Journal of Computational and Applied Mathematics 226 (2009) 92–102

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Arnoldi–Tikhonov regularization methods

Bryan Lewis [a], Lothar Reichel [b,*]

[a] Rocketcalc LLC, 100 W. Crain Ave., Kent, OH 44240, USA
[b] Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA

ARTICLE INFO

Article history:
Received 23 December 2007
Received in revised form 21 January 2008

Keywords:
Ill-posed problem
Inverse problem
Regularization
Arnoldi decomposition
Discrepancy principle

ABSTRACT

Tikhonov regularization for large-scale linear ill-posed problems is commonly implemented by determining a partial Lanczos bidiagonalization of the matrix of the given system of equations. This paper explores the possibility of instead computing a partial Arnoldi decomposition of the given matrix. Computed examples illustrate that this approach may require fewer matrix–vector product evaluations and, therefore, less arithmetic work. Moreover, the proposed range-restricted Arnoldi–Tikhonov regularization method does not require the adjoint matrix and, hence, is convenient to use for problems for which the adjoint is difficult to evaluate.

1. Introduction

This paper considers the solution of linear systems of equations

$$A\boldsymbol{x} = \boldsymbol{b}, \qquad A \in \mathbb{R}^{n\times n}, \quad \boldsymbol{x}, \boldsymbol{b} \in \mathbb{R}^n, \tag{1}$$

with a large square matrix of ill-determined rank. In particular, A is severely ill-conditioned and has many singular values of different orders of magnitude close to zero; some singular values may vanish.

Linear systems of equations with a matrix of ill-determined rank often are referred to as linear discrete ill-posed problems. They may be obtained by discretizing linear ill-posed problems, such as Fredholm integral equations of the first kind with a smooth kernel. This type of integral equations arises in science and engineering when one seeks to determine the cause (the solution) of an observed effect represented by the right-hand side $\boldsymbol{b}$ (the data). Since the entries of $\boldsymbol{b}$ are obtained through observation, they typically are contaminated by measurement error and sometimes also by discretization error. We denote the sum of these errors by $\boldsymbol{e} \in \mathbb{R}^n$. Let $\hat{\boldsymbol{b}} \in \mathbb{R}^n$ be the unavailable error-free right-hand side associated with $\boldsymbol{b}$, i.e.,

$$\boldsymbol{b} = \hat{\boldsymbol{b}} + \boldsymbol{e}. \tag{2}$$

We assume the linear system of equations with the unavailable error-free right-hand side,

$$A\boldsymbol{x} = \hat{\boldsymbol{b}}, \tag{3}$$

to be consistent. Let $\hat{\boldsymbol{x}}$ denote a desired solution of (3), e.g., the solution of minimal Euclidean norm. We seek to determine an approximation of $\hat{\boldsymbol{x}}$ by computing an approximate solution of the available linear system of Eq. (1). Due to the severe ill-conditioning of A and the error $\boldsymbol{e}$ in the right-hand side $\boldsymbol{b}$, straightforward solution of (1) typically does not yield a meaningful approximation of $\hat{\boldsymbol{x}}$.

* Corresponding author.
E-mail addresses: blewis@rocketcalc.com (B. Lewis), reichel@math.kent.edu (L. Reichel).

0377-0427/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2008.05.003

In order to be able to determine a meaningful approximation of $\hat{\boldsymbol{x}}$, one replaces the linear system (1) by a nearby system that is less sensitive to perturbations of the right-hand side, and considers the solution of the latter system an approximation of $\hat{\boldsymbol{x}}$. This replacement is commonly referred to as regularization. The most popular regularization methods are Tikhonov regularization and truncated iteration; see Engl et al. [13], Groetsch [16], Hanke [17], and Hansen [19] for discussions. Tikhonov regularization in its simplest form replaces the linear system of Eqs. (1) by the minimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}\left\{\|A\boldsymbol{x}-\boldsymbol{b}\|^2+\frac{1}{\mu}\|\boldsymbol{x}\|^2\right\}, \tag{4}$$

where $\mu>0$ is a regularization parameter. Throughout this paper $\|\cdot\|$ denotes the Euclidean vector norm or the associated induced matrix norm. The value of μ determines how sensitive the solution $\boldsymbol{x}_\mu$ of (4) is to the error $\boldsymbol{e}$ in $\boldsymbol{b}$ and how close $\boldsymbol{x}_\mu$ is to $\hat{\boldsymbol{x}}$.

The penalty term $\frac{1}{\mu}\|\boldsymbol{x}\|^2$ in (4) may be replaced by $\frac{1}{\mu}\|L\boldsymbol{x}\|^2$, where the matrix L is referred to as a regularization operator. Common choices of L are matrices related to finite difference operators or the identity matrix. The method of this paper requires L to be a square matrix. A variety of square regularization operators are described by Calvetti et al. [10], Hansen and Jensen [20], and Morigi et al. [22]. In order to keep our presentation simple, we discuss only regularization with the identity operator in the present paper, but note that the application of any of the square regularization operators in these references is quite straightforward.

This paper discusses numerical methods that are applicable when the norm of the error in the right-hand side,

$$\varepsilon=\|\boldsymbol{e}\|, \tag{5}$$

or a fairly accurate estimate thereof, is known. The regularization parameter is quite easy to determine by application of the discrepancy principle in this situation, and this simplifies comparison between different solution methods. However, we note that the decompositions described in this paper also can be applied when ε is not available.

Solution methods for Tikhonov minimization problems (4) of small size often first compute the singular value decomposition of A and then determine a suitable value of μ; see, e.g., Engl et al. [13] and Hansen [19]. The present paper is concerned with solution methods for large Tikhonov minimization problems. We propose to reduce the problem (4) to a problem of smaller size by application of a few, say $\ell\ll n$, steps of the Arnoldi process applied to A with initial vector $\boldsymbol{u}_1=A\boldsymbol{b}/\|A\boldsymbol{b}\|$. This yields the decomposition

$$AU_\ell=U_{\ell+1}\bar{H}_\ell, \tag{6}$$

where $U_{\ell+1}=[\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_\ell,\boldsymbol{u}_{\ell+1}]\in\mathbb{R}^{n\times(\ell+1)}$ has orthonormal columns, which span the Krylov subspace

$$\mathbb{K}_\ell(A,A\boldsymbol{b})=\operatorname{span}\{A\boldsymbol{b},A^2\boldsymbol{b},\ldots,A^\ell\boldsymbol{b}\}. \tag{7}$$

The matrix $U_\ell\in\mathbb{R}^{n\times\ell}$ consists of the first ℓ columns of $U_{\ell+1}$. We assume that ℓ is chosen sufficiently small so that $\bar{H}_\ell\in\mathbb{R}^{(\ell+1)\times\ell}$ is an upper Hessenberg matrix with nonvanishing subdiagonal entries. Then $\bar{H}_\ell$ is of rank ℓ. We refer to (6) as a range-restricted Arnoldi decomposition, because $\mathcal{R}(U_\ell)\subset\mathcal{R}(A)$. Here and elsewhere in this paper $\mathcal{R}(M)$ denotes the range of the matrix M and $\mathcal{N}(M)$ denotes its null space.

When the last subdiagonal entry of $\bar{H}_\ell$ vanishes, the Arnoldi process is said to break down. This is a rare event; see [2,4,5,26] for discussions on this situation.

