SIAM J. SCI. COMPUT.
Vol. 30, No. 3, pp. 1430–1458

IMPROVEMENT OF SPACE-INVARIANT IMAGE DEBLURRING BY PRECONDITIONED LANDWEBER ITERATIONS*

PAOLA BRIANZI†, FABIO DI BENEDETTO†, AND CLAUDIO ESTATICO‡

Abstract. The Landweber method is a simple and flexible iterative regularization algorithm, whose projected variant provides nonnegative image reconstructions. Since the method is usually very slow, we apply circulant preconditioners, exploiting the shift invariance of many deblurring problems, in order to accelerate the convergence. This way reasonable reconstructions can be obtained within a few iterations; the method becomes competitive and more robust than other approaches that, although faster, sometimes lead to lower accuracy. Some theoretical analysis of convergence is given, together with numerical validations.

Key words. Landweber, two-level Toeplitz and circulant matrices, preconditioning, regularization

AMS subject classifications. 65F22, 65F10, 45Q05, 15A18

DOI. 10.1137/050636024

1. Introduction. Image deblurring is the process of correcting degradations from a detected image. In the first analysis [4], the process of image formation is described by a Fredholm operator of the first kind; in many applications the blurring system is assumed to be space-invariant, so that the mathematical model is the following:

$$g(x,y) \;=\; \int_{\mathbb{R}^2} \mathcal{K}(x-\theta, y-\xi)\, f_*(\theta,\xi)\, d\theta d\xi \;+\; \omega(x,y), \tag{1.1}$$

where f_* is the (true) input object, $\mathcal{K}$ is the space-invariant integral kernel of the operator, also called the point spread function (PSF), ω is the noise which arises in the process, and g is the observed data.

The image restoration problem is the inversion of (1.1): Given the observed data g, we want to recover (an approximation of) the true data f_*. Its discrete version requires one to invert a linear system, typically of very large size and very sensitive to data error, due to the ill-posed nature of the continuous problem [16].

Space invariance leads to strong algebraic structures in the system matrix; depending on the boundary conditions enforced in the discretization, we find circulant, Toeplitz, or even more complicated structures related to sine/cosine transforms (see [33] for details). Exploiting structures in the inversion algorithm is necessary to face computational issues.

The problem of noise sensitivity is usually addressed by using regularization methods, where a suitable (sometimes more than one) parameter controls the degree of bias in the computed solutions. There are several techniques in literature (such as Tikhonov [16] or truncated SVD [22]), but in large-scale problems the main choice is given by *iterative regularization algorithms*, where the parameter is represented by the

*Received by the editors July 14, 2005; accepted for publication (in revised form) July 23, 2007; published electronically April 9, 2008. This work was partially supported by MIUR, grants 2002014121, 2004015437, and 2006017542.
http://www.siam.org/journals/sisc/30-3/63602.html
†Dipartimento di Matematica, Università di Genova, Via Dodecaneso 35, 16146 Genova, Italy (brianzi@dima.unige.it, dibenede@dima.unige.it).
‡Dipartimento di Matematica e Informatica, Università di Cagliari, Via Ospedale 72, 09124 Cagliari, Italy (estatico@unica.it).

number of iterations: The method works if an early stop prevents the reconstruction of noisy components in the approximated solution [14, 4].

The simplest iterative technique in this class is the Landweber method [27], proposed in 1951 but in agreement with an older work of Cimmino (see the historical overview in [2]); besides its easy implementation, this method presents very good regularization and robustness features. An early and very comprehensive analysis of the method is contained in [34]; a wide variety of even more recent applications can be found in [3, 6, 5, 29, 38]. In section 6 we report further details from the related literature.

In many real problems, the use of a priori information is basic for obtaining a substantial improvement in reconstructions; an important instance in imaging is taking into account nonnegativity constraints. In the past years more attention has been paid to faster Krylov methods such as conjugate gradient applied to normal equations (CGLS) [18] or GMRES [8], but, unfortunately, these methods do not provide nonnegative iterates, independently of the initial guess.

Although the recent literature has proposed specific approaches to enforce sign constraints (see, e.g., [21]), the Landweber method allows for a straightforward extension in order to do so, leading to the *projected Landweber method* [12]. On the other hand, its main disadvantage is that convergence may be very slow in practical applications (see the bibliographic notes in section 6).

In this paper we aim to overcome this disadvantage of the Landweber method by proposing an acceleration technique specifically designed for the space-invariant setting. Following a general idea introduced in [34, 31], we study the effect on this method of structure-based preconditioning techniques recently investigated for conjugate gradient iterations [20, 15]. We prove that such preconditioners can improve the convergence speed of the Landweber method (10 to 20 iterations are often sufficient to obtain a reasonable reconstruction), preserving its regularization capabilities; this way the method becomes more competitive with respect to other algorithms from a computational point of view. The same considerations could be extended to other preconditioning proposals [19, 24].

We stress that the removal of computational disadvantages allows us to emphasize the advantages of Landweber in comparison to other iterative methods:

- *simplicity* (we are able to give a formal proof of convergence and regularization behavior for the nonprojected version);
- *flexibility* (sign or even other constraints are easily incorporated; there are several parameters at our disposal, so a fine-tuning can be performed according to time/accuracy demands);
- *robustness* (little sensitivity to the inaccurate choice of parameters).

It is worth mentioning that the Landweber idea has applications in nonlinear inverse problems, too; in this context, it has been successfully applied to many real problems due to its robustness and strong regularization capabilities (see the survey in [13] and the references therein).

The paper is organized as follows. In section 2 we introduce the Landweber method, the circulant preconditioning technique (section 2.1), and a first convergence analysis in the simplest case of periodic boundary conditions (section 2.2). In section 3 we use this analysis to discuss the choice of parameters which define the preconditioned method. In section 4 we show how a convergence analysis can be performed without the simplifying assumptions made in [34], by developing the case study of Dirichlet boundary conditions. Numerical results are presented in section 5, and final remarks

are given in section 6. Technical details on convergence for suitable parameter values are given in the appendix.

2. Landweber method and preconditioning. The discretization of (1.1), with image size $n = (n_1, n_2)$, reads as $g = Af_* + \omega$, where g, f_*, ω represent the column-ordered vectors of the corresponding quantities and the matrix A discretizes the kernel $\mathcal{K}$ [4].

In order to enforce the same finite length $N = n_1 n_2$ to all of the vectors g, f_*, ω, appropriate "boundary conditions" must be applied; for an exhaustive survey of possible choices, see [33]. This way A is a square $N \times N$ matrix having a multilevel structure depending on the specific choice; for instance, A is a block Toeplitz matrix with Toeplitz blocks in the case of zero (Dirichlet) boundary conditions, and A is a block circulant matrix with circulant blocks if periodic boundary conditions are assumed.

In the discrete setting, given the blurred and noisy image g, we want to recover a suitable approximation f of the true image f_*, by computing a regularized solution of the least squares problem $\min \|Af - g\|_2$ [4]. Since the continuous problem is known to be ill-posed, the matrix A has ill-determined rank, since its smallest singular values accumulate to zero as N increases.

In this paper we deal with the *Landweber method* [27], which is the following iterative method for solving the normal equation $A^*Af = A^*g$. Let f_0 be an arbitrarily chosen initial guess; as we will see later, a recommended choice is $f_0 \equiv 0$. Compute, for $k = 0, 1, 2, \ldots$, the iterate

$$f_{k+1} = f_k + \tau(A^*g - A^*Af_k)\,, \tag{2.1}$$

where τ is a fixed value which should belong to $(0, 2/\|A^*A\|)$ in order to ensure the convergence along every direction.

The Landweber method (2.1) can be studied in several ways. It corresponds to the method of successive approximations for the computation of a fixed point of the operator $G(f) = f + \tau(A^*g - A^*Af)$. Moreover, it is the simplest method which returns the minimum point of the convex operator $H(f) = \frac{1}{2}\|Af - g\|_2^2$, since $A^*g - A^*Af_k = -\nabla H(f_k)$ is the steepest descent direction. By induction, it is simple to verify that

$$f_{k+1} = \tau \sum_{i=0}^{k} (I - \tau A^*A)^i A^*g + (I - \tau A^*A)^{k+1} f_0\,.$$

The Landweber algorithm belongs to the class of Krylov methods [4]. If we consider the fixed initial guess $f_0 \equiv 0$, the iteration can be written as $f_{k+1} = Q_{k,\tau}(A^*A)A^*g$, where $Q_{k,\tau}$ is the polynomial of degree k defined as $Q_{k,\tau}(t) = \tau P_k(\tau t)$, with

$$P_k(s) = \sum_{i=0}^{k} (1-s)^i = \sum_{i=0}^{k} \binom{k+1}{i+1} (-s)^i = \frac{1 - (1-s)^{k+1}}{s} \tag{2.2}$$

if $s \neq 0$, and $P_k(0) = k + 1$.

The method is linear since the polynomial $Q_{k,\tau}$ does not depend on g. We remark that, if $t \in (0, 2/\tau) \supsetneq (0, \|A^*A\|]$, then $Q_{k,\tau}(t) \longrightarrow t^{-1}$ $(k \longrightarrow +\infty)$, and, if $t \in [0, 2/\tau) \supsetneq [0, \|A^*A\|]$, then $|tQ_{k,\tau}(t)| = \left|1 - (1 - t\tau)^k\right| \leq 1$. These two latter properties of $Q_{k,\tau}$ state that the Landweber method is a continuous regularization algorithm, where the number of iterations k plays the role of regularization

parameter [14, Theorem 6.1]. Basically, the first iterations of the method filter out the components of data mainly corrupted by noise; hence, an early stop of the deblurring process improves the stability and gives a good noise filtering. Notice that $Q_{k,\tau}(t) \geq Q_{1,\tau}(\|A^*A\|) = \tau(2 - \tau\|A^*A\|) > 0$ for all $t \in [0, \|A^*A\|]$, which can be useful to improve the numerical stability.

A wide analysis of the convergence properties of the method was first carried out more than thirty years ago [34]. Indeed, by exploiting its simple formulation, the behavior of the iterations along each eigendirection can be successfully assessed.

Here, in order to briefly study the regularization properties of the algorithm, we analyze the convergence of $Q_{k,\tau}(t)$ in a (right) neighborhood of 0. The function $Q_{k,\tau}(t)$ is a polynomial of degree k such that

$$Q_{k,\tau}(0) = \tau(k+1)\,, \qquad Q'_{k,\tau}(0) = -(1/2)\tau^2(k+1)k\,. \tag{2.3}$$

By continuity arguments, the values of $Q_{k,\tau}(t)$ are bounded by $\tau(k+1)$ in a right neighborhood of 0.

The "level" of regularization of the kth iteration is summarized by the values $Q_{k,\tau}(0) = O(k)$ and $Q'_{k,\tau}(0) = O(k^2)$, with $Q'_{k,\tau}(0) < 0$ for $k \geq 1$. At the kth iteration, the approximation of the largest eigenvalues of the Moore–Penrose generalized inverse of A, that is, the approximation of the reciprocal of the smallest nonnull eigenvalues of A^*A, is bounded by $\tau(k+1)$ and decreasing at a rate $O(k^2)$. Furthermore, the kth iteration of the Landweber algorithm has basically the same regularization effects of the Tikhonov regularization method with regularization parameter $\alpha = (\tau(k+1))^{-1} > 0$, where $f_\alpha = (A^*A + \alpha I)^{-1}A^*g$ is the Tikhonov's α-regularized solution of the normal equations [14].

The Landweber method is a linear regularization algorithm, which allows a modification denoted as the *projected Landweber method*, which is very useful to solve inverse problems where some specific constraints on the solution play an important role. For example, the nonnegative constraint $f_* \geq 0$ is very common in image deblurring, whereas many classical regularization methods do not ensure any sign property for computed reconstructions.

The projected variant consists of the following simple modification of (2.1):

$$f_{k+1} = \mathcal{P}_+[f_k + \tau(A^*g - A^*Af_k)]\,, \tag{2.4}$$

where $\mathcal{P}_+$ is the projection onto the nonnegative cone. This leads to a nonlinear algorithm, for which the theoretical understanding is not complete [12]; it is proved that the iterates converge, for exact data, to a minimizer of $\|Af - g\|_2$ among nonnegative vectors.

Other important observed properties are just conjectures at this moment: We have numerical evidence of semiconvergence for noisy data, and the natural initial guess $f_0 = 0$ seems to provide the convergence of f_k to the least squares nonnegative solution having minimal norm [31].

