

# A GLOBAL LANCZOS METHOD FOR IMAGE RESTORATION

A. H. BENTBIB\*, M. EL GUIDE†, K. JBILOU‡, AND L. REICHEL§

**Abstract.** Image restoration often requires the solution of large linear systems of equations with a very ill-conditioned, possibly singular, matrix and an error-contaminated right-hand side. The latter represents the available blur and noise-contaminated image, while the matrix models the blurring. Computation of a meaningful restoration of the available image requires the use of a regularization method. We consider the situation when the blurring matrix has a Kronecker product structure and an estimate of the norm of the desired image is available, and illustrate that efficient restoration of the available image can be achieved by Tikhonov regularization based on the global Lanczos method, and by using the connection of the latter to Gauss-type quadrature rules.

**1. Introduction.** Image restoration is the process of removing blur and noise from an available degraded image to recover an approximation of the unavailable original noise- and blur-free image. Let the original image be represented by the matrix  $\widehat{X} \in \mathbb{R}^{m \times n}$ , whose entries represent pixels. The operator  $\text{vec}$  maps  $\widehat{X}$  to the vector  $\widehat{x} \in \mathbb{R}^{mn}$  defined by stacking the columns of  $\widehat{X}$  from left to right. Blurring is a linear deterministic process which can be modeled by a matrix  $H \in \mathbb{R}^{mn \times mn}$ ; see, e.g., [2, 3, 7, 14, 15] for discussions on image restoration. Thus,  $\widehat{b} = H\widehat{x}$  represents a blurred image associated with  $\widehat{x}$ . The available image, represented by  $b \in \mathbb{R}^{mn}$ , is assumed to be contaminated by both blur and noise. Let the entries of the “noise vector”  $e \in \mathbb{R}^{mn}$  be normally distributed with zero mean and variance  $\delta^2$ . Then  $b$  can be expressed as

$$b = H\widehat{x} + e. \quad (1.1)$$

The aim of image restoration is to determine an approximation of  $\widehat{x}$  by computing an approximate solution of the linear system of equations

$$Hx = b. \quad (1.2)$$

If the system is inconsistent, then we consider it a least-squares problem. Blurring matrices  $H$  typically are very ill-conditioned and may be singular. Since  $b$  is contaminated by the error  $e$ , straightforward solution of (1.2) generally does not yield a useful approximation of  $\widehat{x}$  due to severe propagation of the error  $e$  into the computed solution. Let  $H^\dagger$  denote the Moore–Penrose pseudoinverse of  $H$  and let  $\|\cdot\|_2$  be the Euclidean vector norm. Then the least-squares solution of minimal Euclidean norm of (1.2) is given by  $H^\dagger b$  and, typically,

$$\|\widehat{x}\|_2 \ll \|H^\dagger b\|_2 \approx \|H^\dagger e\|_2.$$

This difficulty can be remedied by replacing (1.2) by a nearby problem, whose solution is less sensitive to the error in  $b$ . The replacement is referred to as regularization. One of the most popular regularization methods is due to Tikhonov [8, 12]. In its simplest

---

\*Laboratory LAMAI, FST Marrakesh, University of Cadi Ayyad, Morocco. E-mail: a.bentbib@uca.ma

†Laboratory LAMAI, FST Semlalia, University of Cadi Ayyad, Marrakesh, Morocco. E-mail: medelguide@gmail.com

‡Université de Lille Nord de France, L.M.P.A, ULCO, 50 rue F. Buisson, BP699, F-62228 Calais-Cedex, France. E-mail: jbilou@lmpa.univ-littoral.fr

§Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: reichel@math.kent.edu. Research supported in part by NSF grant DMS-1115385.

form, Tikhonov regularization replaces the linear system (1.2) by the penalized least-squares problem

$$\min_{x \in \mathbb{R}^{m_n}} \{\|Hx - b\|_2^2 + \mu \|x\|_2^2\}, \quad (1.3)$$

where  $\mu > 0$  is a regularization parameter. The normal equations associated with (1.3) are given by

$$(H^T H + \mu I)x = H^T b, \quad (1.4)$$

where  $I$  is the identity matrix of suitable order and the superscript  $T$  denotes transposition. It follows that (1.3) has the unique solution

$$x_\mu := (H^T H + \mu I)^{-1} H^T b \quad (1.5)$$

for any  $\mu > 0$ . Our computed restoration of the vector  $b$  will be an approximation of a vector of the form (1.5).

**PROPOSITION 1.1.** *Assume that  $H^T b \neq 0$ . Then  $\|x_\mu\|_2$  is a monotonically decreasing function of  $\mu$  and*

$$\lim_{\mu \searrow 0} \|x_\mu\|_2 = \|H^\dagger b\|_2, \quad \lim_{\mu \rightarrow \infty} \|x_\mu\|_2 = 0. \quad (1.6)$$

*Proof.* The representation

$$\|x_\mu\|_2^2 = x_\mu^T x_\mu = b^T H (H^T H + \mu I)^{-2} H^T b \quad (1.7)$$

shows that  $\|x_\mu\|_2$  is a decreasing function of  $\mu > 0$ . The right-hand side limit (1.6) is immediate; the left-hand side limit follows by substituting the singular value decomposition of  $H$  into (1.7).  $\square$

The quality of the computed restoration depends on the choice of  $\mu > 0$ . Several approaches to choosing  $\mu$  are described in the literature, including the discrepancy principle, generalized cross validation, and the L-curve; see, e.g., [19, 21] for overviews and discussions. In this paper, we will assume an estimate of  $\|\hat{x}\|_2$  to be available, and we will use this estimate to determine  $\mu$ . Knowledge of the norm of the desired approximate solution of linear systems (1.2) is available in some applications, see, e.g., Ahmad et al. [1], and this approach to select  $\mu$  has received considerable attention in the literature; see, e.g., [5, 6, 22, 23]. It works well when the relative error  $\|e\|_2 / \|\hat{b}\|_2$  is not too large; see [6] for illustrations.

Many blurring matrices can be represented, or be well approximated, by a Kronecker product; see, e.g., Kamm and Nagy [17, 18] and Van Loan and Pitsianis [25]. We will assume  $H$  to have a Kronecker product structure and present an efficient algorithm for approximating  $x_\mu$  by exploiting this structure. We will show that the Kronecker product structure allows the application of the global Lanczos method. This method replaces the evaluation of the matrix-vector products of the standard Lanczos method by evaluation of matrix-matrix products, which can be executed efficiently on many modern computers. The global Lanczos method is described in [16]. We determine a suitable value of the regularization parameter  $\mu$  by estimating a certain Stieltjes integral by Gauss-type quadrature rules. These rules can be evaluated with the aid of the global Lanczos method. This paper extends the method for determining  $\mu$  described in [6] for the standard Lanczos method to the global Lanczos method.