We seek to determine an approximate solution $\boldsymbol{x}_{\mu,\ell}$ of (4) in the Krylov subspace (7). Substituting $\boldsymbol{x}=U_\ell\boldsymbol{y}$, $\boldsymbol{y}\in\mathbb{R}^\ell$, into (4) and using (6) yields the reduced minimization problem

$$\min_{\boldsymbol{y}\in\mathbb{R}^\ell}\left\{\|\bar{H}_\ell\boldsymbol{y}-U_{\ell+1}^{\mathrm{T}}\boldsymbol{b}\|^2+\frac{1}{\mu}\|\boldsymbol{y}\|^2\right\}, \tag{8}$$

whose solution we denote by $\boldsymbol{y}_{\mu,\ell}$. Then

$$\boldsymbol{x}_{\mu,\ell}=U_\ell\boldsymbol{y}_{\mu,\ell} \tag{9}$$

is an approximate solution of (4). We refer to this approach to determine an approximate solution of (4), and thereby also of (1), as the range-restricted Arnoldi–Tikhonov (RRAT) method.

The numerical solution of the reduced minimization problem (8) and the determination of the regularization parameter μ will be discussed in Section 2. Here we only note that the computation of the decomposition (6) requires the evaluation of $\ell+1$ matrix–vector products with the matrix A, see, e.g., Saad [27, Section 6.3] for several implementations of the Arnoldi process. Adequate solutions of large linear discrete ill-posed problems often can be found in Krylov subspaces (7) of fairly small dimension; see Section 3 for computed examples. In this situation the dominant computational work for determining an appropriate value of μ and computing an associated approximate solution $\boldsymbol{x}_{\mu,\ell}$ of (4) is the evaluation of the $\ell+1$ matrix–vector products required to compute the decomposition (6).

The range-restricted GMRES (RRGMRES) iterative method is a variant of the (standard) GMRES method designed for the solution of large linear discrete ill-posed problems; see [6,26]. It determines iterates in Krylov subspaces of the form (7); the ℓth approximate solution $\boldsymbol{x}_\ell$ satisfies

$$\|A\boldsymbol{x}_\ell-\boldsymbol{b}\|=\min_{\boldsymbol{x}\in\mathbb{K}_\ell(A,A\boldsymbol{b})}\|A\boldsymbol{x}-\boldsymbol{b}\|,\quad \ell=1,2,3,\ldots, \tag{10}$$

with $\boldsymbol{x}_0 = \boldsymbol{0}$. The iterate $\boldsymbol{x}_\ell$ is computed by solving a reduced minimization, which is obtained by substituting $\boldsymbol{x} = U_\ell \boldsymbol{y}$ into the right-hand side of (10) and using (6). We compare in Section 3 approximations of $\hat{\boldsymbol{x}}$ determined by RRAT with iterates in the same solution subspace computed by RRGMRES. The determination of approximations of $\hat{\boldsymbol{x}}$ by carrying out suitably few steps of RRGMRES is an example of regularization by truncated iteration; regularization is achieved by determining an approximate solution $\boldsymbol{x}_\ell$ of (1) in a Krylov subspace (7) of suitably small dimension ℓ; see, e.g., [7,13,17] for further details on this approach to regularization.

Many available methods for the solution of large-scale Tikhonov regularization problems (4) are based on the Lanczos bidiagonalization process instead of the Arnoldi process; see, e.g., Björck [3], Calvetti et al. [8,9], Golub and von Matt [14], Kilmer and O'Leary [21], O'Leary and Simmons [24], and the references therein. These methods carry out ℓ steps of Lanczos bidiagonalization of the matrix A to determine the decompositions

$$AV_\ell = U_{\ell+1}\bar{C}_\ell, \qquad A^{\mathrm{T}}U_\ell = V_\ell C_\ell^{\mathrm{T}}, \tag{11}$$

for a suitable $\ell > 0$. Here the matrices $U_{\ell+1} \in \mathbb{R}^{n\times(\ell+1)}$ and $V_\ell \in \mathbb{R}^{n\times\ell}$ have orthonormal columns, $U_{\ell+1}\boldsymbol{e}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$, $U_\ell \in \mathbb{R}^{n\times\ell}$ consists of the ℓ first columns of $U_{\ell+1}$, $\bar{C}_\ell \in \mathbb{R}^{(\ell+1)\times\ell}$ is lower bidiagonal, and $\boldsymbol{e}_1 = [1, 0, \ldots, 0]^{\mathrm{T}}$ is the first axis vector. The columns of the matrix V_ℓ form a basis of the Krylov subspace

$$\mathbb{K}_\ell(A^{\mathrm{T}}A, A^{\mathrm{T}}\boldsymbol{b}) = \mathrm{span}\{A^{\mathrm{T}}\boldsymbol{b}, A^{\mathrm{T}}AA^{\mathrm{T}}\boldsymbol{b}, \ldots, (A^{\mathrm{T}}A)^{\ell-1}A^{\mathrm{T}}\boldsymbol{b}\}, \tag{12}$$

where we tacitly assume that this space is of dimension ℓ. An approximate solution of the Tikhonov minimization problem (4) is determined in this subspace. Specifically, one computes a solution of the form $\boldsymbol{x}_\ell = V_\ell \boldsymbol{y}_\ell$ by first substituting this expression and the left-hand side decomposition (11) into (4), and then solving the reduced least-squares problem so obtained. We refer to this kind of schemes as Lanczos bidiagonalization-Tikhonov (LBDT) methods.

Each Lanczos bidiagonalization step requires two matrix–vector product evaluations, one with A and one with A^{T}. Assuming that the space (12) is of dimension ℓ, the construction of an orthogonal basis therefore requires 2ℓ matrix–vector product evaluations.

Our interest in the RRAT method stems from the fact that for many linear discrete ill-posed problems the spaces (7) and (12) can be chosen to be of about the same dimension; see, e.g., Examples 3.1 and 3.2 of Section 3. Since the computation of an orthonormal basis for the space (7) requires about half the number of matrix–vector product evaluations than for the space (12) of the same dimension, and these evaluations typically constitute the dominant computational effort required by the methods, Tikhonov regularization based on the Arnoldi process can be cheaper to use than Tikhonov regularization based on Lanczos bidiagonalization. Moreover, the RRAT method is attractive for problems for which matrix–vector products with A are easier to evaluate than matrix–vector products with A^{T}. This situation arises, e.g., when solving large nonlinear problems by Krylov subspace methods; see [11] for a discussion. It also arises when matrix–vector products are evaluated by multi-pole methods. However, the purpose of the present paper is to compare the quality of the approximate solutions and the number of matrix–vector product evaluations required by RRAT, LBDT, and other methods. The computed examples of Section 3 therefore are chosen so that they allow application of all the methods.