Because of this absence of effective mathematical tools for investigating the convergence, the projected version of the method will not be considered in the following theoretical analysis. Besides, the numerical experiments of section 5 will concern both the projected and the nonprojected variants.

It is interesting to make a comparison with the widely used CGLS method [18]. The CGLS method is a nonlinear regularization algorithm, and its kth iteration is $f_{k+1} = P_{k,g}(A^*A)A^*g$, provided that $f_0 \equiv 0$ as before. Here $P_{k,g} = P_{k,g}(t)$ is a polynomial of degree k which depends on the input data g. The value $P_{k,g}(0)$, which

mainly controls the level of regularization at the kth iteration, is usually much greater than $k+1$ [18]. This implies that the CGLS method is faster than the Landweber one. This is confirmed by recalling that the CGLS method is an optimal Krylov method in the sense that the error at any iteration is minimized among all of the Krylov polynomials. On the other hand, in the absence of a reliable stopping rule, the fast convergence speed of the CGLS method may be a negative fact in image deblurring, since it can give rise to a fast amplification of the components of the restored image f_k which are related to the noise on input data [17].

We can summarize that the regularization of the Landweber method is high, whereas the convergence speed is low. In the following, our aim is to improve its convergence speed without losing its very favorable regularization capabilities.

2.1. Circulant regularizing preconditioners. As already noticed, the discrete image deblurring system $A^*Af = A^*g$ has ill-determined rank, since the continuous problem is ill-posed. The Landweber scheme is a suitable algorithm for solving that system, since it is a very effective regularization method. The negative fact is that the method is often quite slow; in several applications, such as astronomical image deblurring, thousands of iterations could be necessary.

Here we improve the convergence speed by means of preconditioning techniques. Basically, in order to speed up the convergence of any iterative method, a preconditioner is often an approximation of the, possibly generalized, inverse of the system matrix.

Following [34], the $N \times N$ linear system $A^*Af = A^*g$ is replaced by an algebraic equivalent system

$$DA^*Af = DA^*g\,, \tag{2.5}$$

where the $N \times N$ matrix D is the preconditioner which approximates the generalized inverse of A^*A. This way the preconditioned version of the method reads as follows:

$$f_{k+1} = \mathcal{P}_+[\tau DA^*g + (I - \tau M)f_k]\,, \tag{2.6}$$

where $M := DA^*A$ is the preconditioned matrix [31].

We stress that the least squares problem underlying (2.5) has been changed by inserting the preconditioner; therefore, the iteration (2.6) does not necessarily converge to the same limit of (2.4), with exact data. In the case of real data, as observed in [31, page 449], we cannot even expect in principle the same behavior of the nonpreconditioned method.

From now on we will consider for the theoretical discussion just the nonprojected variant, for which the operator $\mathcal{P}_+$ does not appear on the right-hand side of (2.6). In this case f_{k+1} linearly depends on f_k, whence we obtain the closed formula for the case $f_0 = 0$

$$f_k = G_kA^*g\ , \qquad G_k := \tau P_{k-1}(\tau M)D\ , \tag{2.7}$$

$P_k(t)$ being the polynomial introduced in (2.2). The new limitation for τ becomes $0 < \tau < 2/\|DA^*A\|$.

If B denotes an approximation of the matrix A, we construct the preconditioner D by computing $D = (B^*B)^\dagger$, where the symbol $\dagger$ denotes the Moore–Penrose generalized inverse.

Since for space-invariant problems A has a two-level Toeplitz-like structure, we look for B inside the matrix algebra $\mathcal{C}$ of block circulant matrices with circulant

blocks (BCCB). The BCCB matrices are very useful in the Toeplitz context since they provide fast diagonalization and matrix-vector multiplication within $O(N \log N)$ operations, via two-dimensional fast Fourier transform (FFT). From now on, we consider the T. Chan optimal approximation $B = B_{opt}$ of the system matrix A; that is, B_{opt} solves the following minimization problem [10]:

$$B_{opt} = \arg\min_{X \in \mathcal{C}} \|A - X\|_{\mathcal{F}} \ , \tag{2.8}$$

where $\|\cdot\|_{\mathcal{F}}$ is the Frobenius norm $\|G\|_{\mathcal{F}}^2 = \sum_{i,j} |(G)_{i,j}|^2$.

Since B_{opt} is the best approximation of A in the space $\mathcal{C}$ of the BCCB matrices, with respect to the Frobenius norm, it "inherits" the spectral distribution of A. This means that, if A has ill-determined rank, the same will hold for B_{opt}. The solution of the preconditioned system $DA^*Af = DA^*g$, with $D = (B_{opt}^* B_{opt})^\dagger$, leads to worse numerical results due to amplification of the components related to the noise of g. Differing from B_{opt}, any useful preconditioner for deblurring should approximate its system matrix only in the subspace less sensitive to data errors.

According to [20], this so-called *signal subspace* corresponds to the largest singular values of A, in the sense that it is spanned by the associated singular vectors. On the other hand, the *noise subspace* is related to the smallest singular values and represents the components where the direct reconstruction is more contaminated by data errors.

In the Toeplitz deblurring context, the problem of locating these two fundamental subspaces was first studied by Hanke, Nagy, and Plemmons [20]. Having fixed a small real value $\alpha > 0$ called the truncation parameter, the signal space can be roughly identified by the eigendirections corresponding to the eigenvalues of B_{opt} greater than α. If the parameter is well chosen, we can believe that the noise space falls into the directions related to the eigenvalues of B_{opt} with an absolute value smaller than α. Therefore the authors proposed in [20] to set equal to 1 all of these eigenvalues; in that way, the convergence speed increases in the signal space only, without fast amplification of the noisy components.

On these grounds, now we extend the approach of [20], by providing a family of different filtering procedures. Given a BCCB matrix G, let $\lambda_1(G), \lambda_2(G), \ldots, \lambda_N(G)$ denote its eigenvalues with respect to the fixed base of eigenvectors collected into the columns of the two-dimensional unitary Fourier matrix. If $\alpha > 0$ is a truncation parameter, we define the regularizing BCCB preconditioner $D = D_\alpha$ as the matrix whose eigenvalues $\lambda_1(D_\alpha), \lambda_2(D_\alpha), \ldots, \lambda_N(D_\alpha)$ are such that

$$\lambda_i(D_\alpha) = \mathcal{F}_\alpha\big(\lambda_i(B_{opt}^* B_{opt})\big) \, , \tag{2.9}$$

where the function $\mathcal{F}_\alpha : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is one of the eight filters of Table 2.1. Notice that the central column prescribes the eigenvalues along the noise space, and the last one is referred to the signal space.

The filter I comes from the Tikhonov regularization [4], and the filter III is that of Hanke, Nagy, and Plemmons [20]. The filter IV was introduced by Tyrtyshnikov, Yeremin, and Zamarashkin in [37], and VIII is the Showalter filter for asymptotic regularization [14]. For the polynomial filter V, we consider an integer $p > 0$.

It is worth noticing two important properties common to all of these filters:

- if the truncation parameter α is small enough, then $\lambda_i(D_\alpha)$ approximates the reciprocal of $\lambda_i(B_{opt}^* B_{opt})$ on the signal space;
- since $\mathcal{F}_\alpha$ is a bounded function in all eight cases, the eigenvalues of D_α have a uniform upper bound independent of the dimension.

TABLE 2.1
Regularizing functions for the optimal-based preconditioners.

$\mathcal{F}_\alpha(t)$	$0 \le t < \alpha$	$t \ge \alpha$
I	$(t+\alpha)^{-1}$	$(t+\alpha)^{-1}$
II	0	t^{-1}
III	1	t^{-1}
IV	α^{-1}	t^{-1}
V	$\alpha^{-(p+1)}t^p$	t^{-1}
VI	$\alpha^{-\frac{\alpha+1}{\alpha}}t^{\frac{1}{\alpha}}$	t^{-1}
VII	$\alpha^{-1}e^{\frac{t-\alpha}{\alpha t}}$	t^{-1}
VIII	$\int_0^{1/\alpha} e^{-ts}ds$	$\int_0^{1/\alpha} e^{-ts}ds$

Further properties of these filters, together with a wide numerical comparison, can be found in [15].

It is worth mentioning that the role played here by B_{opt} can be replaced by other popular basic preconditioners, such as those of Strang or R. Chan type; see [30]. We restrict our investigation to B_{opt} just because this was the original choice made in [20], but we expect similar results in the other cases.

2.2. A simplified convergence analysis: Periodic boundary conditions. The theoretical analysis of preconditioned iterations is in general a complicated task, due to the incomplete knowledge of how the singular vectors of the blurring matrix A and of the preconditioner D are related. We will try to give partial results in section 4, with special emphasis to the case where A is Toeplitz; by now we want to carry on just a preliminary analysis, in order to give useful insights about the role played by the different parameters at our disposal.

For this reason, in the present section we overcome the main difficulty by considering the following simplifying assumption, which will be removed later:

Both A^*A and D are circulant matrices.

It is understood that this assumption is restrictive, but it corresponds to the practical application of *periodic* boundary conditions in the discretization of the continuous space-invariant blurring model, as described in [4, 23, 33]. This is very common in some imaging contexts where the boundary effect is negligible, e.g., in astronomical observations of isolated sources.

With this choice, A can be regarded as the classical Strang [35] circulant approximation of the PSF matrix. In order to avoid a trivial situation, we assume that D is obtained from the T. Chan approximation of the Toeplitz pattern of the PSF, determined *before* applying boundary conditions; therefore B_{opt} is no longer related to the minimizer of (2.8), which in this case would reduce to A itself. As stated in the appendix (Lemma A.1), every $\lambda_i(B_{opt}^* B_{opt})$ considered in (2.9) is a good approximation of the corresponding eigenvalue of A^*A.

In light of the circulant structure, A^*A and D are diagonalized by a common eigenvector basis specifically represented by the discrete Fourier transform F (here we intend the two-dimensional transform, but we will use the same symbol independently of the number of dimensions):

$$A^*A = F\Lambda_A F^* , \qquad D = F\Lambda_D F^* ,$$

$$\Lambda_A = \mathrm{diag}(\lambda_1^A, \ldots, \lambda_N^A) , \qquad \Lambda_D = \mathrm{diag}(\lambda_1^D, \ldots, \lambda_N^D) .$$

In this case D satisfies the Strand condition for regularizing preconditioners (compare [34, eq. (34)]), and therefore the sequence f_k of preconditioned iterations preserves the semiconvergence property [31, Proposition 3.1].

Under this strong assumption, convergence analysis of the preconditioned Landweber method becomes straightforward; we will obtain similar results to those of [34] concerning the "response" of (continuous) iterations, but we will use the language of numerical linear algebra since we are directly interested in the discrete case.

Both the preconditioned matrix M and the iteration matrix G_k, such that $f_k = G_k A^* g$, introduced in (2.7), can be diagonalized in the same way as A^*A and D:

$$M = DA^*A = F\Lambda_M F^* , \qquad \Lambda_M = \Lambda_D \Lambda_A = \operatorname{diag}(\lambda_1^M, \dots, \lambda_N^M) , \qquad \lambda_j^M = \lambda_j^D \lambda_j^A ;$$

$$G_k = \tau P_{k-1}(\tau M) D = F \Lambda_{G_k} F^* , \qquad \Lambda_{G_k} = \tau P_{k-1}(\tau \Lambda_M)\Lambda_D = \operatorname{diag}(\lambda_1^{G_k}, \dots, \lambda_N^{G_k}) .$$

Recalling the closed formula (2.2) for $P_{k-1}(t)$, whenever $\lambda_j^M \neq 0$ we have

$$\lambda_j^{G_k} = \tau P_{k-1}(\tau \lambda_j^M)\lambda_j^D = \tau \, \frac{1 - (1 - \tau \lambda_j^M)^k}{\tau \lambda_j^M} \lambda_j^D = \frac{1 - (1 - \tau \lambda_j^D \lambda_j^A)^k}{\lambda_j^A} . \tag{2.10}$$

Now we are ready to study the components of the kth Landweber approximation f_k and of the generalized solution $f^\dagger$ along the Fourier directions $v_1, \dots, v_N$. On one hand,

$$f_k = G_k A^* g \implies F^* f_k = \Lambda_{G_k} F^* A^* g,$$

and therefore the jth Fourier component of f_k is

$$v_j^* f_k = \lambda_j^{G_k} \gamma_j , \qquad \text{where } \gamma_j = v_j^* A^* g,$$

in agreement with [34, eq. (59)]. On the other hand, we can use for $f^\dagger$ the expression

$$f^\dagger = (A^*A)^\dagger A^* g \implies F^* f^\dagger = \Lambda_A^\dagger F^* A^* g,$$

whence

$$v_j^* f^\dagger = \begin{cases} \gamma_j / \lambda_j^A & \text{if } \lambda_j^A > 0, \\ 0 & \text{if } \lambda_j^A = 0. \end{cases}$$

We obtain the following relation for the relative error on the components lying outside the null space of A:

$$\frac{|v_j^* f_k - v_j^* f^\dagger|}{|v_j^* f^\dagger|} = \frac{|\lambda_j^{G_k} - 1/\lambda_j^A|}{1/|\lambda_j^A|} = |1 - \tau \lambda_j^D \lambda_j^A|^k . \tag{2.11}$$

It is evident that the rate of decay for this error, as the iterations proceed, heavily depends on the value of the parameter τ and on the distance of the preconditioned eigenvalues $\lambda_j^D \lambda_j^A$ from 1. As described in the previous subsection, this distance is reduced just in the *signal space*, where λ_j^D approximates the reciprocal of the optimal circulant preconditioner eigenvalue (which in turn is a good approximation of λ_j^A).

