ELSEVIER

Linear Algebra and its Applications 346 (2002) 177–197

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

A polynomial fit preconditioner for band Toeplitz matrices in image reconstruction

P. Favati [a], G. Lotti [b], O. Menchi [c,*]

[a] *IMC-CNR, Via G. Moruzzi 1, 56124 Pisa, Italy*
[b] *Dipartimento di Matematica, University of Parma, Via M. D'Azeglio 85/A, 43100 Parma, Italy*
[c] *Dipartimento di Informatica, University of Pisa, Corso Italia 40, 56125 Pisa, Italy*

Received 9 July 2001; accepted 25 September 2001

Submitted by R.A. Brualdi

Abstract

The preconditioned conjugate gradient (CG) is often applied in image reconstruction as a regularizing method. When the blurring matrix has Toeplitz structure, the modified circulant preconditioner and the inverse Toeplitz preconditioner have been shown to be effective. We introduce here a preconditioner for symmetric positive definite Toeplitz matrices based on a trigonometric polynomial fit which has the same effectiveness of the previous ones but has a lower cost when applied to band matrices. The case of band block Toeplitz matrices with band Toeplitz blocks (BTTB) corresponding to separable point spread functions (PSFs) is also considered.

Keywords: Preconditioned conjugate gradient (PCG); Image reconstruction

1. Introduction

In many engineering applications it is necessary to numerically solve large structured $m \times m$ dimensional linear systems

$$H_m \mathbf{x} = \mathbf{b} - \boldsymbol{\eta}, \tag{1}$$

where $\boldsymbol{\eta}$ represents unknown noise or measurement errors. In particular for image restoration problems, H_m is the blur operator, $\mathbf{x}$ is the original image and $H_m\mathbf{x}$

* Corresponding author.
E-mail address: menchi@di.unipi.it (O. Menchi).

0024-3795/02/$ - see front matter
PII: S0024-3795(01)00513-4

is the blurred image (see [2]). The matrix H_m is defined by the so-called *point spread function* (PSF), which describes how the imaging system affects the points of the original image. In many imaging systems the PSF is space invariant, band-limited and center-symmetric, so that the matrix H_m turns out to have a symmetric 2-level band Toeplitz structure. In this paper, we assume that these hypotheses hold.

Unlike blurring, which we consider to be a deterministic process, the noise is due to a random process, for which we can only assume the knowledge of some statistical properties. A frequent assumption is that $\boldsymbol{\eta}$ is a Gaussian white noise, having spectral components η_i which are independent stochastic variables with mean 0 and variance σ^2. Moreover, we assume that $\mathbf{b}$ dominates $\boldsymbol{\eta}$, otherwise the reconstruction of the original image would be impossible.

The problem of the image restoration is then that of finding a good approximation of $\mathbf{x}$, given H_m and $\mathbf{b}$. The problem, though well-posed, in the sense that it has a unique solution, is actually the discretization of an ill-posed continuous problem. Hence it inherits a large condition number, and the finer the discretization, the larger the condition number of H_m. It follows that the exact solution of the system $H_m\mathbf{y} = \mathbf{b}$ may differ considerably from $\mathbf{x}$ even if $\boldsymbol{\eta}$ is small. This difficulty can be overcome by employing special techniques, known as *regularization methods*. One of the first regularization methods and probably the most used one is the Tikhonov method, which seeks the solution of a least-squared problem depending on a regularizing parameter. The difficulty of tuning this parameter has suggested a different approach to the problem. In fact, some iterative methods for solving linear systems enjoy a regularizing property known as *semiconvergence*. The classical conjugate gradient (CG) method, which applies to symmetric positive definite matrices, shares the regularizing property.

Since the problem is ill-conditioned, the use of a suitable preconditioner is required. Circulant preconditioners, modified in order to cope with the noise, have shown to be effective [10]. They can be inverted by using FFT with a computational cost $\mathrm{O}(m \log m)$. The same computational cost is reached by the inverse Toeplitz preconditioner [9], based on the symbol function of the Toeplitz coefficient matrix and having regularizing effects. Unfortunately this cost $\mathrm{O}(m \log m)$ holds even for band matrices.

For band matrices a band preconditioner, which can be inverted with the cost $\mathrm{O}(m)$ (that is the same of the single CG iteration), would be preferable. In [11] a general-purpose preconditioner based on the minimum-phase LU factorization is studied which, when applied to band matrices, lowers the computational complexity of the iterative step to $\mathrm{O}(m)$, but the modification of this preconditioner for problems with noise does not appear to be simple.

In this paper, we propose a band preconditioner, which is based on the symbol function of the Toeplitz coefficient matrix and achieves regularizing effects. This preconditioner is effective for the image reconstruction problems with band positive definite Toeplitz matrix and has a computational cost $\mathrm{O}(m)$.

After some preliminaries in Section 2, the idea on which the preconditioner is based is described in Section 3, where we explain in detail how to construct the preconditioner for the 1-D case and examine the eigenvalues of the preconditioned matrix. The preconditioner is then extended to the separable 2-D case in Section 5. Sections 4 (for the 1-D case) and 6 (for the 2-D case) describe some numerical experiments with different problems: the obtained results are compared with those produced by the circulant preconditioner and the inverse Toeplitz preconditioner.

2. Preliminaries

Let us first examine the behavior of CG when applied to solve problem (1) with a symmetric positive definite matrix H_m. It is useful to consider the sets $\mathscr{L} = \{\lambda_i,\ i = 1, \ldots, m\}$ and $\mathscr{U} = \{\mathbf{u}_i,\ i = 1, \ldots, m\}$ of eigenvalues and eigenvectors of H_m, respectively. Following [10], we detect two subspaces

$$\begin{aligned}\mathscr{U}_n &= \mathrm{span}\{\mathbf{u}_i,\ \lambda_i \leqslant \epsilon_1\}, \quad \text{noise subspace of dimension } d_n,\\ \mathscr{U}_s &= \mathrm{span}\{\mathbf{u}_i,\ \lambda_i \geqslant \epsilon_2\}, \quad \text{signal subspace of dimension } d_s,\end{aligned}$$

where ϵ_1 and ϵ_2 are two parameters connected to the magnitudes of noise and of signal, respectively. Between $\mathscr{U}_n$ and $\mathscr{U}_s$ a transient subspace is spanned by the remaining eigenvalues. The subsets $\boldsymbol{\sigma}_n$ and $\boldsymbol{\sigma}_s$ of $\boldsymbol{\sigma} = \{1, \ldots, m\}$ contain the indices corresponding to $\mathscr{U}_n$ and $\mathscr{U}_s$. The hypothesis that $\mathbf{b}$ dominates $\boldsymbol{\eta}$ corresponds to require that $\epsilon_1 < \epsilon_2$ and that the original image belongs essentially to $\mathscr{U}_s$.

Since $\mathscr{U}$ is a basis for $\mathbb{R}^m$, vectors $\mathbf{x}$ and $\boldsymbol{\eta}$ can be written as

$$\mathbf{x} = \sum_{i=1}^{m} x_i \mathbf{u}_i \quad \text{and} \quad \boldsymbol{\eta} = \sum_{i=1}^{m} \eta_i \mathbf{u}_i,$$

hence the exact solution of the system $H_m \mathbf{y} = \mathbf{b}$ can be written as

$$\mathbf{y} = H_m^{-1} \mathbf{b} = \sum_{i=1}^{m} \xi_i\ \mathbf{u}_i, \quad \text{where } \xi_i = \frac{\lambda_i x_i + \eta_i}{\lambda_i}.$$

The components ξ_i for $i \in \boldsymbol{\sigma}_s$ are nearly equal to the corresponding components x_i of $\mathbf{x}$ and the components ξ_i for $i \in \boldsymbol{\sigma}_n$ can be much greater than x_i. Let $\mathbf{x}^{(k)}$ be the vector computed at the kth iteration by applying the CG method with $\mathbf{x}^{(0)} = \mathbf{0}$. It can be shown that

$$\mathbf{x}^{(k)} = \sum_{i=1}^{m} \varphi_i^{(k)} \xi_i\ \mathbf{u}_i, \quad \text{where } \varphi_i^{(k)} = 1 - R_k(\lambda_i),$$

