays, RHESSI exploits a bigrid collimator strategy, which allows a partial transmission of the incident photons depending on their incident direction. More precisely, RHESSI observes X-ray emission from the sun through a set of nine coaligned pairs of rotating modulation collimators (RMCs), and the transmitted radiation is recorded on a set of cooled high-purity germanium detectors [23]. The raw data provided by RHESSI are, therefore, nine count profiles of the detected radiation as a function of time, modulated by the pairs of grids. Several reconstruction algorithms (e.g., CLEAN, Pixon) have been developed since the launch of RHESSI for recovering the X-ray image directly from the modulated counts provided by the detectors, by means of the so-called modulation profiles. However, a growing interest has been devoted recently to an alternative approach, which consists in (a) exploiting the combination between rotation and modulation in order to estimate some image Fourier components at specific spatial frequencies (also known as *visibilities* [34]), and (b) using Fourier-based inversion algorithms to restore the image [8, 10, 33]. The procedure adopted to compute the visibilities from the
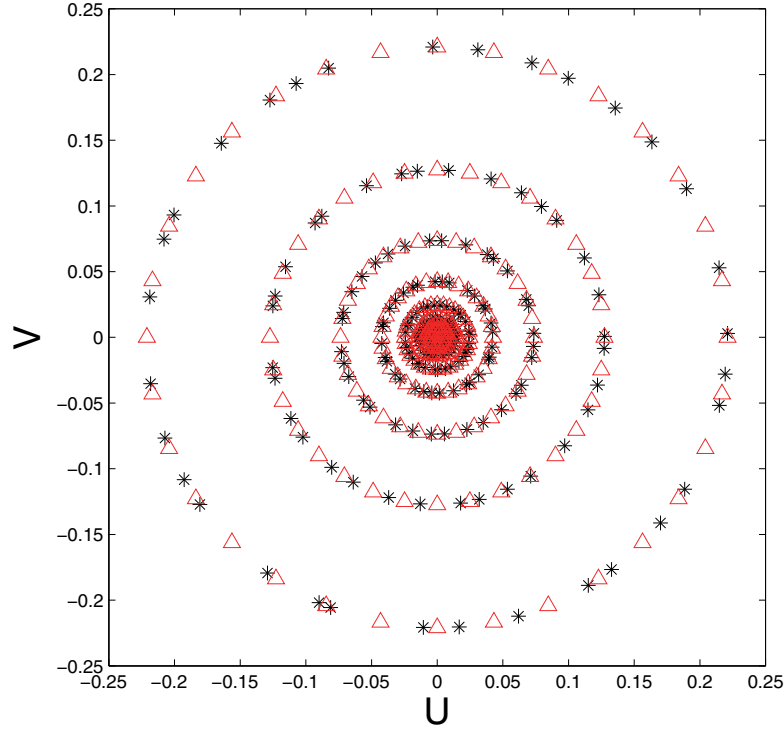
Figure 1. *Example of distribution of the sampling points for the RHESSI imaging problem (black asterisks), together with the bounds of the box constraints $\alpha^{\min}, \alpha^{\max}$ (red triangles).*

modulated count profiles is outlined in Appendix A (see also [23, 42]). In practice, because of the RHESSI instrument design, imaging information is recorded as a set of visibilities (varying from tens to two or three hundreds according to the strength of the signal), measured at spatial frequencies (u, v) arranged around nine concentric circles in the Fourier plane whose radii are determined by the angular resolution of the nine collimators (see Figure 1).

The resulting mathematical model is given by

$$(3.1) \qquad \min_{\substack{f \in \mathbb{R}^{n^2},\ f \geq 0 \\ \alpha \in \mathbb{R}^N,\ \alpha^{\min} \leq \alpha \leq \alpha^{\max}}} J(f, \alpha) \equiv \frac{1}{2} \|A(\alpha)f - g\|_{\mathbb{C}^N}^2,$$

which is a particular case of (2.4) in which the radial coordinate ρ of each sample is known and the angular coordinate α is given by the rotation angle of the spacecraft. The key point of this formulation is that the actual value of the angular coordinate α is not exactly known, since only its existence in the interval $[\alpha^{\min}, \alpha^{\max}]$ is theoretically ensured (for the first mean value theorem for integration; see section A.3). What is typically done in the existing reconstruction algorithms is to associate the computed visibilities to the angular coordinate α given by the middle point of $[\alpha^{\min}, \alpha^{\max}]$, since in most cases this choice provides a good estimate of the real value. Our approach allows us to remove this assumption and to treat the value of $\alpha \in [\alpha^{\min}, \alpha^{\max}]$ as a further unknown of the reconstruction problem.

3.1. Scaling matrix and CBSGP parameters. In a recent paper [8], Bonettini and Prato proposed a nonscaled gradient method, called Space-D, as an iterative regularization algorithm for solving the minimum problem (2.1), showing significant improvements with respect to the other visibility-based image reconstruction methods developed for the RHESSI data analysis. Space-D falls within the scheme described in Algorithm 1, where D_k is set equal to the identity matrix for all k and the steplength parameter is chosen with the alternating rule described in section 2.2. The stopping criterion chosen to terminate the iterations is given by

$$(3.2) \qquad |\psi(x^{(k)}) - \psi(x^{(k-1)})| < 10^{-4}|\psi(x^{(k)})|.$$

The choice of the threshold 10^{-4} guarantees a limited number of iterations at each step and, therefore, a regularization effect on the recovered solution (more details are shown in [9]). Driven by the good results achieved by Space-D in reconstructing RHESSI images, we used the same settings for both steps of the cyclic approach.[2] In the following, the term "Blind-Gradient" will refer to this choice. Moreover, for the nonlinear least squares problem (2.10), we also tried Space-D with a nontrivial scaling matrix D_k. In particular, we used the diagonal matrix suggested by applying a Gauss–Newton approximation to the Hessian $\nabla^2_{\alpha\alpha}J$ defined in (2.7) [35]. The resulting scaling matrix D_k defined in Step 1 of Algorithm 1 is then given by

$$(3.3) \qquad \left[\mathrm{diag}\left(\min\left\{L_2, \ \max\left\{L_1, |A'_\alpha f^{(h+1)}|^2\right\}\right\}\right)\right]^{-1}, \quad 0 < L_1 < L_2,$$

computed in the current iterate α.

Due to the acceleration provided by the presence of the scaling matrix, in this case we performed only one iteration of the SGP algorithm, instead of stopping the procedure when (3.2) was satisfied. Further tests with N_α greater than 1 did not show any significant improvement with respect to the single iteration case. In the following, the term "Blind-Newton" will refer to the CBSGP method in which Space-D is used for the minimization of (2.9), while SGP with the scaling matrix (3.3) is applied to the nonlinear least squares problem (2.10) (the values of L_1 and L_2 have been set equal to 10^{-10} and 10^{10}, respectively).