It is worth noticing that (2.10) do not apply to the directions belonging to the *null space;* instead, we may deduce from (2.3) the relation $P_{k-1}(0) = k$, which implies

$$v_j^* f_k = \tau k \lambda_j^D \gamma_j , \tag{2.12}$$

so that this component is amplified as the iteration count k increases, unless λ_j^D is very small.

A similar behavior can be generally observed along the *noise space*, since under our assumptions these directions correspond to the indices such that $\lambda_j^A < \alpha$, the small threshold used to define the regularizing preconditioner. If this occurs, from the inequalities

$$1 - i\tau\lambda_j^D\alpha \leq (1 - \tau\lambda_j^D\lambda_j^A)^i \leq 1,$$

we obtain the bounds

$$P_{k-1}(\tau\lambda_j^M) = \sum_{i=0}^{k-1}(1 - \tau\lambda_j^D\lambda_j^A)^i \in \left[k\left(1 - \frac{k}{2}\tau\lambda_j^D\alpha\right)\ ,\ k\right];$$

substituting into the relation

$$v_j^* f_k = \lambda_j^{G_k}\gamma_j = \tau P_{k-1}(\tau\lambda_j^M)\lambda_j^D\gamma_j$$

gives

$$\tau k\lambda_j^D|\gamma_j|\left(1 - \frac{k}{2}\tau\lambda_j^D\alpha\right) \leq |v_j^* f_k| \leq \tau k\lambda_j^D|\gamma_j|\ . \tag{2.13}$$

The last relation also gives a dependence between the chosen threshold α and the iteration count k, since both of them affect the lower bound for the amplification of the undesired components.

The different role of the parameters k, τ, α and of the filter (which mainly involves the magnitude of λ_j^D in the noise space) is now made explicit in (2.11), (2.12), and (2.13); we are ready to present a full discussion in the next section.

3. Choice of parameters. The preconditioned Landweber method described so far involves four parameters to choose: the relaxation parameter τ, the regularization parameter k (iteration count), the filter parameter α, and finally the type of filtering function $\mathcal{F}_\alpha$.

As already shown in (2.11), the parameter τ allows us to control the convergence of the iterations towards the solution of the normal equations. More precisely, the relative reconstruction error is the largest one ($\approx 100\%$) along the components for which the value of $\lambda_j^D\lambda_j^A$ is close to 0 or $2\tau^{-1}$, whereas it is the smallest one ($\approx 0\%$) when $\lambda_j^D\lambda_j^A$ is close to τ^{-1}. This implies that the convergence is always slow in the noise space where $\lambda_j^D\lambda_j^A$ is small, but an appropriate choice of τ enables us to "select" the most important subspace of components to be first resolved in the reconstruction process.

We recall that, by using a suitable filtering preconditioner D as explained in section 2.1, the numbers $\lambda_j^D\lambda_j^A$ giving the spectrum of DA^*A can be made clustered at unity in the signal space and clustered at zero in the noise space; in that case, the simplest choice $\tau \equiv 1$ should provide good results in terms of convergence speed. It is worth noticing that for some choices of the filter (as those labeled by I and VIII in Table 2.1) the reciprocal function is only approximated, and therefore the preconditioned spectrum has a slightly different distribution; in these cases another value of τ could provide better results.

In any case, as we will see in section 4, the constant τ must be such that $\tau\lambda_j^M \in (0, 2)$, at least along the signal components. In the case of periodic boundary

conditions, all of the filters listed in Table 2.1 ensure that the choice $\tau = 1$ always verifies the constraint above for α not too large; a proof is given in the appendix. With other boundary conditions (for instance, Dirichlet conditions, which are used in the experiments of section 5), the constraint can be violated by few outliers, but numerical evidence suggests to us the conjecture that the corresponding eigendirections lie in the noise subspace, and in practice, since we often compute few iterations, these diverging components are still negligible with respect to all of the converging ones, especially for large-scaled systems arising in real applications.

Choosing the number of iterations k can heavily affect the performance when a method is not robust. If we underestimate k, we can provide restorations which are not sufficiently accurate because the "signal" component (2.11) of the relative error has not yet reached an acceptable value. If we overestimate k, we perform too many iterations, and this way we do not improve the efficiency of the method; moreover, we can obtain satisfactory restorations, provided that the unwanted components (related to k through the expressions (2.12) and (2.13)) have not been amplified too much.

On the other hand, by the numerical experiments performed in section 5 on simulated data (see, in particular, Figure 5.3), the restoration error is decreasing first and increasing afterwards preserving the property of semiconvergence [4], and, in general, the region of the minimum is large enough; this allows us to choose k in a wide range of values without consequences on the restoration. Therefore we propose to apply for general problems the most simple and efficient stopping rule for estimating the optimal value of k, that is, the *discrepancy principle*; if we have an estimate of the noise level $\epsilon = \|\omega\| = \|Af_* - g\|$, we stop the iterations when the residual $r_k = \|Af_k - g\|$ becomes less than ϵ.

In the case where such an estimate of the noise level ϵ is not at our disposal, [29] gives a stopping criterion connected to the general behavior of the residual r_k, which rapidly decreases in a few iterations and then decreases much more slowly; it looks "natural" to stop the iterations at the beginning of the flat region. The authors find also that the performance of the method does not change significantly for different values of the parameter k in this region (in their specific example, the range was $50 \leq k \leq 100$).

Concerning the choice of α, we recall here the "recipe" of [20]. We consider the discrete Fourier transform (DFT) of the right-hand side g, and we observe that there is an index r at which the Fourier coefficients begin to stagnate; this corresponds to the components where the random error starts to dominate the data vector g. Then we consider the eigenvalues λ_i of $B_{opt}^* B_{opt}$, obtained again by means of a DFT as discussed in section 2.1; we take as approximation of the filter parameter α the magnitude of the eigenvalue λ_r.

In many cases this strategy is difficult to apply, due to a great uncertainty in locating the index r; in a case of doubt, it is better to underestimate its value because this is equivalent to taking a greater value of α in order to prevent the reconstruction of noisy components.

The choice of the type of filtering function $\mathcal{F}_\alpha$ depends on the features of the problem and on the action we want to apply to the high frequency components in the reconstruction. The filters presented in Table 2.1 can be classified in two categories: "smooth" filters and "sharp" filters. For instance, the "smooth" filters I and VIII have the same expression for all t, whereas the "sharp" filtering functions from II to VII have some discontinuity in $t = \alpha$.

The smooth filters give an approximation of the reciprocal function everywhere,

so that the error in (2.11) is somehow reduced in the noise space, too; therefore they allow us to modulate the restoration from high frequencies according to the problem. On the other hand, the sharp filters do not try to invert on the noise space, and so they do not introduce further improvement in the solution; hence they produce a sequence that at first reduces the restoration error and after becomes stable, because for $t \leq \alpha$ the filter functions are slightly varying. This behavior is particularly desirable in the case of a strongly ill-conditioned problem, since it makes the method more robust with respect to a wrong choice of the "critical" parameters k and α.

As we can see in the numerical results of section 5, the nature of the problem may suggest that we use one type of filter instead of another.

4. Convergence analysis: The Toeplitz case. If the assumption of a common eigenvector basis for A and D is dropped, the classical Strand approach does not work, and convergence analysis of the preconditioned Landweber method becomes more involved, if we look at the Fourier components. An alternative way to quantify the acceleration and regularization behavior of iterations is to perform the analysis with respect to the eigenvector basis of the *preconditioned matrix* $M = DA^*A$; this choice needs no particular assumption on the structure of A^*A.

The argument is the classical one for stationary iterative methods; recalling that the iterates satisfy the recurrence relation

$$f_{k+1} = \tau DA^*g + (I - \tau M)f_k \ , \tag{4.1}$$

the generalized solution $f^\dagger$ can be expressed in a fixed-point form:

$$A^*Af^\dagger = A^*g \quad \Rightarrow \quad f^\dagger = \tau DA^*g + (I - \tau M)f^\dagger \ . \tag{4.2}$$

Subtracting (4.2) from (4.1), we obtain for the kth error a recurrence relation leading to the closed formula

$$f_k - f^\dagger = (I - \tau M)^k(f_0 - f^\dagger) = -(I - \tau M)^k f^\dagger \ , \tag{4.3}$$

having assumed the standard choice for the initial guess, that is, $f_0 = 0$.

Let $M = V_M \Lambda_M V_M^{-1}$ be the eigenvalue decomposition of the preconditioned matrix, where Λ_M is real nonnegative because M is symmetrizable (i.e., similar to a symmetric matrix) and semidefinite, but V_M may be not unitary. If we define ϕ_j^k and $\phi_j^\dagger$ as the jth components of f_k and $f^\dagger$, respectively, along the eigenvector basis V_M, by (4.3) we obtain

$$\phi_j^k - \phi_j^\dagger = -(1 - \tau\lambda_j^M)^k \phi_j^\dagger \ ,$$

λ_j^M being the generic eigenvalue of M; this kth componentwise error tends to zero provided that $\tau\lambda_j^M \in (0, 2)$. From this characterization we can draw the following general conclusions:

1. Along the directions where $\lambda_j^M \approx 1$, the relative error on the associated component of $f^\dagger$ decreases in magnitude with a linear rate close to $1 - \tau$; in the case $\tau = 1$, the reconstruction of such components is heavily accelerated.
2. Along the directions where λ_j^M is small enough, the associated component of f_k stays bounded as follows:

$$|\phi_j^k| = \left[1 - (1 - \tau\lambda_j^M)^k\right] |\phi_j^\dagger| = \tau\lambda_j^M \sum_{l=0}^{k-1} (1 - \tau\lambda_j^M)^l |\phi_j^\dagger| \leq \tau k \lambda_j^M |\phi_j^\dagger| \ .$$

Hence such components are poorly reconstructed, provided that the iterations are stopped early; in particular, f_k has no component along the null space of M.

On the other hand, the spectral analysis carried out in [20, Theorem 6.1] for the filter III (but extendable to several other choices) proves that the eigenvalues of M have a unit cluster and accumulate to zero, without giving any insight about the related eigenvectors; hence we are sure that statements in items 1 and 2 do not refer to the empty set.

Anyway, our conclusions are of no practical relevance unless we give an answer to some crucial questions:

- Are the directions considered in item 1 related to the *signal space* (here we desire a fast reconstruction)?
- Are we sure that the *noise space* (where reconstruction is not wanted) falls into the directions considered in item 2?

This delicate matter is crucial, as similarly pointed out in [26], and no exact knowledge is at our disposal. In space-invariant deblurring problems, it is known that signal and noise spaces can be described in terms of low and high frequencies (see, e.g., [20]), but no direct relation between the frequency-related Fourier basis and the eigenvectors of M is known in literature, except for the trivial case of common bases assumed in section 2.2.

In the case where the blurring operator A is a Toeplitz matrix (this occurs when Dirichlet boundary conditions are imposed to the PSF; see [4, 23, 33]), we take some insight from an indirect relation known as *equal distribution*.