R_k being the kth Ritz polynomial [17]. Under the assumptions that $\lambda_i |x_i| \leqslant \lambda_i \|\mathbf{x}\|$ (*discrete Picard condition*) and that $\boldsymbol{\eta}$ is a Gaussian white noise, the oscillations around the zero of the polynomial $R_k(\lambda)$ tend to be smaller when λ keeps away from zero. Moreover, the neighborhood of zero where $R_k(\lambda)$ is near 1 shrinks when k

increases (see [10]). As a consequence $\varphi_i^{(k)}$ is close to 1 for indices i corresponding to the highest eigenvalues λ_i, while it is close to 0 for low values of k and indices i corresponding to the smallest eigenvalues λ_i. Hence it acts as a filter: at the beginning of iteration, for small values of k, only the components corresponding to the signal subspace are allowed to pass, while the components corresponding to the noise subspace are filtered. In this phase, the iteration reconstructs the signal. On the contrary, for high values of k, all the filters are nearly equal to 1 and the noise components interfere. Hence the method must be stopped before it starts to reconstruct the noise. The most used stopping rule prescribes to iterate until the residue becomes lower than a suitable quantity δ related to $\boldsymbol{\eta}$ (see [8]).

When the coefficient matrix is ill-conditioned, as in the present case, the number of iterations required by CG for obtaining a satisfactory result can be large and preconditioning is required to increase the rate of convergence. The preconditioners that are used in the general case are not satisfactory for our problem, because they are designed to reduce the condition number by clustering all the eigenvalues around 1. In this way the signal and the noise subspaces are mixed up and the effect of the noise appears before the image is fully reconstructed.

In the present context, the right preconditioner should reduce the number of iterations required to reconstruct the information from $\mathscr{U}_s$, that is, it should cluster around 1 the d_s greatest eigenvalues, letting the d_n smallest ones away from the cluster. Circulant preconditioners, frequently applied to Toeplitz matrices, can be easily modified in order to cluster only the eigenvalues we are interested in, instead of all the spectrum (see [10]). To this aim an estimate of d_n must be known. A technique to separate $\mathscr{U}_s$ from $\mathscr{U}_n$ is based on the assumption that the Fourier coefficients of $\boldsymbol{\eta}$ have approximately the same magnitude for all the frequencies and they dominate the Fourier coefficients of $\mathbf{b}$ corresponding to $\boldsymbol{\sigma}_n$ (see [10]). In this paper, we assume a rough estimate of d_n and d_s to be available.

In the following, we present in detail a new regularizing preconditioner based on a fit technique. At first, we deal with the 1-D problem, then we extend the same technique to the more interesting 2-D case.

3. The 1-D problem

We consider the problem of approximating the vector $\mathbf{x}$, solution of

$$H_N\mathbf{x} = \mathbf{b} - \boldsymbol{\eta}, \tag{2}$$

where H_N is the symmetric band Toeplitz matrix

$$H_N \equiv (h_{ij}), \quad h_{ij} = \begin{cases} h_{|i-j|} & \text{for } |i-j| \leqslant w, \\ 0 & \text{otherwise,} \end{cases} \qquad i, j = 1, \ldots, N,$$

w being the bandwidth. Matrix H_N is the N-section of a bi-infinite Toeplitz matrix, say H, whose symbol is the symmetric Laurent polynomial of degree w

$$h(z) = h_0 + \sum_{i=1}^{w} h_i (z^i + z^{-i}), \quad h_i \in \mathbb{R}. \tag{3}$$

Let $\widetilde{h}(\theta)$ be the restriction of $h(z)$ on the unit circle of the complex plane, that is

$$\widetilde{h}(\theta) = h(\mathrm{e}^{\mathrm{i}\theta}) = h_0 + 2\sum_{i=1}^{w} h_i \cos i\theta, \quad \theta \in [-\pi, \pi].$$

We assume H to be positive definite, hence $\widetilde{h}(\theta) > 0$ for any θ.

In the following for any bi-infinite Toeplitz matrix T

- $t(z)$ denotes the symbol of T and $\widetilde{t}(\theta)$ denotes its restriction on the unit circle of the complex plane,
- $\mathscr{I}_t$ denotes the interval of $\mathbb{R}$ having endpoints

  $$t_{\inf} = \inf_{\theta} \widetilde{t}(\theta) \quad \text{and} \quad t_{\sup} = \sup_{\theta} \widetilde{t}(\theta),$$

- for any integer $m \geqslant 1$, T_m denotes the m-section of T.

A classical result due to Grenander and Szegö [6] relates the spectrum of the finite sections T_m of a positive definite Toeplitz matrix T with the behavior of its symbol $t(z)$ on the unit circle. It states that all the eigenvalues $\lambda_i(T_m)$ belong to $\mathscr{I}_t$ and that for any function φ continuous on $\mathscr{I}_t$ it holds

$$\lim_{m\to\infty} \frac{1}{m} \sum_{i=0}^{m-1} \left[\varphi\left(\lambda_i(T_m)\right) - \varphi\left(\widetilde{t}(\theta_i)\right)\right] = 0, \quad \text{where } \theta_i = \frac{2\pi i}{m}. \tag{4}$$

Hence the eigenvalues $\lambda_i(T_m)$ are equally distributed as $\widetilde{t}(\theta_i)$.

To construct a good preconditioner P_N for our problem we should know, at least approximatively, the eigenvalues of H_N. This knowledge is not directly available, but can be obtained indirectly from $\widetilde{h}(\theta)$. In fact, let $\mathscr{H} = \{\widetilde{h}(\theta_i),\ \theta_i = 2\pi i/N,\ i = 0, \dots, N-1\}$ be the set of N equally spaced samples of $\widetilde{h}$. From (4), we derive that $\mathscr{H}$ results in a sufficiently good approximation of $\mathscr{L}$ to be used for our purpose. Then we detect in $\mathscr{H}$ two subsets $\mathscr{H}_n$ and $\mathscr{H}_s$, formed by the d_n smallest and the d_s greatest elements of $\mathscr{H}$, respectively.

Our idea is to modify selectively the spectrum of the preconditioned matrix by acting on the symbol. More precisely, we want to find a Laurent symmetric polynomial $p(z)$ of suitable degree with $\widetilde{p}(\theta) > 0$, in such a way that the function $f(z) = p^{-1}(z)h(z)$ has a set $\mathscr{F}$ of samples of $\widetilde{f}$ where

- as many as possible values of $\mathscr{F}_s$ are clustered around 1,
- the values of $\mathscr{F}_n$ do not belong to the cluster around 1 and are as small as possible.

Let F and P be the bi-infinite Toeplitz matrices having symbols $f(z)$ and $p(z)$, respectively. Matrix $P_N^{-1}H_N$ (which in general has no Toeplitz structure) differs from F_N (which has a Toeplitz structure), but it is close enough to have the required clustering [13] and regularizing property.

A similar approach is commonly used to remove the ill-conditioning of a problem by deleting the zeroes of function $h(z)$ on the unit circle (see for example [3,4,13,15,16]).

3.1. Construction of the polynomial $p(z)$

Let ϵ be a truncation parameter and let $\boldsymbol{\tau}$ be a subset of equally spaced elements of $\boldsymbol{\sigma}$ having cardinality κ. Partition $\boldsymbol{\tau} = \boldsymbol{\tau}_1 \cup \boldsymbol{\tau}_2$, where

$$\boldsymbol{\tau}_1 = \left\{ i \in \boldsymbol{\tau} : \widetilde{h}(2\pi i/N) \geqslant \epsilon \right\}.$$

Consider the discrete function $\psi(\theta)$ defined on the nodes $\theta_i = 2\pi i/N$ with $i \in \boldsymbol{\tau}$ as follows:

$$\psi(\theta_i) = \begin{cases} \widetilde{h}(\theta_i) & \text{for } i \in \boldsymbol{\tau}_1, \\ \epsilon & \text{for } i \in \boldsymbol{\tau}_2. \end{cases}$$

We fix a (small) integer μ, with $1 \leqslant \mu \leqslant \kappa - 1$. The function

$$p(z) = p_0 + \sum_{i=1}^{\mu} p_i (z^i + z^{-i}), \quad p_i \in \mathbb{R}, \tag{5}$$

we are looking for, should satisfy the following requirements

- $\widetilde{p}(\theta) = p_0 + 2\sum_{i=1}^{\mu} p_i \cos i\theta$ is a good approximation of $\psi(\theta)$,
- $\widetilde{p}(\theta) > c$ for a constant $0 < c < \epsilon$ and $\widetilde{p}(\theta)^{-1}\widetilde{h}(\theta)$ does not oscillate too much.