The standard Arnoldi decomposition is of the same form as (6),

$$AU'_\ell = U'_{\ell+1}\bar{H}'_\ell \tag{13}$$

with $U'_{\ell+1} \in \mathbb{R}^{n\times(\ell+1)}$ having orthonormal columns, U'_ℓ consisting of the first ℓ columns of $U'_{\ell+1}$, and $\bar{H}'_\ell \in \mathbb{R}^{(\ell+1)\times\ell}$ of upper Hessenberg form. The decomposition (13) differs from the range-restricted Arnoldi decomposition (6) in that $U'_{\ell+1}\boldsymbol{e}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$. The matrix U'_ℓ satisfies

$$\mathcal{R}(U'_\ell) = \mathbb{K}_\ell(A, \boldsymbol{b}) = \mathrm{span}\{\boldsymbol{b}, A\boldsymbol{b}, \ldots, A^{\ell-1}\boldsymbol{b}\},$$

and, generally, $\mathcal{R}(U'_\ell) \not\subset \mathcal{R}(A)$. Thus, typically, the decomposition (13) is not range-restricted. This decomposition is the basis for the (standard) GMRES method; see [27, Section 6.5]. Substituting

$$\boldsymbol{x}'_{\mu,\ell} = U'_\ell \boldsymbol{y}'_{\mu,\ell} \tag{14}$$

into (4) and using the decomposition (13) yields a reduced Tikhonov minimization problem analogous to (8). With $\boldsymbol{y}'_{\mu,\ell}$ the solution of this reduced problem, the vector (14) gives an approximate solution of (4). We refer to this scheme as the Arnoldi–Tikhonov (AT) method; it has been described in [8]. For some problems the AT method works well, but for others the presence of the error-contaminated vector $\boldsymbol{b}$ in the solution subspace can cause a faster propagation of the error $\boldsymbol{e}$ into the computed approximate solution than for the RRAT method. This is illustrated in Section 3.

This paper is organized as follows. Section 2 discusses the determination of the regularization parameter and an approximate solution of (4) that satisfies the discrepancy principle, and Section 3 presents computed examples. Concluding remarks can be found in Section 4.

2. Arnoldi–Tikhonov regularization

The discrepancy $\boldsymbol{b} - A\boldsymbol{x}_\mu$ associated with the solution $\boldsymbol{x}_\mu$ of (4) can be used to determine a suitable value of the regularization parameter μ provided that an estimate ε of the norm of the error $\boldsymbol{e}$ in $\boldsymbol{b}$ is available; cf. (5). Let $\eta > 1$ be a constant, whose size reflects the uncertainty in the estimate ε; if the estimate is accurate, then η is chosen about one; otherwise a larger value of η is used. The *discrepancy principle* prescribes that the regularization parameter μ be chosen so that the discrepancy satisfies

$$\|\boldsymbol{b} - A\boldsymbol{x}_\mu\| = \eta\varepsilon. \tag{15}$$

It can be shown that for any fixed $\eta > 1$ and $\mu = \mu(\varepsilon)$ determined by (15), the solution $\boldsymbol{x}_\mu$ of (4) satisfies $\boldsymbol{x}_\mu \to \hat{\boldsymbol{x}}$ as $\varepsilon \searrow 0$; see, e.g., Engl et al. [13] or Groetsch [16] for proofs in Hilbert space settings.

Introduce the function

$$\phi(\mu) = \|\boldsymbol{b} - A\boldsymbol{x}_\mu\|^2. \tag{16}$$

The following properties of ϕ are shown in [9]. In the proposition and elsewhere I_j denotes the $j \times j$ identity matrix.

Proposition 2.1. *The function (16) allows the representation*

$$\phi(\mu) = \boldsymbol{b}^{\mathrm{T}}(\mu AA^{\mathrm{T}} + I_n)^{-2}\boldsymbol{b}. \tag{17}$$

Assume that $A^{\mathrm{T}}\boldsymbol{b} \neq \boldsymbol{0}$. Then ϕ is strictly decreasing and convex for $\mu \geq 0$. The equation

$$\phi(\mu) = \tau$$

has a unique solution μ_τ, such that $0 < \mu_\tau < \infty$, for any τ that satisfies $\|\boldsymbol{b}_0\|^2 < \tau < \|\boldsymbol{b}\|^2$, where $\boldsymbol{b}_0$ denotes the orthogonal projection of $\boldsymbol{b}$ onto $\mathcal{N}(A^{\mathrm{T}})$. In particular, if A is of full rank, then $\boldsymbol{b}_0 = \boldsymbol{0}$.

Proof. The solution $\boldsymbol{x}_\mu$ of (4) is given by

$$\boldsymbol{x}_\mu = \left(A^{\mathrm{T}}A + \frac{1}{\mu}I_n\right)^{-1} A^{\mathrm{T}}\boldsymbol{b}.$$

This can be seen from the normal equations associated with the minimization problem (4). Substituting this expression into (16) and using the identity

$$I_n - A(A^{\mathrm{T}}A + \mu^{-1}I_n)^{-1}A^{\mathrm{T}} = (\mu AA^{\mathrm{T}} + I_n)^{-1} \tag{18}$$

shows (17). □

Eq. (15) is equivalent to

$$\phi(\mu) = \eta^2\varepsilon^2. \tag{19}$$

We will assume that

$$\|\boldsymbol{b}_0\| < \eta\|\boldsymbol{e}\| < \|\boldsymbol{b}\|$$

holds, where $\boldsymbol{b}_0$ is defined in Proposition 2.1. Then it follows from the proposition that Eq. (19) has a unique positive bounded solution.

We turn to the computation of the solution of the reduced minimization problem (8) for a fixed positive value of μ. This minimization problem can be expressed in the form

$$\min_{\boldsymbol{y}\in\mathbb{R}^\ell} \left\| \begin{bmatrix} \mu^{1/2}\bar{H}_\ell \\ I_\ell \end{bmatrix} \boldsymbol{y} - \begin{bmatrix} \mu^{1/2}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix} \right\|. \tag{20}$$

We evaluate the solution $\boldsymbol{y}_{\mu,\ell}$ by applying a judiciously chosen sequence of Givens rotations to bring the matrix in (20) into upper triangular form, followed by back substitution. The computation of $\boldsymbol{y}_{\mu,\ell}$ in this manner requires $\mathcal{O}(\ell^3)$ arithmetic floating point operations. Moreover, it avoids the solution of the normal equations associated with the minimization problem (8). This is advantageous because the matrix of the normal equations, $\mu\bar{H}_\ell^{\mathrm{T}}\bar{H}_\ell + I_\ell$, has a larger condition number than the matrix in (20). Therefore the solution of the least-squares problem may yield computed solutions of higher accuracy than the solution of the normal equations; see, e.g., Golub and Wilkinson [15] for discussions. A related solution method, also based on the application of Givens rotations for the case when the matrix $\bar{H}_\ell$ is tridiagonal has been described by Eldén [12]. Having determined the solution $\boldsymbol{y}_{\mu,\ell}$ of (20), we obtain an associated approximate solution of (4) from (9).

We determine the regularization parameter μ by application of the discrepancy principle to the discrepancy $\boldsymbol{b} - A\boldsymbol{x}_{\mu,\ell}$ associated with $\boldsymbol{x}_{\mu,\ell}$, i.e., we choose μ so that

$$\|\boldsymbol{b} - A\boldsymbol{x}_{\mu,\ell}\| = \eta\varepsilon. \tag{21}$$

In order to see when this is possible, we introduce, analogously to (16), the function

$$\phi_\ell(\mu) = \|\boldsymbol{b} - A\boldsymbol{x}_{\mu,\ell}\|^2, \tag{22}$$

where $\boldsymbol{x}_{\mu,\ell}$ is given by (9) with U_ℓ determined by (6) and $\boldsymbol{y}_{\mu,\ell}$ the solution of (20). Throughout this section, we assume that $\ell < n$ is small enough so that the matrix $\bar{H}_\ell$ is of full rank. Thus, we would like to solve the equation

$$\phi_\ell(\mu) = \eta^2\varepsilon^2. \tag{23}$$

The following theorem discusses some properties of this equation.