DEFINITION 4.1 (see [39]). *The eigenvectors of the sequence $\{B_n\}$, where B_n is a $n \times n$ matrix, are* distributed like *the sequence of unitary vectors $\{q_k^{(n)}\}$, where $q_k^{(n)} \in \mathbb{C}^n$ for $k = 1, \ldots, n$, if the* discrepancies

$$r_k^{(n)} := \|B_n q_k^{(n)} - \langle q_k^{(n)}, B_n q_k^{(n)} \rangle q_k^{(n)}\|_2$$

are clustered around zero*, in the sense that*

$$\forall \epsilon > 0 : \quad \# \left\{ k \in \{1, \ldots, n\} : \ |r_k^{(n)}| > \epsilon \right\} = o(n) \ .$$

Since the discrepancies are a sort of measure of how much the $\{q_k^{(n)}\}$ behave like the eigenvectors of B_n, our goal is now to show that the eigenvectors of M are distributed like the Fourier vectors and therefore are frequency-related; this way we have a partial positive answer to the questions addressed above.

Clearly, in order to apply Definition 4.1 we should think of M as the element of a sequence of matrices indexed by their total size, which of course is the product of the individual ones along the different directions. Thus we are not sure that the sequence is well-defined for any positive integer n, and it is better to adopt the typical multi-index notation of *multilevel matrices* (see, e.g., [36]), with a slight change in the original Definition 4.1.

From now on, we put $n = (n_1, n_2) \in \mathbb{N}^2$ as the vector of individual dimensions, in the sense that M_n is a (block) two-level matrix of total size $N(n) := n_1 n_2$. By the multi-indices $i = (i_1, i_2)$ and $j = (j_1, j_2)$ we label a single entry of M_n, located at the inner i_2, j_2 position of the block having the indices i_1, j_1 at the outer level. The indices i_l, j_l at each level ($l = 1, 2$) range from 1 to the respective dimension n_l.

The same notation applies to the preconditioners D_n and to the rectangular Toeplitz blurring matrices $A_{m,n} \in \mathbb{R}^{N(m) \times N(n)}$, too. This way we may consider a

(double) sequence $\{M_n\}_{n\in\mathbb{N}^2}$ and adjust the concept of equal distribution to our two-level setting, according to the following new definition.

DEFINITION 4.2. *Consider a sequence $\{B_n\}_{n\in\mathbb{N}^2}$ of two-level matrices as described before and a sequence of unitary vectors*

$$\left\{q_k^{(n)}: \quad n=(n_1,n_2)\in\mathbb{N}^2,\ k\in\{1,\dots,n_1\}\times\{1,\dots,n_2\}\right\}\subseteq\mathbb{C}^{N(n)}\ .$$

We say that the eigenvectors of B_n are distributed like the sequence $\{q_k^{(n)}\}$ if the discrepancies $r_k^{(n)}$ as in Definition 4.1 satisfy

$$\forall\epsilon>0: \quad \#\left\{k\in\{1,\dots,n_1\}\times\{1,\dots,n_2\}:\ |r_k^{(n)}|>\epsilon\right\}=o(N(n))\ .$$

Our vectors $q_k^{(n)}$ will be the two-dimensional Fourier vectors, indexed in such a way that they correspond to the columns of the two-level Fourier matrix

$$F=\left(\frac{1}{\sqrt{n_1}}\exp\left\{\hat{\imath}\,\frac{2\pi(i_1-1)(j_1-1)}{n_1}\right\}\right)_{i_1,j_1=1}^{n_1}\otimes\left(\frac{1}{\sqrt{n_2}}\exp\left\{\hat{\imath}\,\frac{2\pi(i_2-1)(j_2-1)}{n_2}\right\}\right)_{i_2,j_2=1}^{n_2}\ , \tag{4.4}$$

where $\otimes$ denotes the Kronecker (tensor) product and $\hat{\imath}$ is the imaginary unit.

LEMMA 4.3. *For a sequence $\{T_n(a)\}_{n\in\mathbb{N}^2}$ of two-level Toeplitz matrices generated* [36] *by the L^2 bivariate function $a(x)$, the eigenvectors are distributed like the Fourier vectors, in the sense of Definition* 4.2.

Proof. The same statement has been proved in [39] for the one-level case; the proof was a direct consequence of the estimate

$$\|T_n(a)-C_n(a)\|_F^2=o(n)\ , \tag{4.5}$$

where $C_n(a)$ is the optimal circulant preconditioner of $T_n(a)$, denoted as B_{opt} in (2.8). The last result has been extended to the two-level version

$$\|T_n(a)-C_n(a)\|_F^2=o((N(n))$$

in the paper [36], so that the proof can be fully generalized by following the same steps as in [39], without any substantial change. ∎

The one-level equidistribution result has been further extended to L^1 generating functions in [40]; since the proof no longer uses (4.5), we cannot say that Lemma 4.3 holds if $a\in L^1$, too. The result is probably true but is not of interest in image deblurring, where generating functions are related to the PSFs and hence they have a high degree of regularity.

In order to extend the equidistribution result to our matrices $M_n=D_nA_{m,n}^*A_{m,n}$, we must overcome two difficulties: the presence of the first factor D_n (this will not be a great problem, since the Fourier vectors are exactly the eigenvectors of D_n) and the loss of Toeplitz structure in forming the normal equation matrix product $A_{m,n}^*A_{m,n}$. The next lemma takes care of this second topic.

LEMMA 4.4 (see [11]). *Let $\{A_{m,n}\}$ be a sequence of two-level Toeplitz matrices generated by a continuous bivariate function $a(x)$. Then for every $\epsilon>0$ there exist*

two sequences $\{R_n\}, \{E_n\}$ *of* $N(n) \times N(n)$ *matrices and a constant* s *such that, for* m *and* n *large enough,*

$$A^*_{m,n} A_{m,n} = T_n(f) + R_n + E_n \qquad \forall n, m \ ,$$

where $f(x) = |a(x)|^2$, $\|E_n\|_2 < \epsilon$, *and each* R_n *is a low-rank matrix having the following multilevel pattern:*

$$R_n = (R_{i_1,j_1})_{i_1,j_1=1}^{n_1}, \qquad R_{i_1,j_1} = (R_{i,j})_{i_2,j_2=1}^{n_2},$$

where each element $R_{i,j}$ *is nonzero only if all of the indices* i_1, i_2, j_1, j_2 *take the first or the last* s *values.*

The continuity assumption for $a(x)$ is stronger than the L^2 hypothesis used in Lemma 4.3 but again is not restrictive in the image deblurring context.

We end this section by proving the equal distribution for the eigenvectors of our preconditioned matrices with respect to Fourier vectors.

THEOREM 4.5. *Under the same assumptions as in Lemma* 4.4, *the eigenvectors of the matrices* M_n *are distributed like the Fourier vectors, in the sense of Definition* 4.2.

Proof. First we will show that the equidistribution property holds for normal system matrices $A^*_{m,n} A_{m,n}$, too. Notice that for all $B \in \mathbb{C}^{N \times N}$ and for all $q \in \mathbb{C}^N$, as observed, e.g., in [39],

$$\langle q, Bq \rangle = \arg\min_{\lambda \in \mathbb{C}} \|Bq - \lambda q\|_2 \ ,$$

so that in order to prove an equidistribution result for the sequence $\{B_n\}$ it suffices to show an $o(N(n))$ bound for vectors of the form $Bq_k^{(n)} - \lambda q_k^{(n)}$ for any suitable λ in place of the discrepancies $r_k^{(n)}$.

Now let $\lambda := \langle q_k^{(n)}, T_n(f) q_k^{(n)} \rangle$, where $q_k^{(n)}$ is a Fourier vector and f is given by Lemma 4.4. Then for every $\epsilon > 0$ and m, n large enough

$$\begin{aligned} \|A^*_{m,n} A_{m,n} q_k^{(n)} - \lambda q_k^{(n)}\|_2 &= \|T_n(f) q_k^{(n)} + R_n q_k^{(n)} + E_n q_k^{(n)} - \lambda q_k^{(n)}\|_2 \\ &\le r_k^{(n)} + \|R_n q_k^{(n)}\|_2 + \|E_n\|_2 \|q_k^{(n)}\|_2 < \|R_n q_k^{(n)}\|_2 + 2\epsilon \end{aligned} \tag{4.6}$$

except for $o(N(n))$ multi-indices k, where we have denoted by $r_k^{(n)}$ the discrepancy referred to as the Toeplitz matrix $T_n(f)$ and we have applied Lemmas 4.3 and 4.4.

In order to manage the product $R_n q_k^{(n)}$, we must take into account the sparsity pattern of R_n given by Lemma 4.4 and the following general expression for the $j = (j_1, j_2)$th entry of the Fourier vector $q_k^{(n)}$, deduced by (4.4):

$$(q_k^{(n)})_j = \frac{1}{\sqrt{N(n)}} \exp\left\{ 2\pi \hat{\imath} \left[\frac{(j_1 - 1)(k_1 - 1)}{n_1} + \frac{(j_2 - 1)(k_2 - 1)}{n_2} \right] \right\} \ . \tag{4.7}$$

The structure of R_n implies that the product with the Fourier vector involves just the first and the last s values of the indices j_1 and j_2; moreover, only such entries of the product are nonzero. More precisely, by exploiting the usual multi-index notation we obtain

$$\|R_n q_k^{(n)}\|_2^2 = \sum_{i \in I_1 \times I_2} \left| \sum_{j \in I_1 \times I_2} R_{i,j} (q_k^{(n)})_j \right|^2 \ ,$$

where $I_l = \{1, \dots, s\} \cup \{n_l - s + 1, \dots, n_l\}$, $l = 1, 2$. Setting $c := \max_{i,j} |R_{i,j}|$ and using the expression (4.7), we have the bound

$$\left| \sum_{j \in I_1 \times I_2} R_{i,j} (q_k^{(n)})_j \right| \leq c \sum_{j \in I_1 \times I_2} |(q_k^{(n)})_j| = \frac{c}{\sqrt{N(n)}} \#(I_1 \times I_2) = \frac{4cs^2}{\sqrt{N(n)}} ,$$

whence

$$\|R_n q_k^{(n)}\|_2^2 \leq \#(I_1 \times I_2) \frac{16c^2 s^4}{N(n)} = \frac{64c^2 s^6}{N(n)} ,$$

which can be made less than ϵ provided that $N(n)$ is large enough. Substituting into (4.6) and observing that ϵ is arbitrarily small, we have proved the equal distribution between the eigenvectors of $A_{m,n}^* A_{m,n}$ and the Fourier vectors.

Concerning the preconditioned matrices, it suffices to observe that for $\mu := \lambda \cdot \lambda_k^D$

$$\|M_n q_k^{(n)} - \mu q_k^{(n)}\|_2 = \|D_n [A_{m,n}^* A_{m,n} q_k^{(n)} - \lambda q_k^{(n)}]\|_2 \leq \|D_n\|_2 \, \epsilon \tag{4.8}$$

except for $o(N(n))$ values of k, where we have used the property that Fourier vectors are also eigenvectors of D_n. Since $\|D_n\|_2$ equals the maximal eigenvalue of D_n, having a uniform upper bound for regularizing preconditioners (see section 2.1), the inequality (4.8) proves the equal distribution result for the matrices $\{M_n\}$. □

5. Numerical results. In this section we provide some numerical experiments illustrating the effectiveness of the preconditioned Landweber method, both in the basic version and in the projected variant. In particular, the deblurring capabilities of the method will be tested through both synthetic and widely used experimental data. The analysis of section 3 is a useful starting point for the appropriate choice of all of the parameters of the algorithm. Along this direction, the main aim of the present section is to compare the results related to different settings. As already shown (see section 2), our discrete model of image formation is the image blurring with shift invariance, and it can be simply written as follows:

$$g = A f_* + \omega, \tag{5.1}$$

where g is the blurred data, A is the multilevel Toeplitz matrix version of the PSF associated to Dirichlet boundary conditions, f_* is the input true object, and ω is the noise which arises in the process. The deblurring problem is to recover a suitable approximation f of the true object f_*, by means of the knowledge of g, A, and some statistical information about the noise ω.

We consider two different test examples.