The organization of this paper is as follows. Section 2 discusses the Kronecker product structure of  $H$  and Section 3 describes how this structure allows the application of the global Lanczos method for solving the normal equations (1.4). The section also shows how upper and lower bounds for  $\|x_\mu\|_2^2$  can be determined by using the relation between the global Lanczos method and certain Gauss-type quadrature rules. This extends techniques developed by Golub and Meurant [9, 10] based on the relation between the standard Lanczos method and Gauss-type quadrature rules to the global Lanczos method. An algorithm for image restoration is presented in Section 4, and Section 5 describes a few computed examples. These examples illustrate the benefit of exploiting the Kronecker product structure of  $H$ . Concluding remarks can be found in Section 6.

**2. Kronecker structure.** We will assume that the blurring matrix  $H$  is the Kronecker product of the matrices  $H_2 \in \mathbb{R}^{m \times m}$  and  $H_1 = [h_{i,j}^{(1)}] \in \mathbb{R}^{n \times n}$ , i.e.,

$$H = H_1 \otimes H_2 = \begin{bmatrix} h_{1,1}^{(1)}H_2 & h_{1,2}^{(1)}H_2 & \cdots & h_{1,n}^{(1)}H_2 \\ h_{2,1}^{(1)}H_2 & h_{2,2}^{(1)}H_2 & \cdots & h_{2,n}^{(1)}H_2 \\ \vdots & \vdots & \ddots & \vdots \\ h_{n,1}^{(1)}H_2 & h_{n,2}^{(1)}H_2 & \cdots & h_{n,n}^{(1)}H_2 \end{bmatrix}. \quad (2.1)$$

Let the operator  $\text{vec}$  transform a matrix  $A = [a_{i,j}] \in \mathbb{R}^{m \times n}$  to a vector  $a \in \mathbb{R}^{mn}$  by stacking the columns of  $A$  from left to right, i.e.,

$$a = [a_{1,1}, a_{2,1}, \dots, a_{m,1}, a_{1,2}, a_{2,2}, \dots, a_{m,2}, \dots, a_{m,n}]^T. \quad (2.2)$$

We also need the inverse operator,  $\text{mat}$ , which transforms a vector (2.2) to an associated matrix  $A = [a_{i,j}] \in \mathbb{R}^{m \times n}$ . Thus,

$$\text{vec}(A) = a, \quad \text{mat}(a) = A.$$

The Kronecker product satisfies the following relations for matrices  $A, B, C, D, X$  of suitable sizes:

$$\left. \begin{aligned} (A \otimes B)\text{vec}(X) &= \text{vec}(BXA^T), \\ (A \otimes B)^T &= A^T \otimes B^T, \\ (AB) \otimes (CD) &= (A \otimes C)(B \otimes D). \end{aligned} \right\} \quad (2.3)$$

For  $A, B \in \mathbb{R}^{m \times n}$ , we define the inner product

$$\langle A, B \rangle_F := \text{tr}(A^T B), \quad (2.4)$$

where  $\text{tr}(\cdot)$  denotes the trace. We note for future reference that

$$\langle A, B \rangle_F = (\text{vec}(A))^T \text{vec}(B). \quad (2.5)$$

The Frobenius norm is associated with this inner product,

$$\|A\|_F := \langle A, A \rangle_F^{1/2}.$$

It satisfies

$$\|A\|_F = \|\text{vec}(A)\|_2. \quad (2.6)$$

Let  $H_1$  and  $H_2$  be the Kronecker factors of the blurring matrix (2.1) and introduce the linear operator

$$\begin{aligned}\mathcal{A} : \mathbb{R}^{m \times n} &\rightarrow \mathbb{R}^{m \times n} \\ \mathcal{A}(X) &= H_2 X H_1^T.\end{aligned}$$

Its transpose is given by  $\mathcal{A}^T(X) = H_2^T X H_1$ . Define the symmetric linear operator

$$\check{\mathcal{A}}(X) = (\mathcal{A}^T \circ \mathcal{A})(X),$$

where  $\circ$  denotes composition.

**PROPOSITION 2.1.** *Let  $H$  have the Kronecker structure (2.1) and assume that  $G := \mathcal{A}^T(B) \neq O$ . Let  $\mu > 0$ . Then the equation*

$$(\check{\mathcal{A}} + \mu I)(X) = G \tag{2.7}$$

has a unique solution  $X_\mu \in \mathbb{R}^{m \times n}$ . Let  $b := \text{vec}(B)$  and let  $x_\mu$  be given by (1.5) with this vector  $b$ . Then

$$X_\mu = \text{mat}(x_\mu). \tag{2.8}$$

Moreover,  $\|X_\mu\|_F$  is a decreasing function of  $\mu > 0$  with

$$\lim_{\mu \searrow 0} \|X_\mu\|_F = \|H^\dagger b\|_2, \quad \lim_{\mu \rightarrow \infty} \|X_\mu\|_F = 0. \tag{2.9}$$

*Proof.* We have  $\check{\mathcal{A}}(X) = H_2^T H_2 X H_1^T H_1$ . Therefore (2.7) can be written as

$$H_2^T H_2 X H_1^T H_1 + \mu X = H_2^T B H_1.$$

Using the properties (2.3), this equation can be expressed as

$$((H_1 \otimes H_2)^T (H_1 \otimes H_2) + \mu I) \text{vec}(X) = (H_1 \otimes H_2)^T \text{vec}(B),$$

which is the same as (1.4). This establishes (2.8). The properties of  $\|X_\mu\|_F$  now follow from Proposition 1.1.  $\square$

The matrix  $\hat{X} = \text{mat}(\hat{x})$  represents the unavailable blur- and noise-free image that we would like to determine. We assume that an estimate of

$$\Delta := \|\hat{X}\|_F \tag{2.10}$$

is known and will compute an approximation of  $\hat{X}$  by solving the constrained least-squares problem

$$\min_{\|X\|_F \leq \Delta} \|B - \mathcal{A}(X)\|_F. \tag{2.11}$$

The solution of this minimization problem without constraint is equivalent to solving (1.2) when  $H$  is nonsingular. The solution of the unconstrained problem typically is of very large norm due to contamination by propagated error; see Proposition 2.2 below. If this solution has Frobenius norm larger than or equal to  $\Delta$ , then the solution of (2.11) satisfies  $\|X\|_F = \Delta$ . We will assume this to be the case. This solution is described by the following proposition.

PROPOSITION 2.2. *Assume that the solution of the minimization problem (2.11) without constraint is of norm larger than  $\Delta$ . Then the solution of (2.11) is the solution of (2.7) for some  $\mu > 0$ .*

*Proof.* By (2.8) the solution  $X_\mu$  of (2.7) is equivalent to the solution  $x_\mu$  of (1.3) given by (1.5). A proof that  $x_\mu$  is the unique solution of

$$\min_{\|x\|_2=\Delta} \|b - Hx\|_2 \quad (2.12)$$

for a suitable value of  $\mu > 0$  can be established with the aid of Lagrange multipliers; see Golub and von Matt [11]. It is easy to see that  $x_\mu$  also is a solution when the constraint in (2.12) is replaced by  $\|x\|_2 \leq \Delta$ . The problem so obtained is equivalent to (2.11).  $\square$

We remark that if the solution of (2.11) without constraint has Frobenius norm smaller than or equal to  $\Delta$ , then we may choose the regularization parameter  $\mu = 0$ .