Theorem 2.1. *The function (22) allows the representation*

$$\phi_\ell(\mu) = \boldsymbol{b}^{\mathrm{T}}U_{\ell+1}(\mu\bar{H}_\ell\bar{H}_\ell^{\mathrm{T}} + I_{\ell+1})^{-2}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b} + \|(I_n - U_{\ell+1}U_{\ell+1}^{\mathrm{T}})\boldsymbol{b}\|^2, \tag{24}$$

where the matrices $\bar{H}_\ell$ and $U_{\ell+1}$ are defined by (6). Assume that $A\boldsymbol{b} \neq \boldsymbol{0}$ and $\bar{H}_\ell^{\mathrm{T}}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b} \neq \boldsymbol{0}$. Then ϕ_ℓ is strictly decreasing and convex for $\mu \geq 0$ with $\phi_\ell(0) = \|\boldsymbol{b}\|^2$. Moreover, the equation

$$\phi_\ell(\mu) = \tau \tag{25}$$

has a unique solution $\mu_{\tau,\ell}$, such that $0 < \mu_{\tau,\ell} < \infty$, for any τ with

$$\|P_{\mathcal{N}(\bar{H}_\ell^{\mathrm{T}})}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b}\|^2 + \|(I_n - U_{\ell+1}U_{\ell+1}^{\mathrm{T}})\boldsymbol{b}\|^2 < \tau < \|\boldsymbol{b}\|^2, \tag{26}$$

where $P_{\mathcal{N}(\bar{H}_\ell^{\mathrm{T}})}$ denotes the orthogonal projector onto $\mathcal{N}(\bar{H}_\ell^{\mathrm{T}})$.

Proof. If $A\boldsymbol{b} = \boldsymbol{0}$, then we cannot start the Arnoldi process. We therefore have to rule out this case. Substituting (9) into (22) and using that $I_n - U_{\ell+1}U_{\ell+1}^{\mathrm{T}}$ is an orthogonal projector onto the complement of $\mathcal{R}(U_{\ell+1})$ yields

$$\phi_\ell(\mu) = \|U_{\ell+1}^{\mathrm{T}}\boldsymbol{b} - \bar{H}_\ell\boldsymbol{y}_{\mu,\ell}\|^2 + \|(I_n - U_{\ell+1}U_{\ell+1}^{\mathrm{T}})\boldsymbol{b}\|^2. \tag{27}$$

The solution of the minimization problem (8) is given by

$$\boldsymbol{y}_{\mu,\ell} = \left(\bar{H}_\ell^{\mathrm{T}}\bar{H}_\ell + \frac{1}{\mu}I_\ell\right)^{-1}\bar{H}_\ell^{\mathrm{T}}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b},$$

and substitution into (27), using the identity (18) with A replaced by $\bar{H}_\ell$, yields (24).

The matrix $\bar{H}_\ell\bar{H}_\ell^{\mathrm{T}}$ is positive semidefinite. Therefore it follows from the representation (24) that ϕ_ℓ is convex and decreasing for $\mu \geq 0$ with

$$\phi_\ell(0) = \boldsymbol{b}^{\mathrm{T}}U_{\ell+1}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b} + \|(I_n - U_{\ell+1}U_{\ell+1}^{\mathrm{T}})\boldsymbol{b}\|^2 = \|\boldsymbol{b}\|^2.$$

In order to show that ϕ_ℓ is strictly decreasing, we have to make sure that the first term in (24) depends on μ. Since by assumption $\bar{H}_\ell$ is of full rank, the matrix $\bar{H}_\ell\bar{H}_\ell^{\mathrm{T}}$ has precisely one vanishing eigenvalue. Expansion of $(\mu\bar{H}_\ell\bar{H}_\ell^{\mathrm{T}} + I_{\ell+1})^{-1}$ into a geometric series for small $\mu > 0$ shows that we have to require that $U_{\ell+1}^{\mathrm{T}}\boldsymbol{b} \notin \mathcal{N}(\bar{H}_\ell^{\mathrm{T}})$.

Substituting the spectral factorization of $\bar{H}_\ell\bar{H}_\ell^{\mathrm{T}}$ into the first term of (24) yields

$$\lim_{\mu\to\infty}\phi_\ell(\mu) = \|P_{\mathcal{N}(\bar{H}_\ell^{\mathrm{T}})}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b}\|^2 + \|(I_n - U_{\ell+1}U_{\ell+1}^{\mathrm{T}})\boldsymbol{b}\|^2. \tag{28}$$

This establishes the lower bound in (26) and completes the proof. □

The condition $U_{\ell+1}^{\mathrm{T}}\boldsymbol{b} \notin \mathcal{N}(\bar{H}_\ell)$ is satisfied for most linear discrete ill-posed problems. However, there are combinations of A and $\boldsymbol{b}$ for which the condition is violated and the RRAT method cannot be applied.

Example 2.1. Let $A \in \mathbb{R}^{n\times n}$ be the circulant downshift matrix, i.e., the subdiagonal entries of A as well as the $(1, n)$-entry are one and all other entries are zero. Let $\boldsymbol{b} = \boldsymbol{e}_1$. Then the matrix U_ℓ in the Arnoldi decomposition (6) consists of the first ℓ columns of A for any $1 \leq \ell < n$. Hence, $U_\ell^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{0}$. □

We can establish a connection between the RRAT and RRGMRES methods. An analogous connection holds between the AT and GMRES methods.

Corollary 2.1. *Let $\boldsymbol{x}_\ell$ be the ℓth iterate determined by RRGMRES applied to (1) with initial iterate $\boldsymbol{x}_0 = \boldsymbol{0}$. Then*

$$\|A\boldsymbol{x}_\ell - \boldsymbol{b}\|^2 = \phi_\ell(\infty).$$

Proof. Substituting the decomposition (6) into the minimization problem (10) yields

$$\begin{aligned}\|A\boldsymbol{x}_\ell - \boldsymbol{b}\|^2 &= \min_{\boldsymbol{y}\in\mathbb{R}^\ell} \|U_{\ell+1}\bar{H}_\ell \boldsymbol{y} - \boldsymbol{b}\|^2 \\ &= \min_{\boldsymbol{y}\in\mathbb{R}^\ell} \|\bar{H}_\ell \boldsymbol{y} - U_{\ell+1}^{\mathrm{T}}\boldsymbol{b}\|^2 + \|(I_n - U_{\ell+1}U_{\ell+1}^{\mathrm{T}})\boldsymbol{b}\|^2. \end{aligned} \tag{29}$$

Letting $\mu \to \infty$ in the reduced Tikhonov minimization problem (8) and comparing with (29) shows that

$$\boldsymbol{x}_\ell = \lim_{\mu\to\infty} \boldsymbol{x}_{\mu,\ell},$$

where $\boldsymbol{x}_{\mu,\ell}$ is defined by (9). The corollary now follows from (22). □

Assume that $\tau = \eta^2\varepsilon^2$ satisfies the bounds (26). Then Eq. (23) has a unique solution $0 < \mu_\ell < \infty$. We determine μ_ℓ with a zero-finder, such as Newton's method or the cubically convergent zero-finder described in [25]. Newton's method requires that values of ϕ_ℓ and its first derivative with respect to μ be computed at approximations $\mu_\ell^{(j)}$ of μ_ℓ for $j = 0, 1, 2, \ldots$; the cubically convergent zero-finder in [25] in addition demands the values of the second derivative at these points.