- T1. In the first one, the true image f_* is the 256×256 brain section of Figure 5.1 (top left), and the convolution operator A is a Gaussian PSF with standard deviation $\sigma = 5$ shown in Figure 5.1 (top right), with an estimated condition number of $1.2 \cdot 10^{20}$. In this first example, we compare the restorations corresponding to different levels of artificial Gaussian white noise ω, where the relative data error $\|\omega\|/\|Af_*\|$ ranges from 3% to 30%, with $\|\cdot\|$ being the vector 2-norm.
- T2. In the second example, the object f_* to be recovered is the 256×256 satellite image of Figure 5.2 (top left), while the blurring operator and the blurred data are experimental and plotted in Figure 5.2 (top right, bottom). These

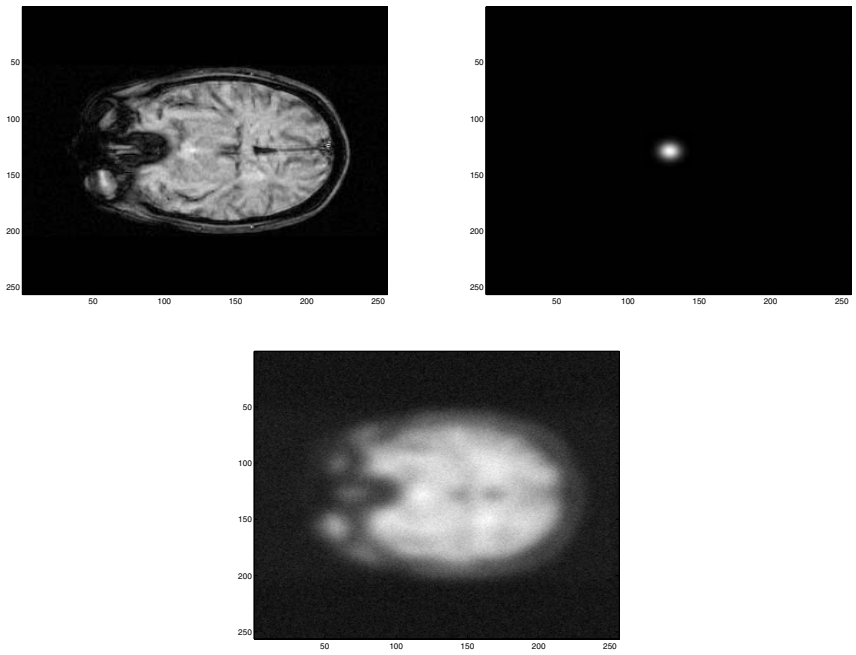

FIG. 5.1. *Test set* 1—*True object, PSF, and synthetic blurred image with relative noise* $\approx 8\%$.

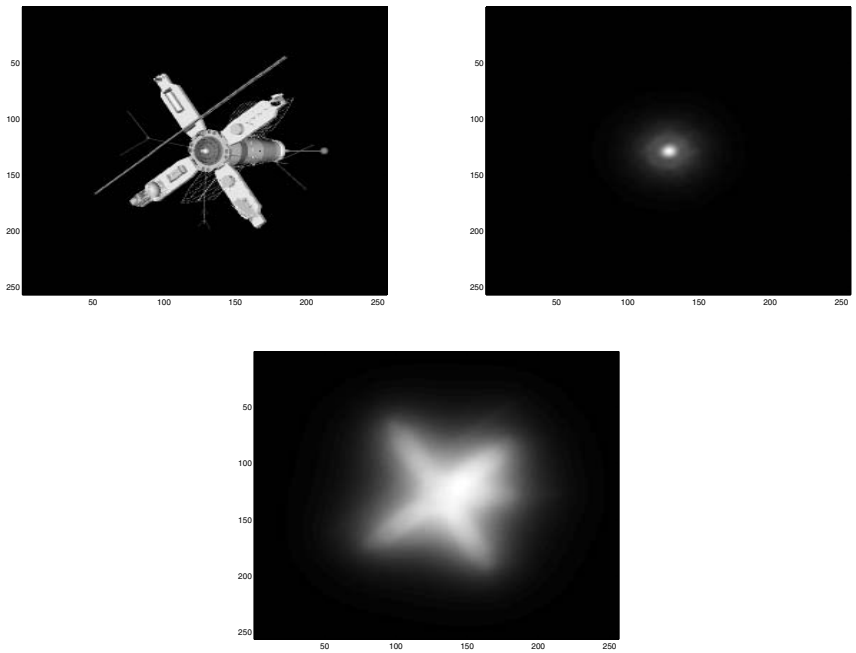

FIG. 5.2. *Test set* 2—*True data, experimental PSF, and blurred image.*

test data have been developed by the U.S. Air Force Phillips Laboratory, Lasers and Imaging Directorate, Kirtland Air Force Base, New Mexico, and they are widely used in the literature [32]. This time A is just moderately ill-conditioned ($\text{cond}(A) \approx 1.3 \cdot 10^6$), but the data image g is corrupted by a noise of unknown distribution, corresponding to $\|\omega\|/\|Af_*\| \approx 5\%$.

In addition, on the same true object f_* and blurring operator A, we consider another corrupted data g, developed by Bardsley and Nagy [1] and related to a statistical model where the noise comes from both Poisson and Gaussian distributions, plus a background contribution, so that the global noise level is $\|\omega\|/\|Af_*\| \approx 1\%$. These two sets of data allow us to numerically compare the features of the method with other techniques proposed in [1, 19, 21, 24, 28].

According to Table 2.1 of section 2.1, we test the following four filters $\mathcal{F}_\alpha$:

F1. $\mathcal{F}_\alpha$ is the Tikhonov filter I;

F2. $\mathcal{F}_\alpha$ is the low pass filter II;

F3. $\mathcal{F}_\alpha$ is the Hanke, Nagy, and Plemmons filter III;

F4. $\mathcal{F}_\alpha$ is the p-polynomial vanishing filter V, with $p = 1$.

The convergence parameter τ of the method is set to 1, providing a good compromise between fast convergence and noise filtering for the preconditioned Landweber method, as discussed in section 3.

We stress that the choice of the regularization parameter α is not a simple task, so we attempt several values in order to select the best one. In test T2 we try to adopt the strategy proposed by Hanke, Nagy, and Plemmons [20] recalled in section 3, which is based on an appropriate estimate of the noise space, and we suggest some improvements and remarks.

The number of iterations can be fairly well controlled by means of the discrepancy principle, even though it can be underestimated with respect to the optimal one; therefore we prefer to present the best achievable results within the first 200 iterations.

All of the experiments have been implemented in MATLAB 6.1 and performed on an IBM PC, with a floating-point precision of 10^{-16}.

5.1. Test 1 (synthetic data with different levels of noise). We test the projected variant of the Landweber method, where each iteration f_k is projected onto the nonnegative cone (see end of section 2), using the four filters for α ranging from 0.005 to 0.1, and we take the best restoration among the first 200 iterations. Table 5.1

TABLE 5.1
Test 1—Best relative restoration errors and number of iterations for 3% and 8% noise.

Data	3% relative noise				8% relative noise			
No prec.	0.1794 200				0.1855 200			
α	F1	F2	F3	F4	F1	F2	F3	F4
0.005	0.2028 1	0.1900 3	0.1898 3	0.1908 2	0.2107 1	0.1968 2	0.1967 2	0.2031 2
0.01	0.1890 2	0.1979 7	0.1969 7	0.1849 4	0.2016 2	0.2024 3	0.2018 3	0.1930 3
0.02	0.1854 3	0.2012 199	0.1847 199	0.1862 13	0.1925 2	0.2074 19	0.1967 139	0.1946 7
0.03	0.1823 7	0.2055 200	0.1798 200	0.1844 43	0.1907 5	0.2090 179	0.1886 200	0.1949 15
0.05	0.1782 16	0.2172 123	0.1796 200	0.1801 200	0.1896 9	0.2188 200	0.1867 200	0.1914 116
0.1	0.1752 49	0.2325 200	0.1794 200	0.1861 200	0.1881 20	0.2335 200	0.1858 200	0.1910 200

shows the values of the minimal relative restoration error (RRE) $\|f_k - f_*\|/\|f_*\|$ and the corresponding iteration number k. The left side of the table reports the results with 3% of relative data error, the right side with 8%.

First of all, we point out that the nonpreconditioned Landweber method is really slower than the preconditioned version. After 200 iterations, the RRE without preconditioning gives the same accuracy provided by the Tikhonov filter F1 within at most 20 iterations (see the columns F1 of Table 5.1 relative to the values $\alpha = 0.05$ for 3% of noise and $\alpha = 0.1$ for 8% of noise).

As expected, when the noise is higher the filtering parameter α has to be larger, and the method has to be stopped earlier. Indeed, both α and k together play the role of regularization parameter, as can be observed by comparing the results on the left and right sides of the table. In particular, a good choice of α should be the compromise between noise filtering and fast convergence.

Small values of α yield low noise filtering; in the first rows of the table, the noise dominates on the restoration process, and the RREs are larger. If this is the case, it is interesting to notice that the filter F4 outperforms the others, because the problem stays ill-conditioned and F4 "cuts" very much on the noisy components. On the other hand, too large values of α do not speed up the convergence, as shown by the latter rows, especially for filters F3 and F4.

The best restorations for this example are given by the Tikhonov filter F1, provided that a good value of α has been chosen. The corresponding convergence histories, that is, the values of all RREs versus the first 200 iterations, are shown in Figure 5.3. In this graph, it is quite evident that small filtering parameters α provide fast convergence and low accuracy, while large filtering parameters α provide high accuracy and low convergence speed. The graph of Figure 5.3 confirms how the improvement provided by the preconditioned Landweber method is high with respect to the nonpreconditioned case. Moreover, the shape of any semiconvergence graph is very regular, without oscillations and fast changes of the concavity, even for low values of the filtering parameter α (see also the analogous Figures 5.7 and 5.9 related to Test 2). This good behavior is not common in other preconditioned iterative strategies which, although faster, may give much lower stability (see, e.g., [28, Figure 3.2]).

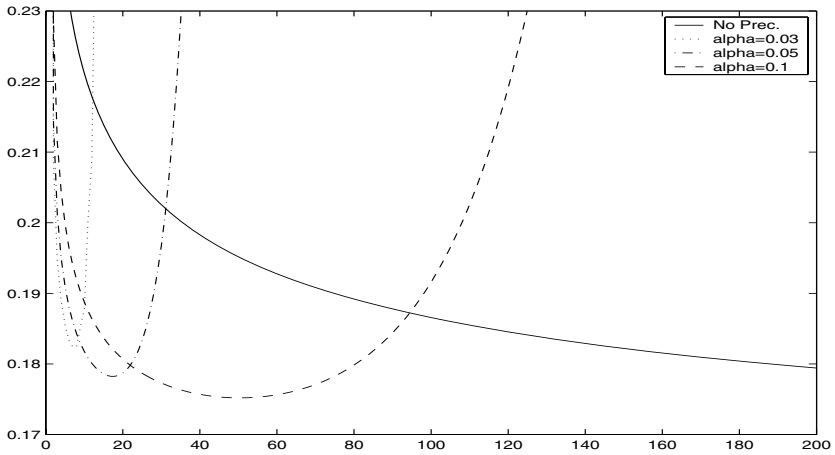

FIG. 5.3. *Test* 1*—RREs versus the first* 200 *iterations; noise* 3%, *filter* $F1$.

We remark that the low pass filter F2 always gives worse results. The reason is that this filter neglects all of the eigenvalues lower than α; in this case, the corresponding components are completely lost, and the reconstruction is done without those pieces of information. All of the other filters take into account such components, although the restoration is slow therein, and hence the results are more accurate.

A comparison between projected and nonprojected versions is shown in Table 5.2, related to filters F1 and F4. The left side concerns the Landweber method with projection on positives, while the right side concerns the method without projection. Notice that the positivity constraint improves the results, since actually the projected variant always leads to the minimal RREs and the maximal convergence speed. We can say that the projection on positives acts as a regularizer, since it reduces the instability due to the noise on the data. In addition, the higher regularizing properties of the projected variant allow us to adopt smaller values of the regularization parameter α, which accelerates the method.

As graphical examples of restorations, the two images on the top of Figure 5.4 are the best restorations after 200 iterations of the nonpreconditioned Landweber method, with projection onto positive values (left) and without projection (right), for input data with 8% of relative noise. The images in the central row are the best restorations with filter F1 and $\alpha = 0.03$: 5 iterations with projection on positives on the left and 3 iterations without projection on the right. The images at the bottom are the best restorations with filter F4, again with $\alpha = 0.03$: 15 iterations with projection on positives on the left and 32 iterations without projection on the right. The images of the projected version on the left are quite better than the images of the classical Landweber method on the right, since the ringing effects are sensibly reduced, especially on the background of the object. It is worth noticing that the corresponding numerical values of the RREs in Table 5.2 are not so different and do not allow one to completely evaluate these qualitative improvements.