The function $\widetilde{p}(\theta)$ can be computed in several ways. We suggest to find a minimum norm approximation, for example, a least squares or a minimax approximation, on the basis $\{1, \cos\theta, \ldots, \cos\mu\theta\}$. We point out that the degree μ of $\widetilde{p}(\theta)$ should be small with respect to κ, in order to avoid oscillations, and that κ should be independent of N, in order to make the cost of the computation of the coefficients $p_0, \ldots, p_\mu$ negligible.

Actually the request on the nodes θ_i to be equally spaced can be released, provided that they are regularly placed in the whole interval $[0, \pi]$. For example they could be selected more sparsely near the eigenvalues corresponding to the transient subspace and more densely elsewhere.

The choice of the truncation parameter ϵ should be made taking into account that $\widetilde{f}(\theta) = \widetilde{p}^{-1}(\theta)\widetilde{h}(\theta)$ must have a value nearly 1 when θ is close to the nodes θ_i, with $i \in \boldsymbol{\tau}_1$, and a value nearly $\widetilde{h}(\theta)/\epsilon$ when θ is close to the nodes θ_i, with $i \in \boldsymbol{\tau}_2$. Thus a suitable value for ϵ is the greatest one for which $\boldsymbol{\tau}_2$ does not contain indices belonging to $\boldsymbol{\sigma}_s$. In this way the growth of the eigenvalues of $\mathscr{U}_n$ is kept as small as possible.

The band preconditioner P_N, outlined above, compares favorably with the circulant preconditioner from the point of view of the computational cost, because it can be factorized with a cost of $\mathrm{O}(n)$. But it is unsuitable when N is large and there is not enough memory space to keep the factorization of P_N. To overcome this problem the matrix P_N is further approximated by a product $Q_N Q_N^{\mathrm{T}}$, where Q_N is a lower triangular Toeplitz matrix with the same bandwidth of P_N. In this way the computational cost of an iteration of PCG is not increased and the memory space to keep Q_N is reduced to O(1). In Section 3.3, we show that the eigenvalues of the preconditioned matrix ${Q_N^{-1}}^{\mathrm{T}} Q_N^{-1} H_N$ so obtained are interlaced, except for a small number, with those of F_N, thus sharing the required clustering and regularizing property.

3.2. *Factorization of the polynomial $p(z)$*

Once the function $p(z)$ is determined, the polynomial

$$q(z) = \sum_{i=0}^{\mu} q_i z^i, \quad q_i \in \mathbb{R},$$

such that

- $q(z)$ has all roots outside the unit circle,
- $q_0 > 0$,
- $q(z)q(z^{-1}) = p(z)$

is sought. The Toeplitz matrix Q having symbol $q(z)$ is banded lower triangular and satisfies $QQ^{\mathrm{T}} = P$. Hence Q is the Choleski factor of P. The polynomial $q(z)$ is called the *factor* of $p(z)$.

A factorization analogous to the present one is the minimum-phase factorization, described in [11] where the preconditioner of the coefficient matrix is constructed by using the factorization of the Laurent polynomial symbol of the matrix into two factors; the first one having only zeros outside the unit circle, the second one only inside the unit circle. We follow [7,12], where the Wiener–Hopf factorization of bi-infinite Toeplitz matrices in the Wiener class is studied. Among the five methods described in [7] for computing $q(z)$, we suggest the third one (called the *roots method*) which is simple and effective in the case of low degree polynomials. It is based on the computation of the roots z_i for $i = 1, \ldots, 2\mu$ of (5). Clearly if z_i is a root also z_i^{-1} is a root, thus for any root lying outside the unit circle, there is a corresponding root in the circle and vice versa. No root lies on the unit circle. Hence the set $\mathbf{z} = \{z_i, |z_i| > 1\}$ contains μ roots. The polynomial $q(z)$ is constructed by taking all the roots in $\mathbf{z}$

$$q(z) = q_\mu \prod_{z_i \in \mathbf{z}} (z - z_i), \quad \text{where } q_\mu^2 = (-1)^\mu p_\mu \Big/ \prod_{z_i \in \mathbf{z}} z_i$$

and

$$\operatorname{sgn}(q_\mu) = \operatorname{sgn}(p_\mu).$$

If the coefficients p_i are severely unbalanced, a scaling technique is suggested before computing the roots (see [7]).

The matrix $Q_N Q_N^{\mathrm{T}}$ is called *fit preconditioner*. Due to the band structure, the computation of the vector ${Q_N^{-1}}^{\mathrm{T}} Q_N^{-1}\mathbf{v}$ for any $\mathbf{v} \neq \mathbf{0}$ requires $2\mu N$ operations. Then any iteration of the PCG costs $(2\mu + 2w + 5)N$ operations [1].

3.3. Eigenvalues of the preconditioned matrix

In order to examine the behavior of the PCG with the fit preconditioner, we need the two following lemmas which relate the spectrum of the preconditioned matrix $(Q_N Q_N^{\mathrm{T}})^{-1} H_N$ effectively used to the spectrum of the matrix F_N. The symbol of F is $f(z) = p^{-1}(z)h(z) = q^{-1}(z)q^{-1}(z^{-1})h(z)$ and can be expressed by a symmetric Laurent series $f(z) = f_0 + \sum_{i=1}^{\infty} f_i(z^i + z^{-i})$.

Lemma 1. *Let* $a(z) = a_0 + \sum_{i=1}^{\infty} a_i(z^i + z^{-i})$, $b(z) = \sum_{i=0}^{\mu} b_i z^i$, $a_i, b_i \in \mathbb{R}$ *be the symbols of the Toeplitz matrices* A (*symmetric*) *and* B (*lower banded*). *The product matrix* $C = BAB^{\mathrm{T}}$ *has symbol* $c(z) = b(z)a(z)b(z^{-1})$. *For any integer* n, *with* $n \geqslant 2\mu$, *the symmetric matrix* $D_n = C_n - B_n A_n B_n^{\mathrm{T}}$ *has rank* $\rho \leqslant 2\mu$.

Proof. Partition matrices A_n and B_n as follows:

$$A_n = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^{\mathrm{T}} & A_{22} \end{bmatrix}, \quad B_n = \begin{bmatrix} B_{11} & O \\ B_{21} & B_{22} \end{bmatrix},$$

where A_{11} and $B_{11} \in \mathbb{R}^{\mu \times \mu}$. Consider the bordered matrices

$$\overline{A}_n = \begin{bmatrix} A_{00} & A_{01} & A_{02} \\ A_{01}^{\mathrm{T}} & A_{11} & A_{12} \\ A_{02}^{\mathrm{T}} & A_{12}^{\mathrm{T}} & A_{22} \end{bmatrix}, \quad \overline{B}_n = \begin{bmatrix} B_{10} & B_{11} & O \\ O^{\mathrm{T}} & B_{21} & B_{22} \end{bmatrix},$$

where A_{00} and $B_{10} \in \mathbb{R}^{\mu \times \mu}$. The n-section C_n of C satisfies $C_n = \overline{B}_n \overline{A}_n \overline{B}_n^{\mathrm{T}}$ and it can be shown by direct computation that

$$D_n = \overline{B}_n \overline{A}_n \overline{B}_n^{\mathrm{T}} - B_n A_n B_n^{\mathrm{T}} = \begin{bmatrix} D_{11} & D_{12} \\ D_{12}^{\mathrm{T}} & O \end{bmatrix},$$

where

$$D_{11} = B_{10}A_{00}B_{10}^{\mathrm{T}} + B_{11}A_{01}^{\mathrm{T}}B_{10}^{\mathrm{T}} + B_{10}A_{01}B_{11}^{\mathrm{T}} \in \mathbb{R}^{\mu \times \mu},$$
$$D_{12} = B_{10}A_{01}B_{21}^{\mathrm{T}} + B_{10}A_{02}B_{22}^{\mathrm{T}} \in \mathbb{R}^{\mu \times (n-\mu)}.$$