**3. The global Lanczos algorithm and Gauss quadrature.** Let  $X_\mu$  denote the solution of (2.7) for  $\mu > 0$ . Introduce the function

$$\phi(\mu) := \|X_\mu\|_F^2. \quad (3.1)$$

We will approximate  $\phi$  to be able to determine an estimate of  $\|X_\mu\|_F$  inexpensively. Our approximation is obtained by expressing (3.1) as a Stieltjes integral, and exploiting the connection between the global Lanczos method and Gauss-type quadrature rules.

PROPOSITION 3.1. *Introduce for  $\mu > 0$  the function*

$$f_\mu(t) := (t + \mu)^{-2}. \quad (3.2)$$

*Then (3.1) can be expressed as*

$$\phi(\mu) = \int f_\mu(t) d\omega(t), \quad (3.3)$$

*where  $d\omega$  is a measure with support on the nonnegative real axis. Moreover,  $\phi$  is decreasing and convex for  $\mu > 0$  with*

$$\lim_{\mu \searrow 0} \phi(\mu) = \|H^\dagger b\|_2^2, \quad \lim_{\mu \rightarrow \infty} \phi(\mu) = 0. \quad (3.4)$$

*Proof.* It follows from (3.1), (2.8), (2.6), and (1.5) that

$$\phi(\mu) = \|x_\mu\|_2^2 = x_\mu^T x_\mu = b^T H(H^T H + \mu I)^{-2} H^T b. \quad (3.5)$$

Substituting the spectral factorization

$$H^T H = U \Lambda U^T,$$

where  $\Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_{mn}] \in \mathbb{R}^{mn \times mn}$  and  $U \in \mathbb{R}^{mn \times mn}$  is orthogonal, into the right-hand side of (3.5), with  $w = [w_1, w_2, \dots, w_{mn}]^T := U^T H^T b$ , gives

$$\phi(\mu) = \sum_{j=1}^{mn} f_\mu(\lambda_j) w_j^2.$$

The right-hand side is a Stieltjes integral, which can be expressed as (3.3). The distribution function  $\omega$  associated with the measure  $d\omega$  can be chosen as a nondecreasing piecewise constant function with nonnegative jumps  $w_j^2$  at the eigenvalues  $\lambda_j$ . Since  $H^T H$  is positive semidefinite, the support of the measure  $d\omega$  lives on the nonnegative real axis.

It follows from (3.5) that the derivatives of  $\phi$  satisfy  $\phi'(\mu) < 0$  and  $\phi''(\mu) > 0$ , which shows that  $\phi$  is decreasing and convex. The limits (3.4) are a consequence of (2.9).  $\square$

It is convenient to define the integral operator

$$\mathcal{I}f := \int f_\mu(t) d\omega(t) \quad (3.6)$$

and the associated inner product

$$[p, q] := \mathcal{I}(pq) \quad (3.7)$$

for polynomials  $p$  and  $q$  of low enough degree. We may define orthogonal polynomials with respect to this inner product and, therefore, also Gauss quadrature rules for the approximation of (3.6). The  $k$ -point Gauss quadrature rule  $\mathcal{G}_k$  is characterized by the property that

$$\mathcal{G}_k p = \mathcal{I}p, \quad \forall p \in \mathbb{P}_{2k-1},$$

where  $\mathbb{P}_{2k-1}$  denotes the set of polynomials of degree at most  $2k-1$ . The remainder term for this quadrature rule can be expressed as

$$\mathcal{I}f_\mu - \mathcal{G}_k f_\mu = \frac{f_\mu^{(2k)}(\xi)}{(2k)!} \int \prod_{j=1}^k (t - t_j)^2 d\omega(t),$$

where  $f_\mu^{(\ell)}$  denotes the derivative of order  $\ell$  of the function (3.2),  $\xi$  is in the convex hull of the support of the measure  $d\omega$ , and  $t_1, t_2, \dots, t_k$  are the nodes of  $\mathcal{G}_k$ . Since  $f_\mu^{(2k)}(\xi)$  is positive, it follows that  $\mathcal{G}_k f < \mathcal{I}f$ .

Let  $\mathcal{R}_{k+1,0}$  denote the  $(k+1)$ -point Gauss–Radau rule for the measure  $d\omega$  with a fixed node  $t_0 = 0$ . Then

$$\mathcal{R}_{k+1,0} p = \mathcal{I}p, \quad \forall p \in \mathbb{P}_{2k}. \quad (3.8)$$

The quadrature error is

$$\mathcal{I}f_\mu - \mathcal{R}_{k+1,0} f_\mu = \frac{f_\mu^{(2k+1)}(\tilde{\xi})}{(2k+1)!} \int (t - t_0) \prod_{j=1}^k (t - \tilde{t}_j)^2 d\omega(t),$$

where  $\tilde{\xi}$  is in the convex hull of the support of the measure  $d\omega$  and the origin, and  $\tilde{t}_1, \tilde{t}_2, \dots, \tilde{t}_k$  are the “free” Gauss–Radau nodes; see, e.g., [9, 10]. In view of that  $f_\mu^{(2k+1)}(\xi)$  is negative and the support of  $d\omega$  lives on the nonnegative real axis,  $\mathcal{R}_{k+1,0} f$  is an upper bound for  $\mathcal{I}f$ . Indeed, one can show that

$$\mathcal{G}_{k-1} f_\mu < \mathcal{G}_k f_\mu < \mathcal{I}f < \mathcal{R}_{k+1,0} f_\mu < \mathcal{R}_{k,0} f_\mu, \quad (3.9)$$

see, e.g., [20] for details.

We conclude that pairs of Gauss and Gauss–Radau quadrature rules  $\mathcal{G}_k f_\mu$  and  $\mathcal{R}_{k+1,0} f_\mu$  yield lower and upper bounds for  $\phi(\mu)$ . We will now describe how these quadrature rules can be evaluated by carrying out  $k$  steps of the global Lanczos method without explicit knowledge of the measure  $d\omega$ .

**3.1. The global Lanczos method.** The global Lanczos method can be applied to reduce a large symmetric matrix to a small tridiagonal one; see [16]. This method differs from the standard Lanczos method in that it uses the inner product (2.4) between matrices. This makes it possible to use matrix-matrix products during the execution of the global Lanczos method. These products can be evaluated efficiently on many computers. Algorithm 1 below executes  $k$  steps of the global Lanczos method applied to  $\check{\mathcal{A}}$  with initial matrix  $V_1 = G/\|G\|_F$  of unit Frobenius norm with  $G = \mathcal{A}^T(B)$ .