As far as the number of cycles h is concerned, the Blind-Gradient algorithm is stopped when criterion (3.2) in either the step on f or the step on α is satisfied at the first iteration. Since the step on α in the Blind-Newton approach is always performed in one iteration, in this case only the step on f is considered for the choice of the number of cycles. Moreover, for both algorithms a maximum number of cycles equal to 20 is imposed.

The remaining parameters are set according to the Space-D algorithm: $\nu = 10^{-4}$, $\mu = 0.4$, $\sigma_0 = 1.3$, $\sigma_{\min} = 10^{-5}$, $\sigma_{\max} = 10^5$, $\tau_1 = 0.5$, and $M_\sigma = 3$. Finally, the iterative procedure is initialized with a constant image $f^{(0)}$ whose flux is given by $\max_{k=1,...,N}\{|g_k|\}$.

3.2. Numerical experiments. In this section we investigate the effectiveness of the CBSGP approach in reconstructing X-ray images of solar flares from RHESSI data. The evaluation of the results will be carried out in comparison with the Space-D algorithm. To

---

[2]We remark that in this situation both of the block-iterative optimization methods described in section 2.1 reduce to the alternate inexact solution of subproblems (2.9)–(2.10).

this aim, we built up three simulated datasets by following a strategy similar to the one proposed in [8]. In particular, we did the following:

1. We considered a real flare event (23 July 2002, 00:29:10–00:30:19 UT), and, starting from the radiation collected by RHESSI in three different energy ranges, we reconstructed the corresponding images with the CLEAN algorithm [23], available in the SolarSoftWare (SSW) system, which is a set of integrated software libraries, databases, and system utilities for solar physics. These images, suitably cleared of possible artifacts introduced by the reconstruction method, have been considered as our target distributions.

2. We selected the detectors from 3 to 9, which is a usual choice when dealing with RHESSI images (see, e.g., [36, 39, 43]), and for the kth subcollimator ($k = 3, \ldots, 9$) we simulated a typical uniform discretization $(\overline{\alpha}_{0,k}, \ldots, \overline{\alpha}_{N_k,k})$ of the rotation angle values. In particular, $(N_3, \ldots, N_9)$ have been set equal to $(32, 32, 32, 30, 22, 12, 6)$, thus leading to a total number of $N = 166$ visibilities.

3. In order to mimic the RHESSI data stacking process, for each bin $[\overline{\alpha}_{i-1,k},\ \overline{\alpha}_{i,k}]$ ($i = 1, \ldots, N_k$), we chose a random angle $\alpha_{i,k}$ and calculated the corresponding visibility through numerical integration of the Fourier transform by means of the SSW routine `hsi_vis_map2vis.pro`. The values $\overline{\alpha}_{i-1,k}, \overline{\alpha}_{i,k}$ have been used as bounds of the box constraints for $\alpha_{i,k}$. We will denote the resulting three synthetic datasets as Sim1, Sim2, and Sim3. The visibilities in these datasets will be characterized by the presence of a systematic source of noise only, due to the fact that in the inversion procedure the exact values of $\alpha_{i,k}$ are unknown.

4. Finally, we corrupted the visibilities by realistic statistical noise through the SSW routine `hsi_vis_randomize.pro`. These latter datasets will be denoted by Sim1_N, Sim2_N, and Sim3_N.

We point out that our way of simulating the visibilities is a very simplified version of the procedure employed by RHESSI and described in Appendix A. We chose this iter since it allows us to know the "true" values of α and, therefore, to evaluate the effectiveness of the semiblind approach. In particular, the choice of angles $\alpha_{i,k}$ according to a uniform random distribution in $[\overline{\alpha}_{i-1,k},\ \overline{\alpha}_{i,k}]$ is arbitrary.

3.3. Results. We first consider the Sim1, Sim2, and Sim3 cases, in which no statistical noise is present on the simulated visibilities. Besides all the parameters described in the previous section, the CBSGP scheme is initialized with a vector $\alpha^{(0)}$ equal to the middle points of each bin. In these settings, the image reconstructed by the Space-D algorithm is exactly $f^{(1)}$ (see Algorithm 2).

In Figure 2 we report the target images (first row) and the reconstructions obtained with Space-D (second row), Blind-Gradient (third row), and Blind-Newton (fourth row). Moreover, in Figure 3 we show the relative reconstruction errors in Euclidean norm of the image (first row) and the array of the angular coordinates (second row) as functions of the cycle number. From the results obtained we can see that the semiblind approach is able to improve the image quality with respect to the Space-D algorithm in all of the considered datasets, thanks to a better estimate of the underlying array α. Although the restored images appear to be very similar, a better separation of the two strongest sources in Sim2 can be noticed, together with a higher resolution in the two northern compact sources in Sim3 (see Figure 2) and
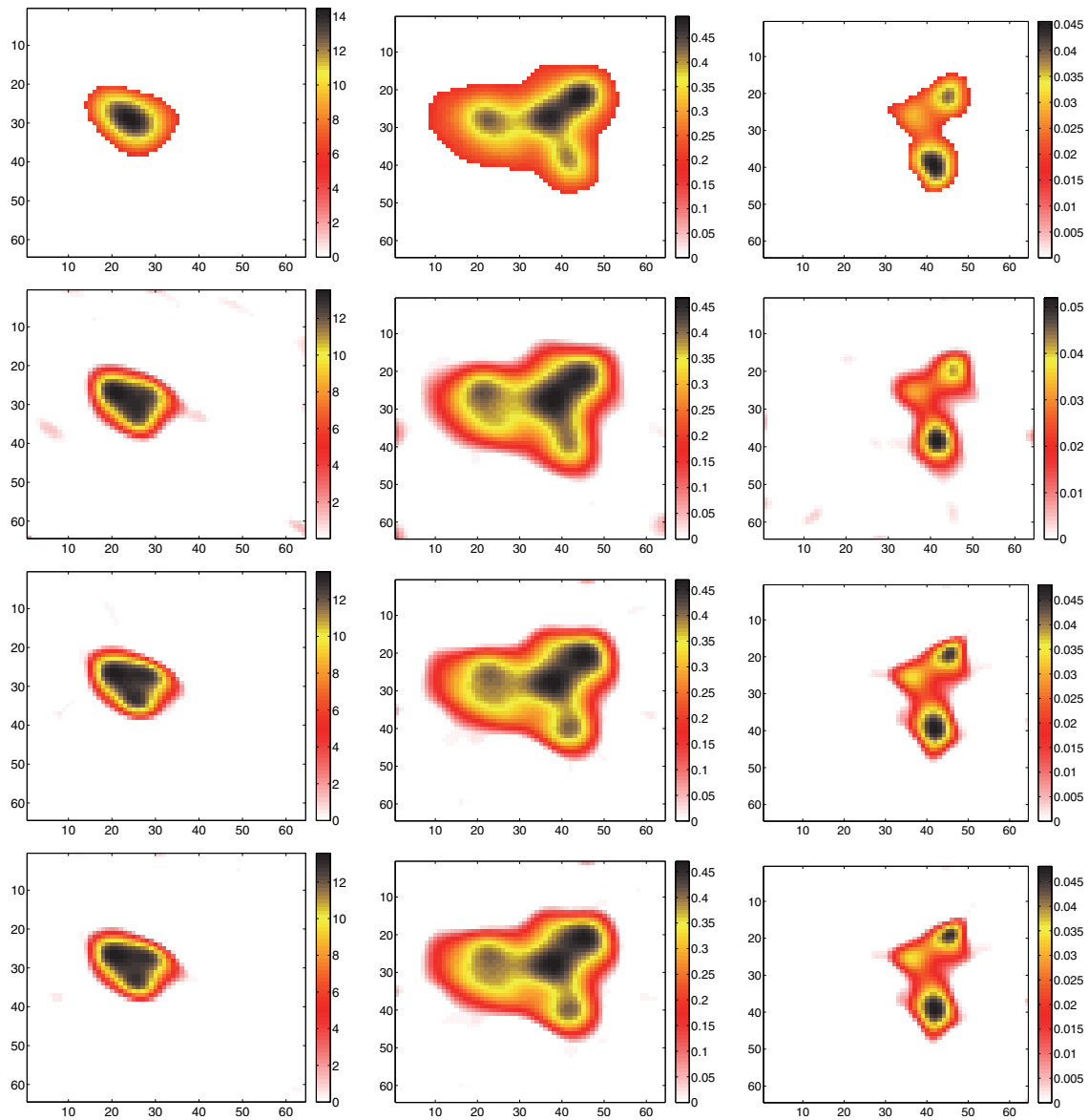
Figure 2. *Results of the three simulated tests considered in this paper with noise-free data: Sim1 (first column), Sim2 (second column), and Sim3 (third column). From the first to the last row, the theoretical image and the reconstructions with Space-D, Blind-Gradient, and Blind-Newton are presented, respectively.*