Hence D_n has rank $\rho \leqslant 2\mu$. □

Lemma 2 (The rank theorem of [14]). *Let* A, W, Y *be* n*-dimensional symmetric matrices*, *with* $A = W + Y$, *and let* α_i, ω_i *and* η_i *be the respective eigenvalues indexed in increasing order. Let* π *and* ν *be the number of positive and negative eigenvalues of* Y ($\rho = \pi + \nu$ *is the rank of* Y). *Then for all* k *such that* $\nu < k \leqslant n - \pi$ *it holds*

$$\omega_{k-\rho} \leqslant \omega_{k-\nu} \leqslant \alpha_k \leqslant \omega_{k+\pi} \leqslant \omega_{k+\rho}.$$

Moreover, for i, j *such that* $1 \leqslant i + j - 1 \leqslant n$ *it holds*

$$\omega_i + \eta_j \leqslant \alpha_{i+j-1} \quad \text{and} \quad \alpha_{n-i-j+2} \leqslant \omega_{n-i+1} + \eta_{n-j+1}$$

(in particular $\omega_1 + \eta_1 \leqslant \alpha_1 \leqslant \omega_1 + \eta_n$*).*

Theorem 1. *Let* H_N *be the matrix of system* (2). *For a fixed integer* μ, *let* $p(z)$ *be the Laurent polynomial obtained in Section* 3.1 *as a suitable approximation of* $h(z)$. *Let* $q(z)$ *be the factor of* $p(z)$ *and* $f(z) = p^{-1}(z)h(z)$. *Then no more than* 2μ *eigenvalues of* $Q_N^{-\mathrm{T}} Q_N^{-1} H_N$ *lie outside of* $\mathscr{I}_f$ *and each of the other eigenvalues is interlaced between two eigenvalues of* F_N.

Proof. For the matrix $Q_N^{-1} H_N Q_N^{-\mathrm{T}}$, similar to the preconditioned matrix $Q_N^{-\mathrm{T}} Q_N^{-1} H_N$, we have

$$Q_N^{-1} H_N Q_N^{-\mathrm{T}} - F_N = Q_N^{-1} \left(H_N - Q_N F_N Q_N^{\mathrm{T}}\right) Q_N^{-\mathrm{T}}.$$

From Lemma 1, applied to the functions $a(z) = f(z)$ and $b(z) = q(z)$, it follows that matrix $H_N - Q_N F_N Q_N^{\mathrm{T}}$ has rank $\rho \leqslant 2\mu$, that is, $Q_N^{-1} H_N Q_N^{-\mathrm{T}} = F_N + Y$, where Y is a symmetric correction of rank ρ. Then by Lemma 2, no more than ρ eigenvalues of the preconditioned matrix $Q_N^{-\mathrm{T}} Q_N^{-1} H_N$ are outside the interval $\mathscr{I}_f$ and the other eigenvalues are interlaced with the eigenvalues of F_N. $\square$

From Theorem 1 it follows that the behavior of the eigenvalues of the preconditioned matrix, except a small number of them, is described by the graphic of the function $f(z)$ on the unit circle, which by construction has a suitable selected clustering with regularizing effect.

4. Experiments with the 1-D problem

The 1-D examples refer to an original image $\mathbf{x}$ of size $N = 64$, defined by

$$x_i = \lfloor 124(1 + \sin\lfloor 15i/N \rfloor) \rfloor, \quad i = 1, \ldots, N.$$

Example 4.1 (*Gaussian PSF*). The $N \times N$ blurring matrix is the N-section of the bi-infinite Toeplitz matrix H whose symbol is the symmetric Laurent polynomial (3) of degree $w = 8$ with coefficients $h_i = \beta \exp(-\alpha i^2)$, $i = 0, \ldots, w$, where $\alpha = 0.2$ and β is the normalizing constant such that $\sum_{i=-w}^{w} h_i = 1$. This matrix corresponds to the 1-D version of a Gaussian PSF. The function $\widetilde{h}(\theta)$ is positive and the condition number of H_N is approximately 10^5.

The noisy image $\mathbf{b}$ is obtained by computing $H_N\mathbf{x} + \boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a vector of randomly generated entries, with normal distribution and mean 0. The entries of $\boldsymbol{\eta}$ are scaled in such a way that $\|\boldsymbol{\eta}\|_2 / \|H_N\mathbf{x}\|_2 = 10^{-4}$. The dimensions of the noise and signal subspaces are estimated to be $d_n = 22$ and $d_s = 40$, respectively. By checking the set $\mathscr{H}$, the estimate 0.01 is found for ϵ.

First the system is solved by CG. At the ith iteration the relative norm 2 error e_i of the computed vector $\mathbf{x}^{(i)}$ with respect to the original image $\mathbf{x}$ is taken. The error history (long dashed line) is plotted in Fig. 2. As expected, the error shows an initial quick decrease until a minimal value is reached, followed by a slow increase, when the noise is reconstructed.

Approximations to $\psi(\theta)$ with various degrees and point numbers have been considered. For example, for $\mu = 5$ and $\kappa = 20$ the least-squared fit of function $\psi(\theta)$ is

$$\begin{aligned}\widetilde{p}(\theta) =& 0.256 + 0.408\ \cos\theta + 0.229\ \cos 2\theta + 0.083\ \cos 3\theta + 0.02\ \cos 4\theta \\ &+ 0.004\ \cos 5\theta.\end{aligned}$$

Function $\widetilde{f}(\theta) = \widetilde{p}(\theta)^{-1}\widetilde{h}(\theta)$ is plotted in Fig. 1 (continuous line) together with function $\widetilde{h}(\theta)$ (dashed line).

The factor $q(z)$ results to be

$$q(z) = 0.256 + 0.324\,z + 0.256\,z^2 + 0.12\,z^3 + 0.028\,z^4 + 0.008\,z^5.$$

Two eigenvalues of the preconditioned matrix $Q_N^{-\mathrm{T}} Q_N^{-1} H_N$ lie outside $\mathscr{I}_f$.

The *PCG* method is then applied with the fit preconditioner so obtained. The error history (continuous line) is compared with that of the non-preconditioned CG (long dashed line) in Fig. 2.

For comparison purpose the same problem is solved using Chan [5] circulant preconditioner, modified as in [10], with truncation parameter $\tau = 0.005$ and the inverse Toeplitz preconditioner of [9] with the same truncation parameter (the experiments show that this is the optimal value to be used for both the methods). The error histories (short dashed line for the circulant preconditioner, dot–dashed line for the inverse Toeplitz preconditioner) are also shown in Fig. 2.

Table 1 summarizes the minimal value obtained and the iteration number required by the non-preconditioned CG, the PCG with the fit preconditioner, the PCG with the circulant preconditioner and the PCG with the inverse Toeplitz preconditioner.

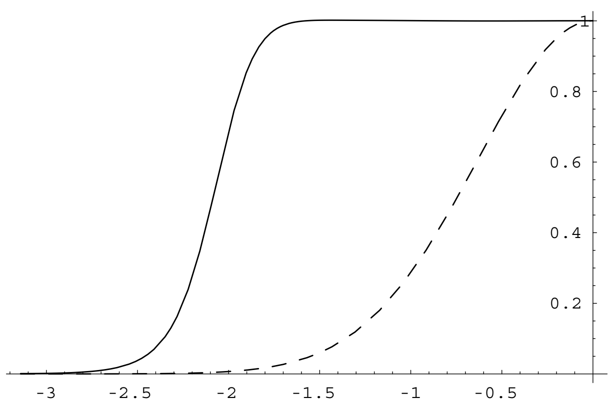

Fig. 1. Functions $\widetilde{f}(\theta) = \widetilde{p}(\theta)^{-1}\widetilde{h}(\theta)$ (continuous line) and $\widetilde{h}(\theta)$ (dashed line).