**Algorithm 1. Global Lanczos method.**

1. Let  $\beta_1 = 0$ ,  $V_0 = O$ , and  $V_1 = G/\|G\|_F$ .
2. for  $j = 1, 2, \dots, k$ 
  - (a)  $W = \check{\mathcal{A}}(V_j) - \beta_j V_{j-1}$ ,
  - (b)  $\alpha_j = \langle V_j, W \rangle_F$ ,
  - (c)  $\check{W} = W - \alpha_j V_j$ ,
  - (d)  $\beta_{j+1} = \|\check{W}\|_F$ ,
  - (e)  $V_{j+1} = \check{W}/\beta_{j+1}$ ,
3. endfor

The zero matrix  $V_0 \in \mathbb{R}^{m \times n}$  does not have to be stored. We assume that all coefficients  $\beta_{j+1}$  are positive. This is the generic situation. Otherwise, the algorithm breaks down, because the recursions cannot be continued. This is a very rare event. We therefore will not dwell on it further. Thus, under the assumption that all generated coefficients  $\beta_{j+1}$  are positive, the algorithm determines the matrices

$$\mathcal{V}_k = [V_1, V_2, \dots, V_k] \in \mathbb{R}^{m \times kn}, \quad \mathcal{V}_{k+1} = [V_1, V_2, \dots, V_{k+1}] \in \mathbb{R}^{m \times (k+1)n},$$

whose ‘‘matrix columns’’  $V_j$  are F-orthonormal, i.e.,

$$\langle V_i, V_j \rangle_F = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases} \quad (3.10)$$

It follows from the recursion formulas of Algorithm 1 and the definition of  $V_1$  that

$$V_j = p_{j-1}(\check{\mathcal{A}})(G), \quad j = 1, 2, \dots, k+1, \quad (3.11)$$

for some polynomials  $p_{j-1}$  of precisely degree  $j-1$ . We refer to these polynomials as global Lanczos polynomials. They satisfy the same recursion relations as the matrix columns  $V_j$ ; see Section 3.2 below. Their explicit form is not required, only their existence. The property (3.11) together with the F-orthonormality (3.10) of the matrix columns shows that the set  $\{V_j\}_{j=1}^{k+1}$  forms an F-orthonormal basis for the global Krylov subspace

$$\mathcal{K}_{k+1}(\check{\mathcal{A}}, G) := \text{span}\{G, \check{\mathcal{A}}(G), \dots, \check{\mathcal{A}}^k(G)\},$$

where  $\check{\mathcal{A}}^j = \check{\mathcal{A}}^{j-1} \circ \check{\mathcal{A}}$  for  $j = 2, 3, \dots, k$ .

The coefficients  $\alpha_1, \alpha_2, \dots, \alpha_k$  and  $\beta_2, \beta_3, \dots, \beta_k$  determined by Algorithm 1 define the symmetric tridiagonal matrix

$$T_k = \begin{bmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & \beta_k & \alpha_k \end{bmatrix} \in \mathbb{R}^{k \times k}, \quad (3.12)$$

and the recurrence formulas of Algorithm 1 can be expressed as

$$[\check{\mathcal{A}}(V_1), \check{\mathcal{A}}(V_2), \dots, \check{\mathcal{A}}(V_k)] = \mathcal{V}_k(I_n \otimes T_k) + \beta_{k+1}[O, \dots, O, V_{k+1}], \quad (3.13)$$

where  $O \in \mathbb{R}^{m \times n}$  denotes the zero matrix.

The scheme of this paper is based on applying the global Lanczos method to  $\check{\mathcal{A}}$  (Algorithm 1). It is possible to develop an analogous scheme based on applying the global Lanczos bidiagonalization method to  $\mathcal{A}$  and  $\mathcal{A}^T$ . A global Lanczos bidiagonalization method is described by Toutounian and Karimi [24]. The latter approach would give a method closely related to the scheme in [6], which is based on the standard Lanczos bidiagonalization method (also referred to as Golub–Kahan bidiagonalization). We apply the global Lanczos method because it requires less computer storage than the global Lanczos bidiagonalization method for the same number of steps, since the latter generates two sets of F-orthonormal vectors while the former only determines one set. Since our aim is to develop a method suitable for large-scale problems, reducing the computer storage required is important.

**3.2. Gauss-type rules associated with the global Lanczos method.** We first show orthogonality of the polynomials  $p_{j-1}$  determined by (3.11). This provides the link between the global Lanczos method and Gauss quadrature.

**PROPOSITION 3.2.** *The polynomials  $p_{j-1}$ ,  $j = 1, 2, \dots$ , defined by (3.11) are orthonormal with respect to the inner product (3.7).*

*Proof.* The definition of the measure  $d\omega$  that determines the functional  $\mathcal{I}$  in (3.7) yields

$$[p_{j-1}, p_{i-1}] = \int p_{j-1}(t)p_{i-1}(t)d\omega(t) = b^T H p_{i-1}(H^T H) p_{j-1}(H^T H) H^T b.$$

Using the linearity of  $\check{\mathcal{A}}$  and the properties (2.3) of the Kronecker product, we obtain that for any polynomial  $p$ ,

$$\text{vec}(p(\check{\mathcal{A}})(G)) = p(H^T H)\text{vec}(G).$$

Moreover,  $\text{vec}(G) = H^T b$ . It follows that

$$\begin{aligned} [p_{j-1}, p_{i-1}] &= (\text{vec}(p_{j-1}(\check{\mathcal{A}})(G)))^T \text{vec}(p_{i-1}(\check{\mathcal{A}})(G)) \\ &= (\text{vec}(V_j))^T \text{vec}(V_i) = \langle V_j, V_i \rangle_F, \end{aligned}$$

where the last inequality is due to (2.5). The orthonormality of the polynomials  $p_{i-1}$  is a consequence of (3.10).  $\square$

It follows from Algorithm 1 or (3.13) that the  $V_j$ 's satisfy the three-term recursion formula

$$\check{\mathcal{A}}(V_j) = \alpha_j V_j + \beta_{j+1} V_{j+1} + \beta_j V_{j-1}, \quad j = 1, 2, \dots, k,$$

and by (3.11) the polynomials  $p_{j-1}$  satisfy the same recursions, with  $p_{-1}(t) = 0$ ,

$$t p_{j-1}(t) = \alpha_j p_{j-1}(t) + \beta_{j+1} p_j(t) + \beta_j p_{j-2}(t), \quad j = 1, 2, \dots, k.$$

These relations can be expressed as

$$\begin{aligned} [p_0(t), p_1(t), \dots, p_{k-1}(t)] T_k &= t[p_0(t), p_1(t), \dots, p_{k-1}(t)] \\ &\quad - \beta_{k+1}[0, \dots, 0, p_k(t)], \end{aligned}$$

which shows that the zeros of  $p_k$  are the eigenvalues of  $T_k$ . These zeros are the nodes of the  $k$ -point Gauss rule  $\mathcal{G}_k$  associated with the measure  $d\omega$ . This Gauss rule applied to the function  $f_\mu$  can be expressed as

$$\mathcal{G}_k f_\mu = \|G\|_F^2 e_1^T f_\mu(T_k) e_1. \quad (3.14)$$

Throughout this paper  $e_j = [0, \dots, 0, 1, 0, \dots, 0]^T$  denotes the  $j$ th axis vector. The simple form of  $f_\mu$  makes it possible to evaluate (3.14) without computing the spectral factorization of  $T_k$ . We will return to this below.