an overall presence of fewer small artifacts. As far as the use of a nontrivial scaling matrix D_k is concerned, in this case we are not able to highlight a clear preference, even if the Blind-Gradient method seems to provide in general a lower reconstruction error.

Further comparisons can be made by analyzing the results obtained in the presence of statistical noise, which are reported in Figures 4 and 5. In particular, we can see the following:

    1. For the Sim1_N and Sim2_N datasets, the CBSGP methodology still leads to some
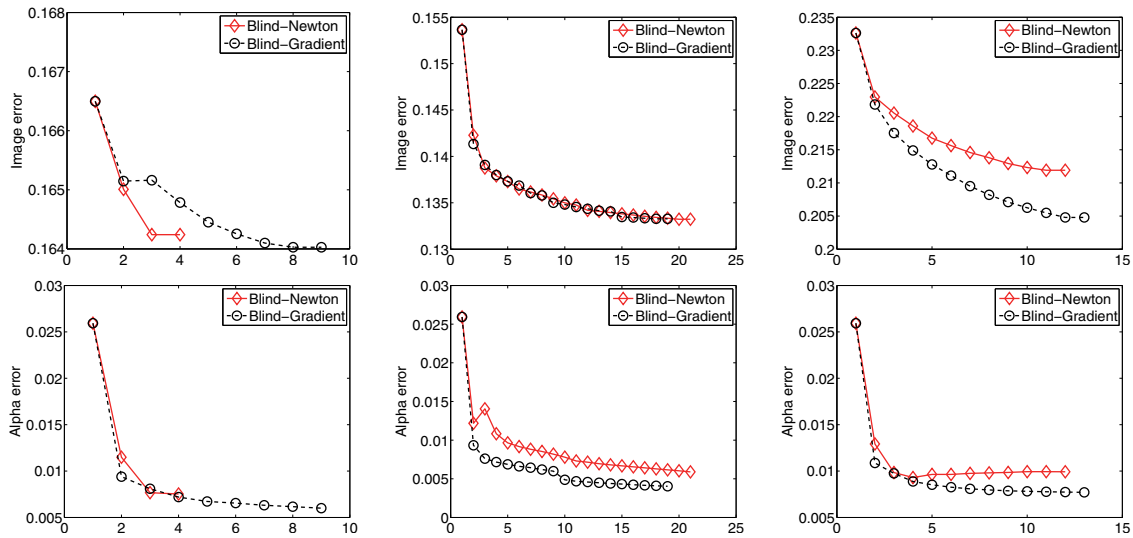
Figure 3. *Relative reconstruction errors on the image (top row) and α (bottom row) as a function of the number of cycles for the Blind-Gradient and Blind-Newton methods in the case of noise-free data. The left (resp., central, right) column refers to the Sim1 (resp., Sim2, Sim3) dataset. The analogous values for the Space-D algorithm correspond to the first point of each plot.*

improvements in the final image, attested to by a lower reconstruction error with respect to an approach with a fixed rotation angle array.

2. Especially in the Sim1_N case, the Blind-Newton algorithm allows us to obtain a better reconstruction when compared with a nonscaled method. We think that the reason behind this behavior might be identified with the possible nonoptimal stopping criterion (3.2) used for the step α in the Blind-Gradient algorithm (see Algorithm 2).

3. The semiblind approach does not lead to significant improvements in the Sim3_N simulation. We point out that, in this case, the target image is the one with the lower radiation. Therefore, the corruption of the Sim3_N input visibilities introduced by the Poissonian-like statistical noise overcomes the systematic error related to a bad choice for α. As a result of this, the regularizing effect achieved by the Space-D algorithm already provides the best reconstruction. It is worth to point out that our semiblind approach seems to "recognize" these situations and automatically arrests the iterations at the second cycle.