TABLE 5.2
Test 1*—Best relative restoration errors and number of iterations for filters F*1 *and F*4.

	With positivity			Without positivity		
Noise	3%	8%	30%	3%	8%	30%
No prec.	0.1794	0.1855	0.2185	0.1817	0.1928	0.2263
	200	200	29	200	156	15

Filter F1	With positivity			Without positivity		
Noise	3%	8%	30%	3%	8%	30%
$\alpha = 0.01$	0.1890	0.2016	0.1866	0.2091	0.2142	0.2083
	2	2	2	1	1	1
$\alpha = 0.03$	0.1823	0.1907	0.2276	0.1906	0.1977	0.2361
	7	5	1	4	3	1
$\alpha = 0.1$	0.1752	0.1881	0.2219	0.1763	0.1929	0.2268
	49	20	3	69	16	2

Filter F4	With positivity			Without positivity		
Noise	3%	8%	30%	3%	8%	30%
$\alpha = 0.01$	0.1849	0.1930	0.2642	0.1891	0.1964	0.2886
	4	3	1	3	2	1
$\alpha = 0.03$	0.1844	0.1949	0.2301	0.1851	0.1949	0.2395
	43	15	2	48	32	2
$\alpha = 0.1$	0.1861	0.1910	0.2240	0.1895	0.1956	0.2293
	200	200	8	200	200	5

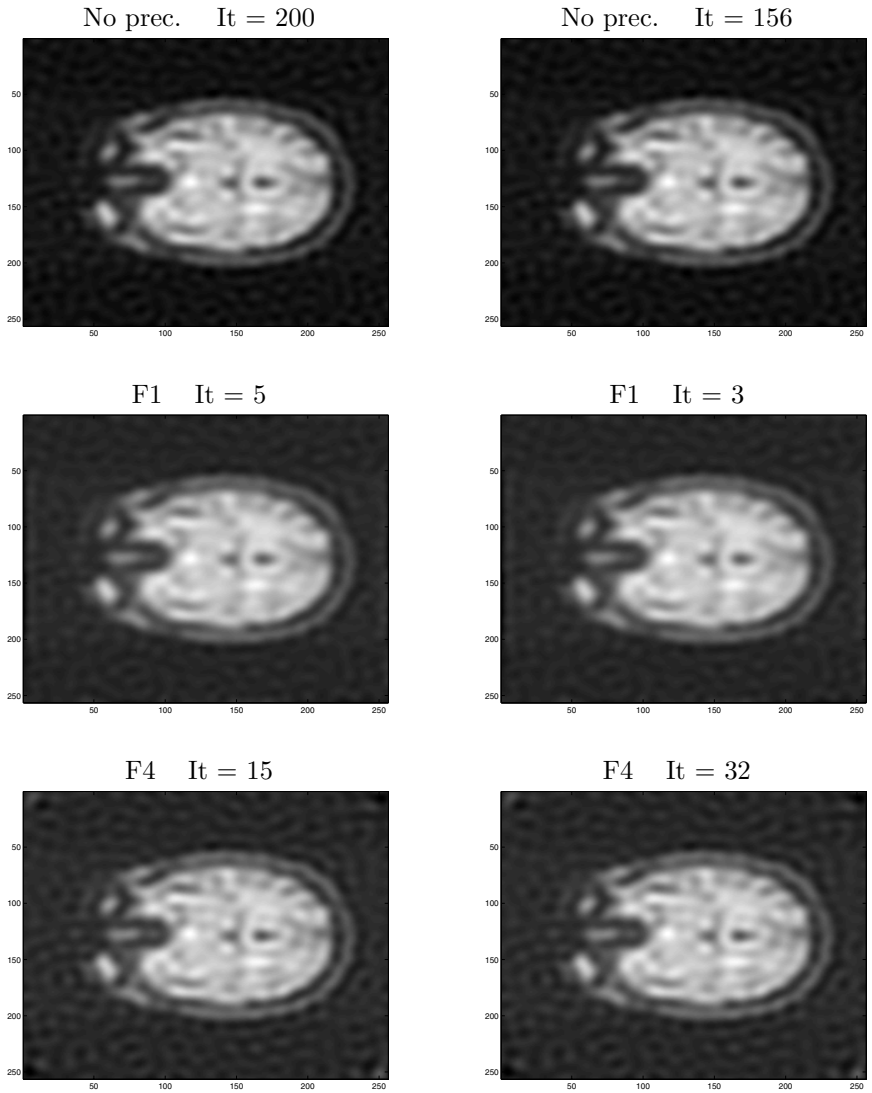

FIG. 5.4. *Test* 1*—Best reconstructions with* $\alpha = 0.03$*, for relative noise* 8%*, within* 200 *iterations. Left: With projection on positives. Right: Without projection.*

5.2. Test 2 (experimental data). The experimental data of the blurred image with about 5% of noise are now used here (see Figure 5.2). Table 5.3 shows the best RREs $\|f_k - f_*\|/\|f_*\|$ and the corresponding iterations k, obtained by using the four preconditioners and several thresholds α with the projected variant of the Landweber method.

In this second test, we try to adopt the strategy proposed by Hanke, Nagy, and Plemmons for the choice of the regularization parameter α. Basically, the Fourier spectral components of the ("unregularized") optimal preconditioner B^*B are compared with the Fourier spectral components of the blurred and noisy image g, in order to estimate the components of g where the noise dominates on the signal. We summarize the procedure by the following three steps: (i) collect in decreasing order the

TABLE 5.3
Test 2—Best relative restoration errors and number of iterations.

No prec.	0.4689 200			
α	F1	F2	F3	F4
0.01	0.5176 2	0.4837 200	0.4351 200	0.5347 3
0.02	0.4924 5	0.5449 200	0.4568 200	0.5132 17
0.03	0.4249 19	0.5581 200	0.4606 200	0.4146 200
0.04	0.3929 37	0.5752 200	0.4630 200	0.4340 200
0.05	0.3758 62	0.5911 200	0.4642 200	0.4496 200
0.06	0.3623 99	0.6171 200	0.4677 200	0.4626 200

eigenvalues (computed by two-dimensional FFT) of the circulant optimal preconditioner B^*B; (ii) look at the Fourier components of the blurred image g with respect to the same ordering of step (i), and choose the first index, say r, such that all of the following harmonics have small widths of approximately constant size; (iii) $\lambda_r(B^*B)$ is the truncation parameter α for the filtered optimal preconditioner D_α.

A graphic example of the three steps is given in Figure 5.5. The graph on the top shows all of the 256^2 Fourier components, and the graph on the bottom is the corresponding zoom onto the first 500. As already mentioned, the procedure is not simple to apply. Indeed, there is a large set of "reasonable" indices where the Fourier components of g start to be approximately constant, which gives rise to very different truncation parameters α. In the figure, the value $\alpha \approx 10^{-3}$ seems to be a good "stagnating" value, corresponding to a noise space generated by the Fourier components of index greater than about 350. Unfortunately, this choice gives a too small parameter, as can be observed in Table 5.3, and the results are unsatisfactory due to the high contribution of the noise in the restoration process. In our test, good values of α are contained between $0.4 \cdot 10^{-1}$ and 10^{-1}, corresponding to a noise space generated by the Fourier components more or less after the first 30–70. Hence, in order to avoid noise amplification, it seems to be better to overestimate the noise space by choosing a truncation parameter α higher than the one located by steps (i)–(iii). In this way, although the convergence may be slower, the restoration is generally good.

Since it is easy to miss the proper value of α, it is worth noticing that sharp filters are not appropriate for such situations, due to their discontinuity with respect to this parameter. This is confirmed by the column F4 of Table 5.3: There is a remarkable gap in the results between the second and the third rows (in particular, a dramatic change occurs from 17 iterations for $\alpha = 0.02$ to 200 iterations for $\alpha = 0.03$). Similarly, other experiments not reported in the table show very irregular performances for the filter F3 when α goes from 0.0087 (RRE of 0.5734 in 3 iterations) to 0.01 (RRE of 0.4351 in 200 iterations).

Concerning the quality of the restorations, we can say that after 200 iterations the RRE without preconditioning is still very high (RRE=0.4689), and the reconstruction is unsatisfactory, as shown on the top left of Figure 5.6. A similar accuracy is provided within about 10 iterations by the Tikhonov filter, with $\alpha = 0.03$, for instance. The Tikhonov filter F1 gives again the best results, and for this filter the choice of the regularization parameter is simpler since it is not very sensitive with respect to

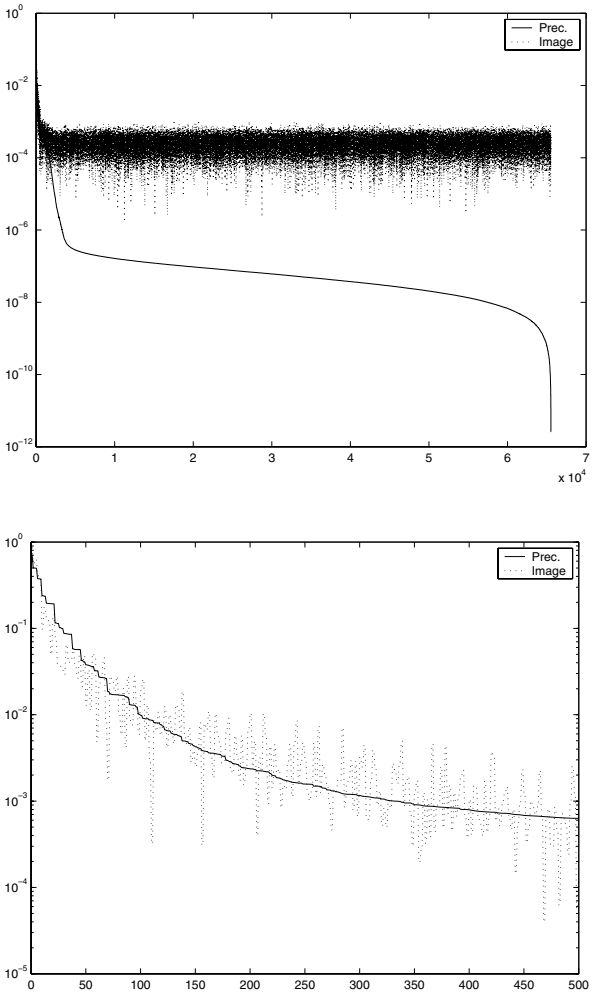

FIG. 5.5. *Test 2—Choice of the regularization parameter α.*

nonoptimal choices. The other choices, especially F2 and F3, do not provide good results, since they bring too much regularization, that is, the convergence speed is too slow along the components related to the smallest eigenvalues.

Notice that the restorations with F1 are good for any α between 0.03 and 0.06. Within few iterations, we obtain very fast and sufficiently good restorations with the low values $\alpha = 0.03$ and $\alpha = 0.04$ (see the corresponding images of Figure 5.6). On the other hand, if we are more interested in low RREs than fast computation, we can adopt a larger α. For instance, by using $\alpha = 0.06$, within about 100 iterations the details of the restored image are accurate, as shown by the image on the center right of Figure 5.6.

The results of this numerical test could be directly compared with those of other solving techniques used in the literature on the same input data set (see, e.g., [19, 21, 24, 28]). For instance, we consider the (preconditioned) conjugate gradient and the (preconditioned) steepest descent methods applied to the normal equations (CGLS/PCGLS and RNSD/PRNSD), as developed by Lee and Nagy [28].