The relation between the standard Lanczos method applied to a symmetric matrix and Gauss quadrature has been exploited by Golub and Meurant [9, 10]; see also [4, 6]. The application of the relation between the global Lanczos method and Gauss quadrature to determine the regularization parameter in Tikhonov regularization generalizes the approach in [6] based on the standard Lanczos method.

The  $(k+1)$ -point Gauss–Radau quadrature rule  $\mathcal{R}_{k+1,0}$  associated with the measure  $d\omega$  and with a preassigned node  $t_0 = 0$  can be represented analogously to (3.14). Define the matrix

$$T_{k+1,0} = \begin{bmatrix} T_k & \beta_{k+1} e_{k+1} \\ \beta_{k+1} e_k^T & \tilde{t}_{k+1,k+1} \end{bmatrix} \in \mathbb{R}^{(k+1) \times (k+1)}$$

by appending a row and a column to the matrix (3.12) to obtain a symmetric matrix with the last subdiagonal entry defined by (3.13) and the last diagonal entry to be determined. We would like this entry to be such that the matrix  $T_{k+1,0}$  has the smallest eigenvalue zero. Note that since  $\beta_{k+1}$  is positive, the eigenvalues of  $T_k$  are in an open interval contained in the convex hull of the support of the measure  $d\omega$ . In particular, all eigenvalues of  $T_k$  are positive.

A simple way to determine  $\tilde{t}_{k+1,k+1}$  so that all eigenvalues of  $T_{k+1,0}$  are nonnegative and one eigenvalue vanishes is to compute the bidiagonal Cholesky factor of  $T_{k+1,0}$  (with  $\tilde{t}_{k+1,k+1}$  chosen large enough to make  $T_{k+1,0}$  positive definite)

$$C_{k+1,0} = \begin{bmatrix} \rho_1 & & & & & \\ \sigma_2 & \rho_2 & & & & \\ & \ddots & \ddots & & & \\ & & \sigma_k & \rho_k & & \\ & & & \sigma_{k+1} & \rho_{k+1} & \end{bmatrix}$$

and set  $\rho_{k+1} = 0$ . This determines the matrix  $C_{k+1,0}$  and the desired matrix  $T_{k+1,0}$  is given by

$$T_{k+1,0} = C_{k+1,0} C_{k+1,0}^T;$$

see, e.g., [4, Proposition 3.1] for a proof. Note that the entries  $\rho_1, \rho_2, \dots, \rho_k$  and  $\sigma_2, \sigma_3, \dots, \sigma_{k+1}$  of  $C_{k+1,0}$  are independent of  $\tilde{t}_{k+1,k+1}$ . The Gauss–Radau rule (3.8) can be evaluated according to

$$\mathcal{R}_{k+1,0} f_\mu = \|G\|_F^2 e_1^T f_\mu(T_{k+1,0}) e_1,$$

analogously to (3.14).

Introduce the functions

$$\phi_k^-(\mu) := \mathcal{G}_k f_\mu, \quad \phi_{k+1}^+(\mu) := \mathcal{R}_{k+1,0} f_\mu. \quad (3.15)$$

Then, in view of (3.9),

$$\phi_{k-1}^-(\mu) < \phi_k^-(\mu) < \phi(\mu) < \phi_{k+1}^+(\mu) < \phi_k^+(\mu). \quad (3.16)$$

Instead of computing  $\phi(\mu)$  for several values of  $\mu$ , which is expensive for large-scale problems, we evaluate the upper and lower bounds (3.15) for desired values of  $\mu$ . This is inexpensive when  $k$  is small. The computation of these bounds is discussed in the following subsection.

**3.3. Evaluation of  $\phi_k^-(\mu)$  and  $\phi_{k+1}^+(\mu)$ .** We outline the evaluation of the functions  $\phi_k^{\mp}(\mu)$  and the derivative of  $\phi_{k+1}^+(\mu)$ , which are needed when determining a suitable value of  $\mu$  by our zero-finder. We have

$$\phi_k^-(\mu) = \|G\|_F^2 e_1^T (T_k + \mu I)^{-2} e_1.$$

For each value of  $\mu > 0$ , we solve the system of equations  $(T_k + \mu I)z_\mu = e_1$  for  $z_\mu \in \mathbb{R}^k$ . Since the matrix  $T_k + \mu I$  is symmetric and positive definite when  $\mu > 0$ , we solve this system with the aid of Cholesky factorization. Then

$$\phi_k^-(\mu) = \|G\|_F^2 z_\mu^T z_\mu.$$

Since  $T_k + \mu I$  is tridiagonal, the evaluation of  $\phi_k^-(\mu)$  requires only  $\mathcal{O}(k)$  arithmetic floating point operations (flops) for each value of  $\mu$ .

We turn to the computation of  $\phi_{k+1}^+(\mu)$ . The evaluation can be carried out similarly as the computation of  $\phi_k^-(\mu)$  described above, with the matrix  $T_k$  replaced by  $T_{k+1,0}$ . However, since the Cholesky factor  $C_{k+1,0}$  of  $T_{k+1,0}$  is available, we outline an approach that uses this factor. We have

$$\phi_{k+1}^+(\mu) = \|G\|_F^2 e_1^T (C_{k+1,0} C_{k+1,0}^T + \mu I)^{-2} e_1. \quad (3.17)$$

Compute the solution of the least-squares problem

$$\min_{z \in \mathbb{R}^{k+1}} \left\| \begin{bmatrix} C_{k+1,0}^T \\ \mu^{1/2} I \end{bmatrix} z - \mu^{-1/2} e_{k+2} \right\|_2^2. \quad (3.18)$$

Denoting the solution by  $z_\mu$ , we have

$$\phi_{k+1}^+(\mu) = \|G\|_F^2 z_\mu^T z_\mu.$$

The structure of the matrix in (3.18) makes it possible to compute  $z_\mu$  by QR factorization of the matrix in only  $\mathcal{O}(k)$  flops for each value of  $\mu > 0$ .

Changing the right-hand side in (3.18) gives the least-squares problem

$$\min_{s \in \mathbb{R}^{k+1}} \left\| \begin{bmatrix} C_{k+1,0}^T \\ \mu^{1/2} I \end{bmatrix} s - \mu^{-1/2} \begin{bmatrix} 0 \\ z_\mu \end{bmatrix} \right\|_2^2,$$

whose solution,  $s_\mu$ , is used to evaluate the derivative

$$\frac{d}{d\mu} \phi_{k+1}^+(\mu) = -2 \|G\|_F^2 z_\mu^T s_\mu. \quad (3.19)$$

This also can be carried out in  $\mathcal{O}(k)$  flops for each value of  $\mu > 0$ .