4. Conclusions and future work. In this paper we introduced an optimization approach for the image restoration from Fourier samples with uncertainties on the data frequencies. This problem can be reformulated as a semiblind deconvolution problem, in which both the object under analysis and the PSF of the measuring instruments have to be gathered from the data. A cyclic block alternating scheme has been employed for the solution of the corresponding optimization problem, resulting in two or three alternate minimization steps, according to the used strategy. For each step, the simple constraints together with recent results on the convergence properties suggested the use of a projected scaled gradient method with a suitable adaptive choice for the steplength parameter.
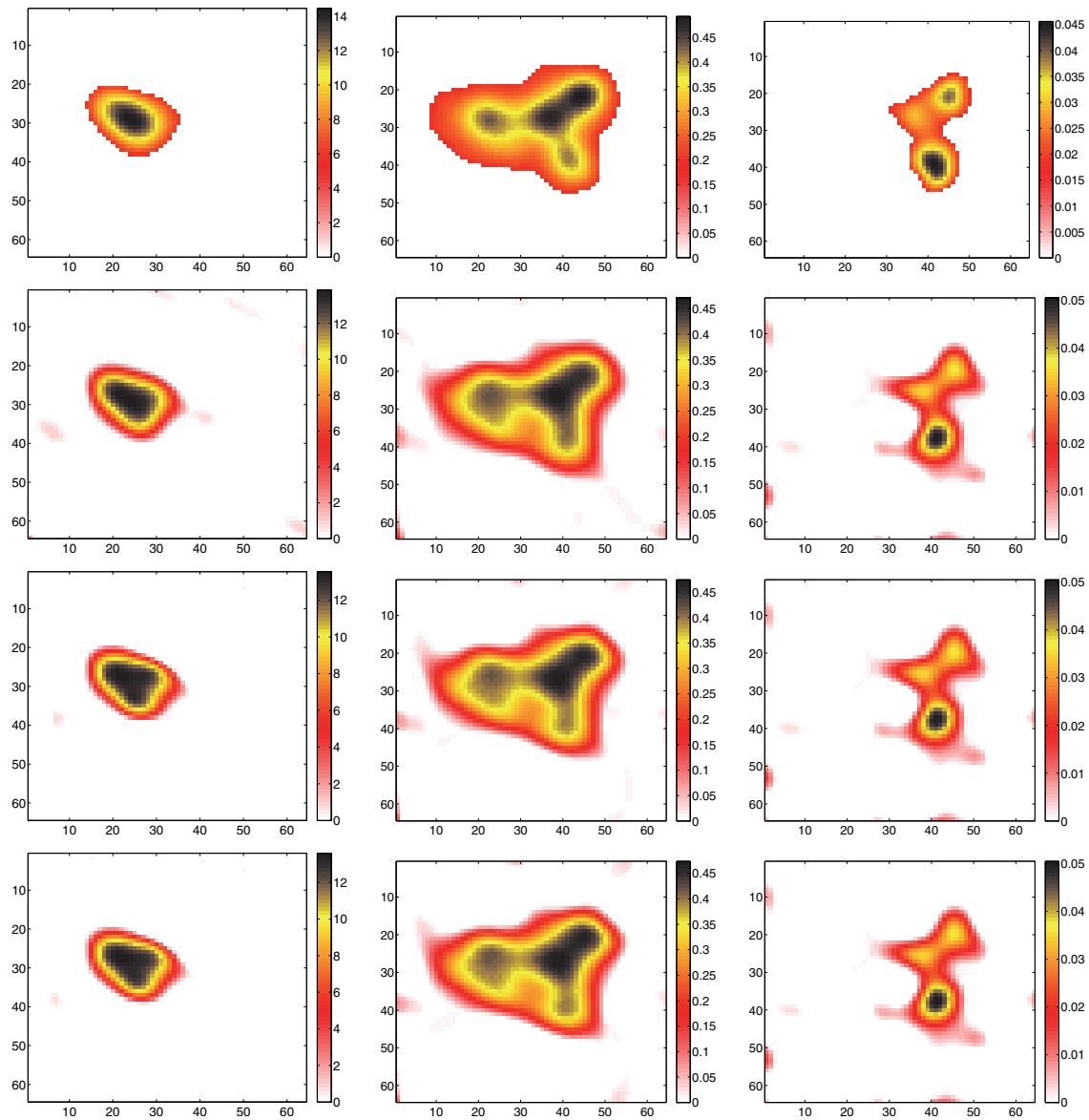
Figure 4. *Results of the three simulated tests considered in this paper with noisy data: Sim1_N (first column), Sim2_N (second column), and Sim3_N (third column). From the first to the last row, the theoretical image and the reconstructions with Space-D, Blind-Gradient, and Blind-Newton are presented, respectively.*

The proposed scheme has been applied to the image reconstruction of X-rays emitted during a solar flare and collected by the NASA RHESSI satellite. Thanks to the RHESSI imaging hardware, Fourier transform values of the unknown image can be estimated from the collected radiation, corresponding to spatial frequencies depending on the spacecraft's rotation angle. The results obtained on simulated datasets showed that, when the unknown dependency of the spatial frequencies on the rotation angle is considered in the inversion
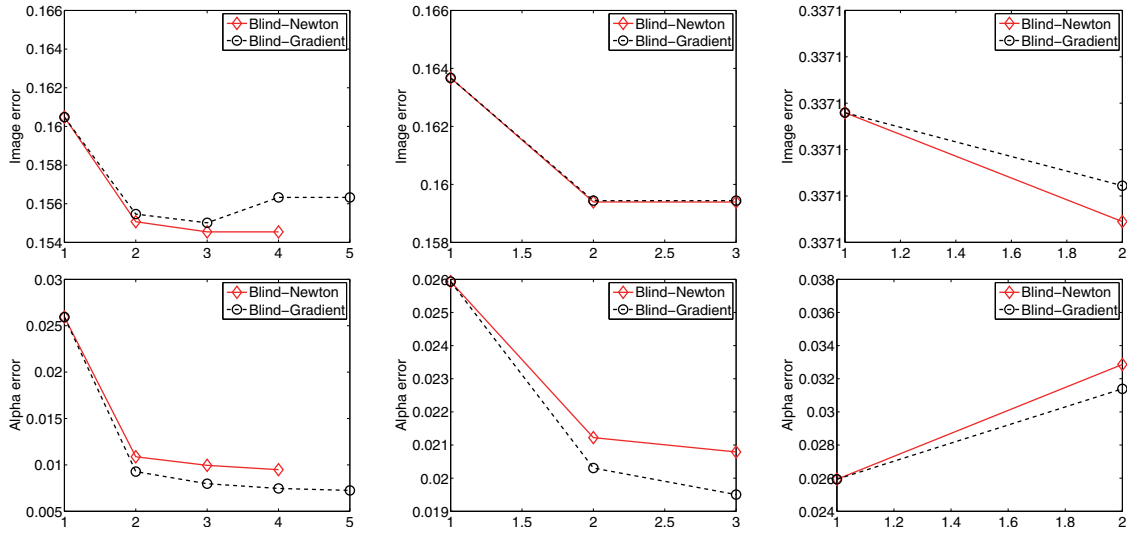
Figure 5. *Relative reconstruction errors on the image (top row) and α (bottom row) as a function of the number of cycles for the Blind-Gradient and Blind-Newton methods in the case of noisy data. The left (resp., central, right) column refers to the Sim1_N (resp., Sim2_N, Sim3_N) dataset. The analogous values for the Space-D algorithm correspond to the first point of each plot.*

procedure, better images can be obtained, as suggested by both a direct observation of the pictures and the lower reconstruction errors registered.

Future work will involve the extension of the semiblind approach to objective functions including regularization penalties on the image, such as the Tikhonov term for smooth objects or an edge-preserving function for sharp contents. With regard to the astronomical application considered in this paper, we would like to compare the CBSPG approach with the reconstruction algorithms available within the RHESSI software that act directly on the measured radiation, without estimating the visibilities. Finally, applications of the methodology described in this paper to the reconstruction of electron images via the Bremsstrahlung integral equation [25, 37] will be also considered.