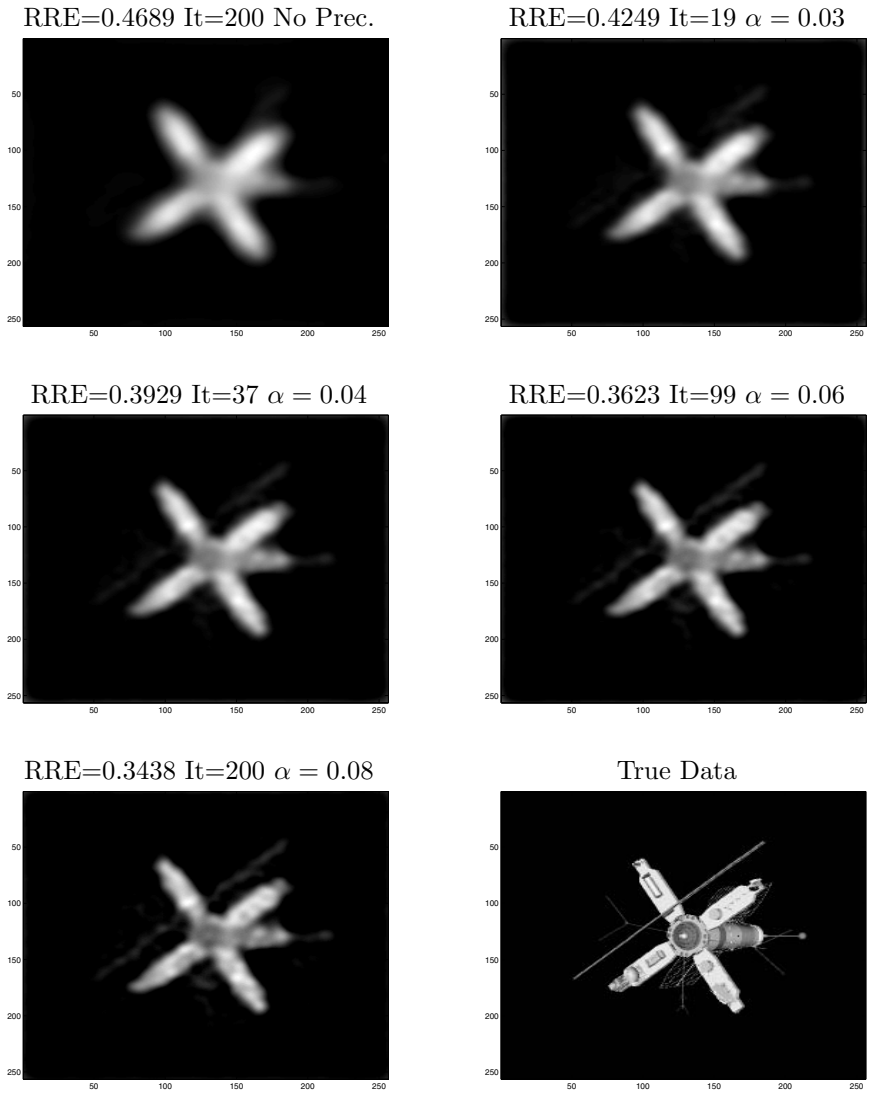

FIG. 5.6. *Test* 2*—Restorations of the Landweber method with positivity, filter* $F1$.

The preconditioner used in [28] corresponds to our BCCB preconditioner with filter F3 and regularization parameter α chosen by a cross-validation technique, which works very well in this case. We recall that the RNSD method is similar to the Landweber one (2.1), since the two methods differ only in the choice of the convergence parameter τ: It changes with any iteration in the former, so that $f_{k+1} = \arg\min_{f=f_k+\tau A^*(g-Af_k)} \|A^*(g-Af)\|_2$, whereas it is simply constant in the latter.

The convergence history of Figure 5.7 can be directly compared with [28, Figure 4.2]. By a direct comparison of the RRE values, we can say that the preconditioned Landweber method is approximately fast as the CGLS method without preconditioning acceleration. In addition, as expected by virtue of its better choice of τ, the RNSD method is faster than the (unpreconditioned) Landweber one (for instance, RNSD reaches after 50 iterations the RRE value of about 0.50 instead of our value of

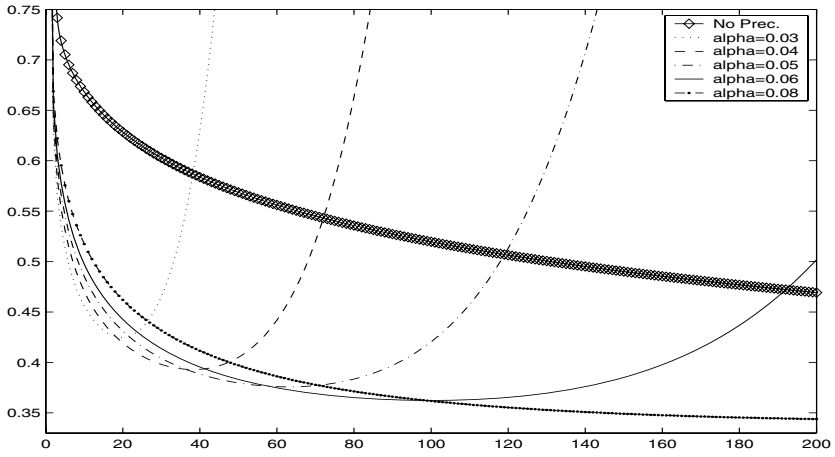

FIG. 5.7. *Test* 2*—RREs versus the first* 200 *iterations; experimental blurred data with* 5% *of noise, filter F*1.

0.56), and the same is true for the corresponding preconditioned versions. Concerning the graphical quality of the restored images, comparing the output of Figure 5.6 (see, for instance, the top right image) and [28, Fig. 4.3], the situation is different, since the projection on positives of the proposed method substantially reduces the artifacts due to ringing effects, as already noticed at the end of Test 1. These results show that the method is not competitive by considering the convergence speed only, but the quality of the restorations is generally good in a moderate number of iterations, with high regularization effectiveness.

In addition, we notice that the projected Landweber method with regularizing preconditioners presented here allows us to recover better restorations than other widely used deblurring algorithms; with a large value of the regularization parameter the convergence is not fast, but the restored image is very good (RRE=0.3510 at iteration 158 with $\alpha = 0.07$; RRE=0.3438 at iteration 200 with $\alpha = 0.08$). In this case, the method could be very favorable when the speed is not crucial. The image on the bottom left of Figure 5.6, relative to $\alpha = 0.08$, where RRE=0.3438, seems better than others provided in the previous literature on the same example (consider, e.g., [19, 21, 24]).

As a final test, we compare the projected Landweber algorithm with the constrained methods proposed by Bardsley and Nagy in [1]. In particular, we consider the same blurred data developed and used in [1] as a simulation of a charge-coupled-device (CCD) camera image of the satellite taken from a ground-based telescope. These input data are shown in Figure 5.8 (top), where the noise comes from both Gaussian and Poisson distributions plus a background contribution so that the relative noise is about 1% (see [1, section 4] for details). The PSF and the true image are the same as in the previous test (see Figure 5.2).

The first method used in [1] is basically a nonnegatively constrained RNSD method, where at any iteration the convergence parameter τ is the result of a more involved minimization procedure which guarantees the nonnegativity of the restoration [25]. A BCCB preconditioned version of this first method is also considered (RNSD/PRNSD). The second method of [1] is a weighted version of the previous nonnegatively constrained RNSD method (WRNSD), where the weight at any pixel

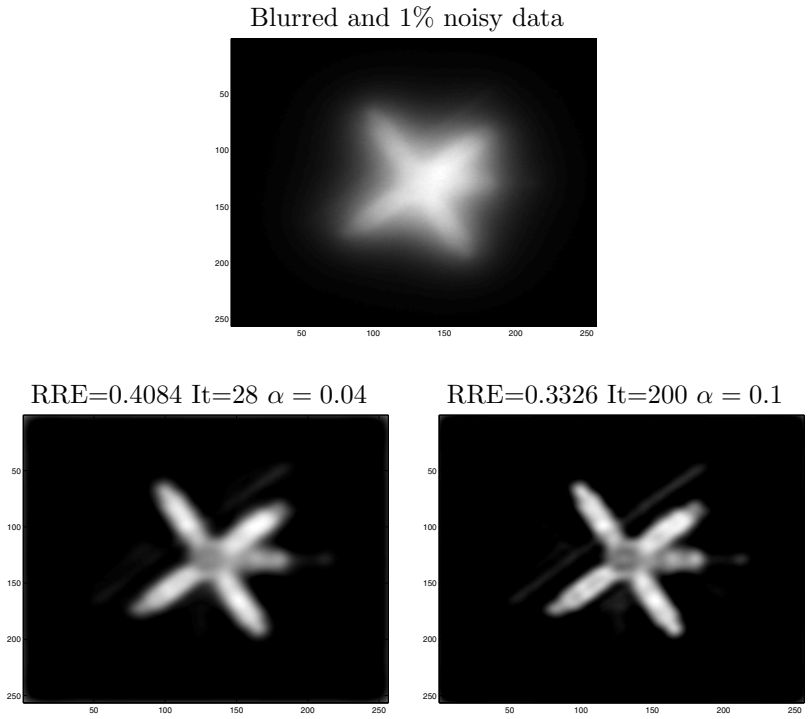

FIG. 5.8. *Test* 2*—Input data of* [1] *and restorations of the Landweber method with positivity, filter F*1.

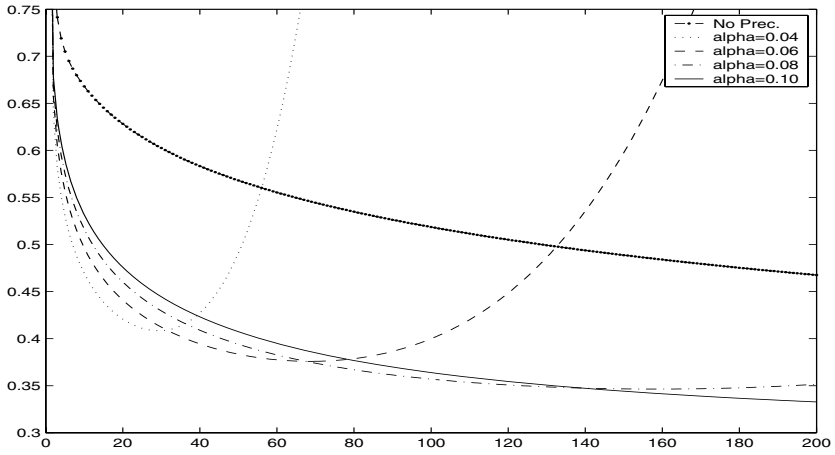

FIG. 5.9. *Test* 2*—RREs versus the first* 200 *iterations, filter F1; satellite blurred data with* 1% *of noise of* [1].

depends on the values of the blurred data. As similarly noticed in the previous test, the convergence history of Figure 5.9 shows that the unpreconditioned Landweber algorithm is slower than all of these more sophisticated methods, as can be seen by a direct comparison with [1, Figure 4.4 (left)]. The preconditioned version is a bit faster than both RNSD and WRNSD (about 70 iterations are needed by these two

methods to achieve an RRE value of 0.4), but it is much slower than the preconditioned RNSD method. On the other hand, the convergence curves of Figure 5.9 show again that the projected Landweber algorithm has a very regular semiconvergence behavior, stronger than the other approaches. When speed is not a basic aim, the method becomes competitive due to its very high regularization capabilities also for the preconditioned version; in particular, the restorations are quite good, and the application of a stopping rule may be simplified. For instance, Figures 5.8 (bottom left) and 5.8 (bottom right) show two different restorations, the first related to a good compromise between speed and accuracy and the second with high accuracy.

6. Conclusions. On account of its reliability and easiness of implementation, we have proposed a regularizing preconditioned version of the Landweber algorithm for space-invariant image deblurring, which speeds up the convergence without losing the regularization effectiveness.

Our theoretical contribution has proceeded along two main directions.

First, a basic analysis of the method in a simplified setting related to periodic boundary conditions has been given, in agreement with the theory developed by Strand [34] but described here in terms of linear algebra language. This way, the convergence properties of the algorithm have been easily addressed to provide practical rules about the choice of all of the several parameters of the method.

Second, Dirichlet boundary conditions have been considered as a case study of other settings not considered in [34]. In particular, by proving that the eigenvectors of the preconditioned matrix are distributed like the Fourier vectors, we have shown that the algorithm is able to speed up the convergence among all of the eigendirections related to the signal space only. These arguments also extend the results of [20] and the subsequent literature, where signal and noise spaces were described in terms of low and high frequencies, regardless of the eigenvectors of the preconditioned matrix.

The numerical results have confirmed the main properties of the method, that is, robustness and flexibility; on these grounds, the method may become a valid tool for the solution of inverse problems arising in real applications.

More precisely, in seismology the Landweber method has been used to estimate the source time function [3] because it provides numerically stable and physically reasonable solutions by introducing positivity, causality, and support constraints. In astronomy the projected Landweber (PL) method is considered as a routine method for the analysis of large binocular telescope images [5], together with Tikhonov regularization and ordered subsets-expectation maximization (OS-EM); in addition, as stated in [7] Landweber should be preferred to others because it is more flexible and modifiable (e.g., by taking into account the support of localized objects) in order to obtain superresolution effects. Again, the PL method is the routine method used for the restoration of chopped and nodded images in thermal infrared astronomy [6]; it is able to determine the response function in the chirp-pulse microwave computerized tomography [29] in a reasonable iteration number (as discussed in section 3, without problems of parameter estimation); even for more ill-conditioned problems in multiple-image deblurring [38], Landweber is more robust than the other methods in the literature.