We first discuss the evaluation of $\phi_\ell(\mu)$. The rightmost term in (24) can be computed as

$$\gamma_\ell = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} - \sum_{j=1}^{\ell+1} (\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{b})^2. \tag{30}$$

In order to evaluate the other term in (24), introduce

$$\boldsymbol{z}_{\mu,\ell} = (\mu\bar{H}_\ell\bar{H}_\ell^{\mathrm{T}} + I_{\ell+1})^{-1}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b}.$$

Then

$$\phi_\ell(\mu) = \boldsymbol{z}_{\mu,\ell}^{\mathrm{T}}\boldsymbol{z}_{\mu,\ell} + \gamma_\ell.$$

We compute the vector $\boldsymbol{z}_{\mu,\ell}$ by solving the least-squares problem

$$\min_{\boldsymbol{z}\in\mathbb{R}^{\ell+1}} \left\| \begin{bmatrix} \mu^{1/2}\bar{H}_\ell^{\mathrm{T}} \\ I_{\ell+1} \end{bmatrix} \boldsymbol{z} - \begin{bmatrix} \boldsymbol{0} \\ U_{\ell+1}^{\mathrm{T}}\boldsymbol{b} \end{bmatrix} \right\|. \tag{31}$$

This requires $\mathcal{O}(\ell^3)$ arithmetic floating point operations for each value of μ, similarly as the solution of the least-squares problem (20). The evaluation of γ_ℓ is independent of μ.

The derivative-values can be determined similarly. We have

$$\phi_\ell'(\mu) = -2\,\boldsymbol{z}_{\mu,\ell}^{\mathrm{T}}\boldsymbol{w}_{\mu,\ell}, \qquad \phi_\ell''(\mu) = 6\,\boldsymbol{w}_{\mu,\ell}^{\mathrm{T}}\boldsymbol{w}_{\mu,\ell},$$

where

$$\boldsymbol{w}_{\mu,\ell} = (\mu\bar{H}_\ell\bar{H}_\ell^{\mathrm{T}} + I_{\ell+1})^{-1}\bar{H}_\ell\bar{H}_\ell^{\mathrm{T}}\boldsymbol{z}_{\mu,\ell}.$$

Hence, we may compute $\boldsymbol{w}_{\mu,\ell}$ by solving a least-squares problem analogous to (31) with the vector $U_{\ell+1}^{\mathrm{T}}\boldsymbol{b}$ replaced by $\bar{H}_\ell\bar{H}_\ell^{\mathrm{T}}\boldsymbol{z}_{\mu,\ell}$. Another approach to the evaluation of the derivatives of ϕ_ℓ, which is somewhat cheaper when ℓ is large, is described in [25].

We would like to avoid that approximations $\mu_\ell^{(j)}$ of μ_ℓ determined by the zero-finder are significantly larger than μ_ℓ, because the condition number of the matrix in (31) grows monotonically with μ. Since typically no positive lower bound for the solution μ_ℓ of (23) is available, we use the initial value $\mu_\ell^{(0)} = 0$ in the numerical examples of Section 3 and note that

$$\phi_\ell(0) = \|\boldsymbol{b}\|^2, \qquad \phi_\ell'(0) = -2\,\|\bar{H}_\ell^{\mathrm{T}}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b}\|^2, \qquad \phi_\ell''(0) = 6\,\|\bar{H}_\ell\bar{H}_\ell^{\mathrm{T}}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b}\|^2.$$

Corollary 2.2. *Assume that the range-restricted Arnoldi process (6) breaks down at step k. Then the sequence* $\{s_\ell\}_{\ell=0}^k$ *defined by*

$$\begin{aligned} s_0 &= \|\boldsymbol{b}\|^2, \\ s_\ell &= \|P_{\mathcal{N}(\bar{H}_\ell^{\mathrm{T}})}U_{\ell+1}^{\mathrm{T}}\boldsymbol{b}\|^2 + \|(I_n - U_{\ell+1}U_{\ell+1}^{\mathrm{T}})\boldsymbol{b}\|^2, \quad 0 < \ell < k, \\ s_k &= 0, \end{aligned} \tag{32}$$

is decreasing.

Proof. It follows from the proof of Theorem 2.1, and specifically from (28), that $s_\ell = \phi_\ell(\infty)$ for $1 \le \ell < k$. In view of Corollary 2.1, the s_ℓ are the square of the norm of the discrepancy associated with the ℓth iterate $\boldsymbol{x}_\ell$ determined by RRGMRES applied to (1) with initial iterate $\boldsymbol{x}_0 = \boldsymbol{0}$. It follows from (10) and the nesting $\mathbb{K}_{\ell-1}(A, A\boldsymbol{b}) \subset \mathbb{K}_\ell(A, A\boldsymbol{b})$ that $s_{\ell-1} \ge s_\ell$ for $1 \le \ell < k$, where we define $\mathbb{K}_0(A, A\boldsymbol{b}) = \{\boldsymbol{0}\}$.

When A is nonsingular, it follows from [2, Theorem 2.3] that the discrepancy vanishes at breakdown. For A singular, this may not be the case. □

Corollary 2.2 allows us to discuss the selection of ℓ in (8). When determining the regularization parameter by solving (23), the minimum number of Arnoldi steps that allows a solution is the smallest index ℓ, such that $s_\ell < \eta^2\varepsilon^2$. To simplify the computations, we ignore the first term in (32) and let $\ell_{\min}$ be the smallest integer such that

$$\gamma_\ell < \eta^2\varepsilon^2, \tag{33}$$

where γ_ℓ is defined by (30). We refer to $\ell_{\min}$ as the smallest number of Arnoldi steps. Note that γ_ℓ can be evaluate quite inexpensively; when increasing ℓ by one the only vector-operation required is the evaluation of one inner product. We remark that in all our computations, Eq. (23) had a solution for $\ell = \ell_{\min}$. However, when this is not the case, we increase $\ell_{\min}$.

Numerical experience indicates that for some linear discrete ill-posed problems the quality of the computed solution may improve by choosing the number of Arnoldi steps ℓ somewhat larger than $\ell_{\min}$. The following algorithm summarizes the computations. The input parameter ℓ_+ specifies the number of Arnoldi steps to be carried out in addition to $\ell_{\min}$.

Algorithm 1 (*Range-Restricted Arnoldi–Tikhonov Algorithm*).

Input: $\boldsymbol{b} \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$, ε, η, ℓ_+;

Output: Approximate solution $\boldsymbol{x}_{\mu,\ell}$, regularization parameter μ_ℓ, total number of Arnoldi steps $\ell = \ell_{\min} + \ell_+$;

1. Compute the decomposition (6) with $\ell = \ell_{\min} + \ell_+$, where $\ell_{\min}$ is the smallest value of ℓ, such that (33) holds.

2. Solve Eq. (23) for μ_ℓ. If there is no solution because we have ignored the first term in (32), then increase $\ell_{\min}$ by one and go to 2.