Appendix A. RMC and Fourier transform. In this section we describe how the RMC-based devices provide samples of the Fourier transform of the radiation emitted from a source (for the basic concepts of RHESSI imaging, see also [23]). First, we explain how the rotation of two parallel grids modulates the radiation flux emitted by a point source located at infinity. For this purpose, we deduce the analytical expression of the effective area coming from the superposition of the two grids of the collimator. In the following, we rewrite this expression in a mapping geometry setting obtained through the introduction of a coordinate system in the image plane. Finally, we show how the previous considerations can be used to deduce the values of the Fourier transform of the emitted radiation for specific frequencies.

The content of the following sections is well known in the RHESSI community and has been outlined in several astronomical works and conferences. Here we provide a purely mathematical formalization of the procedure, with the aim of allowing scientists with a mathematical
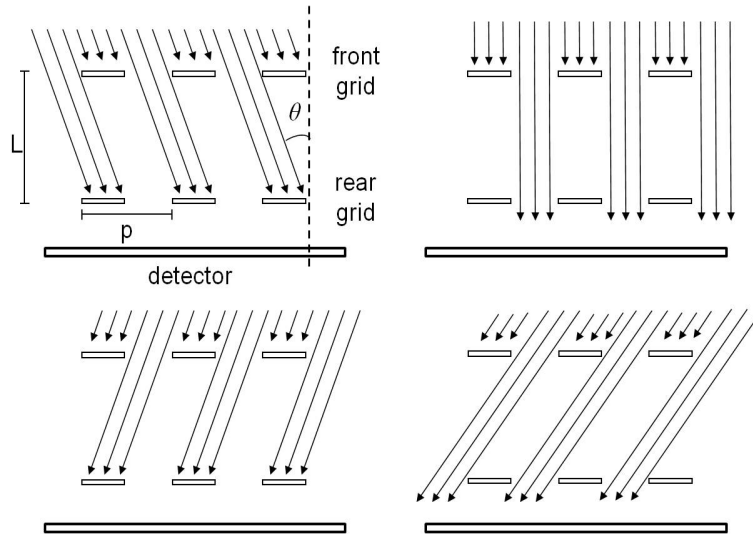
Figure 6. *Collimator schematic profile: the detector is below the two parallel grids. The figure illustrates the four cases of incident angles θ of the incoming radiation equal to −p/2L (top left), 0 (top right), p/2L (bottom left), and p/L (bottom right).*

background to appreciate the relation between the general framework described in section 2 and the application to the RHESSI imaging problem.

A.1. The transmission probability. For a given subcollimator, let p (pitch) be the space between each slat of the grids, L the distance between the front and the rear grids (see Figure 6), and $q \in [0, 1]$ the ratio between the width of a slit and the pitch of the grids. For simplicity, first we consider the ideal case $q = \frac{1}{2}$ when the slits are exactly equal to the slats. Moreover, we consider radiation energies that are sufficiently low so that the slats can be considered opaque, but also sufficiently high so that the diffraction effects can be considered negligible.

We observe that the fraction of the radiation flux reaching the detector depends on the incidence angle θ. In particular, when θ is equal to 0, i.e., when the radiation is orthogonal to the grids, the grids are perfectly aligned, so 50% of the flux will be rejected, while the other 50% will reach the detector. If we assume $\theta \approx \tan(\theta)$ (which is allowed since $L \gg p$), we obtain the same flux modulation when $\theta = kp/L$, $k \in \mathbb{Z}$ (see Figure 6). On the contrary, we will have the full rejection of the flux when the slits of the rear grid are exactly aligned to the slats of the front grid, i.e., for angles $\theta = (2k + 1)p/2L$, $k \in \mathbb{Z}$. Therefore, the fraction of the incoming radiation flux that will reach the detector (called *transmission probability*) is a piecewise linear function of the incidence angle θ given by

$$(A.1) \qquad P(\theta) = \begin{cases} 0 & \text{if } -1 \leq \theta \leq -2q, \\ \frac{1}{2}\theta + q & \text{if } -2q \leq \theta \leq 0, \\ -\frac{1}{2}\theta + q & \text{if } 0 \leq \theta \leq 2q, \\ 0 & \text{if } 2q \leq \theta \leq 1 \end{cases}$$

for $q \in [0, \frac{1}{2}]$ and by

(A.2)
$$P(\theta) = \begin{cases} 2q - 1 & \text{if } -1 \le \theta \le -2(1-q), \\ \frac{1}{2}\theta + q & \text{if } -2(1-q) \le \theta \le 0, \\ -\frac{1}{2}\theta + q & \text{if } 0 \le \theta \le 2(1-q), \\ 2q - 1 & \text{if } 2(1-q) \le \theta \le 1 \end{cases}$$

for $q \in [\frac{1}{2}, 1]$. In these expressions the angle θ is given in units of $p/2L$, and, consequently, $P(\theta)$ is periodic with period 2 in the interval $[-1, 1]$ (see Figure 7).

By introducing a new variable $\phi = 2\pi L\theta/p$, which allows us to consider angles in the usual intervals $(-\pi + 2k\pi, \pi + 2k\pi)$, we can write the cosine expansion of the transmission probability function as follows:

(A.3)
$$P(\phi) = \left(\frac{c_0}{2} + c_1 \cos(\phi) + c_2 \cos(2(\phi)) + c_3 \cos(3(\phi)) + \cdots \right).$$

By straightforward application of their definition, the Fourier coefficients c_k will become

(A.4)
$$c_0 = \int_{-1}^{1} P(\theta)d\theta = 2q^2,$$

(A.5)
$$c_k = \int_{-1}^{1} P(\theta)\cos(k\pi\theta)d\theta = \frac{2\sin^2(k\pi q)}{(k\pi)^2}, \qquad k \in \mathbb{Z}_+.$$

We observe that the Fourier coefficients c_k decay as $\frac{1}{k^2}$ so that the higher harmonics have a small weight. We also point out that for realistic detectors the coefficients c_k have significant uncertainty; therefore, (A.5) provides only an approximation of the true coefficients.

The detected flux coming from a point source of intensity f_0 and incidence angle θ will be proportional to the fraction $P(\theta)$ of the detector area, which is called *effective area*. More details are given in the following section.

A.2. The imaging geometry. In (A.3) the independent variable ϕ has been built as a function of the incident angle between the incoming radiation and the collimator axis. In this section we move the reference frame from the collimator to the image plane. To this aim, we imagine looking at the image plane from the detector's point of view. At a given instant, the pair of grids will be placed in a certain way, and a certain effective area will be visible from the detector. In Figure 8 we schematically graph this configuration by displaying the segments in which the transmission is maximum, which essentially correspond to the peaks of the profiles in Figure 7.

Let us denote by $\Pi(t)$ the orthogonal projection on the image plane of the collimator axis, and by $(x_{coll}(t), y_{coll}(t))$ the coordinates of $\Pi(t)$ with respect to a fixed coordinate system (X, Y) on the image plane. At the time of reconstruction, the user is asked to choose the map center Ξ, with coordinates (x_{map}, y_{map}) with respect to the same coordinate system, the pixel size, and the field of view. We denote by $r_0(t)$ the vector connecting $\Pi(t)$ with Ξ (see Figure 8).