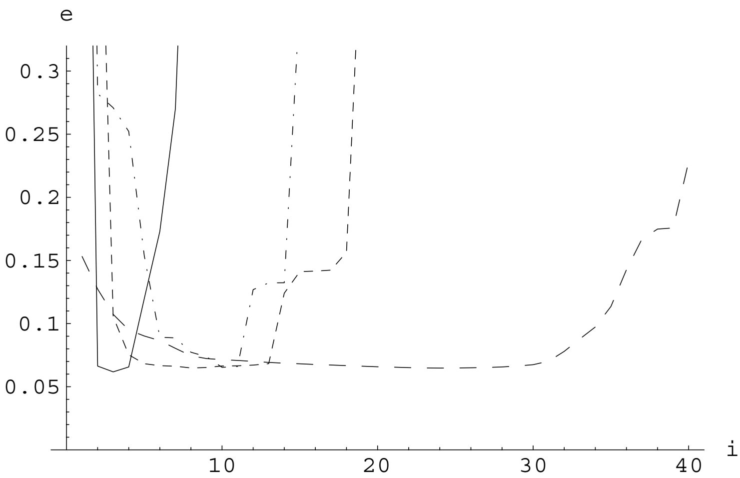

Fig. 2. *Example* 4.1: Error histories for non-preconditioned CG (long dashed line), PCG with the fit preconditioner (continuous line), PCG with the circulant preconditioner (short dashed line) and PCG with the inverse Toeplitz preconditioner (dot–dashed line).

Table 1
Example 4.1: Minimal error values and iteration numbers for non-preconditioned CG, PCG with the fit preconditioner, PCG with the circulant preconditioner and PCG with the inverse Toeplitz preconditioner

Method	Minimal error	Iteration
CG	0.0648	24
Fit preconditioner	0.0618	3
Circulant preconditioner	0.0648	8
Toeplitz preconditioner	0.0653	10

Example 4.2 (*Diffraction in incoherent illumination PSF*). The blurring matrix H_N is the N-section of the matrix H whose symbol is the symmetric Laurent polynomial (3) of degree $w = 8$ with coefficients $h_i = \beta(\sin(\alpha i)/i)^2$, $i = 0, \ldots, w$, where $\alpha = 1.5145$ and β is the normalizing constant. This matrix corresponds to the 1-D version of the PSF which describes the diffraction effects caused by a system of lenses in a spatially incoherent illumination [2]. For the chosen value of α the function $\widetilde{h}(\theta)$ is positive and the condition number of H_N is approximately 600. The noise vector $\boldsymbol{\eta}$ is scaled in such a way that $\|\boldsymbol{\eta}\|_2/\|H_N\mathbf{x}\|_2 = 10^{-3}$. The dimensions of the noise and signal subspaces are estimated to be $d_n = 7$ and $d_s = 40$, respectively. By checking the set $\mathscr{H}$, the estimate 0.3 is found for ϵ. The least-squared fit to $\psi(\theta)$ obtained for $\mu = 5$ and $\kappa = 20$ is

$$\widetilde{p}(\theta) = 0.548 + 0.333\,\cos\theta + 0.765\,\cos 2\theta + 0.0005\,\cos 3\theta + 0.023\,\cos 4\theta + 0.011\,\cos 5\theta$$

and the factor $q(z)$ results to be

$$q(z) = 0.704 + 0.220\,z + 0.547\,z^2 - 0.0046\,z^3 + 0.0138\,z^4 + 0.008\,z^5.$$

Fig. 3 shows the error histories of the non-preconditioned CG (long dashed line), the *PCG* method with the fit preconditioner (continuous line), the *PCG* method with the modified circulant preconditioner (short dashed line) and the inverse Toeplitz preconditioner (dot–dashed line). The value $\tau = 0.07$ for the circulant preconditioner and the value $\tau = 0.05$ for the inverse Toeplitz preconditioner have been used as truncation parameters (from the experiments these values result to be the best ones).

Table 2 summarizes the minimal value obtained and the iteration number required by the non-preconditioned CG, the PCG with the fit preconditioner, the PCG with the circulant preconditioner and the PCG with the inverse Toeplitz preconditioner.

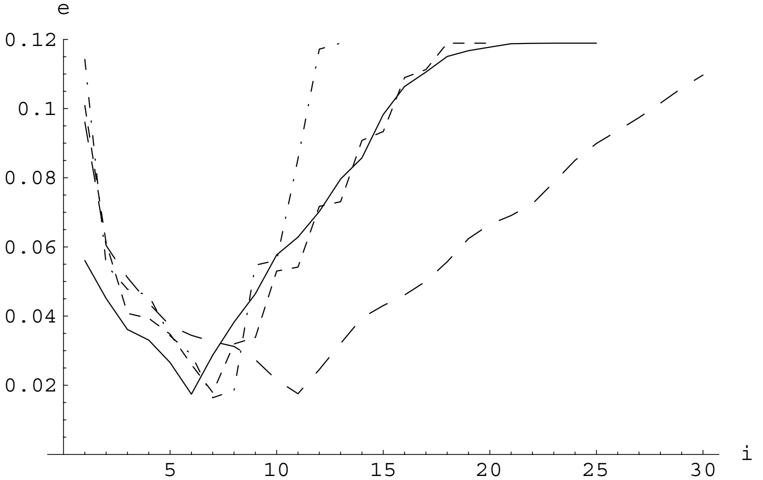

Fig. 3. *Example* 4.2: Error histories for non-preconditioner CG (long dashed line), PCG with the fit preconditioner (continuous line), PCG with the circulant preconditioner (short dashed line) and PCG with the inverse Toeplitz preconditioner (dot–dashed line).

Table 2
Example 4.2: Minimal error values and iteration numbers for non-preconditioned CG, PCG with the fit preconditioner, PCG with the circulant preconditioner and PCG with the inverse Toeplitz preconditioner

Method	Minimal error	Iteration
CG	0.0175	11
Fit preconditioner	0.0174	6
Circulant preconditioner	0.0179	7
Toeplitz preconditioner	0.0164	7

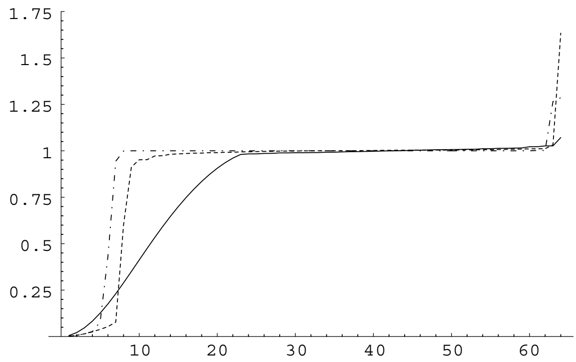

Fig. 4. Eigenvalues of the fit preconditioned matrix (continuous line), of the circulant preconditioned matrix (dashed line) and of the inverse Toeplitz preconditioned matrix (dot–dashed line).

By examining the previous examples (especially the second one), it appears that the values, which should be taken as truncation parameters by both the circulant preconditioner and the inverse Toeplitz preconditioner to enjoy the regularizing property, are much smaller than the value ϵ required by the fit preconditioner. This can be explained by the different way the preconditioners cluster the eigenvalues (see Fig. 4).

In fact the eigenvalues of the circulant preconditioned matrix (dashed line) and those of the inverse Toeplitz preconditioner (dot–dashed line) show two neatly separated clusters around 1 and around 0. All the eigenvalues smaller than τ are in the cluster around 0. On the contrary the two clusters of eigenvalues of the fit preconditioned matrix (continuous line) are connected by an intermediate increasing set. A too small value for ϵ would reduce the cluster around 0, letting some noise eigenvalues to escape from it. This different behavior must be taken into consideration in the 2-D case, as we will see later. As noted in [10], if it is not clear where the noise subspace ends, an overestimation of the noise subspace is better than an underestimation. Experiments conducted with higher values of the truncation parameters show a comparable performance from the point of view of the minimum error obtained. Of course the iteration number increases for higher truncation values.