3. Compute the approximate Tikhonov solution $\boldsymbol{x}_{\mu,\ell}$ determined by (9) and (20). □

3. Computed examples

We compare the RRAT method implemented by Algorithm 1 to the Tikhonov regularization methods described in [8,9] and to RRGMRES when applied to three widely studied linear discrete ill-posed problems. The matrices in all examples are of ill-determined rank. We determine a regularization parameter μ that satisfies the discrepancy principle (15) in the various Tikhonov methods. Iterations with RRGMRES are terminated as soon as the discrepancy (10) falls below $\eta\varepsilon$, i.e., we regularize by truncated iteration. We also report results when $\ell_+ > 0$ additional iterations are carried out.

The evaluation of matrix–vector products with the matrix A in (1) constitutes the bulk of the computational work in each example. We compare the relative computational efficiency of the methods by counting the number of matrix–vector product evaluations required to compute each solution. The desired solution $\hat{\boldsymbol{x}}$ of the error-free system (3) is known for each example. Let $\tilde{\boldsymbol{x}}$ denote the solution determined by one of the Tikhonov methods or RRGMRES. We use the relative error norm, $\|\tilde{\boldsymbol{x}} - \hat{\boldsymbol{x}}\|/\|\hat{\boldsymbol{x}}\|$, to compare the quality of the approximate solutions determined by the methods. The examples of this section show that RRAT can out-perform other solution methods both with respect to quality and computational efficiency.

All computations are carried out with GNU/Octave or Matlab with about 16 significant decimal digits. We display two significant digits in the tabulated results. The norm of the error ε in the first two examples is obtained by selecting a relative error $\varepsilon/\|\hat{\boldsymbol{b}}\|$ of 1%, which corresponds to a signal-to-noise ratio of about 92 dB. The quantity ε is known accurately in every example, and we therefore set η to a value close to 1; specifically, we let $\eta = 1.01$.

Example 3.1. We consider the inversion of the Laplace transform

$$\int_0^\infty \exp(-\sigma\tau)\hat{x}(\sigma)\mathrm{d}\sigma = \hat{b}(\tau), \quad \tau \geq 0, \tag{34}$$

with right-hand side $\hat{b}$ and solution $\hat{x}$ given by

$$\hat{b}(\tau) = \frac{1}{\tau + 1/2}, \qquad \hat{x}(\tau) = \exp(-\tau/2).$$

We use the Matlab code ilaplace from [18] to discretize the integral equation (34) by a 100-point Gauss–Laguerre quadrature rule. The equations so obtained are required to be satisfied at the collocation points $\tau_j = j/10$, $1 \leq j \leq 100$. This determines the matrix $A \in \mathbb{R}^{100\times 100}$ and the entries $\hat{\boldsymbol{b}}_j = \hat{b}(\tau_j)$ of the right-hand side vector $\hat{\boldsymbol{b}} \in \mathbb{R}^{100}$ of (3). We compute the perturbed right-hand side $\boldsymbol{b}$ of (1) by adding an "error vector" $\boldsymbol{e} \in \mathbb{R}^{100}$ with normally distributed pseudorandom entries with zero mean and variance chosen so that $\varepsilon = 4.1 \times 10^{-2}$ to $\hat{\boldsymbol{b}}$; cf. (2). A summary of the computed results is shown in Table 1. Fig. 1 displays the computed approximate solutions produced by RRAT and RRGMRES for $\ell_+ = 1$, as well as the exact solution $\hat{\boldsymbol{x}}$.

Table 1 shows RRAT with $1 \leq \ell_+ \leq 2$ in Algorithm 1 to give much better approximations of $\hat{\boldsymbol{x}}$ than any of the other methods in this comparison. This also is illustrated by Fig. 1.

Fig. 2 shows the effect of the regularization parameter μ on the relative error in approximate solutions determined by RRAT implemented by Algorithm 1 with $\ell_+ = 1$. The figure also displays the relative error of the RRGMRES solution in the same Krylov subspace and the orthogonal projection of the solution of the noise-free exact solution $\hat{\boldsymbol{x}}$ into this Krylov subspace. The latter is the best possible approximation of $\hat{\boldsymbol{x}}$ in the subspace. □

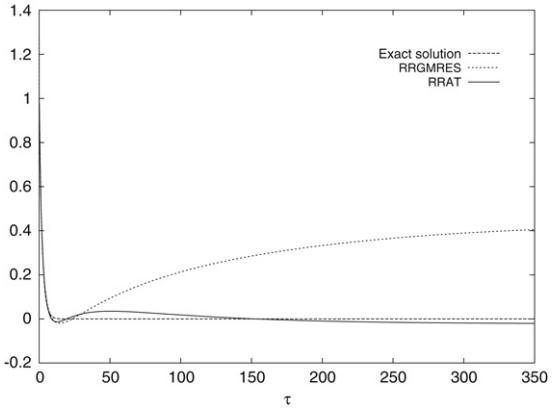

Fig. 1. Example 3.1: The approximate solutions $\boldsymbol{x}_{\mu,6} \in \mathbb{K}_6(A, A\boldsymbol{b})$ determined by RRAT with $\ell_+ = 1$ and $\boldsymbol{x}_6 \in \mathbb{K}_6(A, A\boldsymbol{b})$ determined by RRGMRES with $\ell_+ = 1$ are shown. The jth components of the vectors $\boldsymbol{x}_{\mu,6}$ and $\boldsymbol{x}_6$ approximate the desired solution $\hat{\boldsymbol{x}}$ at the jth zero of the Laguerre polynomial of degree 100. The vector $\hat{\boldsymbol{x}}$, referred to as the exact solution, also is displayed.

Table 1
Example 3.1: Computational work and quality of the approximate solutions determined by the RRAT and RRGMRES methods for several values of ℓ_+, the AT method in [8], and the LBDT method in [9] applied to the solution of a discretization of (34)

Method	ℓ_+	Matrix–vector Products	Relative error in computed approximate solution
RRGMRES	0	6	1.2×10^{0}
RRGMRES	1	7	9.3×10^{-1}
RRGMRES	2	8	1.0×10^{0}
RRGMRES	3	9	2.8×10^{0}
RRAT	0	6	9.1×10^{-1}
RRAT	1	7	8.7×10^{-2}
RRAT	2	8	9.2×10^{-2}
RRAT	3	9	1.4×10^{-1}
AT	0	5	8.7×10^{-1}
LBDT	0	14	1.3×10^{-1}

Table 2
Example 3.2: Comparison of the RRAT and RRGMRES methods for several values of ℓ_+, the AT method in [8], and the LBDT method in [9] applied to the solution of a discretization of (35)

Method	ℓ_+	Matrix–vector products	Relative error in computed approximate solution
RRGMRES	0	4	5.6×10^{-2}
RRGMRES	1	5	1.9×10^{-1}
RRGMRES	2	6	1.5×10^{0}
RRGMRES	3	7	1.9×10^{2}
RRAT	0	4	4.7×10^{-2}
RRAT	1	5	1.4×10^{-1}
RRAT	2	6	1.6×10^{-1}
RRAT	3	7	1.7×10^{-1}
AT	0	3	1.2×10^{-1}
LBDT	0	8	1.6×10^{-1}

Example 3.2. Consider the Fredholm integral equation of the first kind,

$$\int_0^{\pi/2} \kappa(\sigma, \tau)\hat{x}(\sigma)\mathrm{d}\sigma = \hat{b}(\tau), \quad 0 \le \tau \le \pi, \tag{35}$$

with $\kappa(\sigma, \tau) = \exp(\sigma \cos(\tau))$, $\hat{b}(\tau) = 2\sinh(\tau)/\tau$, and solution $\hat{x}(\tau) = \sin(\tau)$, which is discussed by Baart [1]. We use the Matlab code baart from [18] to discretize (35) by a Galerkin method with orthonormal box functions to produce the linear system (3) with $A \in \mathbb{R}^{200\times 200}$ and $\hat{\boldsymbol{b}} \in \mathbb{R}^{200}$. The perturbed right-hand side $\boldsymbol{b}$ in (1) is determined by adding an "error vector" $\boldsymbol{e} \in \mathbb{R}^{200}$ with normally distributed pseudorandom components with mean zero and variance such that $\varepsilon = 2.9 \times 10^{-2}$ to $\hat{\boldsymbol{b}}$.