**4. Computation of an approximate solution of specified norm.** We describe a method for determining the regularization parameter  $\mu$  for Tikhonov regularization (1.3). Specifically, we would like to determine  $\mu > 0$  so that the computed solution  $X_\mu$  satisfies

$$\min_{\|X\|_F=\Delta} \|B - \mathcal{A}(X)\|_F \quad (4.1)$$

with  $\Delta > 0$  defined by (2.10). The constraint imposes that the restored image should have the same “energy” as the target image. Since this section is closely related to [6, Section 3], we only outline our approach and refer to [6] for further details.

Assume that  $0 < \Delta < \|H^\dagger b\|_2$ . Then, by Proposition 3.1, the equation

$$\phi(\mu) = \Delta^2 \quad (4.2)$$

has a unique solution  $0 < \mu_\Delta < \infty$ , which determines the desired solution  $X_{\mu_\Delta}$ . However, this equation is expensive to solve for large-scale problems, i.e., when  $mn$  is large. Similarly to the approach in [6], we therefore instead solve

$$\phi_{k+1}^+(\mu) = \Delta^2. \quad (4.3)$$

**PROPOSITION 4.1.** *The function  $\phi_{k+1}^+(\mu)$  defined by (3.15) is strictly decreasing and convex for  $\mu > 0$ . Equation (4.3) has a unique solution  $\mu_{k+1}$  such that  $0 < \mu_{k+1} < \infty$  for any positive finite  $\Delta^2$ .*

*Proof.* The fact that  $\phi_{k+1}^+$  is strictly decreasing and convex follows from the representation (3.17), which also shows that  $\lim_{\mu \rightarrow \infty} \phi_{k+1}^+(\mu) = 0$ . Moreover, since all subdiagonal entries of  $T_{k+1,0}$  are nonvanishing, so do all subdiagonal entries of the Cholesky factor  $C_{k+1,0}$ . Therefore, by (3.17),  $\lim_{\mu \searrow 0} \phi_{k+1}^+(\mu) = \infty$ .  $\square$

Thus, whenever (4.2) has a unique bounded positive solution  $\mu_\Delta$ , then so does (4.3). We denote the solution of the latter by  $\mu_{k+1}$ . Since both  $\phi(\mu)$  and  $\phi^+(\mu)$  are decreasing convex functions, it follows from (3.16) that

$$\mu_\Delta < \mu_{k+1} < \mu_k.$$

We apply the monotonically and quadratically convergent zero-finder described by Golub and von Matt [11, equations (75)–(78)] to compute approximations  $\mu_{k+1}^{(j)}$ ,  $j = 1, 2, \dots$ , of the zero  $\mu_{k+1}$  of

$$h_{k+1}^+(\mu) := \phi_{k+1}^+(\mu) - \Delta^2. \quad (4.4)$$

This zero-finder requires repeated evaluation of the function (3.17) and its derivative (3.19).

We would like to compute a value of  $\mu$  such that

$$\eta^2 \Delta^2 \leq \phi(\mu) \leq \Delta^2, \quad (4.5)$$

where the constant  $0 \leq \eta \leq 1$  determines how close to  $\Delta$  we require the norm of the computed approximate solution of (4.1) to be. We therefore calculate a value of  $\mu$  such that

$$\eta^2 \Delta^2 \leq \phi_k^-(\mu), \quad \phi_{k+1}^+(\mu) \leq \Delta^2. \quad (4.6)$$

This value satisfies (4.5). Using an initial approximation  $\mu_{k+1}^{(0)}$  of the desired zero,  $\mu_{k+1}$ , of (4.4), we first seek to satisfy the left-hand side inequality of (4.6). It may not

be possible to satisfy this inequality for small values of  $k$ . For instance,  $\phi_k^-(\mu)$  may be negative for all  $\mu > 0$  when  $k$  is small. We increase  $k$  by one when the left-hand side inequality of (4.6) cannot be satisfied, until  $k$  is large enough so that we can satisfy this inequality; see [6] for details. Turning to the right-hand side inequality of (4.6), we use the zero-finder to determine  $\mu_k^{(p)}$  such that

$$\frac{1}{10}(\eta^2 - 1)\Delta^2 + \Delta^2 \leq \phi_{k+1}^+(\mu_k^{(p)}) \leq \Delta^2.$$

In order for  $\eta^2\Delta^2 \leq \phi_k^-(\mu_{k+1}^{(p)})$  to hold, we may have to increase  $k$  further; see [6]. Having computed  $\hat{\mu} := \mu_k^{(p)}$ , we determine the associated solution  $\hat{y}_k$  of  $(T_k + \hat{\mu}I)y = \|G\|_F e_1$  by solving the least-squares problem

$$\min_{y \in \mathbb{R}^{k+1}} \left\| \begin{bmatrix} C_k \\ \hat{\mu}^{1/2}I \end{bmatrix} y - \hat{\mu}^{-1/2}\|G\|_F e_{k+1} \right\|_2^2.$$

Finally, our approximate solution of (4.1) is determined by

$$X_{\hat{\mu},k} = \mathcal{V}_k(I \otimes \hat{y}_k).$$

**Algorithm 2. Outline of solution method for (4.1).**

1. Set  $k = 2$ ;  $k$  is the number of global Lanczos steps.  
Let  $V_1 := G/\|G\|_F$ ;
2. Determine the orthonormal basis  $\{V_j\}_{j=1}^{k+1}$ , the tridiagonal matrix  $T_k$ , and the last subdiagonal entry  $\beta_{k+1}$  of  $T_{k+1,0}$  using Algorithm 1.
3. Determine an approximation of the zero  $\hat{\mu} = \mu_k$  of the function  $\phi_{k+1}^+$  as described above. This may require  $k$  to be increased, in which case one returns to step 2.
4. Determine  $\hat{y}_k$  and  $X_{\hat{\mu},k}$  as described above.

**5. Numerical results.** This section presents some representative numerical experiments. All computations were carried out using the MATLAB environment on an Intel(R) Core (TM) 2 Duo CPU T5750 computer with 3 GB of RAM. The computations were carried out with approximately 15 decimal digits of accuracy. To secure that the global Lanczos block vectors  $V_j$  are numerically orthonormal, we reorthogonalize them with global QR factorization.

Let  $\hat{x} = \text{vec}(\hat{X})$  denote the error-free exact solution of the linear system of equations (1.1) and define

$$e = \text{vec}(E), \quad \hat{b} = \text{vec}(\hat{B}), \quad b = \text{vec}(B),$$

where  $\hat{B} = H_2\hat{X}H_1^T$  and  $B = \hat{B} + E$ . The error matrix  $E$  has normally distributed entries with zero mean and is normalized to correspond to a specific noise level

$$\nu = \frac{\|E\|_F}{\|\hat{B}\|_F}.$$

To determine the effectiveness of our approach, we evaluate the relative error

$$\frac{\|\hat{X} - X_k\|_F}{\|\hat{X}\|_F}$$

of the computed approximate solution  $X_k = X_{\hat{\mu},k}$  of equation (2.7) obtained with Algorithm 2. The first two examples are concerned with the solution of Fredholm integral equations of the first kind. Discretization gives matrices that are severely ill-conditioned. The last two examples discuss image restoration problems.