We remark that, at each instant, both the coordinates (x_{map}, y_{map}), which never change in the reference system fixed on the image plane, and $(x_{coll}(t), y_{coll}(t))$, which change in a
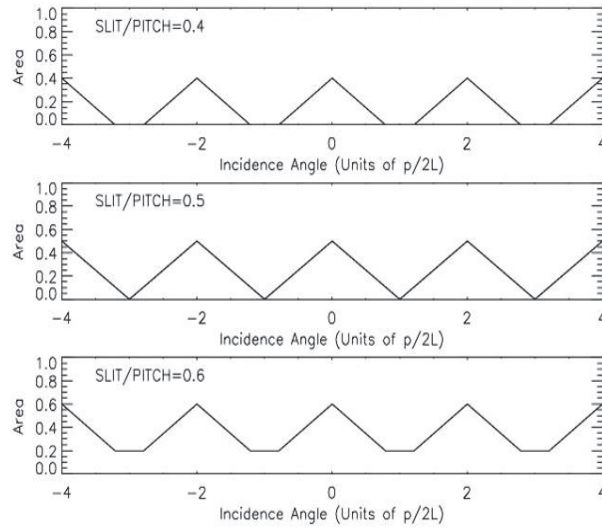
Figure 7. *Transmission probability as a function of the incidence angle θ for $q = 0.4$ (top), $q = 0.5$ (middle), and $q = 0.6$ (bottom).*
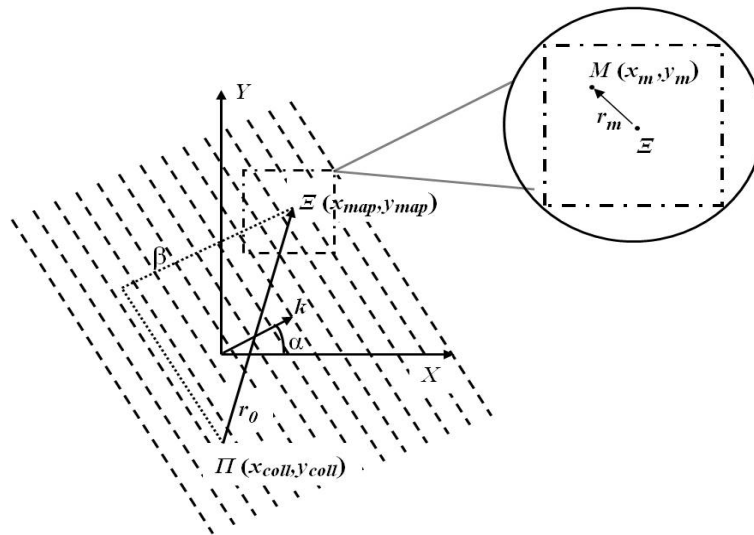

Figure 8. *The imaging geometry. The dashed lines represent the points of maximum transmission of the subcollimators. The inset highlights the image to be reconstructed. Further details are discussed in the text.*

predictable way as the collimator rotates, are known. Therefore, the vector $r_0(t)$ can be calculated as a function of time (or, equivalently, of rotation angle). Moreover, each pixel M of the image will have coordinates (x_m, y_m) and will be connected with the map center through a vector r_m.

We also denote by $\alpha(t)$ the angle between the X-axis and the direction orthogonal to the slats, and by $k(t) = \frac{2\pi}{p}(\cos(\alpha(t)), \sin(\alpha(t)))$ the vector with magnitude $\frac{2\pi}{p}$ pointing in this direction.

According to the approximation $\phi \approx \tan(\phi)$, the incidence angle $\phi_m(t)$ of a point source placed in M is the component of the vector connecting M and $\Pi(t)$ along the direction $k(t)$. Thus, we can approximate the incidence angle as

$$\phi_m(t) = (r_m + r_0(t)) \cdot k(t) = (x_m - x_{coll}(t), y_m - y_{coll}(t)) \cdot \frac{2\pi}{p}(\cos(\alpha(t)), \sin(\alpha(t)))$$

$$= \frac{2\pi}{p}[(x_m - x_{coll}(t))\cos(\alpha(t)) + (y_m - y_{coll}(t))\sin(\alpha(t))]$$

$$\text{(A.6)} \qquad = \underbrace{\frac{2\pi}{p}[(x_m - x_{map})\cos(\alpha(t)) + (y_m - y_{map})\sin(\alpha(t))]}_{=:\phi'_m}$$

$$\underbrace{- \frac{2\pi}{p}[(x_{coll}(t) - x_{map})\cos(\alpha(t)) + (y_{coll}(t) - y_{map})\sin(\alpha(t))]}_{=:\beta(t)},$$

where $\beta(t)$ is called the *aspect phase* and describes the component of $r_0(t)$ orthogonal to the grids. Then, the detected flux due to the source in the point M can be written as

$$\text{(A.7)} \quad C(\phi_m(t)) = Kf_0\left(\frac{c_0}{2} + c_1\cos(\phi_m(t)) + c_2\cos(2(\phi_m(t))) + c_3\cos(3(\phi_m(t))) + \cdots\right),$$

where K is a hardware-dependent constant (accounting, e.g., for the detector's live time and the energy-dependent subcollimator's transmission). As stated in [23], the parameters K and c_i ($i \in \mathbb{N}$) have fixed values in the case of ideal grids, while for RHESSI they become slowly varying functions of time and are determined through a self-calibrating preprocessing.

If we consider only a finite number N_H of harmonics, at a given time t, the modulation of the incident radiation provided by each subcollimator leads to a measured count profile in the form

$$\text{(A.8)} \qquad C(t) = K\int f(x,y)\left(\frac{c_0}{2} + \sum_{j=1}^{N_H} c_j\cos(j\phi(x,y,t))\right)dxdy.$$

A.3. Visibilities and data stacking. A visibility is the value of the two-dimensional Fourier transform of a spatial flux distribution $f(x,y)$ at a specific frequency (u,v) in the Fourier space:[3]

$$\text{(A.9)} \qquad V(u,v) = \int f(x,y)e^{2\pi i(u(x-x_{map})+v(y-y_{map}))}dxdy, \qquad (u,v) \in \mathbb{R}^2.$$

If we compare (A.8) and (A.9), and if we recall the relation $\phi(x,y,t) = \phi'(x,y,t) - \beta(t)$ shown