Table 2 summarizes the computed results. Fig. 3 shows the effect of μ on the relative error norm of approximate solutions determined by RRAT with $\ell_+ = 0$. □

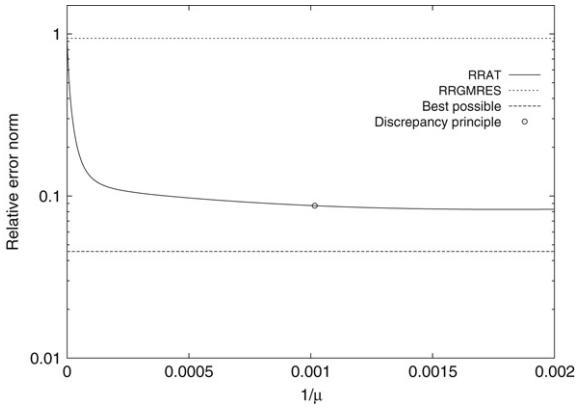

Fig. 2. Example 3.1: The effect of μ on approximate solutions $\boldsymbol{x}_{\mu,6} \in \mathbb{K}_6(A, A\boldsymbol{b})$ computed by RRAT with $\ell_+ = 1$. The discrepancy principle yields the computed solution marked by "o". The top straight line shows the relative error norm for the RRGMRES solution $\boldsymbol{x}_6 \in \mathbb{K}_6(A, A\boldsymbol{b})$ obtained with $\ell_+ = 1$. The bottom straight line shows the relative error of the best approximation of $\hat{\boldsymbol{x}}$ in $\mathbb{K}_6(A, A\boldsymbol{b})$.

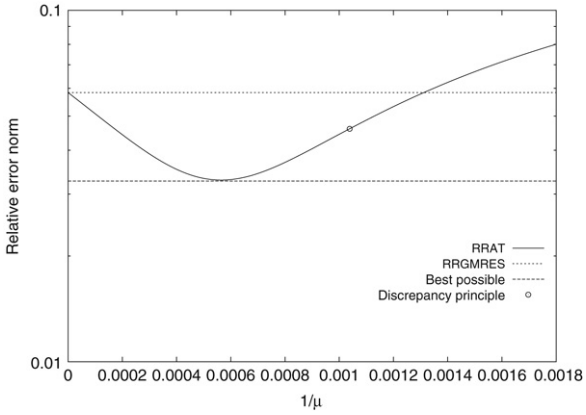

Fig. 3. Example 3.2: The effect of μ on the approximate solutions $\boldsymbol{x}_{\mu,3} \in \mathbb{K}_3(A, A\boldsymbol{b})$ computed by RRAT with $\ell_+ = 0$. The discrepancy principle yields the computed solution marked by "o". The top straight line shows the relative error norm for the RRGMRES solution $\boldsymbol{x}_3 \in \mathbb{K}_3(A, A\boldsymbol{b})$ obtained with $\ell_+ = 0$. The bottom straight line shows the relative error of the best approximation of $\hat{\boldsymbol{x}}$ in $\mathbb{K}_3(A, A\boldsymbol{b})$.

Table 3
Example 3.3: Summary of results for the image restoration problem. RRAT and RRGMRES are applied with $\ell_+ = 0$

Method	Matrix–vector products	Relative error in computed approximate solution
RRGMRES	11	4.0×10^{-1}
RRAT	11	4.1×10^{-1}
AT	6	6.1×10^{-1}
LBDT	60	4.1×10^{-1}

Example 3.3. We consider a two-dimensional image restoration problem from the RestoreTools Matlab package [23]. Our task is to deblur a 256×256-pixel image of a satellite degraded by spatially invariant blur and additive noise. This restoration problem was developed by the US Air Force Phillips Laboratory, Kirtland Air Force Base, New Mexico. The deblurring problem can be modeled by a linear system of Eq. (1) with $n = 2^{16}$. The components of the vectors $\hat{\boldsymbol{x}}$ and $\boldsymbol{b}$ are the lexicographically-ordered pixel values of the exact and distorted images, respectively. The matrix A is a discrete blurring operator representing convolution against a 256×256-pixel kernel referred to as a discrete point spread function. We efficiently compute matrix–vector products without explicitly forming A by using the discrete point spread function and the fast discrete Fourier transform. The "error vector" $\boldsymbol{e} \in \mathbb{R}^{2^{16}}$ has norm $\varepsilon = 3.3 \times 10^{-4}$, which corresponds to a signal-to-noise ratio of about 61dB. This vector models additive "noise" in the contaminated image represented by $\boldsymbol{b}$.

Table 3 compares RRAT, RRGMRES, the AT method in [8], and the LBDT method in [9]. Fig. 4 displays the noise- and blur-free image, the contaminated image, as well as restored images determined by the RRAT, RRGMRES, AT, and LBDT methods. Since the entries of the computed solutions $\tilde{\boldsymbol{x}}$ represent pixel values, they should be nonnegative. However, the methods in our comparison produce unconstrained approximate solutions of (1); in particular the vectors $\tilde{\boldsymbol{x}}$ may contain