Indeed, other widely used strategies, such as CGLS, GMRES, or MRNSD [38] methods, are less reliable; although much faster, these methods may lead to less accuracy in the restorations with respect to the proposed method if the signal-to-noise ratio is not large or in severely ill-posed problems (as formerly noticed by [17]). For instance, in these faster methods it is essential to choose the filtering preconditioner

with very high accuracy and stop the iterations really very close to the optimal point, and we know how both of these tasks are as crucial as they are difficult to solve. On the contrary, the regularizing preconditioned Landweber method with positivity allows us to operate at higher safety, since it is much less sensitive to nonoptimal settings of the parameters.

Appendix A. Convergence for $\tau = 1$. As observed in section 3, $\tau = 1$ is the best parameter choice in order to speed up the preconditioned Landweber method. In any case, we must ensure that this value is suitable for convergence along the directions related to the signal; this is true if the preconditioned eigenvalue λ_j^M is less than 2 for any spectral component of interest.

In this appendix we give a formal proof of the inequality $\lambda_j^M < 2$ for all j, assuming periodic boundary conditions.

In what follows, we will use the same notation as in section 2.2 concerning $\lambda_j^A, \lambda_j^D, \lambda_j^M$: notice that all of these eigenvalues implicitly depend on $n = (n_1, n_2)$. The first step consists of proving a "spectral equivalence" between A^*A and $B_{opt}^* B_{opt}$ (from which the definition of the preconditioner D comes).

LEMMA A.1. *Let $\lambda_j^B := \lambda_j(B_{opt}^* B_{opt})$, B_{opt} being the optimal T. Chan approximation of the PSF. Then*

$$\lim_{n\to\infty} \max_{j=1,\dots,N} |\lambda_j^A - \lambda_j^B| = 0 \ .$$

Proof. Since we are imposing periodic boundary conditions, as observed in section 2.2 both A and B_{opt} are square circulant matrices, representing Strang or T. Chan approximations of the PSF matrix. It is well known (see, e.g., [9, Lemma 5] or [36, Theorem 7.1]) that they are spectrally equivalent in the following sense:

$$\lim_{n\to\infty} \max_{j=1,\dots,N} |\lambda_j(A) - \lambda_j(B_{opt})| = 0 \ . \tag{A.1}$$

The circulant algebraic structure of the matrices involved yields the relations

$$\lambda_j^A = \lambda_j(A^*A) = |\lambda_j(A)|^2 \ , \quad \lambda_j^B = \lambda_j(B_{opt}^* B_{opt}) = |\lambda_j(B_{opt})|^2 \ ;$$

thus

$$\begin{aligned} |\lambda_j^A - \lambda_j^B| &= |\lambda_j(A)\overline{\lambda_j(A)} - \lambda_j(B_{opt})\overline{\lambda_j(B_{opt})}| \\ &\le |\lambda_j(A)|\,|\overline{\lambda_j(A)} - \overline{\lambda_j(B_{opt})}| + |\overline{\lambda_j(B_{opt})}|\,|\lambda_j(A) - \lambda_j(B_{opt})| \\ &\le (\|A\|_2 + \|B_{opt}\|_2) \cdot |\lambda_j(A) - \lambda_j(B_{opt})| \ . \end{aligned}$$

Applying (A.1) and the uniform boundedness of A and B_{opt}, the thesis follows. ∎

THEOREM A.2. *Under the assumption of periodic boundary conditions, if the threshold parameter α is sufficiently small, then any filter listed in Table 2.1 determines a preconditioner D such that $\|DA^*A\|_2 < 2$.*

Proof. Since

$$\|DA^*A\|_2 = \max_{j=1,\dots,N} |\lambda_j^D \lambda_j^A|$$

and $\lambda_j^D = \mathcal{F}_\alpha(\lambda_j^B)$ can change its expression according to the comparison between λ_j^B and α, we distinguish two cases.

Case $\lambda_j^B \geq \alpha$. It is easy to check that the inequality $\lambda_j^D \leq 1/\lambda_j^B$ holds for all of the filters I–VIII of Table 2.1.

Set $\epsilon = \alpha$, and apply Lemma A.1; for n large enough, we have $\lambda_j^A < \lambda_j^B + \alpha$ for all j and then

$$0 < \lambda_j^D \lambda_j^A \leq \frac{\lambda_j^A}{\lambda_j^B} < \frac{\lambda_j^B + \alpha}{\lambda_j^B} = 1 + \frac{\alpha}{\lambda_j^B} \leq 2 \ .$$

Case $\lambda_j^B < \alpha$. As before, $\lambda_j^A < \lambda_j^B + \alpha$ for large n, and in this case we obtain $\lambda_j^A < 2\alpha$. Examining the eight filters of Table 2.1, it can be observed that λ_j^D always has an upper bound: 1 for filter III and α^{-1} for the others. It follows that

$$0 < \lambda_j^D \lambda_j^A < \begin{cases} 2\alpha & \text{for filter III,} \\ 2 & \text{for filters I, II, IV–VIII.} \end{cases}$$

Hence the thesis is proved with the only restriction $\alpha \leq 1$ for filter III. ∎

It is worth noticing that the maximal eigenvalue is attained for the indices j such that $\lambda_j^B \approx \alpha$ or even smaller: If the same situation occurs for other boundary conditions, we could justify the conjecture that large outliers (not belonging to the interval $(0, 2)$) correspond to eigendirections outside the signal subspace, as confirmed by the experiments where the semiconvergence property of the method is preserved.

Acknowledgment. The authors express their thanks to Prof. J. Nagy, who kindly provided the blurred data for Test 2 of section 5.

REFERENCES

[1] J. Bardsley and J. Nagy, *Covariance preconditioned iterative methods for nonnegatively constrained astronomical imaging*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 1184–1197.

[2] M. Benzi, *Gianfranco Cimmino's contributions to numerical mathematics*, in Atti del Seminario di Analisi Matematica, Volume Speciale: Ciclo di Conferenze in Memoria di Gianfranco Cimmino, 2004, Dipartimento di Matematica dell'Università di Bologna, Tecnoprint, Bologna, 2005, pp. 87–109.

[3] M. Bertero, D. Bindi, P. Boccacci, M. Cattaneo, C. Eva, and V. Lanza, *Application of the projected Landweber method to the estimation of the source time function in seismology*, Inverse Problems, 13 (1997), pp. 465–486.

[4] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, Institute of Physics, Bristol, UK, 1998.

[5] M. Bertero and P. Boccacci, *Image restoration methods for the large binocular telescope*, Astron. Astrophys. Suppl. Ser., 147 (2000), pp. 323–332.

[6] M. Bertero, P. Boccacci, F. Di Benedetto, and M. Robberto, *Restoration of chopped and nodded images in infrared astronomy*, Inverse Problems, 15 (1999), pp. 345–372.

[7] M. Bertero and C. De Mol, *Super-resolution by data inversion*, in Progress in Optics XXXVI, E. Wolf, ed., Elsevier, Amsterdam, 1996, pp. 129–178.

[8] D. Calvetti, B. Lewis, and L. Reichel, *On the regularizing properties of the GMRES method*, Numer. Math., 91 (2002), pp. 605–625.

[9] R. Chan and X.-Q. Jin, *A family of block preconditioners for block systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 1218–1235.

[10] T. F. Chan, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 766–771.

[11] F. Di Benedetto, *Solution of Toeplitz normal equations by sine transform based preconditioning*, Linear Algebra Appl., 285 (1998), pp. 229–255.

[12] B. Eicke, *Iteration methods for convexly contrained ill-posed problems in Hilbert space*, Numer. Funct. Anal. Optim., 13 (1992), pp. 413–429.

[13] H. W. Engl and P. Kügler, *Nonlinear inverse problems: Theoretical aspects and some industrial applications*, in Multidisciplinary Methods for Analysis Optimization and Control of Complex Systems, Math. Ind. 7, Springer, Berlin, 2004, pp. 3–48.

[14] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.
[15] C. Estatico, *A classification scheme for regularizing preconditioners, with application to Toeplitz systems*, Linear Algebra Appl., 397 (2005), pp. 107–131.
[16] C. W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Res. Notes Math., 105, Pitman, Boston, 1984.
[17] M. Hanke, *Accelerated Landweber iterations for the solutions of ill-posed problems*, Numer. Math., 60 (1991), pp. 341–373.
[18] M. Hanke, *A second look at Nemirovskii's analysis of the conjugate gradient method*, Beiträge zur angewandten Analysis und Informatik, E. Schock ed., Shaker Verlag, Aachen, Germany, 1994, pp. 123–135.
[19] M. Hanke and J. Nagy, *Inverse Toeplitz preconditioners for ill-posed problems*, Linear Algebra Appl., 284 (1998), pp. 137–156.
[20] M. Hanke, J. Nagy, and R. Plemmons, *Preconditioned iterative regularization for ill-posed problems*, in Numerical Linear Algebra (Kent, OH, 1992), de Gruyter, Berlin, 1993, pp. 141–163.
[21] M. Hanke, J. Nagy, and C. Vogel, *Quasi-Newton approach to nonnegative image restorations*, Linear Algebra Appl., 316 (2000), pp. 223–236.
[22] P. C. Hansen, *Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank*, SIAM J. Sci. Statist. Comput., 11 (1990), pp. 503–518.
[23] A. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
[24] J. Kamm and J. Nagy, *Kronecker product and SVD approximations in image restoration*, Linear Algebra Appl., 284 (1998), pp. 177–192.
[25] L. Kaufman, *Maximum likelihood, least squares, and penalized least squares for PET*, IEEE Trans. Med. Imag., 12 (1993), pp. 200–214.
[26] M. E. Kilmer and D. P. O'Leary, *Pivoted Cauchy-like preconditioners for regularized solution of ill-posed problems*, SIAM J. Sci. Comput., 21 (1999), pp. 88–110.
[27] L. Landweber, *An iteration formula for Fredholm integral equations of the first kind*, Amer. J. Math., 73 (1951), pp. 615–624.
[28] K. Lee and J. Nagy, *Steepest descent, CG and iterative regularization of ill-posed problems*, BIT, 43 (2003), pp. 1003–1017.
[29] M. Miyakawa, K. Orikasa, M. Bertero, P. Boccacci, F. Conte, and M. Piana, *Experimental validation of a linear model for data reduction in Chirp-Pulse Microwave CT*, IEEE Trans. Med. Imag., 21 (2002), pp. 385–395.
[30] M. Ng, *Iterative Methods for Toeplitz Systems*, Oxford University Press, London, 2004.
[31] M. Piana and M. Bertero, *Projected Landweber method and preconditioning*, Inverse Problems, 13 (1997), pp. 441–464.
[32] M. C. Roggemann and B. Welsh, *Imaging through Turbulence*, CRC Press, Boca Raton, FL, 1996.
[33] S. Serra-Capizzano, *A note on antireflective boundary conditions and fast deblurring models*, SIAM J. Sci. Comput., 25 (2003), pp. 1307–1325.
[34] O. N. Strand, *Theory and methods related to the singular-function expansion and Landweber's iteration for integral equations of the first kind*, SIAM J. Numer. Anal., 11 (1974), pp. 798–825.
[35] G. Strang, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.
[36] E. E. Tyrtyshnikov, *A unifying approach to some old and new theorems on distribution and clustering*, Linear Algebra Appl., 232 (1996), pp. 1–43.
[37] E. E. Tyrtyshnikov, A. Y. Yeremin, and N. L. Zamarashkin, *Clusters, preconditioners, convergence*, Linear Algebra Appl., 263 (1997), pp. 25–48.
[38] R. Vio, J. Nagy, L. Tenorio, and W. Wamsteker, *A simple but efficient algorithm for multiple-image deblurring*, Astron. Astrophys., 416 (2004), pp. 403–410.
[39] N. L. Zamarashkin and E. E. Tyrtyshnikov, *Distribution of eigenvalues and singular values of Toeplitz matrices under weakened conditions on the generating function*, Russian Acad. Sci. Sb. Math., 188 (1997), pp. 1191–1201.
[40] N. L. Zamarashkin and E. E. Tyrtyshnikov, *On the distribution of eigenvectors of Toeplitz matrices with weakened requirements on the generating function*, Russian Math. Surveys, 522 (1997), pp. 1333–1334.