**Example 1.** Let the nonsymmetric matrices  $H_1$  and  $H_2$ , both of size  $1500 \times 1500$ , be determined by the MATLAB programs `baart` and `foxgood` in the Regularization Tools package by Hansen [13]. Specifically, we let  $H_1 = \text{baart}(1500)$  and  $H_2 = \text{foxgood}(1500)$ . The computed condition numbers of these matrices are  $\kappa(H_1) = 2 \times 10^{18}$  and  $\kappa(H_2) = 3 \times 10^{13}$ . Since  $\kappa(H) = \kappa(H_1)\kappa(H_2)$ , the matrix  $H$  is numerically singular. We determine the right-hand side matrix  $\widehat{B}$  so that the exact solution is the matrix  $\widehat{X}$  with all entries unity. Table 5.1 displays the computed regularization parameters and the relative error in approximate solutions determined by Algorithm 2 with  $\eta = 0.99$  for different noise levels, as well as the number of iterations required to satisfy the stopping criterion.

TABLE 5.1  
Results for Example 1.

| Noise level | Iterations | Regularization parameter | Relative error        |
|-------------|------------|--------------------------|-----------------------|
| 0.001       | 8          | $2.0993 \times 10^{-6}$  | $8.54 \times 10^{-2}$ |
| 0.01        | 8          | $2.9595 \times 10^{-6}$  | $8.79 \times 10^{-2}$ |

**Example 2.** In this example, we consider the Fredholm integral equation

$$\int_{\Omega} \int_{\Omega} K(x, y, s, t) f(s, t) ds dt = g(x, y), \quad (x, y) \in \Omega', \quad (5.1)$$

where  $\Omega = [0, \pi/2] \times [0, \pi/2]$  and  $\Omega' = [0, \pi] \times [0, \pi]$ . The kernel is given by

$$K(x, y, s, t) = k_1(x, s)k_1(y, t), \quad (x, y) \in \Omega', \quad (s, t) \in \Omega,$$

where

$$g(x, y) = g_1(x)g_1(y)$$

and

$$k_1(s, x) = \exp(\cos(x)), \quad g_1(s) = 2\sinh(s)/s.$$

We use the code `baart` from Regularization Tools [13] to discretize (5.1) by a Galerkin method with orthonormal box functions as test and trial functions to obtain  $H_1$  of size  $1500 \times 1500$ , and let  $H = H_1 \otimes H_1$ . From the output of the code `baart` we determine a scaled approximation  $\widehat{X} \in \mathbb{R}^{1500 \times 1500}$  of the exact solution  $f(t, s) = \sin(t)\sin(s)$ . Table 5.2 shows the computed regularization parameters and the relative error in approximate solutions determined by Algorithm 2 with  $\eta = 0.99$  and different noise levels, as well as the number of iterations required to satisfy the stopping criterion. Figure 5.1 displays the computed approximate solution  $X_8$  obtained when the noise level  $\nu$  of the available data (right-hand side) is 0.01.

TABLE 5.2  
Results for Example 2.

| Noise level | Iterations | Regularization parameter | Relative error        |
|-------------|------------|--------------------------|-----------------------|
| 0.001       | 8          | $9.4480 \times 10^{-7}$  | $4.57 \times 10^{-2}$ |
| 0.01        | 8          | $1.7235 \times 10^{-6}$  | $4.81 \times 10^{-2}$ |

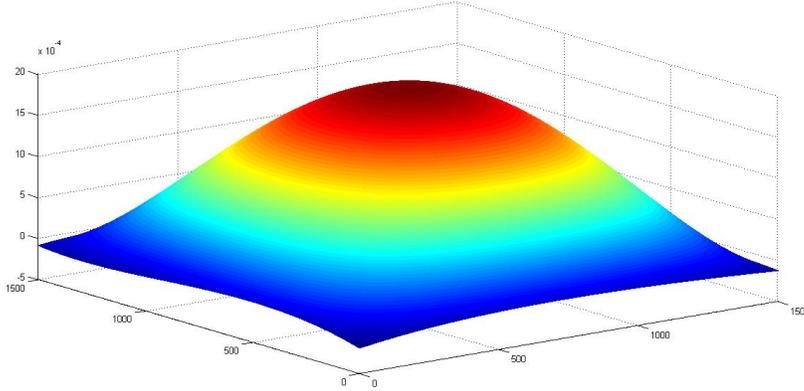


FIG. 5.1. *Example 2: Approximation  $X_8$  of  $\hat{X}$  determined by Algorithm 2 for noise level 0.01.*

**Example 3.** We consider an image restoration problem. Then the vector  $\hat{b}$  in (1.1) represents the blurred image,  $H$  is a blurring matrix,  $\hat{x}$  represents the exact image and  $e$  is a noise vector. In general, the blurring matrix  $H$  is determined by the point spread function (PSF) [14], which defines how each pixel is blurred, and the boundary conditions, which specify our assumptions on the scene just outside our image. We assume that the horizontal and vertical components of the PSF can be separated. Then  $H$  can be expressed as a Kronecker product  $H = H_1 \otimes H_2$ ; see [14]. The blurred image then is given by  $H_2 \hat{X} H_1^T$ . Also when  $H$  cannot be written as a tensor product of two matrices, it may be possible to approximate  $H$  quite accurately by such a product. The factors can be determined by solving the minimization problem

$$\{\hat{H}_1, \hat{H}_2\} = \arg \min_{H_1, H_2} \| H - H_1 \otimes H_2 \|_F;$$

see [25].

The original image in this example is the **peppers** image of dimension  $256 \times 256$  from MATLAB; it is shown in the left-hand side of Figure 5.2. The blurring matrix  $H$  is given by  $H = H_1 \otimes H_2 \in \mathbb{R}^{256^2 \times 256^2}$ , where  $H_1 = H_2 = [h_{ij}]$  are Toeplitz matrices of dimension  $256 \times 256$  with

$$h_{ij} = \begin{cases} \frac{1}{2r-1}, & |i-j| \leq r, \\ 0, & \text{otherwise.} \end{cases}$$

The blurring matrix  $H$  models uniform blur. In our example we set  $r = 5$ . The blurred and noisy image of Figure 5.2 is determined by adding noise of noise level  $\nu = 0.001$  to the blurred image  $H_2 \hat{X} H_1^T$ . Table 5.3 summarizes the results obtained by Algorithm 2 with  $\eta = 0.997$ . The table also reports results obtained with the method proposed in [6]. This method utilizes the connection between (standard) Golub–Kahan bidiagonalization and Gauss quadrature rules for solving large ill-conditioned linear systems of equations (1.2). We refer to this method as GKB. It determines the regularization parameter analogously to Algorithm 2, and uses a similar stopping criterion, but does not exploit the structure of  $H$ . Table 5.3 shows Algorithm 2 to require less CPU-time (in seconds) than the GKB method in [6]. The norm of the

error in the restored images determined by Algorithm 2 and the GKB method are about the same.