5. The 2-D problem

We consider now the case where the coefficient matrix of the system is a band block Toeplitz matrix with band Toeplitz blocks. Such a matrix form is commonly called BTTB. We restrict ourselves to the separable case, that is, we consider the problem of approximating the vector $\mathbf{x}$ solution of

$$H_m\mathbf{x} = \mathbf{b} - \boldsymbol{\eta},$$

where $H_m = K_N \otimes J_N$ has size $m = N^2$, and K and J are symmetric positive definite band Toeplitz matrices, whose symbols are the Laurent polynomials $k(z)$ and $j(v)$ of degrees w_k and w_j, respectively. Hence H_m has symbol $h(z, v) = k(z)j(v)$.

A generalization of the classical Grenader and Szegö theorem states that the eigenvalues of H_m belong to the interval $\mathscr{I}_h$ having endpoints

$$h_{\inf} = \inf_{\theta} \widetilde{k}(\theta) \inf_{\eta} \widetilde{j}(\eta) \quad \text{and} \quad h_{\sup} = \sup_{\theta} \widetilde{k}(\theta) \sup_{\eta} \widetilde{j}(\eta),$$

where $\widetilde{k}(\theta) = k(\mathrm{e}^{\mathrm{i}\theta})$ and $\widetilde{j}(\eta) = j(\mathrm{e}^{\mathrm{i}\eta})$, with $\theta,\ \eta \in [-\pi, \pi]$, are the restrictions of $k(z)$ and $j(v)$ on the unit circle. Moreover, the eigenvalues $\lambda_i(H_m) = \lambda_r(K_N)\lambda_s(J_N)$, $r, s = 1, \ldots, N$, are equally distributed as $\widetilde{k}(\theta_r)\widetilde{j}(\eta_s)$, $\theta_r = 2\pi r/N$ and $\eta_s = 2\pi s/N$.

5.1. Construction of the fit preconditioner

As in the 1-D case the set $\mathscr{H}$ which approximates the set of eigenvalues of H_m is considered. It is formed by all the products $\widetilde{k}(\theta_r)\widetilde{j}(\eta_s)$, $r, s = 1, \ldots, N$. In the set $\mathscr{H}$ the two subsets $\mathscr{H}_n$ and $\mathscr{H}_s$, corresponding to the eigenvalues of the noise and of the signal subspaces, are detected. Two values ϵ_k and ϵ_j are determined in such a way that the set of the products $\widetilde{k}(\theta_r)\widetilde{j}(\eta_s)$ with $\widetilde{k}(\theta_r) > \epsilon_k$ and $\widetilde{j}(\eta_s) > \epsilon_j$ has approximately the cardinality of $\mathscr{H}_s$. Consequently the subsets of indices $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ for the function $\widetilde{k}(\theta)$ and $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ for the function $\widetilde{j}(\eta)$ are determined (see Section 3.1). Two functions $\psi(\theta)$ and $\varphi(\eta)$ are considered such that

$$\psi(\theta_i) = \begin{cases} \widetilde{k}(\theta_i) & \text{for } i \in \boldsymbol{\tau}_1, \\ \epsilon_k & \text{for } i \in \boldsymbol{\tau}_2, \end{cases} \qquad \varphi(\eta_i) = \begin{cases} \widetilde{j}(\eta_i) & \text{for } i \in \boldsymbol{\rho}_1, \\ \epsilon_j & \text{for } i \in \boldsymbol{\rho}_2. \end{cases}$$

Two polynomials $p(z)$ of degree μ_p and $r(v)$ of degree μ_r are found, with $\widetilde{p}(\theta)$ and $\widetilde{r}(\eta)$ approximating $\psi(\theta)$ and $\varphi(\eta)$, respectively. The factors $q(z)$ and $s(v)$ of $p(z)$ and $r(v)$, respectively, are computed and the associated band lower triangular Toeplitz matrices Q and S are used for the preconditioner.

The matrix $V_N V_N^{\mathrm{T}}$, with $V_N = Q_N \otimes S_N$ is called *fit preconditioner*. Matrix V_N is a block banded lower triangular matrix with blocks of the same structure. We have

$$(Q_N^{-\mathrm{T}} Q_N^{-1} K_N) \otimes (S_N^{-\mathrm{T}} S_N^{-1} J_N) = V_N^{-\mathrm{T}} V_N^{-1} (K_N \otimes J_N).$$

Hence the cost of an iteration is $(2(\mu_p + \mu_r + w_k + w_j) + 5)N^2$ operations.

Theorem 2. *Let $V_N V_N^{\mathrm{T}}$ be the fit preconditioner of the matrix $K_N \otimes J_N$. Let $\mathscr{I}_f$ be the interval having endpoints*

$$f_{\inf} = \inf_{\theta} \left[\widetilde{p}^{-1}(\theta) \widetilde{k}(\theta) \right] \inf_{\eta} \left[\widetilde{r}^{-1}(\eta) \widetilde{j}(\eta) \right],$$

$$f_{\sup} = \sup_{\theta} \left[\widetilde{p}^{-1}(\theta) \widetilde{k}(\theta) \right] \sup_{\eta} \left[\widetilde{r}^{-1}(\eta) \widetilde{j}(\eta) \right].$$

Then at most $\mathrm{O}(N)$ *of the eigenvalues of the preconditioned matrix lie outside $\mathscr{I}_f$.*

The proof follows immediately by applying to the block case the same considerations developed in Theorem 1.

6. Experiments with the 2-D problem

We apply the fit preconditioner to two different 2-D images. Having checked that the 2-level circulant preconditioner and the inverse Toeplitz preconditioner have similar performances, in this case, we compare the fit preconditioner only with the circulant preconditioner.

Example 6.1 (*Gaussian PSF*). For this example a 256×256 image of a letter by Galileo has been blurred with the separable 2-D Gaussian PSF (see Fig. 5 for the original image and Fig. 6 for the blurred noisy image). The blurring matrix is $H_m = K_N \otimes J_N$, where both K_N and J_N are N-sections of Gaussian matrices with bandwidth $w = 8$. K_N corresponds to the symbol having coefficients $k_i = \beta \exp(-\alpha i^2)$, $i = 0, \ldots, w$, with $\alpha = 0.2$, and J_N corresponds to the symbol having coefficients $j_i = \beta \exp(-\alpha i^2)$, $i = 0, \ldots, w$, with $\alpha = 0.5$. In both cases β is the normalizing constant. The condition number of H_m is about 7.8×10^6. The noise vector $\boldsymbol{\eta}$ is scaled in such a way that $\|\boldsymbol{\eta}\|_2 / \|H_m \mathbf{x}\|_2 = 10^{-3.5}$. The dimensions of the noise and signal subspaces are estimated to be $d_n = 25\,000$ and $d_s = 10\,000$. By checking the set $\mathscr{H}$, the values $\epsilon_k = 0.2$ and $\epsilon_j = 0.4$ are determined.

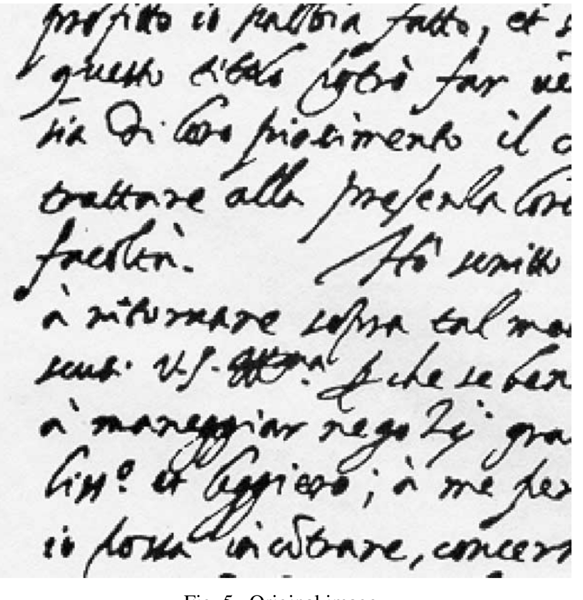

Fig. 5. Original image.