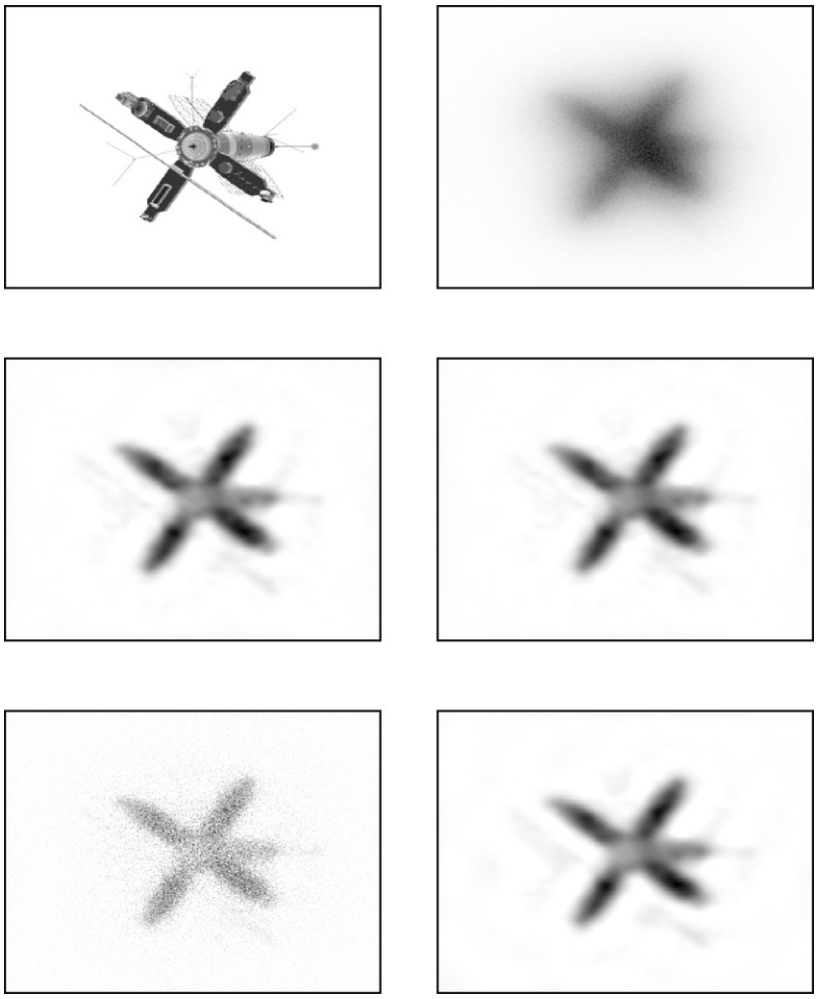

Fig. 4. Example 3.3: Clockwise from top left: Exact image represented by $\hat{\boldsymbol{x}}$, blurred and noisy image represented by $\boldsymbol{b}$, restored images determined by RRAT, LBDT, AT, and RRGMRES.

tiny negative entries. We set these entries to zero before display. Each displayed image is scaled for display purposes into an 8-bit gray-scale range.

Both RRGMRES and RRAT compute approximate solutions that are nearly optimal for this example in the sense that they are close to the orthogonal projection of the exact solution into the range-restricted subspace. The LBDT method computes an approximate solution of similar quality, but requires many more matrix–vector product evaluations. The AT method satisfies the discrepancy principle in a subspace of smaller dimension than RRAT, but produces an approximate solution of lower quality than the other methods. □

Our numerical experience with RRAT indicates that it often is beneficial and seldom detrimental to set $\ell_+ = 1$. Recall that RRAT gives the best approximations of $\hat{\boldsymbol{x}}$ in Examples 3.1 and 3.2 for $\ell_+ = 1$. In Example 3.3, RRAT determines computed solutions of about the same accuracy for $\ell_+ = 0$ and $\ell_+ = 1$; we therefore do not report the latter. RRGMRES on the other hand can yield much worse approximations of $\hat{\boldsymbol{x}}$ when $\ell_+ = 1$ than when $\ell_+ = 0$; cf. Example 3.2. Therefore RRGMRES generally should not be applied with $\ell_+ = 1$.

4. Conclusions

A new Tikhonov regularization method based on the range-restricted Arnoldi process is proposed. For some standard problems, the method compares favorably to Tikhonov regularization based on partial Lanczos bidiagonalization and to Tikhonov regularization based the standard Arnoldi process. The new method also can yield better approximations of $\hat{\boldsymbol{x}}$ than RRGMRES.

Acknowledgement

The second author was supported by an OBR Research Challenge Grant.

References

[1] M.L. Baart, The use of auto-correlation for pseudo-rank determination in noisy ill-conditioned least-squares problems, IMA J. Numer. Anal. 2 (1982) 241–247.
[2] J. Baglama, L. Reichel, Augmented GMRES-type methods, Numer. Linear Algebra Appl. 14 (2007) 337–350.
[3] Å. Björck, A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations, BIT 28 (1988) 659–670.
[4] P.N. Brown, H.F. Walker, GMRES on (nearly) singular systems, SIAM J. Matrix Anal. Appl. 18 (1997) 37–51.
[5] D. Calvetti, B. Lewis, L. Reichel, GMRES-type methods for inconsistent systems, Linear Algebra Appl. 316 (2000) 157–169.
[6] D. Calvetti, B. Lewis, L. Reichel, On the choice of subspace for iterative methods for linear discrete ill-posed problems, Int. J. Appl. Math. Comput. Sci. 11 (2001) 1069–1092.
[7] D. Calvetti, B. Lewis, L. Reichel, On the regularizing properties of the GMRES method, Numer. Math. 91 (2002) 605–625.
[8] D. Calvetti, S. Morigi, L. Reichel, F. Sgallari, Tikhonov regularization and the L-curve for large, discrete ill-posed problems, J. Comput. Appl. Math. 123 (2000) 423–446.
[9] D. Calvetti, L. Reichel, Tikhonov regularization of large linear problems, BIT 43 (2003) 263–283.
[10] D. Calvetti, L. Reichel, A. Shuibi, Invertible smoothing preconditioners for linear discrete ill-posed problems, Appl. Numer. Math. 54 (2005) 135–149.
[11] T.F. Chan, K.R. Jackson, Nonlinearly preconditioned Krylov subspace methods for discrete Newton algorithms, SIAM J. Sci. Statist. Comput. 5 (1984) 533–542.
[12] L. Eldén, Algorithms for the regularization of ill-conditioned least squares problems, BIT 17 (1977) 134–145.
[13] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Kluwer, Dordrecht, 1996.
[14] G.H. Golub, U. von Matt, Tikhonov regularization for large scale problems, in: G.H. Golub, S.H. Lui, F. Luk, R. Plemmons (Eds.), Workshop on Scientific Computing, Springer, New York, 1997, pp. 3–26.
[15] G.H. Golub, J.H. Wilkinson, Note on iterative refinement on least-squares solution, Numer. Math. 9 (1966) 139–148.
[16] C.W. Groetsch, The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind, Pitman, Boston, 1984.
[17] M. Hanke, Conjugate Gradient Type Methods for Ill-Posed Problems, Longman, Essex, 1995.
[18] P.C. Hansen, Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems, Numer. Algorithms 6 (1994) 1–35. Software is available in Netlib at http://www.netlib.org.
[19] P.C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems, SIAM, Philadelphia, 1997.
[20] P.C. Hansen, T.K. Jensen, Smoothing-norm preconditioning for regularizing minimum-residual methods, SIAM J. Matrix Anal. Appl. 29 (2006) 1–14.
[21] M.E. Kilmer, D.P. O'Leary, Choosing regularization parameters in iterative methods for ill-posed problems, SIAM J. Matrix Anal. Appl. 22 (2001) 1204–1221.
[22] S. Morigi, L. Reichel, F. Sgallari, Orthogonal projection regularization operators, Numer. Algorithms 44 (2007) 99–114.
[23] J.G. Nagy, K. Palmer, L. Perrone, Iterative methods for image deblurring: A Matlab object oriented approach, Numer. Algorithms 36 (2004) 73–93. Software is available at http://www.mathcs.emory.edu/~nagy/RestoreTools/index.html.
[24] D.P. O'Leary, J.A. Simmons, A bidiagonalization-regularization procedure for large-scale discretizations of ill-posed problems, SIAM J. Sci. Statist. Comput. 2 (1981) 474–489.
[25] L. Reichel, A. Shyshkov, A new zero-finder for Tikhonov regularization, BIT, in press.
[26] L. Reichel, Q. Ye, Breakdown-free GMRES for singular systems, SIAM J. Matrix Anal. Appl. 26 (2005) 1001–1021.
[27] Y. Saad, Iterative Methods for Sparse Linear Systems, 2nd edition, SIAM, Philadelphia, 2003.