---

[3]Consistently with the definition of visibility given in [33, 39], we define the Fourier transform with a positive sign in the exponent.
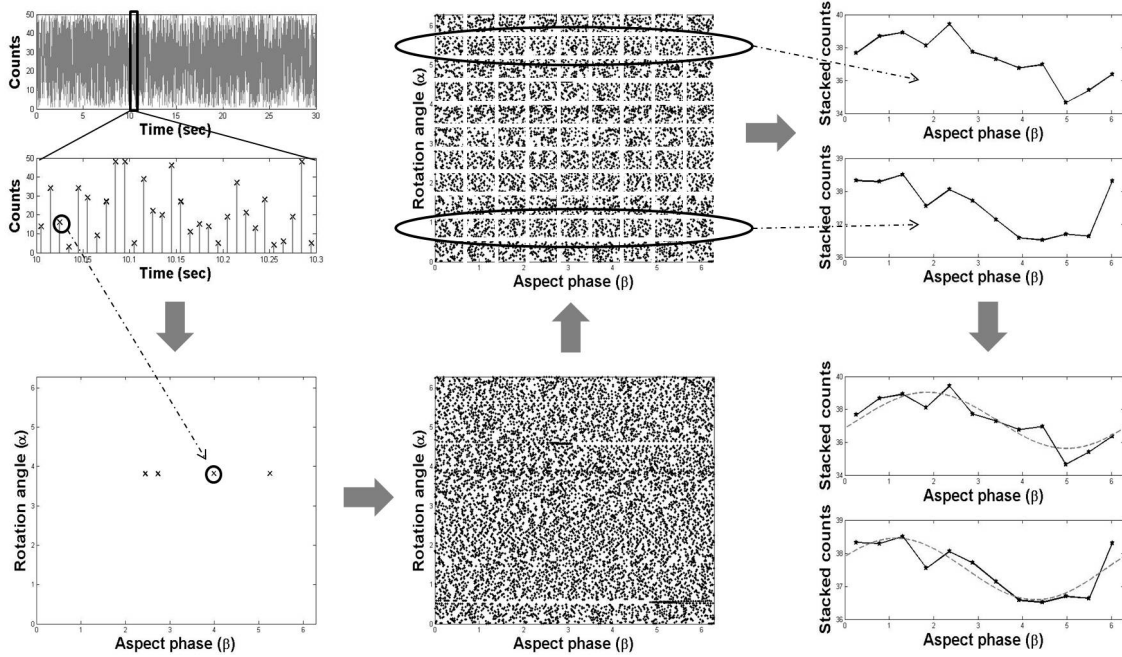
Figure 9. *Schematic representation of the data stacking process:* (a) *counts are discretized into time bins (top left panel);* (b) *an (α,β) map is populated (bottom left and middle panels);* (c) *the α and β ranges are discretized into bins (top middle panel);* (d) *for each bin* $[\overline{\alpha}_{k-1}, \overline{\alpha}_k]$ *(k = 1, ..., N), the stacked counts as a function of β are fitted with* (A.12) *(top and bottom right panels).*

in (A.6), then we have

$$C(t) = K\frac{c_0}{2} \int f(x,y)dxdy + K\sum_{j=1}^{N_H} c_j \mathcal{R}(\alpha^{(j)}(t)) \cos(j\beta(t))$$

(A.10)
$$+ K\sum_{j=1}^{N_H} c_j \mathcal{I}(\alpha^{(j)}(t)) \sin(j\beta(t)),$$

where $\mathcal{R}(\alpha^{(j)})$, $\mathcal{I}(\alpha^{(j)})$ are the real and imaginary parts of $V(u_{\alpha^{(j)}}, v_{\alpha^{(j)}})$ and

(A.11)
$$(u_{\alpha^{(j)}}, v_{\alpha^{(j)}}) = \left(\frac{j\cos(\alpha)}{p}, \frac{j\sin(\alpha)}{p}\right), \quad j = 1, \ldots, N_H.$$

Equation (A.10) shows the continuous relation linking the count profiles with the visibilities at specific spatial frequencies. Since in practice we have to deal with discrete quantities, the following procedure for translating counts into visibilities is adopted (see Figure 9):

1. Each count bin detected by RHESSI is labeled with the corresponding angle α and phase β. For a given time range, a plot of the counts as functions of α and β is generated.

2. The α and β ranges are discretized into two vectors

$$(\overline{\alpha}_0, \ldots, \overline{\alpha}_N), \quad (\overline{\beta}_0, \ldots, \overline{\beta}_M),$$

and a matrix is created by summing up all the counts belonging to the same cell generated by the discretization.

3. For each bin $[\overline{\alpha}_{k-1}, \overline{\alpha}_k]$ $(k = 1, \ldots, N)$, the corresponding row of the matrix is fitted with the function

$$(A.12) \qquad F + \sum_{j=1}^{N_H} R_j \cos(j\beta) + \sum_{j=1}^{N_H} I_j \sin(j\beta),$$

with $\beta = (\beta_1, \ldots, \beta_M)$, $\beta_i \in [\overline{\beta}_{i-1}, \overline{\beta}_i]$ $(i = 1, \ldots, M)$, and, since the parameters K, c_j are known $(j = 1, \ldots, N_H)$, from R_j, I_j and by means of (A.10) the value of $V(u_{\alpha_k}^{(j)}, v_{\alpha_k}^j)$ follows, where $\alpha_k \in [\overline{\alpha}_{k-1}, \overline{\alpha}_k]$.

For the specific case of the RHESSI data, the choice implemented in the SSW routine `hsi_visibility_fit.pro` is to set $N_H = 2$, although only the visibilities obtained through R_1, I_1 are provided to the user.

REFERENCES

[1] G. R. AYERS AND J. C. DAINTY, *Iterative blind deconvolution method and its application*, Opt. Lett., 13 (1988), pp. 547–549.
[2] J. BARZILAI AND J. M. BORWEIN, *Two point step size gradient methods*, IMA J. Numer. Anal., 8 (1988), pp. 141–148.
[3] F. BENVENUTO AND A. FERRARI, *Joint myopic deconvolution*, Inverse Problems, 26 (2010), 105011.
[4] M. BERTERO, D. BINDI, P. BOCCACCI, M. CATTANEO, C. EVA, AND V. LANZA, *A novel blind-deconvolution method with an application to seismology*, Inverse Problems, 14 (1998), pp. 815–833.
[5] M. BERTERO AND P. BOCCACCI, *Introduction to Inverse Problems in Imaging*, Institute of Physics, Bristol, UK, 1998.
[6] D. P. BERTSEKAS, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1999.
[7] S. BONETTINI, *Inexact block coordinate descent methods with application to the nonnegative matrix factorization*, IMA J. Numer. Anal., 37 (2011), pp. 1431–1452.
[8] S. BONETTINI AND M. PRATO, *Nonnegative image reconstruction from sparse Fourier data: A new deconvolution algorithm*, Inverse Problems, 26 (2010), 095001.
[9] S. BONETTINI, R. ZANELLA, AND L. ZANNI, *A scaled gradient projection method for constrained image deblurring*, Inverse Problems, 25 (2009), 015002.
[10] S.-C. BONG, J. LEE, D. E. GARY, AND H. S. YUN, *Spatio-spectral maximum entropy method. I. Formulation and test*, Astrophys. J., 636 (2006), pp. 1159–1165.
[11] R. BRACEWELL, *The Fourier Transform and Its Applications*, McGraw-Hill, New York, 2000.
[12] A. CICHOCKI, R. ZDUNEK, A. H. PHAN, AND S.-I. AMARI, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*, John Wiley & Sons, Chichester, UK, 2009.
[13] J.-M. CONAN, L. M. MUGNIER, T. FUSCO, V. MICHAU, AND G. ROUSSET, *Myopic deconvolution of adaptive optics images by use of object and point-spread function power spectra*, Appl. Opt., 37 (1998), pp. 4614–4622.
[14] J. C. CHRISTOU, D. BONNACINI, N. AGEORGES, AND F. MARCHIS, *Myopic deconvolution of adaptive optics images*, Messenger, 97 (1999), pp. 14–22.
[15] G. DESIDERÀ, B. ANCONELLI, M. BERTERO, P. BOCCACCI, AND M. CARBILLET, *Application of iterative blind deconvolution to the reconstruction of LBT LINC-NIRVANA images*, Astron. Astrophys., 452 (2006), pp. 727–734.
[16] D. A. FISH, A. M. BRINICOMBE, E. R. PIKE, AND J. G. WALKER, *Blind deconvolution by means of the Richardson–Lucy algorithm*, J. Opt. Soc. Amer. A, 12 (1995), pp. 58–65.
[17] G. FRASSOLDATI, L. ZANNI, AND G. ZANGHIRATI, *New adaptive stepsize selections in gradient methods*, J. Ind. Manag. Optim., 4 (2008), pp. 299–312.