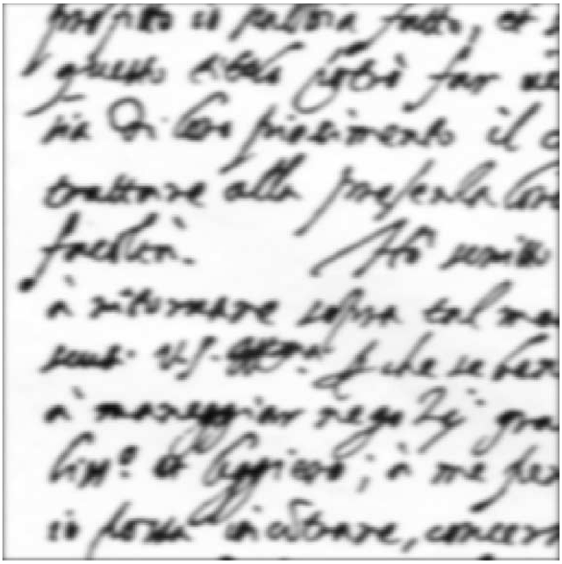

Fig. 6. Blurred image.

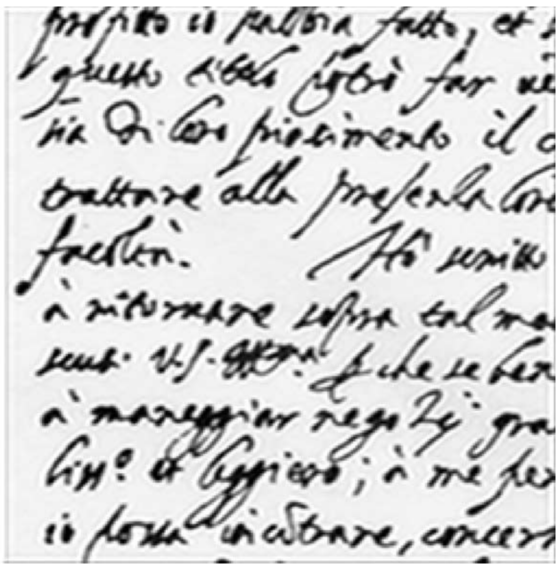

Fig. 7. Fit PCG 1^0 iteration.

Fig. 9 shows the error histories of the non-preconditioned CG (long dashed line), the *PCG* method with the fit preconditioner (continuous line) and the *PCG* method with the modified circulant preconditioner (short dashed line). The value $\tau =$

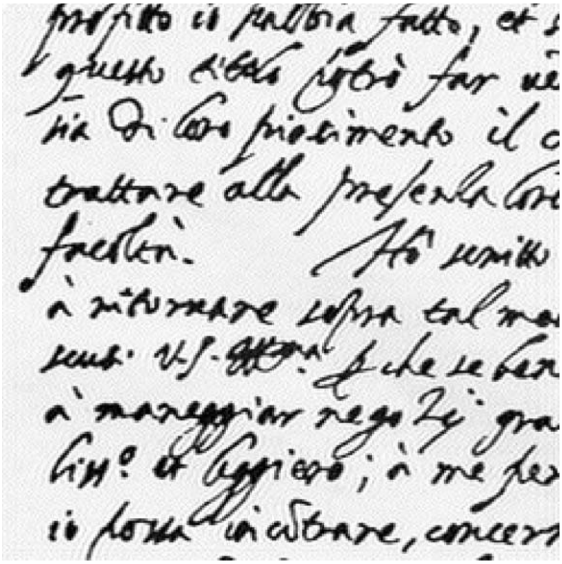

Fig. 8. Fit PCG 5^0 iteration.

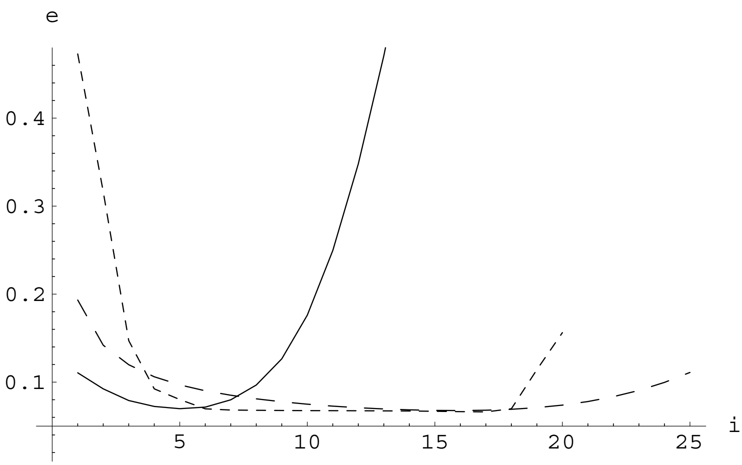

Fig. 9. *Example* 6.1: Error histories for non-preconditioned CG (long dashed line), PCG with the fit preconditioner (continuous line) and PCG with the circulant preconditioner (short dashed line).

0.004 (which from the experiments results to be the best one) has been used as the truncation parameter for the circulant preconditioner.

Figs. 7 and 8 show the images reconstructed after the first and the fifth iterations of PCG with the fit preconditioner. Table 3 summarizes the minimal value obtained and the iteration number required by the non-preconditioned CG, the PCG with the fit preconditioner and the PCG with the circulant preconditioner.

Table 3
Example 6.1: Minimal error values and iteration numbers for non-preconditioned CG, PCG with the fit preconditioner and PCG with the circulant preconditioner

Method	Minimal error	Iteration
CG	0.068	16
Fit preconditioner	0.070	5
Circulant preconditioner	0.048	6

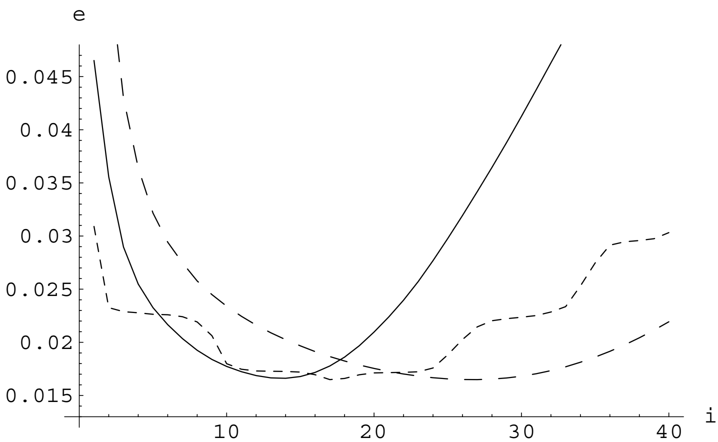

Fig. 10. *Example* 6.2: Error histories for non-preconditioned CG (long dashed line), PCG with the fit preconditioner (continuous line) and PCG with the circulant preconditioner (short dashed line).

Example 6.2 (*Diffraction in incoherent illumination PSF*). For this example a 110×110 image of Saturn has been blurred with the separable 2-D PSF of the diffraction in incoherent illumination (see Fig. 11 for the original image and Fig. 12 for the blurred noisy image). The blurring matrix is $H_m = K_N \otimes K_N$, where K_N is the N-section of the matrix K with bandwidth $w = 8$, whose symbol is the Laurent polynomial considered in Example 4.2, that is, $h(z, v) = k(z)k(v)$, where the coefficients of k are $k_i = \beta(\sin(\alpha i)/i)^2,\ i = 0, \dots, w,\ w = 8$, with $\alpha = 1.5145$ and β is the normalizing constant. The condition number of H_m is about 3.6×10^4. The noise vector $\boldsymbol{\eta}$ is scaled in such a way that $\|\boldsymbol{\eta}\|_2/\|H_m\mathbf{x}\|_2 = 10^{-4}$. The dimensions of the noise and signal subspaces are estimated to be $d_n = 4000$ and $d_s = 3000$. By checking the set $\mathscr{H}$, the value $\epsilon_k = 0.5$ is determined.