TABLE 5.3  
Results for Example 3.

| Method      | Iterations | $\mu$                 | Relative error        | CPU time (seconds) |
|-------------|------------|-----------------------|-----------------------|--------------------|
| GKB         | 110        | $2.65 \times 10^{-4}$ | $6.79 \times 10^{-2}$ | 13.29              |
| Algorithm 2 | 110        | $2.65 \times 10^{-4}$ | $6.79 \times 10^{-2}$ | 4.22               |



FIG. 5.2. Example 3: Original image (left), degraded image (center), and restored image (right) for noise of level 0.001.

**Example 4.** The original image of this example is the `iograyBorder` image of dimension  $256 \times 256$  from MATLAB; it is shown in the left-hand side of Figure 5.3. The blurring matrix is given by  $H = H_1 \otimes H_2 \in \mathbb{R}^{256^2 \times 256^2}$ , where  $H_1$  and  $H_2$  are the same matrices as in Example 3. In this example we set  $r = 5$ . The blurred and noisy image of Figure 5.3 is determined similarly as in Example 3; the noise level is 0.001. Table 5.4 summarizes the results obtained by Algorithm 2 with  $\eta = 0.997$ . The table also compares the number of iterations and the quality of the computed restored images obtained with Algorithm 2 and the GKB method. We can see that both methods yield restorations of the same quality and require about the same number of iterations.

TABLE 5.4  
Results for Example 4.

| Method      | Iterations | $\mu$                 | Relative error        |
|-------------|------------|-----------------------|-----------------------|
| GKB         | 76         | $2.89 \times 10^{-4}$ | $6.58 \times 10^{-2}$ |
| Algorithm 2 | 75         | $2.90 \times 10^{-4}$ | $6.58 \times 10^{-2}$ |

**6. Conclusion.** A new method for the inexpensive computation of a suitable value of the regularization parameter and an associated regularized solution of large linear discrete ill-posed problems is described. The method is based on global Lanczos tridiagonalization and applies Gauss-type quadrature rules to estimate pertinent quantities. Computed examples illustrate the effectiveness of the method.



FIG. 5.3. *Example 4: Original image (left), degraded image (center), and restored image (right) for noise of level 0.001.*

#### REFERENCES

- [1] G. F. AHMAD, D. H. BROOKS, AND R. S. MACLEOD, *An admissible solution approach to inverse electrocardiography*, Ann. Biomed. Eng., 26 (1998), pp. 278–292.
- [2] H. ANDREWS AND B. HUNT, *Digital Image Restoration*, Prentice-Hall, Engelwood Cliffs, 1977.
- [3] M. BERTERO AND P. BOCCACCI, *Introduction to Inverse Problems in Imaging*, IOP Publishing, London, 1998.
- [4] D. CALVETTI, G. H. GOLUB, AND L. REICHEL, *Estimation of the L-curve via Lanczos bidiagonalization*, BIT, 39 (1999), pp. 603–619.
- [5] D. CALVETTI, B. LEWIS, L. REICHEL, AND F. SGALLARI, *Tikhonov regularization with nonnegativity constraint*, Electron. Trans. Numer. Anal., 18 (2004), pp. 153–173.
- [6] D. CALVETTI AND L. REICHEL, *Tikhonov regularization with a solution constraint*, SIAM J. Sci. Comput., 26 (2004), pp. 224–239.
- [7] B. CHALMOND, *Modeling and Inverse Problems in Image Analysis*, Springer, New York, 2003.
- [8] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
- [9] G. H. GOLUB AND G. MEURANT, *Matrices, moments and quadrature*, in Numerical Analysis 1993, D. F. Griffiths and G. A. Watson, eds., Longman, Essex, 1994, pp. 105–156.
- [10] G. H. GOLUB AND G. MEURANT, *Matrices, Moments and Quadrature with Applications*, Princeton University Press, Princeton, 2010.
- [11] G. H. GOLUB AND U. VON MATT, *Quadratically constrained least squares and quadratic problems*, Numer. Math., 59 (1991), pp. 561–580.
- [12] C. W. GROETSCH, *The Theory of Tikhonov Regularization for Fredholm Integral Equations of the First Kind*, Pitman, Boston, 1984.
- [13] P. C. HANSEN, *Regularization tools version 4.0 for MATLAB 7.3*, Numer. Algorithms, 46 (2007), pp. 189–194.
- [14] P. C. HANSEN, J. G. NAGY, AND D. P. O’LEARY, *Deblurring Images: Matrices, Spectra, and Filtering*, SIAM, Philadelphia, 2006.
- [15] A. K. JAIN, *Fundamentals of Digital Image Processing*, Prentice-Hall, Engelwood Cliffs, 1989.
- [16] K. JBILOU, H. SADOK, AND A. TINZEFTE, *Oblique projection methods for linear systems with multiple right-hand sides*, Electron. Trans. Numer. Anal., 20 (2005), pp. 119–138.
- [17] J. KAMM AND J. G. NAGY, *Kronecker product and SVD approximations in image restoration*, Linear Algebra Appl., 284 (1998), pp. 177–192.
- [18] J. KAMM AND J. G. NAGY, *Kronecker product approximations for restoration image with reflexive boundary conditions*, SIAM J. Matrix Anal. Appl., 25 (2004), pp. 829–841.
- [19] S. KINDERMANN, *Convergence analysis of minimization-based noise level-free parameter choice rules for linear ill-posed problems*, Electron. Trans. Numer. Anal., 38 (2011), pp. 233–257.
- [20] G. LOPEZ LAGOMASINO, L. REICHEL, AND L. WUNDERLICH, *Matrices, moments, and rational quadrature*, Linear Algebra Appl., 429 (2008), pp. 2540–2554.
- [21] L. REICHEL AND G. RODRIGUEZ, *Old and new parameter choice rules for discrete ill-posed problems*, Numer. Algorithms, 63 (2013), pp. 65–87.
- [22] M. ROJAS AND D. C. SORENSEN, *A trust-region approach to regularization of large-scale discrete forms of ill-posed problems*, SIAM J. Sci. Comput., 23 (2002), pp. 1842–1860.
- [23] M. ROJAS AND T. STEIHAUG, *An interior-point trust-region-based method for large-scale non-negative regularization*, Inverse Problems, 18 (2002), pp. 1291–1307.

- [24] F. TOUTOUNIAN AND S. KARIMI, *Global least squares method (Gl-LSQR) for solving general linear systems with several right-hand sides*, Appl. Math. Comput., 178 (2006), pp. 452–460.
- [25] C. F. VAN LOAN AND N. P. PITSIANIS, *Approximation with Kronecker products*, in Linear Algebra for Large Scale and Real Time Applications, M. S. Moonen and G. H. Golub, eds., Kluwer, Dordrecht, 1993, pp. 293–314.