[18] G. Golub and V. Pereyra, *Separable nonlinear least squares: The variable projection method and its applications*, Inverse Problems, 19 (2003), pp. R1–R26.

[19] L. Grippo and M. Sciandrone, *Globally convergent block-coordinate techniques for unconstrained optimization*, Optim. Methods Softw., 10 (1999), pp. 587–637.

[20] L. Grippo and M. Sciandrone, *On the convergence of the block nonlinear Gauss-Seidel method under convex constraints*, Oper. Res. Lett., 26 (2000), pp. 127–136.

[21] W. W. Hager, A. M. Bernard, and H. Zhang, *An affine-scaling interior-point CBB method for box-constrained optimization*, Math. Program., 119 (2009), pp. 1–32.

[22] P. C. Hansen, J. G. Nagy, and D. P. O'Leary, *Deblurring Images: Matrices, Spectra and Filtering*, SIAM, Philadelphia, 2006.

[23] G. J. Hurford, E. J. Schmahl, R. A. Schwartz, A. J. Conway, M. J. Aschwanden, A. Csillaghy, B. R. Dennis, et al., *The RHESSI imaging concept*, Solar Phys., 210 (2002), pp. 61–86.

[24] S. M. Jefferies and J. C. Christou, *Restoration of astronomical images by iterative blind deconvolution*, Astrophys. J., 415 (1993), pp. 862–874.

[25] E. P. Kontar, J. C. Brown, A. G. Emslie, W. Hajdas, G. D. Holman, G. J. Hurford, J. Kašparová, et al., *Deducing electron properties from hard X-ray observations*, Space Sci. Rev., 159 (2011), pp. 301–355.

[26] S. Krucker, M. Battaglia, P. J. Cargill, L. Fletcher, H. S. Hudson, A. L. MacKinnon, S. Masuda, et al., *Hard X-ray emission from the solar corona*, Astron. Astrophys. Rev., 16 (2008), pp. 155–208.

[27] D. Kundur and D. Hatzinakos, *Blind image deconvolution*, IEEE Signal Processing Mag., 13 (1996), pp. 43–64.

[28] D. Kundur and D. Hatzinakos, *Blind image deconvolution revisited*, IEEE Signal Processing Mag., 13 (1996), pp. 61–63.

[29] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, *Understanding and evaluating blind deconvolution algorithms*, in Proceedings of the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2009, pp. 1964–1971.

[30] C. J. Lin, *Projected gradient methods for nonnegative matrix factorization*, Neural Comput., 19 (2007), pp. 2756–2779.

[31] R. P. Lin, B. R. Dennis, G. J. Hurford, D. M. Smith, A. Zehnder, P. R. Harvey, D. W. Curtis, et al., *The Reuven Ramaty High-Energy Solar Spectroscopic Imager (RHESSI)*, Solar Phys., 210 (2002), pp. 3–32.

[32] I. Loris, M. Bertero, C. De Mol, R. Zanella, and L. Zanni, *Accelerating gradient projection methods for ℓ_1-constrained signal recovery by steplength selection rules*, Appl. Comput. Harmon. Anal., 27 (2009), pp. 247–254.

[33] A. M. Massone, A. G. Emslie, G. J. Hurford, M. Prato, E. P. Kontar, and M. Piana, *Hard X-ray images of solar flares using interpolated visibilities*, Astrophys. J., 703 (2009), pp. 2004–2016.

[34] J. D. Monnier, *Optical interferometry in astronomy*, Rep. Progr. Phys., 66 (2003), pp. 789–857.

[35] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed., Springer, New York, 2006.

[36] M. Piana, A. M. Massone, G. J. Hurford, M. Prato, A. G. Emslie, E. P. Kontar, and R. A. Schwartz, *Electron flux spectral imaging of solar flares through regularized analysis of hard X-ray source visibilities*, Astrophys. J., 665 (2007), pp. 846–855.

[37] M. Prato, *Regularization methods for the solution of inverse problems in solar X-ray and imaging spectroscopy*, Arch. Comput. Methods Engrg., 16 (2009), pp. 109–160.

[38] M. Prato, R. Cavicchioli, L. Zanni, P. Boccacci, and M. Bertero, *Efficient deconvolution methods for astronomical imaging: Algorithms and IDL-GPU codes*, Astron. Astrophys., 539 (2012), A133.

[39] M. Prato, M. Piana, A. G. Emslie, G. H. Hurford, E. P. Kontar, and A. M. Massone, *A regularized visibility-based approach to astronomical imaging spectroscopy*, SIAM J. Imaging Sci., 2 (2009), pp. 910–930.

[40] M. J. D. Powell, *On search directions for minimization algorithms*, Math. Programming, 4 (1973), pp. 193–201.

[41] J. L. Starck, E. Pantin, and F. Murtagh, *Deconvolution in astronomy: A review*, Publ. Astron. Soc. Pac., 114 (2002), pp. 1051–1069.

[42] T. Teng, J. Wu, and J. Chang, *Simulation research on the solar hard X-ray imaging telescope*, Chin. Astron. Astrophys., 36 (2012), pp. 198–210.

[43] Y. Xu, A. G. Emslie, and G. J. Hurford, *RHESSI hard X-ray imaging spectroscopy of extended sources and the physical properties of electron acceleration regions in solar flares*, Astrophys. J., 673 (2008), pp. 576–585.

[44] R. Zanella, P. Boccacci, L. Zanni, and M. Bertero, *Efficient gradient projection methods for edge-preserving removal of Poisson noise*, Inverse Problems, 25 (2009), 045010.