Fig. 10 shows the error histories of the non-preconditioned CG (long dashed line), the *PCG* method with the fit preconditioner (continuous line) and the *PCG* method with the modified circulant preconditioner (short dashed line). The value $\tau = 0.01$ (which from the experiments results to be the best one) has been used as the truncation parameter for the circulant preconditioner.

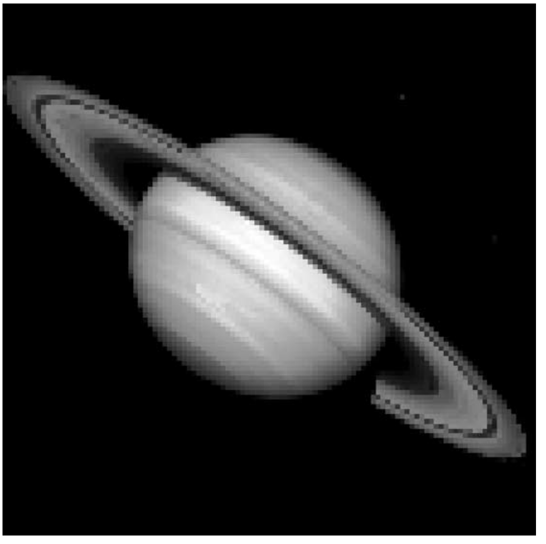

Fig. 11. Original image.

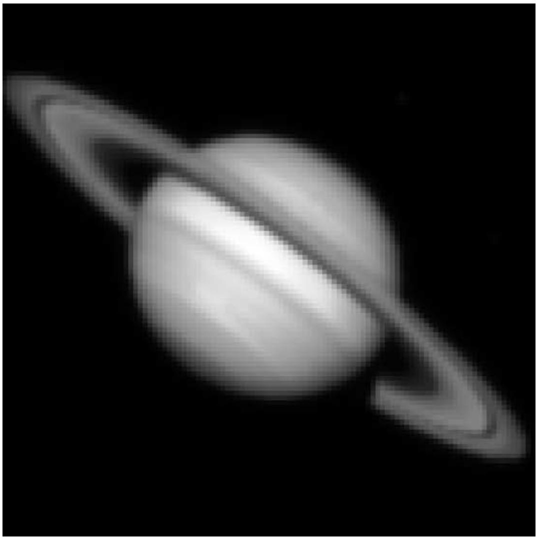

Fig. 12. Blurred image.

Figs. 13 and 14 show the images reconstructed after the first and the 14th iterations of PCG with the fit preconditioner. Table 4 summarizes the minimal value

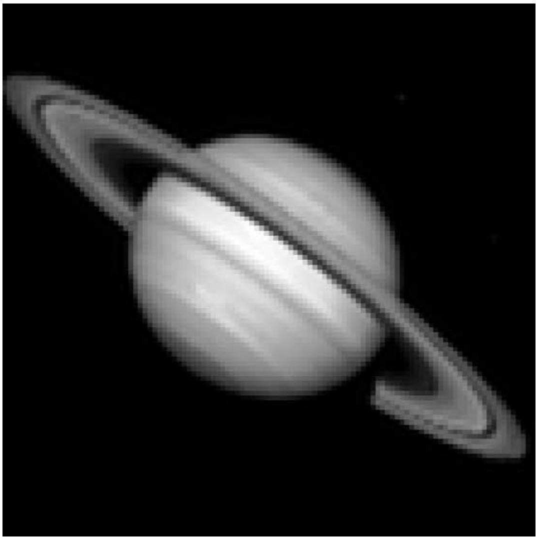

Fig. 13. Fit PCG 1^0 iteration.

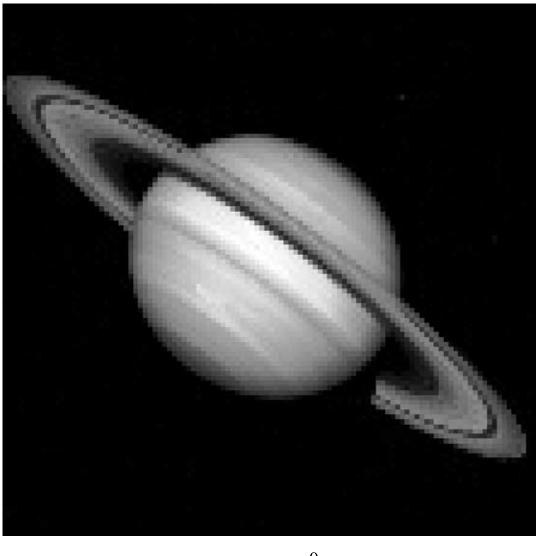

Fig. 14. Fit PCG 14^0 iteration.

obtained and the iteration number required by the non-preconditioned CG, the PCG with the fit preconditioner and the PCG with the circulant preconditioner.

Table 4
Example 6.2: Minimal error values and iteration numbers for non-preconditioned CG, PCG with the fit preconditioner, PCG with the circulant preconditioner and PCG with the inverse Toeplitz preconditioner

Method	Minimal error	Iteration
CG	0.0165	27
Fit preconditioner	0.0166	14
Circulant preconditioner	0.0165	17

References

[1] R. Barrett, M. Berry, T.F. Chan, J. Demmel, J.M. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, H. Van der Vorst, TEMPLATES for the Solution of Linear Systems: Building Blocks for Iterative Methods, SIAM, Philadelphia, PA, 1993.

[2] M. Bertero, P. Boccacci, Introduction to Inverse Problems in Imaging, Institute of Physics Publishing, 1998.

[3] R.H. Chan, W.K. Ching, Iterative methods for queueing models with batch arrivals, in: W.J. Stewart (Ed.), Computations with Markov Chains, Kluwer Academic Publishers, Norwell, MA, 1995, pp. 81–93.

[4] R.H. Chan, M.K. Ng, A.M. Yip, A survey of preconditioners for ill-conditioned Toeplitz systems, in: V. Olshevsky (Ed.), Structured Matrices in Mathematics, Computer Science and Engineering II: Part II, CONM/281, AMS, Providence, RI, 2001.

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

[6] U. Grenander, G. Szegö, Toeplitz Forms and their Applications, second ed., Chelsea, New York, 1984.

[7] T.N.T. Goodman, C.A. Micchelli, G. Rodriguez, S. Seatzu, Spectral factorization of Laurent polynomials, Adv. Comput. Math. 7 (1997) 429–454.

[8] M. Hanke, Iterative regularization techniques in image restoration, Mathematical Methods in Inverse Problems for Partial Differential Equations, Springer, Berlin, 1998.

[9] M. Hanke, J. Nagy, Inverse Toeplitz preconditioner for Ill-posed problems, Linear Algebra Appl. 284 (1994) 137–156.

[10] M. Hanke, J. Nagy, R. Plemmons, Preconditioned iterative regularization for ill-posed problems, in: L. Reichel, A. Ruttan, R.S. Varga (Eds.), Numerical Linear Algebra and Scientific Computing, De Gruyter, Berlin, 1993, pp. 141–163.

[11] T.K. Ku, C.C. Jay Kuo, A minimum-phase LU factorization preconditioner for Toeplitz matrices, SIAM J. Sci. Statist. Comput. 13 (1992) 1470–1487.

[12] C. van der Mee, G. Rodriguez, S. Seatzu, LDU factorization results for bi-infinite and semi-infinite scalar and block Toeplitz matrices, Calcolo 33 (1996) 307–335.

[13] M. Miranda, P. Tilli, Asymptotic spectra of Hermitian block Toeplitz matrices and preconditioning results, SIAM J. Matrix Anal. Appl. 21 (3) (2000) 867–881.

[14] B.N. Parlett, The Symmetric Eigenvalue Problem, SIAM's Classics in Applied Mathematics vol. 20, 1997, SIAM, Philadelphia, PA.

[15] S. Serra, Preconditioning strategies for asymptotically ill-conditioned block Toeplitz systems, BIT 34 (1994) 579–594.

[16] S. Serra, Optimal quasi-optimal and superlinear band-Toeplitz preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems, Math. Comp. 66 (1997) 651–665.

[17] A. van der Sluis, H. van der Vorst, The rate of convergence of conjugate gradients, Numer. Math. 48 (1986) 543–560.